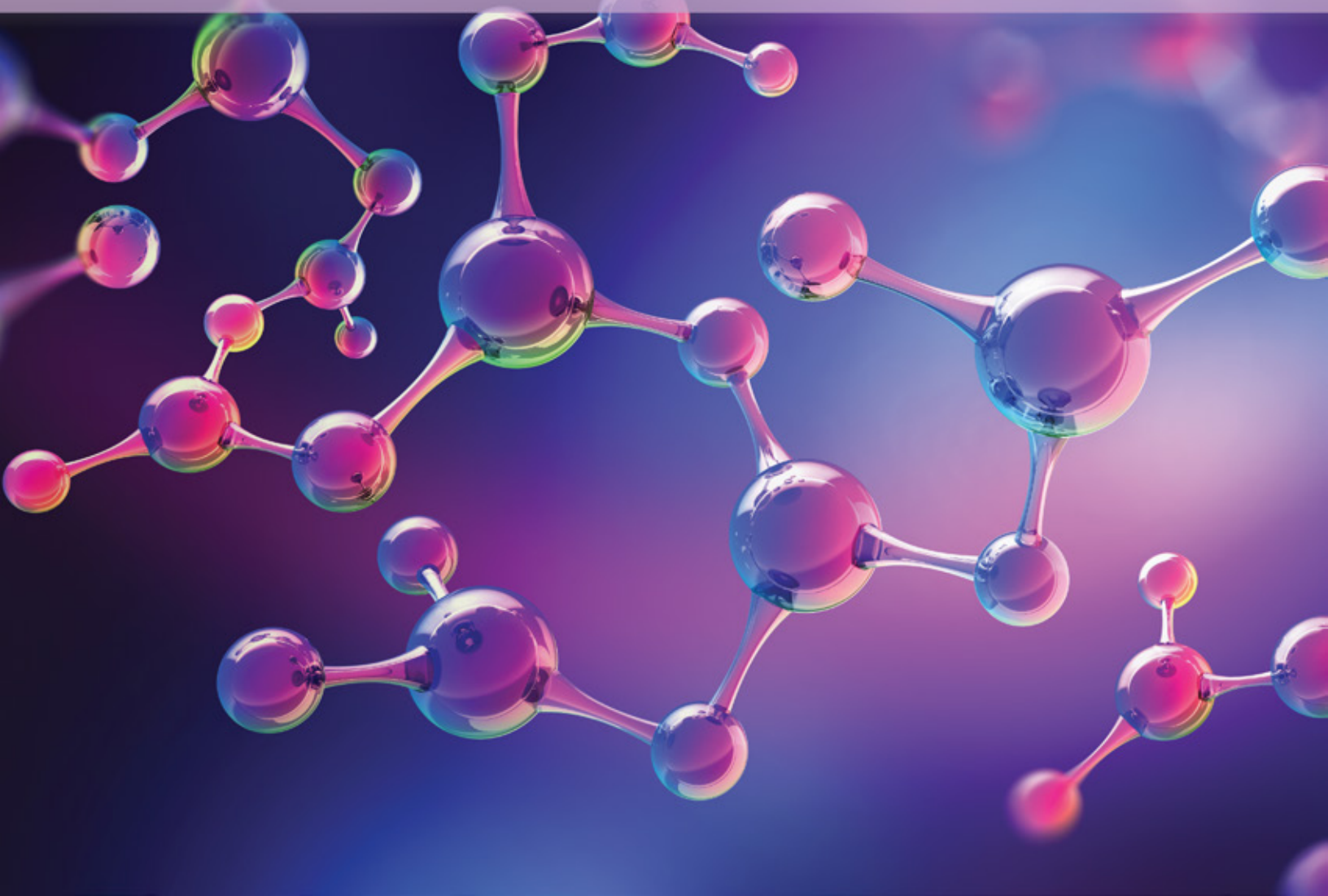
NINTH EDITION

CLINICAL CHEMISTRY

Principles, Techniques, and Correlations

Michael L. Bishop | Edward P. Fody
Carleen Van Siclen | James March Mistler | Michelle Moy



NINTH EDITION

CLINICAL CHEMISTRY

Principles, Techniques, and Correlations

Michael L. Bishop, MS, MLS(ASCP)
Adjunct Faculty, Medical Laboratory Scientist Program, Santa Fe College, Alachua FL
Retired Faculty, Medical Technology Program, Duke University Medical Center, Durham, North Carolina
Retired Program Director and University Department Chair, Medical Laboratory Science and Medical Laboratory Technician Program, Keiser University, Orlando, Florida

Edward P. Fody, MD
Clinical Professor, Department of Pathology, Microbiology and Immunology, Vanderbilt University School of Medicine, Nashville, Tennessee
Pathologist, Holland Hospital, Holland, Michigan

Carleen Van Siclen, MS,MLS(ASCP)^{CM},TS(ABB)
Assistant Professor of Laboratory Medicine and Pathology, MCSHS Medical Laboratory Science Program
Department of Laboratory Medicine and Pathology, Mayo Clinic, Jacksonville, Florida

James March Mistler, ABD, MS, MLS(ASCP)^{CM}
Program Director and Fulltime Lecturer, Department of Medical Laboratory Science, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts

Michelle Moy, MAdEd, MT(ASCP) SC
Assistant Professor of Medical Laboratory Science, Program Director, Biomedical Science, Department of Natural Science and Mathematics, College of Arts and Sciences, Madonna University, Livonia, Michigan

JONES & BARTLETT
LEARNING

World Headquarters
Jones & Bartlett Learning
25 Mall Road
Burlington, MA 01803
978-443-5000
info@jblearning.com
www.jblearning.com

Jones & Bartlett Learning books and products are available through most bookstores and online booksellers. To contact Jones & Bartlett Learning directly, call 800-832-0034, fax 978-443-8000, or visit our website, www.jblearning.com.

Substantial discounts on bulk quantities of Jones & Bartlett Learning publications are available to corporations, professional associations, and other qualified organizations. For details and specific discount information, contact the special sales department at Jones & Bartlett Learning via the above contact information or send an email to specialsales@jblearning.com.

23886-0

Production Credits

Vice President, Product Management: Marisa R. Urbano
Vice President, Content Strategy and Implementation: Christine Emerton
Director, Product Management: Matthew Kane
Product Manager: William Lawrensen
Content Strategist: Carol Brewer Guerrero
Content Coordinator: Samantha Gillespie
Director, Project Management and Content Services: Karen Scott
Manager, Project Management: Kristen Rogers
Project Manager: Daniel Stone
Senior Digital Project Specialist: Angela Dooley
Marketing Manager: Mark Adamiak
Composition: Exela Technologies
Project Management: Exela Technologies
Cover Design: Kristin Parker
Senior Media Development Editor: Troy Liston
Rights & Permissions Manager: John Rusk
Rights Specialist: Rebecca Damon
Rights Specialist: Liz Kincaid
Cover Image (Title Page, Part Opener, Chapter Opener): © Yurchanka Siarhei/Shutterstock
Printing and Binding: LSC Communications

Library of Congress Cataloging-in-Publication Data

Names: Bishop, Michael L., editor. | Fody, Edward P., editor. | Siclen, Carleen Van, editor. | Mistler, James March, editor. | Moy, Michelle, editor.
Title: Clinical chemistry: principles, techniques, and correlations / [edited by] Michael L. Bishop, Edward P. Fody, Carleen Van Siclen, James March Mistler, Michelle Moy.
Other titles: Clinical chemistry (Bishop)
Description: Ninth edition. | Burlington, MA : Jones & Bartlett Learning, [2023] | Includes bibliographical references and index.
Identifiers: LCCN 2021034082 | ISBN 9781284238860 (paperback)
Subjects: LCSH: Clinical chemistry.
Classification: LCC RB40 .C576 2023 | DDC 616.07/56–dc23
LC record available at https://lccn.loc.gov/2021034082

6048

Printed in the United States of America
26 25 24 23 22 10 9 8 7 6 5 4 3 2 1

Disclaimer

Care has been taken to confirm the accuracy of the information presented and to describe generally accepted practices. However, the authors, editors, and publisher are not responsible for errors or omissions or for any consequences from application of the information in this book and make no warranty, expressed or implied, with respect to the currency, completeness, or accuracy of the contents of the publication. Application of this information in a particular situation remains the professional responsibility of the practitioner; the clinical treatments described and recommended may not be considered absolute and universal recommendations.

The authors, editors, and publisher have exerted every effort to ensure that drug selection and dosage set forth in this text are in accordance with the current recommendations and practice at the time of publication. However, in view of ongoing research, changes in government regulations, and the constant flow of information relating to drug therapy and drug reactions, the reader is urged to check the package insert for each drug for any change in indications and dosage and for added warnings and precautions. This is particularly important when the recommended agent is a new or infrequently employed drug.

Some drugs and medical devices presented in this publication have Health Canada clearance for limited use in restricted research settings. It is the responsibility of the healthcare provider to ascertain the Health Canada status of each drug or device planned for use in his or her clinical practice.

Brief Contents

Contents

To all Clinical Laboratory practitioners, educators, and healthcare professionals for their previous and continuing extraordinary commitment, service, and professionalism during the COVID-19 pandemic.

MLB, EPF, CVS, JMM, MM

In memory of my mother and father, Betty Beck Bishop and William Stewart Bishop, Sr., for support, guidance, and encouragement.

To Sheila, Chris, and Carson for their support, patience, and inspiration.

MLB

To Nancy, my wife, for continuing support and dedication.

EPF

To Gary, my husband, for his support of my professional goals and to all the laboratory professionals, including my students, who have contributed to my knowledge and passion for lifelong learning.

CVS

To my husband, Keith, for everything.

JMM

To my college mentors: Pete Gebauer and Herb Miller, I thank you for believing in me.

In memory of my mother SG (1940–2021)

MM

Foreword

Many years ago, I wrote the Foreword to some earlier editions of this text. A ninth edition seems like an unbelievably long time until I reflect that this year is the 40th anniversary of the paper that introduced a multi-rule Shewhart control chart,[1] more commonly known as "Westgard Rules." That paper was written early in my career, but now in my retirement we have updated that approach to provide "Westgard Sigma Rules" in order to customize the QC design on the basis of the quality required by a test and the Sigma performance observed for a method.[2] Even well-established "standard" laboratory practices need periodic review and updating to keep current with the improvements in testing processes. Likewise, this 9th edition of the standard clinical chemistry text reflects the latest knowledge and improvements for laboratory science. That is a testament to the authors' commitment and dedication to providing an up-to-date knowledge base for the professionals in clinical laboratory science.

I am writing this on the one-year anniversary of the declaration of a global pandemic, a year during which over half a million Americans died of COVID-19. This pandemic has revealed the importance of laboratory testing for the health of the nation. Laboratory testing has often been viewed as a behind-the-scenes service in health care. During the pandemic, laboratory testing has been center stage as an essential service for assessing the state of disease, diagnosing those with infection, monitoring those under treatment, and monitoring the immunity and the health of the community.

Laboratory scientists were on the front line in introducing new diagnostic tests, validating their performance, and implementing testing in many diverse settings, including central laboratories, clinic laboratories, and point-of-care settings, including drive-through testing services. Understanding the performance of qualitative tests brought new importance to ideas such as clinical sensitivity, clinical specificity, and the predictive value of laboratory tests. That also meant new protocols for validating new tests to characterize test performance, including adaptations for the nature of molecular tests, such as the real-time Reverse Transcription Polymerase Chain Reaction (rRT-PCR) methods that were critical in the early diagnosis and management of patients. Antibody tests flooded the market and required care and attention by laboratories, especially during the early phases when the FDA exercised very limited control of the companies introducing the new tests. Antigen tests emerged later and more slowly but were critical for providing more widespread diagnostic testing. All in all, this short time period has provided the lessons of a lifetime and demonstrated the importance of what you will be learning in your studies.

This new edition of *Clinical Chemistry: Principles, Techniques, and Correlations* continues its mission of addressing the formal educational needs of students in clinical laboratory science, as well as the ongoing needs of professionals in the field. It facilitates the educational process by identifying the learning objectives, focusing on key concepts and ideas, and applying the theory through case studies. It covers the basics of laboratory testing, as well as many special areas of testing. And it is still possible to carry this text with you to class, to the laboratory, to the office, or home to study!

Having personally worked with some of the editors and contributors, I know they have high standards both in the laboratory and in the classroom. Their interests and background provide an excellent balance between the academic and the practical, ensuring that students are exposed to a well-developed base of knowledge that has been carefully refined by experience.

For the many students for whom this book is intended, let me offer some advice from my close friend and mentor, Hagar the Horrible. It seems his young Viking son was embarking on a voyage to the real world of work. Needing advice, he asked "How do I get to the top?" Hagar's response, "You have to start at the bottom and work your way up." After pondering this for a moment, his son then asked, "How do I get to the bottom?" Hagar replied, "You have to

know somebody." The people you need to know are the authors of this book, as well as the instructors in your courses and your bench teachers in the laboratory. You need to seek them out to profit from their learning and experiences. They are the professionals who know the state of laboratory practice, possess the current knowledge of the field, and are dedicated to helping you become a successful laboratory scientist.

—James O. Westgard,
Madison, WI

References

1. Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem* 1981;27:493–501.
2. Westgard JO, Westgard SA. Establishing evidence-based statistical quality control practices. *Am J Clin Pathol* 2019;151:364–370.

Preface

The events with the worldwide pandemic have placed an extraordinary burden on our healthcare system. Facing staffing, PPE, and diagnostic supply shortages, healthcare professionals stepped up with effort, critical process evaluation, and extraordinary dedication to provide quality patient care with compassion and empathy. Initially, the nightly news became a presentation of CDC guidelines, mask mandates, business shutdowns, travel restrictions, metrics, trends, positivity rates, and hospitalization and death statistics. Months later, the metrics related to more positive information—initial results of vaccine clinical trials, emergency use authorizations, vaccine shipments, and "shots in arms." Through it all, the healthcare system functioned as effectively as possible due to individual efforts and interdisciplinary teamwork. Healthcare professionals have improved communication with each other, as well as with the patient and their families. Collaborative efforts between healthcare disciplines are emerging across the patient care spectrum landscape.

Since the initial idea for this textbook was discussed in a meeting of the Biochemistry/Urinalysis section of ASMT (now ASCLS) in the late 1970s, the only constant has been change and the never wavering commitment of the clinical laboratory professionals. Now almost 45 years since the initiation of this effort, the editors have had the privilege of completing the ninth edition with another diverse team of dedicated clinical laboratory professionals. In this era of focusing on metrics, the editors would like to share the following information. The 401 contributions in the 9 editions and supporting material represent 115 clinical laboratory science education programs, 83 clinical laboratories, 28 medical device companies, 4 government agencies, and 3 professional societies representing 40 states and territories. One hundred and sixty-four contributors were clinical laboratory scientists with advanced degrees. These contributors have produced 289 chapters citing 12,054 references for a total of 5,708 pages that included 2,158 figures and 691 case studies. With today's global focus, the previous editions of the text have been translated into at least six languages. By definition, a profession is a calling requiring specialized knowledge and intensive academic preparation to define its scope of practice and produce its own literature. The clinical laboratory science profession has evolved significantly over these past four-and-a-half decades.

Clinical chemistry continues to be one of the most rapidly advancing areas of laboratory medicine. New technologies and analytical techniques have been introduced, with a dramatic impact on the practice of clinical chemistry and laboratory medicine. In addition, the healthcare system itself is rapidly changing. There is ever-increasing emphasis on improving the quality of patient care, individualized medicine, patient outcomes, financial responsibility, and total quality management. Now, more than ever, clinical laboratorians need to be concerned with disease correlations, result interpretations, problem solving, quality assurance, and cost-effectiveness. Laboratory professionals need to know not only the *how* of tests but more importantly be able to communicate the *what, why*, and *when* to the patient and the healthcare team. The editors of *Clinical Chemistry: Principles, Techniques, and Correlations* have designed the ninth edition to be an even more valuable resource to both students and practitioners.

The ninth edition of *Clinical Chemistry: Principles, Techniques, and Correlations* is comprehensive, up-to-date, and easy to understand for students at all entry levels. It is also intended to be a practically organized resource for both instructors and practitioners. The editors have tried to maintain the book's readability and further improve its content while rearranging content and focusing on the scaffolding provided by the ASCLS MLT and MLS Entry Level Curriculum and the ASCP BOC guidelines. Because clinical laboratorians use their interpretative and analytic skills in the practice of clinical chemistry, an effort has been made to maintain an appropriate balance between analytic principles, techniques, and the correlation of results with disease states.

In this edition, the editors have maintained features in response to requests from our readers,

students, instructors, and practitioners. Ancillary materials have been updated and expanded. Chapters now include current, more frequently encountered case studies modelled after the nursing PICOT initiative in a structured, unfolding style. To provide a thorough, up-to-date study of clinical chemistry, all chapters have been updated and reviewed by professionals who practice clinical chemistry and laboratory medicine on a daily basis. The basic principles of the analytic procedures discussed in the chapters reflect the most recent or commonly performed techniques in the clinical chemistry laboratory. Detailed procedures have been omitted because of the variety of equipment and commercial kits used in today's clinical laboratories. Instrument manuals and analyte package inserts are the most reliable reference for detailed instructions on current analytic procedures. All chapter material has been updated, improved, and rearranged for better continuity and readability.

The **Navigate 2 Advantage** digital access contains additional case studies, review questions, teaching resources, teaching tips, student laboratory procedures, and teaching aids for instructors and students; it is included with the purchase of this textbook and is also available for separate purchase from the publisher.

One last piece of advice to make you successful in the field of clinical laboratory science:

Work with compassion, empathy, and professionalism until you no longer have to introduce yourself.*

Michael L. Bishop
Edward P. Fody
Carleen Van Siclen
James March Mistler
Michelle Moy

*Modified from Harvey Specter in *Suits*.

New to This Edition

Medical laboratory science students need a strong foundation in applied chemistry to meet the requirements of certifying bodies and accreditation organizations that ensure students are prepared for employment.

This textbook provides clear explanations that balance analytic principles, techniques, and correlation of results with coverage of disease states, helping students develop interpretive and analytic skills for their future careers.

Updates to this edition include:

- Chapter content based on the ASCLS Entry Level Curriculum and current ASCP Content Guidelines
- Reorganization of chapter order to reflect clinical chemistry flow in most courses today
- Over 60 unique case studies that evolve throughout the chapters
- NEW Chapter 13: Basic Endocrinology
- NEW Chapter 24: Pregnancy and Prenatal Testing
- Reference range table is included as an Appendix in the printed book and online.

A map of how the textbook correlates to the ASCLS curriculum and ASCP guidelines is provided as an instructor resource.

Each chapter opens with a **Chapter Outline, Key Terms**, and **Chapter Objectives** that correlate to the ASCLS entry-level curriculum and current ASCP content guidelines.

CHAPTER 9

Carbohydrates

Vicki S. Freeman

CHAPTER OUTLINE

KEY TERMS

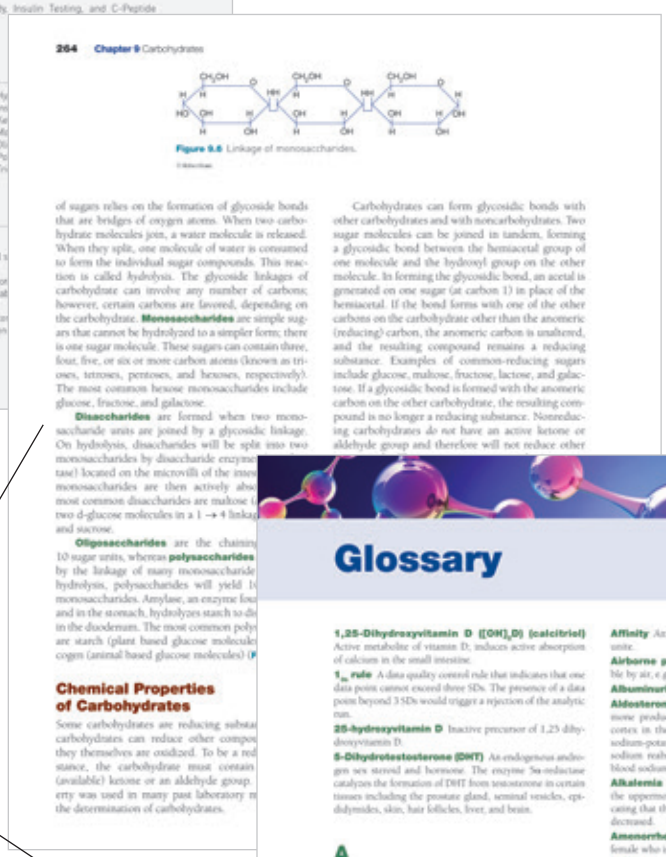
Albuminuria
Carbohydrates
Diabetes mellitus
Disaccharides
Embden-Meyerhof pathway
Glucagon
Gluconeogenesis
Glucose
Glycogen
Glycogenesis
Glycogenolysis
Glycolysis
Glycosylated hemoglobin
Hemoglobin A1c
Hyperglycemic

CHAPTER OBJECTIVES

At the end of this unit of study, the clinical laboratorian should be able to:

- Classify carbohydrates into their respective groups.
- Discuss the metabolism of carbohydrates in the body and the mode of action of hormones in carbohydrate metabolism.
- Differentiate the types of diabetes by clinical symptoms and laboratory findings according to the American Diabetes Association.
- Explain the clinical
- Relate expected labor to the following metab
  - Ketoacidosis
  - Hypersomolar co
- Distinguish between hypoglycemia

260



264 Chapter 9 Carbohydrates

Figure 9.6 Linkage of monosaccharides.

of sugars relies on the formation of glycoside bonds that are bridges of oxygen atoms. When two carbohydrate molecules join, a water molecule is released. When they split, one molecule of water is consumed to form the individual sugar compounds. This reaction is called *hydrolysis*. The glycoside linkages of carbohydrate can involve any number of carbons; however, certain carbons are favored, depending on the carbohydrate. **Monosaccharides** are simple sugars that cannot be hydrolyzed to a simpler form; there is one sugar molecule. These sugars can contain three, four, five, or six or more carbon atoms (known as trioses, tetroses, pentoses, and hexoses, respectively). The most common hexose monosaccharides include glucose, fructose, and galactose.

Disaccharides are formed when two monosaccharide units are joined by a glycosidic linkage. On hydrolysis, disaccharides will be split into two monosaccharides by disaccharide enzyme
tase) located on the microvilli of the intes
monosaccharides are then actively abs
most common disaccharides are maltose
two d-glucose molecules in a 1 → 4 linka
and sucrose.

Oligosaccharides are the chainin
10 sugar units, whereas **polysaccharides**
by the linkage of many monosaccharide
hydrolysis, polysaccharides will yield 1
monosaccharides. Amylase, an enzyme fou
and in the stomach, hydrolyzes starch to di
in the duodenum. The most common poly
are starch (plant based glucose molecule
cogen (animal based glucose molecules) (

Chemical Properties of Carbohydrates

Some carbohydrates are reducing substa
carbohydrates can reduce other compou
they themselves are oxidized. To be a red
stance, the carbohydrate must contain
(available) ketone or an aldehyde group.
erty was used in many past laboratory
the determination of carbohydrates.

Carbohydrates can form glycosidic bonds with other carbohydrates and with noncarbohydrates. Two sugar molecules can be joined in tandem, forming a glycosidic bond between the hemiacetal group of one molecule and the hydroxyl group on the other molecule. In forming the glycosidic bond, an acetal is generated on one sugar (at carbon 1) in place of the hemiacetal. If the bond forms with one of the other carbons on the carbohydrate other than the anomeric (reducing) carbon, the anomeric carbon is unaltered, and the resulting compound remains a reducing substance. Examples of common-reducing sugars include glucose, maltose, fructose, lactose, and galactose. If a glycosidic bond is formed with the anomeric carbon on the other carbohydrate, the resulting compound is no longer a reducing substance. Nonreducing carbohydrates *do not* have an active ketone or aldehyde group and therefore will not reduce other

Key Terms are also highlighted within the chapter and defined in the book's Glossary.

Glossary

1,25-Dihydroxyvitamin D ([OH]$_2$D) (calcitriol) Active metabolite of vitamin D; induces active absorption of calcium in the small intestine.

1$_{in}$ rule A data quality control rule that indicates that one data point cannot exceed three SDs. The presence of a data point beyond 3 SDs would trigger a rejection of the analytic run.

25-hydroxyvitamin D Inactive precursor of 1,25 dihydroxyvitamin D.

5-Dihydrotestosterone (DHT) An endogenous androgen sex steroid and hormone. The enzyme 5α-reductase catalyzes the formation of DHT from testosterone in certain tissues including the prostate gland, seminal vesicles, epididymides, skin, hair follicles, liver, and brain.

A

Accuracy How close the measured value is to the true value due to systematic error, which can be either constant or proportional.

Acidemia A condition in which the pH of blood is below the lower limit of the reference range (<7.35), indicating that the hydrogen-ion concentration in the blood is increased.

Activation energy The excess energy needed to form the transition state of a reaction.

Activators Inorganic cofactors, such as metal ions, needed for enzyme activity.

Active transport Use of energy to move ions or substances across cell membranes.

Acute coronary syndrome (ACS) A progression of pathologic conditions involved in ischemic heart disease, including erosion and rupture of coronary artery plaques, activation of platelets, and thrombi. This progression ranges from unstable angina to extensive tissue necrosis in acute myocardial infarction.

Acute kidney injury (AKI) A sudden, sharp decline in renal function as a result of an acute toxic or hypoxic insult to the kidneys.

Adrenocorticotropin hormone (ACTH) A peptide hormone secreted by the anterior pituitary. It stimulates the cortex of the adrenal glands to produce adrenal cortical hormones.

Affinity Attraction or force causing two substances to unite.

Airborne pathogens Any infectious agent transmissible by air, e.g., tuberculosis, virus particle, etc.

Albuminuria The presence of albumin in the urine.

Aldosterone The main mineralocorticoid steroid hormone produced by the zona glomerulosa of the adrenal cortex in the adrenal gland. This hormone controls the sodium-potassium pump, the primary mechanism for sodium reabsorption in the kidney and regulator of the blood sodium and potassium levels.

Alkalemia A condition in which blood pH is greater than the uppermost limit of the reference range (>7.45), indicating that the hydrogen ion concentration in the blood is decreased.

Amenorrhea Temporary cessation of menstruation in a female who is past menarche but not yet in menopause.

Amines Hormones that are derived directly from amino acids.

Amino acid Simple organic compounds that serve as the building blocks of proteins; contain at least one amine functional group, one carboxyl function group, and a unique R group.

Aminoacidopathies Inborn errors of metabolism that inhibit the body's ability to metabolize specific amino acids.

Ammonia A compound consisting of nitrogen and hydrogen. Formula: NH_3 or H_3N.

Amniocentesis Puncture of the amniotic sac to obtain fluid for analysis.

Amniotic fluid (AF) A fluid in which the fetus is suspended; it provides a cushioning medium for the fetus and serves as a matrix for influx and efflux of constituents.

Amperometry The measurement of amperes. It is the unit of measure for electric current. The reduction of oxygen produces a current that is proportional to the amount of oxygen present in the sample.

Amphoteric A molecule that is both an acid and a base.

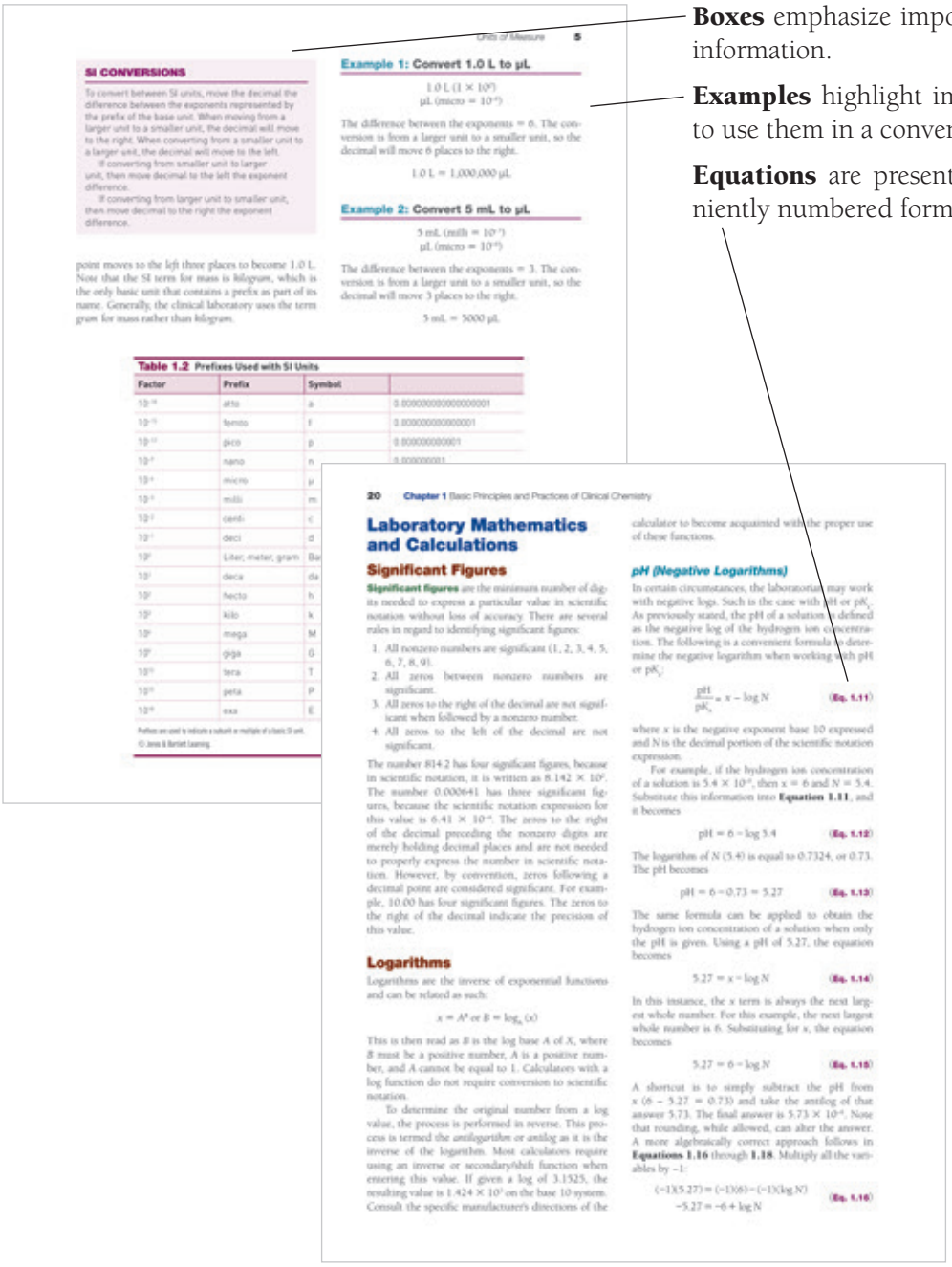
Analyte Substance of interest being measured.

Analytic Introduced during the phase of processing and assaying the specimen in the clinical laboratory.

Analytic measurement range (AMR) Also known as linear or dynamic range. Range of analyte concentrations that can be directly measured without dilution, concentration, or other pretreatment.

781

Case Studies with patient visuals progress through the chapter and pose critical-thinking questions, prompting students to synthesize and apply their new knowledge. A case study answer key is available to instructors.



Boxes emphasize important points and additional information.

Examples highlight important formulas and how to use them in a convenient, numbered format.

Equations are presented throughout in a conveniently numbered format.

Figures and **Tables** provide dynamic visuals and populate the new edition throughout, including illustrations, photos, charts, and graphs.

Table 5.4 Competitive Binding Assay Example

Ag	+	Ag*	+	Ab	→	AgAb	+	Ag*Ab	+	Ag*
CONCENTRATION OF REACTANTS						CONCENTRATION OF PRODUCTS				
Ag		Ag*		Ab		AgAb		Ag*Ab		Ag*
0		200		100		0		100		100
50		200		100		20		80		120
100		200		100		34		66		134
200		200		100		50		50		150
400		200		100		66		34		166

SAMPLE CALCULATIONS

Dose of [Ag]	% B	B/F
0	$\frac{100}{200} = 50$	$\frac{100}{100} = 1$
50	$\frac{80}{200} = 40$	$\frac{80}{200} = 67$
100	$\frac{66}{200} = 33$	$\frac{66}{134} = 33$
200	$\frac{50}{200} = 25$	$\frac{50}{150} = 33$
400	$\frac{34}{200} = 17$	$\frac{34}{166} = 20$

Ag, unlabeled antigen; Ag*, labeled antigen; Ab, antibody; AgAb, antigen-antibody complex; Ag*Ab, labeled antibody complex.
© Jones & Bartlett Learning.

equally to the Ab. As the concentration of Ag increases in a competitive assay, the amount of tracer that complexes with the binding reagent decreases. If the tracer is of low molecular weight, free tracer is often measured. If the tracer is of high molecular weight, the bound tracer is measured. The data may be plotted in one of three ways: bound/free versus the arithmetic dose of unlabeled Ag, percentage bound versus the log dose of unlabeled Ag, and logit bound/B_0 versus the log dose of the unlabeled Ag (**Figure 5.11**).

The bound fraction can be expressed in several different formats. Bound/free is counts per minute (CPM) of the bound fraction compared with the CPM of the free fraction. Percent bound (% B) is the CPM of the bound fraction compared with the CPM of maximum binding of the tracer (B_0) multiplied by 100. Logit B/B_0 transformation is the natural log of $(B/B_0)/(1 - B/B_0)$. When B/B_0 is plotted

Table 21.1 Kidney Functions

Urine formation
Fluid and electrolyte balance
Regulation of acid–base balance
Excretion of the waste products of protein metabolism
Excretion of drugs and toxins
Secretion of hormones
Renin
Erythropoietin
1,25-Dihydroxyvitamin D_3
Prostaglandins

© Jones & Bartlett Learning.

the body by way of the urethra. The highlighted section in Figure 21.1 shows the arrangement of **nephrons** in the kidney; nephrons are functional units of the kidney that can only be seen microscopically. Each kidney contains approximately 1 million nephrons. Each nephron is a complex apparatus composed of five basic parts as shown in **Figure 21.2**.

These five parts are:

- The **glomerulus**—a capillary tuft surrounded by the expanded end of a renal **tubule** known as Bowman's capsule. Each glomerulus has an afferent arteriole that carries the blood in and an efferent arteriole carrying the blood out. The efferent arteriole branches into peritubular capillaries that supply the tubule.
- The proximal convoluted tubule—located in the cortex.
- The long loop of Henle—composed of the thin descending limb, which spans the medulla, and the ascending limb, which is located in both the medulla and the cortex, composed of a region that is thin and then thick.
- The distal convoluted tubule—located in the cortex.

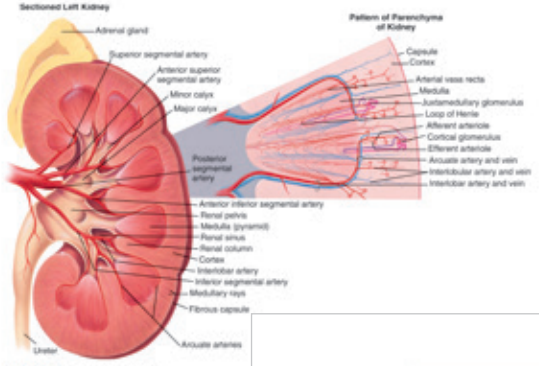


Figure 21.1 Anatomy of the kidney.

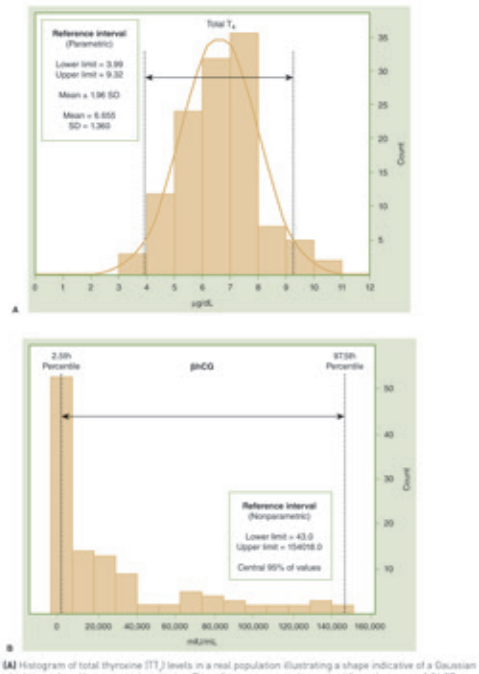
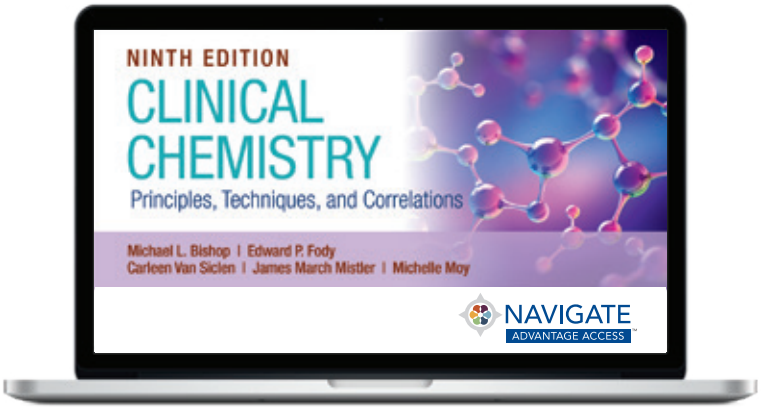


Figure 3.14 **(A)** Histogram of total thyroxine (TT_4) levels in a real population illustrating a shape indicative of a Gaussian distribution, which is analyzed by parametric statistics. The reference range is determined from the mean ± 1.96 SDs. **(B)** Histogram of β-human chorionic gonadotropin (βhCG) levels in a population of pregnant women demonstrating non-Gaussian data and nonparametric determination of the reference range. The reference range is determined from percentiles to include the central 95% of values, although the selection of a wide range of gestational ages makes this a poor population for a reference range study. It does demonstrate the application of nonparametric intervals.

Student Resources



To support your learning, review the chapter learning objectives and complete the online activities. The **Navigate 2 Advantage Access** included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!

- eBook with embedded assessments
- Case Studies
- Review Questions
- Flashcards
- Reference Range Table
- General Reference Tables
- Supplemental Chapter
  - Molecular Theory and Techniques

Instructor Resources

Instructor resources, available to qualified instructors, include the following:

- Learning Objectives mapped to:
  - ASCLS Entry-Level Curriculum (MLS and MLT)
  - Current ASCP Board of Certification Content Guidelines (MLS and MLT)
- Slides in PowerPoint format
- Teaching Resources
- Test Bank (Available in LMS-compatible formats)
- Student Lab Procedures
- Image Bank
- Answer Key to Case Studies
- Answer Key to Eighth Edition Case Studies
- Answer Key to Review Questions
- Sample Syllabus

Acknowledgments

A project as large as this requires the assistance and support of many clinical laboratorians. The editors wish to express their appreciation to the contributors of all the editions of *Clinical Chemistry: Principles, Techniques, and Correlations*—the dedicated laboratory professionals and educators whom the editors have had the privilege of knowing and exchanging ideas with over the years. These individuals were selected because of their expertise in particular areas and their commitment to the education of clinical laboratorians. Many have spent their professional careers in the clinical laboratory, at the bench, teaching students, or consulting with clinicians. In these frontline positions, they have developed a perspective of what is important for the next generation of clinical laboratorians.

We extend appreciation to our students, colleagues, teachers, and mentors in the profession who have helped shape our ideas about clinical chemistry practice and education. Also, we want to thank the many companies and professional organizations that provided product information and photographs or granted permission to reproduce diagrams and tables from their publications. The ASCLS Entry Level Curriculum and CLSI documents have also been important sources of information. These documents are directly referenced in the appropriate chapters.

The editors would like to acknowledge the contribution and effort of all individuals to previous editions. Their efforts provided the framework for many of the current chapters. Finally, we gratefully acknowledge the cooperation and assistance of the staff at Jones & Bartlett Learning for their advice and support.

The editors are continually striving to improve future editions of this book. We again request and welcome our readers' comments, criticisms, and ideas for improvement.

MLB, EPF, CVS, JMM, MM

Contributors

Karen K. Apolloni, MSA, MLS(ASCP)CM
Director (Academic), Clinical Laboratory Science Program
Wayne State University
Detroit, Michigan

Nikola A. Baumann, PhD, DABCC
Co-Director of Central Clinical Laboratory and Central Processing
Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, Minnesota

Heather McNasby, MT(ASCP), MSHS
Manager of Clinical Laboratory Services
Nemours Children's Hospital
Orlando, Florida

Takara L. Blamires, MS, MLS(ASCP)CM
Associate Professor, MLS Undergraduate Program Director
Department of Pathology
University of Utah School of Medicine
Salt Lake City, Utah

Michelle R. Campbell, MS, MLS(ASCP)CMMBCM
Development Technologist
Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, Minnesota
Adjunct Instructor
Department of Allied Health Sciences, Division of Clinical Laboratory Science
University of North Carolina School of Medicine
Chapel Hill, North Carolina

Janelle M. Chiasera
Dean and Professor
School of Health Sciences
Quinnipiac University
Hamden, Connecticut

Kathryn Dugan, MEd, MT(ASCP)
Assistant Professor
Medical and Clinical Laboratory Sciences
Auburn University at Montgomery
Montgomery, Alabama

Vicki S. Freeman, PhD, MASCP, MLS(ASCP)CMSCCM, FAACC
Professor
Clinical Laboratory Sciences
Associate Dean for Faculty Development
School of Health Professions
University of Texas Medical Branch
Galveston, Texas

Jie Gao, PhD, MLS(ASCP)CM
Assistant Professor
Department of Clinical and Diagnostic Science
University of Alabama at Birmingham
Birmingham, Alabama

Brandy Greenhill, DrPH, MLS(ASCP)CM
Professor
University of Texas MD Anderson School of Health Professions
Houston, Texas

Matthew P.A. Henderson, PhD, FCACB, FCCMG
Laboratory Head of Biochemistry
Newborn Screening Ontario
Assistant Professor
Department of Pediatrics
Children's Hospital of Eastern Ontario, University of Ottawa
Ottawa, Ontario
CANADA

Laura M. Hickes
Chemistry and Applications Support Manager
Roche Diagnostics
Indianapolis, Indiana

Kristina Jackson Behan, PhD, MLS(ASCP)
Professor
Medical Laboratory Sciences
University of West Florida
Pensacola, Florida

Stephanie L. Jacobson, MS, MLS(ASCP)CM
MLS Online Instructor
Allied and Population Health
Medical Laboratory Scientist
Monument Health
Rapid City, South Dakota
South Dakota State University
Brookings, South Dakota

Floyd Josephat, EdD, MT(ASCP)
Professor and Program Director
Clinical and Diagnostic Sciences
Graduate Faculty II
University of Alabama at Birmingham
Birmingham, Alabama

Brooke M. Katzman, PhD
Co-Director of Hospital Clinical Laboratory and Point of Care
Department of Laboratory Medicine and Pathology
Mayo Clinic
Rochester, Minnesota

Kathleen M. Kenwright, EdD, MLS(ASCP)CM, MB
Medical Laboratory Science Program Director
Diagnostic and Health Sciences
Associate Professor
University of Tennessee Health Science Center
Memphis, Tennessee

John E. Lee Sang, MD
Laboratory Director Clinical Pathology Associates
Pathology Department
San Marcos, Texas
Clinical Pathology Associates
San Marcos, Texas

Cyndy B. Levtzow, BS, MT(ASCP), LSSBB
Retired; Previous Clinical Core Supervisor and Assistant Administrative Director
Formerly UNC Hospitals, McLendon Clinical Laboratories
Chapel Hill, North Carolina

Shawn H.W. Luby, MS, MLS(ASCP)CM
Assistant Professor
Division of Clinical Laboratory Science
University of North Carolina School of Medicine
Chapel Hill, North Carolina

James March Mistler, ABD, MS, MLS(ASCP)CM
Program Director and Fulltime Lecturer
Department of Medical Laboratory Science
University of Massachusetts Dartmouth
North Dartmouth, Massachusetts

J. Marvin McBride, MD, MBA
Clinical Assistant Professor
Department of Medicine
University of North Carolina at Chapel Hill School of Medicine
Chapel Hill, North Carolina

Christopher R. McCudden, PhD, FACB, FCACB
Clinical Biochemist
Division of Biochemistry
Associate Professor
Department of Pathology & Laboratory Medicine
The Ottawa Hospital, University of Ottawa
Ottawa, Ontario
CANADA

Scott Moore, DO, MLS(ASCP)CM, DipIBLM
Assistant Professor of Medical Laboratory Sciences
Department of Medical Laboratory Sciences
Weber State University
Ogden, Utah

Teresa A. Mortier, PhD, MT(ASCP)
Assistant Professor
School of Health Sciences
Eastern Michigan University
Ypsilanti, Michigan

Yukari Nishizawa-Brennen, PhD, MLS(ASCP)CM, MT(Japan)
Assistant Professor
Biomedical Laboratory Diagnostics Program
Michigan State University
East Lansing, Michigan

Peter P. O'Brien, MBA, MT(ASCP)
Program Director Medical Laboratory Technician Program
Allied Health Sciences
Jackson State Community College
Jackson, Tennessee

Khushbu Patel, PhD, DABCC
Director of Clinical Chemistry and Point-of-Care
Children's Hospital of Philadelphia Department of Pathology and Laboratory Medicine
Assistant Professor of Pathology and Laboratory Medicine
Perelman School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania

Heather L. Phillips, PhD, MLS(ASCP)CM, MT(AMT)
Director Austin Peay State University COVID-19 Testing Laboratory
COVID-19 Molecular Laboratory
Assistant Professor
Allied Health Sciences—Medical Laboratory Sciences
Austin Peay State University
Clarksville, Tennessee

Tracey G. Polsky, MD, PhD
Chief, Division of Community Pathology
Children's Hospital of Philadelphia Department of Pathology and Laboratory Medicine
Assistant Professor of Pathology and Laboratory Medicine
Perelman School of Medicine
University of Pennsylvania
Philadelphia, Pennsylvania

Gerald D. Redwine, PhD, MT(ASCP)
Associate Professor
Clinical Laboratory Science Program
Texas State University
San Marcos, Texas

Kyle B. Riding, PhD, MLS(ASCP)CM
Assistant Professor of Medicine
Burnett School of Biomedical Science, College of Medicine
University of Central Florida
Orlando, Florida

Amanda Rivera-Begeman, DO
Pathologist, Chief of Pathology Medical Education
Pathology Department
Clinical Pathology Associates
Dell Medical School at the University of Texas at Austin
Austin, Texas

Michael W. Rogers, MBA, MT(ASCP)
Laboratory Technical Consultant
Public Health Laboratory
Independent Contractor
Winston Salem, North Carolina

Jane R. Semler, MS, MT(ASCP)
Professor
Allied Health Sciences
Coordinator, MLT to MLS Program
Austin Peay State University
Clarksville, Tennessee

Rebbecca D. Silva, MS, MT(ASCP)
Department Chair, Assistant Professor
Medical Laboratory Technology Program
New England Institute of Technology
East Greenwich, Rhode Island

Donna J. Spannaus-Martin, PhD, MLS(ASCP)CM
Professor, Medical Laboratory Sciences Program
University of Minnesota
Minneapolis, Minnesota

Christopher Swartz, PhD, MLS(ASCP)
Assistant Professor
Department of Health and Clinical Sciences
University of Kentucky, College of Health Sciences
Lexington, Kentucky

Yun L. Trull, PhD, DABCC
Director of Clinical Chemistry
Pathology and Laboratory Medicine
Allegheny Health Network
Pittsburgh, Pennsylvania

Carleen Van Siclen, MS, MLS(ASCP)CM, TS(ABB)
Assistant Professor of Laboratory Medicine and Pathology
MCSHS Medical Laboratory Science Program
Mayo Clinic
Jacksonville, Florida

Elizabeth Warning, MS, MLS(ASCP)CM, AHI (AMT)
Associate Professor, Educator
Clinical & Health Information Sciences
Program Director, MLS Campus Based Program
College of Allied Health Sciences
University of Cincinnati
Cincinnati, Ohio

Monte S. Willis, MD, PhD, MBA, FASCP, FCAP, FAHA
Medical Director, Clinical Pathology, Allegheny Health Network, Allegheny General Hospital Core Laboratory
Pathology and Laboratory Medicine
Patient Experience Officer, Pathology Institute
Allegheny Health Network
Pittsburgh, Pennsylvania

Ancillary Contributors

Brenda C. Barnes, PhD, MT(ASCP)SBB
Director, Medical Laboratory Science Program
Professor
Allen College
Waterloo, Iowa

Leigh A. Belair, MEd, MT(ASCP)
Program Director, MLT Program of Maine
Associate Professor
University of Maine at Presque Isle
Presque Isle, Maine

Michael L. Bishop, MS, MLS(ASCP)
Adjunct Faculty, Medical Laboratory Scientist Program
Santa Fe College
Alachua, Florida
Retired Faculty, Medical Technology Program
Duke University Medical Center
Durham, North Carolina
Retired Program Director and University Department Chair, Medical Laboratory Science and Medical Laboratory Technician Program
Keiser University
Orlando, Florida

Jie Gao, PhD, MLS(ASCP)CM
Assistant Professor
Department of Clinical and Diagnostic Science
University of Alabama at Birmingham
Birmingham, Alabama

James March Mistler, ABD, MS, MLS(ASCP)CM
Program Director and Fulltime Lecturer
Department of Medical Laboratory Science
University of Massachusetts Dartmouth
North Dartmouth, Massachusetts

Michelle Moy, MAdEd, MT(ASCP) SC
Assistant Professor of Medical Laboratory Science
Program Director, Biomedical Science
Department of Natural Science and Mathematics, College of Arts and Sciences
Madonna University
Livonia, Michigan

Carmen Pierce, MS CLS (ASCP)CM
Full Time Lecturer
University of Massachusetts Dartmouth
Dartmouth, Massachusetts

Carleen Van Siclen, MS, MLS(ASCP)CM, TS(ABB)
Faculty, MCSHS Medical Laboratory Science Program
Assistant Professor, Mayo Clinic College of Medicine, Department of Laboratory Medicine and Pathology
Mayo Clinic
Jacksonville, Florida

Reviewers

Laurie B. Alloway, MSES, MLS(ASCP)CM, SC(ASCP)CM, EMT
Associate Teaching Professor
Medical Laboratory Sciences Dept.
Wichita State University
Wichita, Kansas

Nancy D. Cruz-Sanchez, MS, MLS(ASCP)CM
Lecturer
Clinical Laboratory Science Program
The University of Texas at El Paso
El Paso, Texas

Lorraine Doucette, MS, MLS(ASCP)CM
Assistant Professor, MLS Program Director
Department of Medical and Research Technology
University of Maryland School of Medicine
Baltimore, MD

Jie Gao, PhD, MLS(ASCP)CM
Assistant Professor
Department of Clinical and Diagnostic Science
University of Alabama at Birmingham
Birmingham, Alabama

Daniela Gill, MSHS, MLS(ASCP)CM
Clinical Assistant Professor
University of Texas Rio Grande Valley
Edinburg, Texas

Basic Principles and Practice in Clinical Chemistry

Basic Principles and Practices of Clinical Chemistry

Kathryn Dugan and Elizabeth Warning

CHAPTER OUTLINE

KEY TERMS

Analyte
Anhydrous
Arterial blood
Beer's law
Buffer
Calibration
Centrifugation
Cerebrospinal fluid (CSF)
Colligative property
Conductivity
Deionized water
Delta absorbance
Density
Desiccant
Dilution
Distilled water
Equivalent weight
Erlenmeyer flasks
Filtration
Graduated cylinder
Griffin Beaker
Hemolysis
Henderson-Hasselbalch equation
Hydrate
Hygroscopic
Icterus
International unit
Ionic strength
Linearity
Lipemia
Molality
Molarity
Normality
One-point calibration
Osmotic pressure
Oxidized

Oxidizing agent
Percent solution
pH
Pipette
Primary standard
Reagent-grade water
Redox potential
Reduced
Reducing agent
Reverse osmosis
Serial dilution
Serum
Significant figures
Solute
Solution
Solvent
Specific gravity
Standard reference materials (SRMs)
Système International d'Unités (SI)
Thermistor
Valence
Volumetric
Whole blood

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Convert results from one unit format to another using the SI and traditional systems.
- Describe the classifications used for reagent-grade water.
- Identify the varying chemical grades used in reagent preparation and indicate their correct use.
- Define primary standard and standard reference materials.
- Describe the following terms that are associated with solutions and, when appropriate, provide the respective units: percent, molarity, normality, molality, saturation, colligative properties, redox potential, and conductivity.
- Compare and contrast osmolarity and osmolality.
- Define a buffer and give the formula for pH and p*K* calculations.
- Use the Henderson-Hasselbalch equation to determine the missing variable when given either the p*K* and pH or the p*K* and concentration of the weak acid and its conjugate base.
- List and describe the types of thermometers used in the clinical laboratory.
- Classify the type of pipette when given an actual pipette or its description.
- Demonstrate the proper use of a measuring and volumetric pipette.
- Describe two ways to calibrate a pipetting device.
- Define a desiccant and discuss how it is used in the clinical laboratory.
- Describe how to properly care for and balance a centrifuge.
- Correctly perform the laboratory mathematical calculations provided in this chapter.
- Identify and describe the types of samples used in clinical chemistry.
- Outline the general steps for processing blood samples.
- Apply Beer's law to determine the concentration of a sample when the absorbance or change in absorbance is provided.
- Identify the preanalytic variables that can adversely affect laboratory results as presented in this chapter.

CASE STUDY 1.1, PART 1

Meet Miles, a 25-year-old graduate who accepted his first job offer working in the chemistry department at a large medical center. Miles and Mía were classmates in college and often support each other on technical issues, even though they work at different facilities within the same health system.

© dotshock/Shutterstock.

CASE STUDY 1.2, PART 1

Meet Mía, a 35-year-old graduate who is also newly hired and works as a generalist in a small community hospital. Mía received a rainbow of tubes from the emergency department. She handed her coworker the lavender- and blue-top tubes and placed the 8.0-mL plain red-top tube and the 3.5-mL plasma separator tube in the centrifuge. She placed the heparinized whole blood specimen on the mixer and logged in to the laboratory information system to receive the specimens. Once the specimens were accessioned, she ran a STAT profile on the Nova pHOx analyzer using the whole blood specimen, and the results were autoverified.

© Ariel Skelley/DigitalVision/Getty Images.

The primary purpose of a clinical chemistry laboratory is to perform analytic procedures that yield accurate and precise information, aiding in patient diagnosis and treatment. The achievement of reliable results requires that the clinical laboratorian be able to correctly use basic supplies and equipment and possess an understanding of fundamental concepts critical to any analytic procedure. The topics in this chapter include units of measure, basic laboratory supplies, and introductory laboratory mathematics, plus a brief discussion of specimen collection, processing, and reporting.

Units of Measure

Any meaningful *quantitative* laboratory result consists of two components: the first component represents the number related to the actual test value, and the second is a label identifying the units. The unit defines the physical quantity or dimension, such as mass, length, time, or volume.[1] There are a few laboratory tests that do not have units, but whenever possible, units should be used.

The **Système International d'Unités (SI)** was adopted in 1960. It is preferred in scientific literature and clinical laboratories and is the only system employed in many countries. This system was devised to provide the global scientific community with a uniform method of describing physical quantities. The SI system units (referred to as *SI units*) are based on the metric system. Several subclassifications exist within the SI system, one of which is the *basic unit*. There are seven basic units (**Table 1.1**), with length (meter), mass (kilogram), and quantity of a substance (mole) being the units most frequently encountered. Derived units are another subclassification of the SI system. A derived unit is a mathematical function describing one of the basic units. An example of an SI-derived unit is meters per second (m/s), which is used to express velocity. Some non-SI units are so widely used that they have become acceptable for use within the SI system (Table 1.1). These include units such as hour, minute, day, gram, liter, and plane angles expressed as degrees. The SI system uses standard prefixes to indicate a decimal fraction or multiples of that basic unit (**Table 1.2**).[1] For example, 0.001 liter can be expressed using the prefix *milli*, or 10^{-3}, and since it requires moving the decimal point three places to the right, it can then be written as 1 milliliter, or abbreviated as 1 mL. It may also be written in scientific notation as 1×10^{-3} L. Likewise, 1000 liters would use the prefix of kilo (10^3) and could be written as 1 kiloliter or expressed in scientific notation as 1×10^3 L. Table 1.2 indicates prefixes that are frequently used in clinical laboratories. Prefixes smaller than the basic unit have a negative exponent (deci: 10^{-1}), and prefixes larger than the base unit have a positive exponent (kilo: 10^3). When converting between prefixes, note the relationship between the two prefixes based on whether you are changing to a smaller or larger prefix. When converting from a larger to smaller, the decimal will move to the right. For example, converting one liter (1.0×10^0 or 1.0) to milliliters (1.0×10^{-3} or 0.001), the starting unit (L) is larger than milliliters, by a factor of 1000, or 10^3. This means that the decimal place moves to the *right* three places, so 1.0 liter (L) equals 1000 milliliters (mL). The opposite is also true. When converting to a larger unit, the decimal place moves to the left. For example, converting 1000 milliliters (mL) to 1.0 liter (L), the decimal

Table 1.1 SI Units

Base Quantity	Name	Symbol
Length	Meter	m
Mass	Kilogram	kg
Time	Second	s
Electric current	Ampere	A
Thermodynamic temperature	Kelvin	K
Amount of substance	Mole	mol
Luminous intensity	Candela	cd
Selected Derived		
Frequency	Hertz	Hz
Force	Newton	N
Celsius temperature	Degree Celsius	°C
Catalytic activity	Katal	kat
Selected Accepted Non-SI		
Minute (time)	(60 s)	min
Hour	(3600 s)	h
Day	(86,400 s)	d
Liter (volume)	($1\ dm^3 = 10^{-3}\ m^3$)	L
Angstrom	($0.1\ nm = 10^{-10}\ m$)	Å

SI CONVERSIONS

To convert between SI units, move the decimal the difference between the exponents represented by the prefix of the base unit. When moving from a larger unit to a smaller unit, the decimal will move to the right. When converting from a smaller unit to a larger unit, the decimal will move to the left.

If converting from smaller unit to larger unit, then move decimal to the left the exponent difference.

If converting from larger unit to smaller unit, then move decimal to the right the exponent difference.

point moves to the *left* three places to become 1.0 L. Note that the SI term for mass is *kilogram*, which is the only basic unit that contains a prefix as part of its name. Generally, the clinical laboratory uses the term *gram* for mass rather than *kilogram*.

Example 1: Convert 1.0 L to µL

$$1.0\ \text{L}\ (1 \times 10^{0})$$
$$\mu\text{L}\ (\text{micro} = 10^{-6})$$

The difference between the exponents = 6. The conversion is from a larger unit to a smaller unit, so the decimal will move 6 places to the right.

$$1.0\ \text{L} = 1{,}000{,}000\ \mu\text{L}$$

Example 2: Convert 5 mL to µL

$$5\ \text{mL}\ (\text{milli} = 10^{-3})$$
$$\mu\text{L}\ (\text{micro} = 10^{-6})$$

The difference between the exponents = 3. The conversion is from a larger unit to a smaller unit, so the decimal will move 3 places to the right.

$$5\ \text{mL} = 5000\ \mu\text{L}$$

Table 1.2 Prefixes Used with SI Units

Factor	Prefix	Symbol	
10^{-18}	atto	a	0.000000000000000001
10^{-15}	femto	f	0.000000000000001
10^{-12}	pico	p	0.000000000001
10^{-9}	nano	n	0.000000001
10^{-6}	micro	µ	0.000001
10^{-3}	milli	m	0.001
10^{-2}	centi	c	0.01
10^{-1}	deci	d	0.1
10^{0}	Liter, meter, gram	Basic unit	1.0
10^{1}	deca	da	10
10^{2}	hecto	h	100
10^{3}	kilo	k	1000
10^{6}	mega	M	1,000,000
10^{9}	giga	G	1,000,000,000
10^{12}	tera	T	1,000,000,000,000
10^{15}	peta	P	1,000,000,000,000,000
10^{18}	exa	E	1,000,000,000,000,000,000

Prefixes are used to indicate a subunit or multiple of a basic SI unit.

Example 3: Convert 5.3 mL to dL

$$5.3 \text{ mL (milli} = 10^{-3})$$
$$\text{dL (deci} = 10^{-1})$$

The conversion is moving from a smaller unit to a larger unit, so the decimal place will move two places to the left.

$$5.3 \text{ mL} = 0.053 \text{ dL}$$

Reporting of laboratory results is often expressed in terms of substance concentration (e.g., moles) or the mass of a substance (e.g., mg/dL, g/dL, g/L, mmol/L, and IU) rather than in SI units. These traditional units can cause confusion during interpretation and conversion to SI units: examples of conversions can be found later in the chapter. As with other areas of industry, the laboratory and the rest of medicine are moving toward adopting universal standards promoted by the International Organization for Standardization, often referred to as ISO. This group develops standards of practice, definitions, and guidelines that can be adopted by everyone in a given field, providing for more uniform terminology. Many national initiatives have recommended common units for laboratory test results, but none have been widely adopted.[2] As with any transition, the clinical laboratorian should be familiar with all the terms currently used in their field and how to convert these to SI units.

Reagents

In today's highly automated laboratory, there is little need for reagent preparation by the laboratorian. Most instrument manufacturers make the reagents in a ready-to-use form or "kit" in which all necessary reagents and respective storage containers are prepackaged as a unit, requiring only the addition of water or buffer for reconstitution. A heightened awareness of the hazards of certain chemicals and the numerous regulatory agency requirements has caused clinical chemistry laboratories to eliminate massive stocks of chemicals and opt instead for the ease of using prepared reagents. Periodically, the laboratorian may still need to prepare reagents or solutions, especially in hospital laboratories involved in research and development, biotechnology applications, specialized analyses, or method validation.

Chemicals

Analytic chemicals exist in varying grades of purity: Reagent grade or analytic reagent (AR); ultrapure, chemically pure (CP); United States Pharmacopeia (USP); National Formulary (NF); and technical or commercial grade.[3] Chemicals with AR designation are suitable for use in most analytic laboratory procedures. A committee of the American Chemical Society (ACS) established specifications for AR grade chemicals, and chemical manufacturers must either meet or exceed these requirements. The labels on reagents should clearly state the actual impurities for each chemical lot or list the maximum allowable impurities. The label should also include one of the following designations: AR or ACS or *For laboratory use* or *ACS Standard-Grade Reference Materials*. Ultrapure chemicals have additional purification steps for use in specific procedures such as chromatography, immunoassays, molecular diagnostics, standardization, or other techniques that require extremely pure chemicals. These reagents may have designations of HPLC (high-performance liquid chromatography) or chromatographic on their labels.

Because USP- and NF-grade chemicals are used to manufacture drugs, the limitations established for this group of chemicals are based only on the criterion of not being injurious to individuals. Chemicals in this group may be pure enough for use in most chemical procedures, but the purity standards they meet are not based on the needs of the laboratory and may or may not meet all assay requirements.

Reagent designations of CP or ultrapure grade indicate that the impurity limitations are not stated, and preparation of these chemicals is not uniform. It is not recommended that clinical laboratories use these chemicals for reagent preparation unless further purification or a reagent blank is included. Technical or commercial grade reagents are used primarily in manufacturing and should never be used in the clinical laboratory.

Organic reagents also have varying grades of purity that differ from those used to classify inorganic reagents. These grades include a practical grade with some impurities; CP, which approaches the purity level of reagent-grade chemicals; spectroscopic (spectrally pure) and chromatographic grade organic reagents; and reagent grade (ACS), which is certified to contain impurities below established ACS levels. Other than the purity aspects of the chemicals, laws related to the Occupational Safety and Health Administration (OSHA)[4] require manufacturers to indicate any physical or biologic health

hazards and precautions needed for the safe use, storage, and disposal of any chemical. Manufacturers are required to provide a Safety Data Sheet (SDS). A copy of the SDS must be readily available to ensure the safety of laboratorians.

Reference Materials

Unlike other areas of chemistry, clinical chemistry is involved in the analysis of biochemical by-products found in *biological* fluids, such as serum, plasma, or urine. For this reason, traditionally defined standards used in analytical chemistry do not readily apply in clinical chemistry.

A **primary standard** is a highly purified chemical that can be measured directly to have an *exact* known concentration and purity. The ACS has purity tolerances for primary standards; because most biologic constituents are unavailable within these tolerance limitations, the National Institute of Standards and Technology (NIST) has certified **standard reference materials (SRMs)** that are used in place of ACS primary standard materials.[5–7]

These SRMs are assigned a value after analysis using state-of-the-art methods and equipment. The chemical composition of these substances is then certified; however, they may not have the purity of a primary standard. Because each substance has been characterized for certain chemical or physical properties, it can be used in place of an ACS primary standard in clinical work and is often used to verify **calibration** or accuracy/bias assessments. Many manufacturers use a NIST SRM when producing calibrator and standard materials. These materials are considered "traceable to NIST" and may meet certain accreditation requirements. Standard reference materials are used for **linearity** studies to determine the relationship between the standard's concentration and the instrument result. Linearity studies are required when a new test or new test methodology is introduced. There are SRMs for a number of routine analytes, hormones, drugs, and blood gases, with others being added.[5] Calibration of an instrument is a process that pairs an analytical signal with a concentration value of an analyte. When performing a calibration, a series of calibrators with known concentrations of a specific analyte are used. The instrument is programmed with the known concentrations and will adjust the analytic signal to match the given concentration. Calibrators can be purchased as a kit or made by diluting a known stock solution.

Water Specifications[8]

Water is the most frequently used reagent in the laboratory. Tap water is unsuitable for laboratory applications. Most procedures, including reagent and control preparation, require water that has been substantially purified, known as **reagent-grade water**. There are various water purification methods including distillation, ion exchange, reverse osmosis, ultrafiltration, ultraviolet light, sterilization, and ozone treatment. According to the Clinical and Laboratory Standards Institute (CLSI), reagent-grade water is classified into one of six categories based on the specifications needed for its use rather than the method of purification or preparation.[9] These categories include clinical laboratory reagent water (CLRW), special reagent water (SRW), instrument feed water, water supplied by method manufacturer, autoclave and wash water, and commercially bottled purified water. Each category has a specific acceptable limit. The College of American Pathologists requires laboratories to define the specific type of water required for each of its testing procedures and requires water quality testing at least annually. Water quality testing routinely includes monitoring microbial colony-forming units/mL and may also include other parameters.

Distilled water has been purified to remove almost all organic materials, using a technique of distillation where water is boiled and vaporized. Many impurities do not rise in the water vapor and will remain in the boiling apparatus so that the water collected after condensation has less contamination. Water may be distilled more than once, with each distillation cycle removing additional impurities. Ultrafiltration and nanofiltration, like distillation, are excellent in removing particulate matter, microorganisms, and any pyrogens or endotoxins.

Deionized water has some or all ions removed, although organic material may still be present, so it is neither pure nor sterile. Generally, deionized water is purified from previously treated water, such as prefiltered or distilled water. Deionized water is produced using either an anion- or a cation-exchange resin, followed by replacement of the removed ions with hydroxyl or hydrogen ions. A combination of several ion-exchange resins will produce different grades of deionized water. A two-bed system uses an anion resin followed by a cation resin. The different resins may be in separate columns or in the same column. This process is excellent at removing dissolved ionized solids and dissolved gases.

Reverse osmosis is a process that uses pressure to force water through a semipermeable membrane, producing a filtered product. Reverse osmosis may be used for the pretreatment of water, however, it does not remove dissolved gases.

Filtration can remove particulate matter from municipal water supplies before any additional treatments. Filtration cartridges can be composed of glass, cotton, or activated charcoal, which removes organic materials and chlorine. Some have submicron filters (≤0.2 μm), which remove any substances larger than the filter's pores, including bacteria. The use of these filters depends on the quality of the municipal water and the other purification methods used. For example, hard water (containing calcium, iron, and other dissolved elements) may require prefiltration with a glass or cotton filter rather than activated charcoal or submicron filters, which quickly become clogged and are expensive to use. The submicron filter may be better suited after distillation, deionization, or reverse osmosis treatment.

Ultraviolet oxidation, which removes some trace organic material or sterilization processes at specific wavelengths, can destroy bacteria when used as part of a system but may leave behind some residual products. This technique is often followed by other purification processes.

Reagent-grade water can be obtained by initially filtering to remove particulate matter, followed by reverse osmosis, deionization, and a 0.2-μm filter or more restrictive filtration process. Autoclave wash water is acceptable for glassware washing but not for analysis or reagent preparation. SRW is used for specific techniques like the HPLC, molecular diagnostics, or mass spectrophotometry, which may require specific parameters for the analysis. All SRW should meet CLRW standards and, depending on the application, CLRW should be stored in a manner that reduces any chemical or bacterial contamination and for short periods.

Testing procedures to determine the quality of reagent-grade water include measurements of resistance, pH, colony counts on selective and nonselective media for the detection of bacterial contamination, chlorine, ammonia, nitrate or nitrite, iron, hardness, phosphate, sodium, silica, carbon dioxide, chemical oxygen demand, and metal detection. Some accreditation agencies[10] recommend that laboratories document culture growth, pH, and specific resistance on water used in reagent preparation. Resistance is measured because pure water, devoid of ions, is a poor conductor of electricity and has increased resistance. The relationship of water purity to resistance is linear; generally, as purity increases, so does resistance. This one measurement does not suffice for determination of true water purity because a nonionic contaminant may be present that will have little effect on resistance. Reagent water meeting specifications from other organizations, such as the American Society for Testing and Materials (ASTM), may not be equivalent to those established by the CLSI, so care should be taken to meet the assay procedural requirements for water type.

Solution Properties

In clinical chemistry, substances found in biologic fluids, including serum, plasma, urine, and spinal fluid, are quantified. A substance that is dissolved in a liquid is called a **solute**; a biologic solute is also known as an **analyte**. The liquid in which the solute is dissolved—for example, a biologic fluid—is the **solvent**. Together, solute and solvent represent a **solution**. Any chemical or biologic solution can be described by its basic properties, including concentration, saturation, colligative properties, redox potential, conductivity, density, pH, and ionic strength.

Concentration

The analyte concentration in solution can be expressed in many ways. Concentration is commonly expressed as *percent solution*, *molarity*, *molality*, or *normality*. These are non-SI units, however; the SI unit for the amount of a substance is the *mole*. Examples of concentration calculations are provided later in this chapter.

Percent solution is expressed as the amount of solute per 100 total units of solution. Three expressions of percent solutions are weight per weight (w/w), volume per volume (v/v), and weight per volume (w/v). Weight per weight (% w/w) refers to the number of grams of solute per 100 g of solution. Volume per volume (% v/v) is used for liquid solutes and gives the milliliters of solute in 100 mL of solution. For v/v solutions, it is recommended that grams per deciliter (g/dL) be used instead of % v/v. Weight per volume (% w/v) is the most commonly used percent solution in the clinical laboratory and is defined as the number of grams of solute in 100 mL of solution. Weight per volume is not the same as molarity, and care must be taken to not confuse the two. Examples of percent solution calculations can be found later in this chapter.

Molarity (M) is expressed as the number of moles per 1 L of solution. One mole of a substance equals its gram molecular weight (gmw), so the customary units of molarity (M) are moles/liter. The SI representation for the traditional molar concentration is moles

of solute per volume of solution, with the volume of the solution given in liters. The SI expression for concentration should be represented as moles per liter (mol/L), millimoles per liter (mmol/L), micromoles per liter (µmol/L), or nanomoles per liter (nmol/L). The common concentration term *molarity* is not an SI unit for concentration. Molarity depends on volume, and any significant physical changes that influence volume, such as changes in temperature and pressure, will also influence molarity. Calculations can be found in the Laboratory Mathematics and Calculations section of this chapter.

Molality (m) represents the amount of solute per 1 kg of solvent. Molality is sometimes confused with molarity; however, it can be easily distinguished because molality is always expressed in terms of moles per kilogram (weight per weight) and describes moles per 1000 g (1 kg) of solvent. Note that the common abbreviation (m) for molality is a lowercase "m," while the uppercase "M" refers to molarity. Molality is not influenced by temperature or pressure because it is based on mass rather than volume.

Normality is the least likely of the four concentration expressions to be encountered in clinical laboratories, but it is often used in chemical titrations and chemical reagent classification. It is defined as the number of gram equivalent weights per 1 L of solution. An **equivalent weight** is equal to the gmw of a substance divided by its valence. The **valence** is the number of units that can combine with or replace 1 mole of hydrogen ions for acids and hydroxyl ions for bases and the number of electrons exchanged in oxidation–reduction reactions. Normality is always equal to or greater than the molarity of the compound. Calculations can be found later in this chapter. Normality was previously used for reporting electrolyte values, expressed as milliequivalents per liter (mEq/L); however, this convention has been replaced with millimoles per liter (mmol/L). The College of American Pathologists (CAP) currently requires chloride to be reported in mmol/L. Because the four main electrolytes, Na^+, K^+, CO_2^- (HCO_3^-), and Cl^-, all have a valence of 1, the concentration reported will remain the same whether the unit is mEq/L or mmol/L.

Solution saturation gives little specific information about the concentration of solutes in a solution. A solution is considered *saturated* when no more solvent can be dissolved in the solution. Temperature, as well as the presence of other ions, can influence the solubility constant for a solute in a given solution and thus affect the saturation. Routine terms in the clinical laboratory that describe the extent of saturation are *dilute*, *concentrated*, *saturated*, and *supersaturated*. A *dilute solution* is one in which there is relatively little solute or one that has a lower solute concentration per volume of solvent than the original, such as when making a dilution. In contrast, a *concentrated solution* has a large quantity of solute in solution. A solution in which there is an excess of undissolved solute particles can be referred to as a *saturated solution*. As the name implies, a *supersaturated solution* has an even greater concentration of undissolved solute particles than a saturated solution of the same substance. Because of the greater concentration of solute particles, a supersaturated solution is thermodynamically unstable. The addition of a crystal of solute or mechanical agitation disturbs the supersaturated solution, resulting in crystallization of any excess material out of solution. An example is when measuring serum osmolality by freezing point depression.

Colligative Properties

Colligative properties are those properties related to the number of solute particles per solvent molecules, not on the type of particles present. The behavior of particles or solutes in solution demonstrates four properties: **osmotic pressure**, vapor pressure, freezing point, and boiling point. These are called **colligative properties**. *Osmotic pressure* is the pressure that opposes osmosis when a solvent flows through a semipermeable membrane to establish equilibrium between compartments of differing concentration. *Vapor pressure* is the pressure exerted by the vapor when the liquid solvent is in equilibrium with the vapor. *Freezing point* is the temperature at which the first crystal (solid) of solvent forms in equilibrium with the solution. *Boiling point* is the temperature at which the vapor pressure of the solvent reaches atmospheric pressure (usually 1 atmosphere).

The osmotic pressure of a dilute solution is directly proportional to the concentration of the molecules in solution. The expression for concentration is the osmole. One osmole of a substance equals the molarity or molality multiplied by the number of particles, not the kind of particle, at dissociation. If molarity is used, the resulting expression would be termed osmolarity; if molality is used, the expression changes to osmolality. Osmolality is preferred since it depends on the weight rather than volume and is not readily influenced by temperature and pressure changes. When a solute is dissolved in a solvent, the colligative properties change in a predictable manner for each osmole of substance present. In the clinical setting, freezing point and vapor pressure depression can be measured as a function

of osmolality. Freezing point is preferred since vapor pressure measurements can give inaccurate readings when some substances, such as alcohols, are present in the samples.

Redox Potential

Redox potential, or *oxidation–reduction potential*, is a measure of the ability of a solution to accept or donate electrons. Substances that donate electrons are called **reducing agents**; those that accept electrons are considered **oxidizing agents**. The mnemonic—LEO (lose electrons **oxidized**) the lion says GER (gain electrons **reduced**)—may prove useful when trying to recall the relationship between reducing/oxidizing agents.

Conductivity

Conductivity is a measure of how well electricity passes through a solution. A solution's conductivity quality depends principally on the number of respective charges of the ions present. *Resistivity*, the reciprocal of conductivity, is a measure of a substance's resistance to the passage of electrical current. The primary application of resistivity in the clinical laboratory is for assessing the purity of water. Resistivity (resistance) is expressed as ohms and conductivity is expressed as $ohms^{-1}$.

pH and Buffers

Buffers are weak acids or bases and their related salts that minimize changes in the hydrogen ion concentration. Hydrogen ion concentration is often expressed as pH. A lowercase *p* in front of certain letters or abbreviations operationally means the "negative logarithm of" or "inverse log of" that substance. In keeping with this convention, the term **pH** represents the negative or inverse log of the hydrogen ion concentration. Mathematically, pH is expressed as

$$\mathrm{pH} = \log\left(\frac{1}{[\mathrm{H^+}]}\right) \qquad \text{(Eq. 1.1)}$$

$$\mathrm{pH} = -\log[\mathrm{H^+}]$$

where $[H^+]$ equals the concentration of hydrogen ions in moles per liter (M). The pH scale ranges from 0 to 14 and is a convenient way to express hydrogen ion concentration.

Unlike a strong acid or base, which dissociates almost completely, the dissociation constant for a weak acid or base solution (like a buffer) tends to be very small, meaning little dissociation occurs.

The dissociation of acetic acid (CH_3COOH), a weak acid, can be illustrated as follows:

$$[\mathrm{HA}] \leftrightarrow [\mathrm{A^-}] + [\mathrm{H^+}] \qquad \text{(Eq. 1.2)}$$

$$[\mathrm{CH_3COOH}] \leftrightarrow [\mathrm{CH_3COO^-}] + [\mathrm{H^+}]$$

HA = weak acid, A^- = conjugate base, H^+ = hydrogen ions, [] = concentration of item in the bracket.

Sometimes, the conjugate base (A^-) will be referred to as a "salt" since, physiologically, it will be associated with some type of cation, such as sodium (Na^+).

The dissociation constant, K_a, for a weak acid may be calculated using the following equation:

$$K_a = \frac{[\mathrm{A^-}][\mathrm{H^+}]}{[\mathrm{HA}]} \qquad \text{(Eq. 1.3)}$$

Rearrangement of this equation reveals

$$[\mathrm{H^+}] = K_a \times \frac{[\mathrm{HA}]}{[\mathrm{A^+}]} \qquad \text{(Eq. 1.4)}$$

Taking the log of each quantity and then multiplying by minus 1 (–1), the equation can be rewritten as

$$-\log[\mathrm{H^+}] = -\log K_a \times -\log\frac{[\mathrm{HA}]}{[\mathrm{A^-}]} \qquad \text{(Eq. 1.5)}$$

By convention, lowercase *p* means "negative log of"; therefore, $-\log[H^+]$ may be written as pH, and $-K_a$ may be written as pK_a. The equation now becomes

$$\mathrm{pH} = \mathrm{p}K_a - \log\frac{[\mathrm{HA}]}{[\mathrm{A^-}]} \qquad \text{(Eq. 1.6)}$$

Eliminating the minus sign in front of the log of the quantity $\frac{[\mathrm{HA}]}{[\mathrm{A^-}]}$ results in an equation known as the **Henderson-Hasselbalch equation**, which mathematically describes the dissociation characteristics of weak acids (pK_a) and bases (pK_b) and the effect on pH:

$$\mathrm{pH} = \mathrm{p}K_a + \log\frac{[\mathrm{A^-}]}{[\mathrm{HA}]} \qquad \text{(Eq. 1.7)}$$

When the ratio of $[A^-]$ to [HA] is 1, the pH equals the p*K* and the buffer has its greatest buffering capacity. The dissociation constant K_a, and therefore the pK_a, remains the same for a given substance. Any changes in pH are solely due to the ratio of conjugate base $[A^-]$

concentration to weak acid [HA] concentration. Refer to Chapter 12, *Blood Gases, pH, and Buffer Systems,* for more information.

Ionic strength is another important aspect of buffers, particularly in separation techniques. **Ionic strength** is the concentration or activity of ions in a solution or buffer. Increasing ionic strength increases the ionic cloud surrounding a compound and decreases the rate of particle migration. It can also promote compound dissociation into ions effectively increasing the solubility of some salts, along with changes in current, which can also affect electrophoretic separation.

Laboratory Equipment

In today's clinical chemistry laboratory, there are many different types of equipment in use. Most manual techniques have been replaced by automation, but it is still necessary for the laboratorian to be knowledgeable in the operation and use of common laboratory equipment. The following is a brief discussion of the composition and general use of common equipment found in a clinical chemistry laboratory, including heating units, thermometers, pipettes, flasks, beakers, balances, and centrifuges.

Heating Units

Heat blocks and water baths are common heating units within the laboratory. The temperature of these heating units must be monitored daily when in use. The predominant practice for temperature measurement uses the Celsius (°C) scale; however, Fahrenheit (°F) and Kelvin (°K) scales are also used.[11] The SI designation for temperature is the Kelvin scale. **Table 1.3** gives the conversion formulas between Fahrenheit and Celsius scales, and Appendix C (found in the Navigate 2 digital component) lists the various conversion formulas.

All analytic reactions occur at an optimal temperature. Some laboratory procedures, such as enzyme determinations, require precise temperature control, whereas others work well over a wide range of temperatures. Reactions that are temperature dependent use some type of heating/cooling cell, heating/cooling block, or water/ice bath to provide the correct temperature environment. Laboratory refrigerator temperatures are often critical and need periodic verification. Thermometers can be an integral part of an instrument or need to be placed in the device for temperature maintenance and monitoring. Several types of temperature devices are currently used in the clinical laboratory, including liquid-in-glass and electronic (**thermistor**) devices. Regardless of which type is being used, all temperature-reading devices must be calibrated for accuracy. Liquid-in-glass thermometers use a colored liquid (red or other colored material), encased in plastic or glass, measuring temperatures between 20°C and 400°C. Visual inspection of the liquid-in-glass thermometer should reveal a continuous line of liquid, free from separation or bubbles. If separation or bubbles are present, then replace the thermometer.

Liquid-in-glass thermometers should be calibrated against a NIST-certified or NIST-traceable thermometer for critical laboratory applications.[11] NIST has an SRM thermometer with various calibration points (0°C, 25°C, 30°C, and 37°C) for use with liquid-in-glass thermometers. Gallium, another SRM, has a known melting point and can also be used for thermometer verification.

As automation advances and miniaturizes, the need for an accurate, fast-reading electronic thermometer (thermistor) has increased and is now routinely incorporated in many devices. The advantages of a thermistor over the more traditional liquid-in-glass thermometers are size and millisecond response time. Similar to the liquid-in-glass thermometers, the thermistor can be calibrated against an SRM thermometer.

Table 1.3 Common Temperature Conversions

Celsius (Centigrade) to Fahrenheit	°C (9/5) + 32 (multiply Celsius temperature by 9; divide the answer by 5, then add 32)
Fahrenheit to Celsius (Centigrade)	(°F – 32)5/9 (subtract 32 and divide the answer by 9; then multiply that answer by 5)

Glassware and Plasticware

Until recently, laboratory supplies (e.g., pipettes, flasks, beakers) consisted of some type of glass and could be correctly termed *glassware*. As plastic material was refined and made available to manufacturers, plastic has been increasingly used to make laboratory supplies. A brief summary of the types and uses of glass and plastic commonly seen today in laboratories can be found in the Navigate 2 digital component. Regardless of design, most laboratory supplies must satisfy certain tolerances of accuracy

and fall into two classes of precision tolerance, either Class A or Class B as given by ASTM.[12,13] Those that satisfy Class A ASTM precision criteria are stamped with the letter "A" on the glassware and are preferred for laboratory applications. Class B glassware generally have twice the tolerance limits of Class A, even if they appear identical, and are often found in student laboratories where durability is needed. Vessels holding or transferring liquid are designed either *to contain* (TC) or *to deliver* (TD) a specified volume. The major difference is that TC devices do not deliver the volume measured when the liquid is transferred into a container, whereas the TD designation means that the labware will deliver the amount measured.

Glassware used in the clinical laboratory usually fall into one of the following categories: Kimax/Pyrex (borosilicate), Corex (aluminosilicate), high silica, Vycor (acid and alkali resistant), low actinic (amber colored), or flint (soda lime) glass used for disposable material.[14] Glassware routinely used in clinical chemistry should consist of high thermal borosilicate or aluminosilicate glass. The manufacturer is the best source of information about specific uses, limitations, and accuracy specifications for glassware.

Plasticware is beginning to replace glassware in the laboratory setting; high resistance to corrosion and breakage, as well as varying flexibility, has made plasticware appealing. Relatively inexpensive, it allows most items to be completely disposable after each use. The major types of resins frequently used in the clinical chemistry laboratory are polystyrene, polyethylene, polypropylene, Tygon, Teflon, polycarbonate, and polyvinyl chloride. Again, the individual manufacturer is the best source of information concerning the proper use and limitations of any plastic material.

In most laboratories, glass or plastic that is in direct contact with biohazardous material is usually disposable. If not, it must be decontaminated according to appropriate protocols. Should the need arise, cleaning of glass or plastic may require special techniques. Immediately rinsing glass or plastic supplies after use, followed by washing with a detergent designed for cleaning laboratory supplies and several distilled water rinses, may be sufficient. Presoaking glassware in soapy water is highly recommended whenever immediate cleaning is impractical. Many laboratories use automatic dishwashers and dryers for cleaning. Detergents and temperature levels should be compatible with the material and the manufacturer's recommendations. To ensure that all detergent has been removed from the labware, multiple rinses with appropriate grade water is recommended. Check the pH of the final rinse water and compare it with the initial pH of the prerinse water. Detergent-contaminated water will have a more alkaline pH as compared with the pH of the prerinse water. Visual inspection should reveal spotless vessel walls. Any biologically contaminated labware should be disposed of according to the precautions followed by the laboratory.

Some determinations, such as those used in assessing heavy metals or assays associated with molecular testing, require scrupulously clean or disposable glassware. Other applications may require plastic rather than glass because glass can absorb metal ions. It is suggested that disposable glass and plastic be used whenever possible.

Dirty reusable pipettes should be placed, with the pipette tips up, immediately in a specific pipette soaking/washing/drying container. This container should have soapy water high enough to cover the entire pipette. For each final water rinse, fresh reagent-grade water should be used; if possible, designate a pipette container for final rinses only. Cleaning brushes are available to fit almost any size glassware and are recommended for any articles that are washed routinely.

Although plastic material is often easier to clean because of its nonwettable surface, it may not be appropriate for some applications involving organic solvents or autoclaving. Brushes or harsh abrasive cleaners should not be used on plasticware. Many initial cleaning procedures, described in Appendix J (found in the Navigate 2 digital component), can be adapted for plasticware. Ultrasonic cleaners can help remove debris coating the surfaces of glass or plasticware. Properly cleaned laboratory glass and plasticware should be completely dried before using.

Laboratory Glassware

Flasks, beakers, and graduated cylinders are used to hold solutions. Volumetric and Erlenmeyer flasks are two types of containers in general use in the clinical laboratory.

A **volumetric** flask is calibrated to hold one exact volume of liquid (TC). The flask has a round, lower portion with a flat bottom and a long, thin neck with an etched calibration line. Volumetric flasks are used to bring a given reagent to its final volume with the recommended diluent. When bringing the bottom of the meniscus to the calibration mark, a pipette should be used for adding the final drops of diluent to ensure

maximum control is maintained and the calibration line is not missed.

Erlenmeyer flasks and **Griffin beakers** are designed to hold different volumes rather than one exact amount. Because Erlenmeyer flasks and Griffin beakers are often used in reagent preparation, flask size, chemical inertness, and thermal stability should be considered. The Erlenmeyer flask has a wide bottom that gradually evolves into a smaller, short neck. The Griffin beaker has a flat bottom, straight sides, and an opening as wide as the flat base, with a small spout in the lip.

Graduated cylinders are long, cylindrical tubes usually held upright by an octagonal or circular base. The cylinder has horizontal calibration marks and is used to measure volumes of liquids. Graduated cylinders do not have the accuracy of volumetric labware. The sizes routinely used are 10, 25, 50, 100, 500, 1000, and 2000 mL.

All laboratory glassware used for critical measurements should be Class A whenever possible to maximize accuracy and precision and thus decrease calibration time (**Figure 1.1** illustrates representative laboratory glassware).

Pipettes

Pipettes are a type of laboratory equipment used to transfer liquids; they may be reusable or disposable. Although pipettes may transfer any volume, they are usually used for volumes of 20 mL or less; larger volumes are usually transferred or dispensed using automated pipetting devices. **Table 1.4** outlines the pipette classification.

Similar to other laboratory equipment, pipettes are designed to contain (TC) or to deliver (TD) a particular volume of liquid. The major difference is the amount of liquid needed to wet the interior surface of the pipette and the amount of any residual liquid left in the pipette tip. Most manufacturers stamp *TC* or *TD* near the top of the pipette to alert the user as to the type of pipette. Like other TC-designated labware, a TC pipette holds or contains a particular volume but does not dispense that exact volume, whereas a TD pipette will dispense the volume indicated.

When using either pipette, the tip must be immersed in the intended transfer liquid to a level that will allow the tip to remain in solution after the volume of liquid has entered the pipette—without touching the vessel walls. The pipette is held upright, not at an angle (**Figure 1.2**). Using a pipette bulb or similar device, a slight suction is applied to the opposite end until the liquid enters the pipette and the meniscus is brought above the desired graduation line (**Figure 1.3A**), and suction is then stopped. While the meniscus level is held in place, the pipette tip is raised slightly out of the solution and wiped with a laboratory tissue to remove any adhering liquid. The liquid is allowed to drain until the bottom of the meniscus touches the desired calibration mark (**Figure 1.3B**). With the pipette held in a vertical position and the tip against the side of the receiving vessel, the pipette contents are allowed to drain into the vessel (e.g., test tube, cuvette, or flask). A *blowout pipette* has a continuous etched ring or two small,

Table 1.4 Pipette Classification

I.	Design
	A. To contain (TC)
	B. To deliver (TD)
II.	Drainage characteristics
	A. Blowout
	B. Self-draining
III.	Type
	A. Measuring or graduated
	1. Serologic
	2. Mohr
	3. Bacteriologic
	4. Ball, Kolmer, or Kahn
	5. Micropipette
	B. Transfer
	1. Volumetric
	2. Ostwald-Folin
	3. Pasteur pipettes
	4. Automatic macropipettes or micropipettes

Figure 1.1 Laboratory glassware.

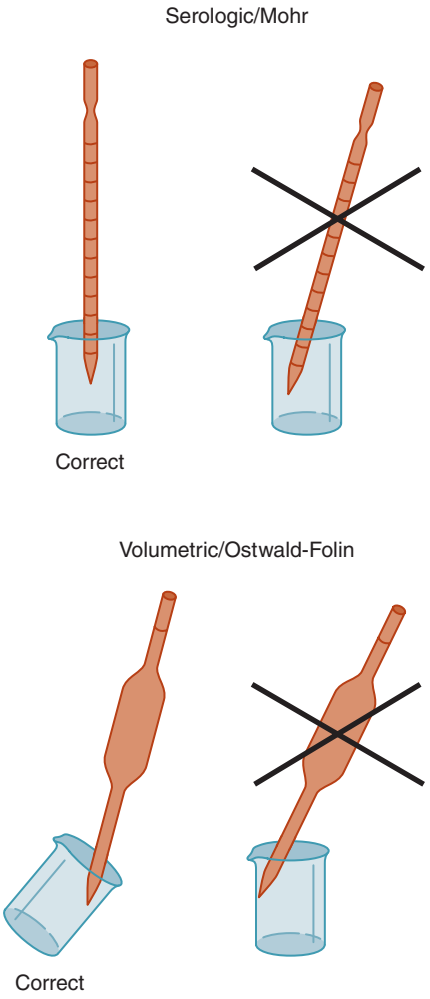


Figure 1.2 Correct and incorrect pipette positions.

close, continuous rings located near the top of the pipette. This means that the last drop of liquid should be expelled into the receiving vessel. Without these markings, a pipette is *self-draining*, and the user allows the contents of the pipette to drain by gravity. The tip of the pipette should not be in contact with the accumulating fluid in the receiving vessel during drainage. With the exception of the Mohr pipette, the tip should remain in contact with the side of the vessel for several seconds after the liquid has drained. The pipette is then removed (Figure 1.2).

Measuring or graduated pipettes are capable of dispensing several different volumes. Measuring pipettes are used to transfer reagents or make dilutions and can be used to repeatedly transfer a particular solution. The markings at the top of a measuring or graduated pipette indicate the volume(s) it is designed to dispense. Because the graduation lines located on the pipette may vary, the increments will be indicated on the top of each pipette. For example, a 5-mL pipette can be used to

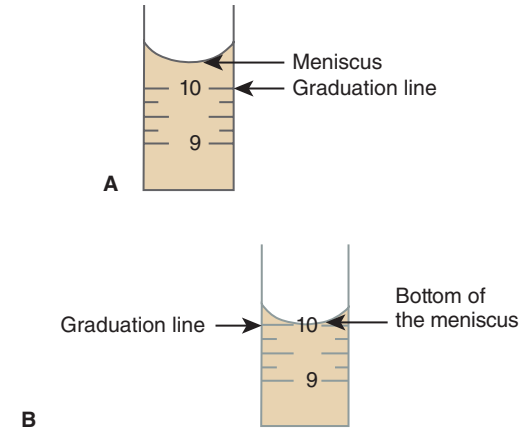


Figure 1.3 Pipetting technique. **(A)** Meniscus is brought above the desired graduation line. **(B)** Liquid is allowed to drain until the bottom of the meniscus touches the desired calibration mark.

measure 5, 4, 3, 2, or 1 mL of liquid, with further graduations between each milliliter. The pipette is designated as 5 in 1/10 increments (**Figure 1.4**) and could deliver any volume in tenths of a milliliter, up to 5 mL. Another pipette, such as a 1-mL pipette, may be designed to dispense 1 mL and have subdivisions of hundredths of a milliliter. The subgroups of measuring or graduated pipettes are Mohr, serologic, and micropipettes. A *Mohr pipette* does not have graduations to the tip. It is a self-draining pipette, but the tip should not be allowed to touch the vessel while the pipette is draining. A *serologic pipette* has graduation marks to the tip and is generally a blowout pipette. A *micropipette* is a pipette with a total holding volume of less than 1 mL; it may be designed as either a Mohr or a serologic pipette.

Transfer pipettes are designed to dispense one volume without further subdivisions. Ostwald-Folin pipettes are used with biologic fluids having a viscosity greater than that of water. They are blowout pipettes, indicated by two etched, continuous rings at the top. The volumetric pipette is designed to dispense or transfer aqueous solutions and is always self-draining. The bulb-like enlargement in the pipette stem easily

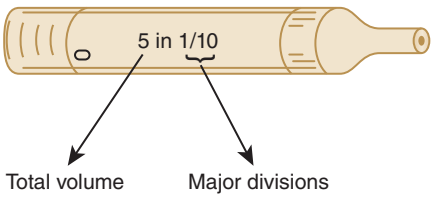


Figure 1.4 Volume indication of a pipette.

identifies the volumetric pipette. This type of pipette usually has the greatest degree of accuracy and precision and should be used when diluting standards, calibrators, or quality control material. They should only be used once prior to cleaning. Disposable transfer pipettes may or may not have calibration marks and are used to transfer solutions or biologic fluids without consideration of a specific volume. These pipettes should not be used in any quantitative analytic techniques (**Figure 1.5**).

The *automatic pipette* is the most routinely used pipette in today's clinical chemistry laboratory. Automatic pipettes come in a variety of types including fixed volume, variable volume, and multichannel. The term *automatic*, as used here, implies that the mechanism that draws up and dispenses the liquid is an integral part of the pipette. It may be a fully automated/self-operating, semiautomatic, or completely manually operated device. Automatic and semiautomatic pipettes have many advantages, including safety, stability, ease of use, increased precision, the ability to save time, and less cleaning required because the pipette tips are

Figure 1.5 Disposable transfer pipettes.

disposable. **Figure 1.6** illustrates many common automatic pipettes. A pipette associated with only one volume is termed a *fixed* volume, and models able to select different volumes are termed *variable*; however, only one volume may be used at a time. The available range of pipette volumes is 1 µL to 5000 mL. A pipette with a capability of

Figure 1.6 **(A)** Adjustable volume pipette. **(B)** Fixed volume pipette with disposable tips. **(C)** Multichannel pipette. **(D)** Multichannel pipette in use.

less than 1 mL is considered a *micropipette*, and a pipette that dispenses greater than 1 mL is called an *automatic macropipette*. Multichannel pipettes are able to attach multiple pipette tips to a single handle and can then be used to dispense a fixed volume of fluid to multiple wells, such as to a multiwell microtiter plate. In addition to classification by volume delivery amounts, automatic pipettes can also be categorized according to their mechanism: air-displacement, positive displacement, and dispenser pipettes. An *air-displacement pipette* relies on a piston for creating suction to draw the sample into a disposable tip that must be changed after each use. The piston does not come in contact with the liquid. A *positive-displacement pipette* operates by moving the piston in the pipette tip or barrel, much like a hypodermic syringe. It does not require a different tip for each use. Because of carryover concerns, rinsing and blotting between samples may be required. *Dispensers* and *dilutor/dispensers* are automatic pipettes that obtain the liquid from a common reservoir and dispense it repeatedly. The dispensing pipettes may be bottle-top, motorized, handheld, or attached to a dilutor. The dilutor often combines sampling and dispensing functions. Many automated pipettes use a wash between samples to eliminate carryover problems. However, to minimize carryover contamination with manual or semiautomatic pipettes, careful wiping of the tip may remove any liquid that adhered to the outside of the tip before dispensing any liquid. Care should be taken to ensure that the orifice of the pipette tip is not blotted, drawing sample from the tip. Another precaution in using manually operated semiautomatic pipettes is to move the plunger in a continuous and steady manner. Pipettes should be operated according to the manufacturer's directions.

Disposable, one-use pipette tips are designed for use with air-displacement pipettes. The laboratorian should ensure that the pipette tip is seated snugly onto the end of the pipette and free from any deformity. Plastic tips used on air-displacement pipettes can vary. Different brands can be used for one particular pipette, but they do not necessarily perform in an identical manner. Tips for positive-displacement pipettes are made of straight columns of glass or plastic. These tips must fit snugly to avoid carryover and can be used repeatedly without being changed after each use. As previously mentioned, these devices may need to be rinsed and dried between samples to minimize carryover.

Class A pipettes do not need to be recalibrated by the laboratory. Automatic pipetting devices, as well as non–Class A materials, do need recalibration.[15,16] Calibration of pipettes is done to verify accuracy and precision of the device and may be required by the laboratory's accrediting agency. A gravimetric method (see the Navigate 2 digital component resources for this procedure) can accomplish this task by delivering and weighing a solution of known specific gravity, such as water. A currently calibrated analytic balance and at least Class 2 weights should be used.[17] Deviation from the chosen volume is calculated based on the type of pipette tested. Pipettes that fall outside of the maximum allowable error will need to be adjusted following the manufacturer's instructions. Although gravimetric validation is the most desirable method,[18,19] pipette calibration may also be accomplished by using photometric methods, particularly for automatic pipetting devices. When a spectrophotometer is used, the molar absorptivity of a compound, such as potassium dichromate, is obtained. After an aliquot of diluent is pipetted, the change in concentration will reflect the volume of the pipette. Another photometric technique used to assess pipette accuracy compares the absorbances of dilutions of potassium dichromate, or another colored liquid with appropriate absorbance spectra, using Class A volumetric labware versus equivalent dilutions made with the pipetting device.

These calibration techniques are time consuming and impractical for use in daily checks. It is recommended that pipettes be checked initially and subsequently three or four times per year, or as dictated by the laboratory's accrediting agency. Many companies offer calibration services; the one chosen should also satisfy any accreditation requirements. A quick, daily check for many larger volume automatic pipetting devices involves the use of volumetric flasks. For example, a bottle-top dispenser that routinely delivers 2.5 mL of reagent may be checked by dispensing four aliquots of the reagent into a 10-mL Class A volumetric flask. The bottom of the meniscus should meet with the calibration line on the volumetric flask.

Syringes

Syringes are sometimes used for transfer of small volumes (< 500 µL) in blood gas analysis or in separation techniques such as chromatography or electrophoresis (**Figure 1.7**). The syringes are glass and have fine barrels. The plunger is often made of a fine piece of wire. Tips are not used when syringes are used for injection of sample into a gas chromatographic or

Figure 1.7 Microliter glass syringe.

high-pressure liquid chromatographic system. In electrophoresis work, however, disposable Teflon tips may be used.

Desiccators and Desiccants

Many compounds combine with water molecules to form loose chemical crystals. The compound and the associated water are called a **hydrate**. When the water of crystallization is removed from the compound, it is said to be **anhydrous**. Substances that take up water on exposure to atmospheric conditions are called **hygroscopic**. Materials that are very hygroscopic can remove moisture from the air as well as from other materials. These materials make excellent drying substances and are sometimes used as **desiccants** (drying agents) to keep other chemicals from becoming hydrated. Closed and sealed containers that include desiccant material are referred to as desiccators and may be used to store more hygroscopic substances. Many sealed packets or shipping containers, often those that require refrigeration, include some type of small packet of desiccant material to prolong storage.

Balances

A properly operating balance is essential in producing high-quality reagents and standards. However, because many laboratories discontinued in-house reagent preparation, balances may no longer be as widely used. Balances are classified according to their design, number of pans (single or double), and whether they are mechanical or electronic or classified by operating ranges.

Analytic and electronic balances are currently the most popular in the clinical laboratory. Analytic

Figure 1.8 Analytic balance.

balances (**Figure 1.8**) are required for the preparation of any primary standard. It has a single pan enclosed by sliding transparent doors, which minimize environmental influences on pan movement, tared weighing vessel, and sample. An optical scale allows the operator to visualize the mass of the substance. The weight range for many analytic balances is from 0.01 mg to 160 g.

Electronic balances (**Figure 1.9**) are single-pan balances that use an electromagnetic force to counterbalance the weighed sample's mass. Their measurements equal the accuracy and precision of any available mechanical balance, with the advantage of a fast response time (< 10 seconds).

Test weights used for calibrating balances should be selected from the appropriate ANSI/ASTM Classes 1 through 4.[19] Weighing instruments will need to be calibrated and adjusted periodically due to wear and tear from frequent use. Mechanisms for automatic adjustments are built into many newer instruments. These instruments will test and adjust the sensitivity of the device. Periodic verification is still necessary to assure the performance of that device. The frequency of calibration is dictated by the

Figure 1.9 Electronic top-loading balance.

accreditation/licensing guidelines for a specific laboratory. Balances should be kept clean and be located in an area away from heavy traffic, large pieces of electrical equipment, and open windows to prevent inaccurate readings. The level checkpoint should always be corrected before weighing occurs.

Centrifuges

Centrifugation is a process in which centrifugal force is used to separate serum or plasma from the blood cells as the blood samples are being processed; to separate a supernatant from a precipitate during an analytic reaction; to separate two immiscible liquids, such as a lipid-laden sample; or to expel air. When samples are not properly centrifuged, small fibrin clots and cells can cause erroneous results during analysis. The centrifuge separates the mixture based on mass and density of the component parts. It consists of a head or rotor, carriers, or shields that are attached to the vertical shaft of a motor or air compressor and enclosed in a metal covering. The centrifuge always has a lid, with new models having a locking lid for safety. Many models include a brake or a built-in tachometer, which indicates speed, and some centrifuges are refrigerated.

Centrifugal force depends on three variables: mass, speed, and radius. The speed is expressed in revolutions per minute (rpm), and the centrifugal force generated is expressed in terms of relative centrifugal force (RCF) or gravities (*g*). The speed of the centrifuge is related to the RCF by the following equation:

$$RCF = 1.118 \times 10^{-5} \times r \times (rpm)^2$$

where 1.118×10^{-5} is a constant, determined from the angular velocity, and r is the radius in centimeters, measured from the center of the centrifuge axis to the bottom of the test tube shield or bucket. Centrifuge classification is based on several criteria, including benchtop (**Figure 1.10A**) or floor model; refrigeration; rotor head (e.g., fixed angle, hematocrit, cytocentrifuge, swinging bucket [**Figure 1.10B**], or angled); or maximum speed attainable (i.e., ultracentrifuge).

Centrifuge maintenance includes daily cleaning of any spills or debris, such as blood or glass, and

Figure 1.10 **(A)** Benchtop centrifuge. **(B)** Swinging-bucket rotor.

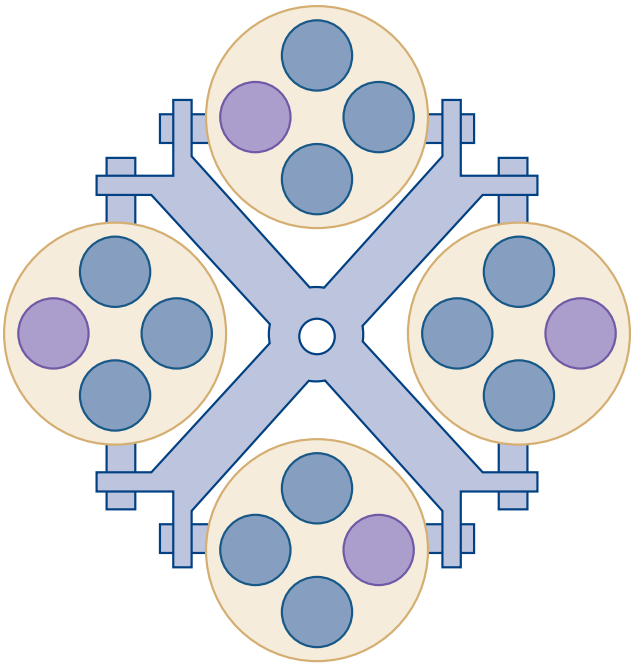


Figure 1.11 Properly balanced centrifuge. *Colored circles* represent counterbalanced positions for sample tubes.

ensuring that the centrifuge is properly balanced and free from any excessive vibrations. Balancing the centrifuge load is critical (**Figure 1.11**). Many newer centrifuges will automatically decrease their speed if the load is not evenly distributed, but more often, the centrifuge will shake and vibrate or make more noise than expected. A centrifuge needs to be balanced by equalizing both the volume and weight distribution across the centrifuge head. Many laboratories will have "balance" tubes of routinely used volumes and tube sizes, which can be used to match those from patient samples. A good rule of thumb is one

Figure 1.12 Properly loaded centrifuge.

of even placement and "opposition" (**Figure 1.12**). Exact positioning of tubes depends on the design of the centrifuge holders.

The centrifuge cover should remain closed until the centrifuge has come to a complete stop to avoid any aerosol production. It is recommended that the timer, brushes (if present), and speed be periodically checked. The brushes, which are graphite bars attached to a retainer spring, create an electrical contact in the motor. The specific manufacturer's service manual should be consulted for details on how to change brushes and on lubrication requirements. The speed of a centrifuge is easily checked using a tachometer or strobe light. The hole located in the lid of many centrifuges is designed for speed verification using these devices but may also represent an aerosol biohazard if the hole is uncovered. Accreditation agencies require periodic verification of centrifuge speeds.

CASE STUDY 1.2, PART 2

Recall Mía, the new graduate.

1. How should Mía place the chemistry tubes in the centrifuge?
2. If the centrifuge starts vibrating, what is the first troubleshooting step Mía should take?
3. If the rubber cap came off the tube during centrifugation, how should Mía decontaminate the centrifuge?

Laboratory Mathematics and Calculations

Significant Figures

Significant figures are the minimum number of digits needed to express a particular value in scientific notation without loss of accuracy. There are several rules in regard to identifying significant figures:

1. All nonzero numbers are significant (1, 2, 3, 4, 5, 6, 7, 8, 9).
2. All zeros between nonzero numbers are significant.
3. All zeros to the right of the decimal are not significant when followed by a nonzero number.
4. All zeros to the left of the decimal are not significant.

The number 814.2 has four significant figures, because in scientific notation, it is written as 8.142×10^2. The number 0.000641 has three significant figures, because the scientific notation expression for this value is 6.41×10^{-4}. The zeros to the right of the decimal preceding the nonzero digits are merely holding decimal places and are not needed to properly express the number in scientific notation. However, by convention, zeros following a decimal point are considered significant. For example, 10.00 has four significant figures. The zeros to the right of the decimal indicate the precision of this value.

Logarithms

Logarithms are the inverse of exponential functions and can be related as such:

$$x = A^B \text{ or } B = \log_A (x)$$

This is then read as B is the log base A of X, where B must be a positive number, A is a positive number, and A cannot be equal to 1. Calculators with a log function do not require conversion to scientific notation.

To determine the original number from a log value, the process is performed in reverse. This process is termed the *antilogarithm* or *antilog* as it is the inverse of the logarithm. Most calculators require using an inverse or secondary/shift function when entering this value. If given a log of 3.1525, the resulting value is 1.424×10^3 on the base 10 system. Consult the specific manufacturer's directions of the calculator to become acquainted with the proper use of these functions.

pH (Negative Logarithms)

In certain circumstances, the laboratorian may work with negative logs. Such is the case with pH or pK_a. As previously stated, the pH of a solution is defined as the negative log of the hydrogen ion concentration. The following is a convenient formula to determine the negative logarithm when working with pH or pK_a:

$$\frac{\text{pH}}{pK_a} = x - \log N \qquad \textbf{(Eq. 1.11)}$$

where x is the negative exponent base 10 expressed and N is the decimal portion of the scientific notation expression.

For example, if the hydrogen ion concentration of a solution is 5.4×10^{-6}, then $x = 6$ and $N = 5.4$. Substitute this information into **Equation 1.11**, and it becomes

$$\text{pH} = 6 - \log 5.4 \qquad \textbf{(Eq. 1.12)}$$

The logarithm of N (5.4) is equal to 0.7324, or 0.73. The pH becomes

$$\text{pH} = 6 - 0.73 = 5.27 \qquad \textbf{(Eq. 1.13)}$$

The same formula can be applied to obtain the hydrogen ion concentration of a solution when only the pH is given. Using a pH of 5.27, the equation becomes

$$5.27 = x - \log N \qquad \textbf{(Eq. 1.14)}$$

In this instance, the x term is always the next largest whole number. For this example, the next largest whole number is 6. Substituting for x, the equation becomes

$$5.27 = 6 - \log N \qquad \textbf{(Eq. 1.15)}$$

A shortcut is to simply subtract the pH from x (6 – 5.27 = 0.73) and take the antilog of that answer 5.73. The final answer is 5.73×10^{-6}. Note that rounding, while allowed, can alter the answer. A more algebraically correct approach follows in **Equations 1.16** through **1.18**. Multiply all the variables by –1:

$$\begin{aligned}(-1)(5.27) &= (-1)(6) - (-1)(\log N)\\ -5.27 &= -6 + \log N\end{aligned} \qquad \textbf{(Eq. 1.16)}$$

Solve the equation for the unknown quantity by adding a positive 6 to both sides of the equal sign, and the equation becomes

$$6 - 5.27 = \log N$$
$$0.73 = \log N \qquad \text{(Eq. 1.17)}$$

The result is 0.73, which is the antilogarithm value of N, which is 5.37, or 5.4:

$$\text{Antilog } 0.73 = N;\ N = 5.37 = 5.4 \qquad \text{(Eq. 1.18)}$$

The hydrogen ion concentration for a solution with a pH of 5.27 is 5.4×10^{-6}. Many scientific calculators have an inverse function that allows for more direct calculation of negative logarithms.

Concentration

A description of each concentration term is provided at the beginning of this chapter. The following discussion focuses on the basic mathematical expressions needed to prepare reagents of a stated concentration.

Percent Solution

A percent solution is determined in the same manner regardless of whether weight/weight, volume/volume, or weight/volume units are used. *Percent* implies "parts per 100," which is represented as percent (%) and is independent of the molecular weight of a substance.

Example 1.1: Weight/Weight (w/w)

To make up 250 g of a 5% aqueous solution of hydrochloric acid (using 12 M HCl), multiply the total amount by the percent expressed as a decimal. The 5% aqueous solution can be expressed as

$$5\% = \frac{5}{100} = 0.050 \qquad \text{(Eq. 1.19)}$$

Therefore, the calculation becomes

$$0.050 \times 250 \text{ g} = 12.5 \text{ g of 12M HCl} \qquad \text{(Eq. 1.20)}$$

Another way of arriving at the answer is to set up a ratio so that

$$\text{Desired solution concentration} = \text{Final product of 12 M HCl}$$
$$\frac{5}{100} = \frac{x}{250} \qquad \text{(Eq. 1.21)}$$
$$x = 12.5$$

Example 1.2: Weight/Volume (w/v)

The most frequently used term for a percent solution is weight per volume, which is often expressed as grams per 100 mL of the diluent. To make up 1000 mL of a 10% (w/v) solution of NaOH, use the preceding approach. Restate the w/v as a fraction:

$$10\% = \frac{10 \text{ g}}{100 \text{ mL}} = 0.10$$

Next, set up a ratio and solve for x

$$\frac{10\text{g}}{100 \text{ mL}} = \frac{x}{1000 \text{ mL}} \qquad \text{(Eq. 1.22)}$$
$$x = 100\text{g}$$

Lastly, add 100 g of 10% NaOH to a 1000-mL volumetric Class A flask and add sufficient volume of reagent-grade water to the calibration mark.

Example 1.3: Volume/Volume (v/v)

If both chemicals in a solution are in the liquid form, the volume per unit volume is used to give the volume of solute present in 100 mL of solution. To make up 50 mL of a 2% (v/v) concentrated hydrochloric acid solution, a similar approach is used. The (v/v) is restated as a fraction:

$$2\% = \frac{2 \text{ mL}}{100 \text{ mL}} = 0.02$$

Then, the calculation becomes

$$0.02 \times 50 \text{ mL} = 1 \text{ mL}$$

or using a ratio

$$\frac{2 \text{ mL}}{100 \text{ mL}} = \frac{x}{50 \text{ mL}} \qquad \text{(Eq. 1.23)}$$
$$x = 1 \text{ mL}$$

Therefore, add 40 mL of reagent-grade water to a 50-mL Class A volumetric flask, add 1 mL of concentrated HCl, mix, and dilute up to the calibration mark with reagent-grade water. Remember, always add acid to water!

Molarity

Molarity (M) is routinely expressed in units of moles per liter (mol/L) or sometimes millimoles per milliliter (mmol/mL). Remember that 1 mol of a substance is equal to the gram molecular weight (gmw) of that

substance. When trying to determine the amount of substance needed to yield a particular concentration, initially decide what final concentration units are needed. For molarity, the final units will be moles per liter (mol/L) or millimoles per milliliter (mmol/mL). The second step is to consider the existing units and the relationship they have to the final desired units. Essentially, try to put as many units as possible into like terms and arrange so that the same units cancel each other out, leaving only those needed in the final answer. To accomplish this, it is important to remember what units are used to define each concentration term. It is key to understand the relationship between molarity (moles/liter), moles, and gmw. While molarity is given in these examples, the approach for molality is the same except that one molal is expressed as one mole of solute per kilogram of solvent. For water, one kilogram is proportional to one liter, so molarity and molality are equivalent.

Example 1.4

How many grams are needed to make 1 L of a 2 M solution of HCl?

Step 1: What *units* are needed in the final answer? *Answer:* Grams per liter (g/L).

Step 2: Assess other mass/volume terms used in the problem. In this case, moles are also needed for the calculation: How many grams are equal to 1 mole? The gmw of HCl, which can be determined from the periodic table, will be equal to 1 mole. For HCl, the gmw is 36.5g/mol, so the equation may be written as

$$\frac{36.5\ \text{g HCl}}{\cancel{\text{mol}}} \times \frac{2\ \cancel{\text{mol}}}{\text{L}} = \frac{73\ \text{g HCl}}{\text{L}} \qquad \textbf{(Eq. 1.24)}$$

Cancel out like units, and the final units should be grams per liter. In this example, 73 g HCl per liter is needed to make a 2 M solution of HCl.

Example 1.5

A solution of NaOH is contained within a Class A 1-L volumetric flask filled to the calibration mark. The content label reads 24 g of NaOH. Determine the molarity.

Step 1: What *units* are needed? *Answer:* Moles per liter (mol/L).

Step 2: The units that exist are grams and L. NaOH may be expressed as moles and grams. The calculated gmw of NaOH is 40 g/mol. Rearrange the equation so that grams can be canceled, and the remaining units reflect those needed in the answer, which are mole/L.

Step 3: The equation becomes

$$\frac{24\ \cancel{\text{g}}\ \text{NaOH}}{\text{L}} \times \frac{1\text{mol}}{40\ \cancel{\text{g}}\ \text{NaOH}} = 0.6\frac{\text{mol}}{\text{L}} \qquad \textbf{(Eq. 1.25)}$$

By canceling out like units and performing the appropriate calculations, the final answer of 0.6 M or 0.6 mol/L.

Example 1.6

Prepare 250 mL of a 4.8 M solution of HCl.

Step 1: Start with what is given: 4.8 moles/L

Step 2: Determine the gmw of HCl ([H = 1] + [Cl = 35.5] = 36.5 g/mol)

Step 3: Set up the equation, cancel out like units, and perform the appropriate calculations:

$$\frac{4.8\ \text{mol}}{1\ \text{L}} \times \frac{36.5\ \text{g HCl}}{1\ \text{mol}} \times \frac{1\ \text{L}}{1000\ \text{mL}} \times 250\ \text{mL} = 43.8\ \text{g HCl} \qquad \textbf{(Eq. 1.26)}$$

In a 250-mL Class A volumetric flask, add 200 mL of reagent-grade water. Add 43.8 g of HCl and mix. Dilute up to the calibration mark with reagent-grade water.

Normality

Normality (N) is expressed as the number of equivalent weights per liter (Eq/L) or milliequivalents per milliliter (mmol/mL). Equivalent weight is equal to gmw divided by the valence (*V*). Normality has often been used in acid–base calculations because an equivalent weight of a substance is also equal to its combining weight (or the weight that will combine with or displace 1 mole of hydrogen). Another advantage in using equivalent weight is that an equivalent weight of one substance is equal to the equivalent weight of any other chemical.

Example 1.7

Give the equivalent weight, in grams, for each substance listed below.

1. NaCl (gmw = 58 g/mol, valence = 1)

$$\frac{58\ \text{g}}{\text{L}} = 58\ \text{g per equivalent weight} \qquad \textbf{(Eq. 1.27)}$$

2. H_2SO_4 (gmw = 98 g/mol, valence = 2)

$$\frac{98\text{ g}}{2} = 49\text{ g per equivalent weight} \qquad \textbf{(Eq. 1.28)}$$

Example 1.8

What is the normality of a 500 mL solution that contains 7 g of H_2SO_4?

Step 1: What units are needed? *Answer:* Normality expressed as equivalents per liter (Eq/L).

Step 2: Start with what is given: 7 g/500 mL

Step 3: Calculate the gmw of H_2SO_4 (98 g/mol) and determine the valence (2)

Step 4: Add a conversion factor to convert mL to L (1000 mL/1 L)

Step 5: Cancel out like terms and calculate the result in Eq/L.

This equation is:

$$\frac{7\text{ g } H_2SO_4}{500\text{ mL}} \times \frac{1\text{ Eq}}{49\text{ g } H_2SO_4} \times \frac{1000\text{ mL}}{1\text{ L}} \qquad \textbf{(Eq. 1.29)}$$
$$= 0.285\text{ Eq/L} = 0.285\text{ N}$$

Example 1.9

What is the normality of a 0.5 M solution of H_2SO_4? Continuing with the previous approach, the final equation is

$$\frac{0.5\text{ mol } H_2SO_4}{L} \times \frac{98\text{ g } H_2SO_4}{\text{mol } H_2SO_4} \times \frac{1\text{ Eq } H_2SO_4}{49\text{ g } H_2SO_4}$$
$$= 1\text{ Eq/L} = 1\text{N} \qquad \textbf{(Eq. 1.30)}$$

When converting between molarity and normality, the following conversion formula may be applied:

$$M \times V = N \qquad \textbf{(Eq. 1.31)}$$

where V is the valence of the compound. Using this formula, **Example 1.9** becomes

$$0.5\text{ M} \times 2 = 1\text{N} \qquad \textbf{(Eq. 1.32)}$$

Example 1.10

What is the molarity of a 2.5 N solution of HCl? This problem may be solved in several ways. One way is to use the stepwise approach in which existing units are exchanged for units needed. The equation is

$$\frac{2.5\text{ Eq HCl}}{\text{L}} \times \frac{36.5\text{ g} \times \text{HCl}}{1\text{ Eq}} \times \frac{1\text{ mol HCl}}{36.5\text{ g HCl}} \qquad \textbf{(Eq. 1.33)}$$
$$= 2.5\text{ mol/L HCl}$$

The second approach is to use the normality-to-molarity conversion formula. The equation now becomes

$$M \times V = 2.5\text{ N}$$
$$V = 1$$
$$M = \frac{2.5\text{ N}}{1} = 2.5\text{ N} \qquad \textbf{(Eq. 1.34)}$$

When the valence of a substance is 1, the molarity will equal the normality. As previously mentioned, normality either equals or is greater than the molarity.

Although there are various methods to calculate laboratory mathematical problems, the technique of using conversion factors and canceling like units is used in most clinical chemistry calculation, regardless of whether the problem requests molarity and normality or exchanging one concentration term for another. However, it is necessary to recall the interrelationship between all the units in the expression.

Specific Gravity

Density is expressed as mass per unit volume of a substance. The **specific gravity** is the ratio of the density of a material when compared with the density of pure water at a given temperature and allows the laboratorian a means of expressing density in terms of volume. The units for density are grams per milliliter (g/mL). Specific gravity is often used with very concentrated materials, such as commercial acids (e.g., sulfuric and hydrochloric acids).

The density of a concentrated acid can also be expressed in terms of an assay or percent purity. The actual concentration is equal to the specific gravity multiplied by the assay or per*cent purity value* (expressed as a decimal) stated on the label of the container.

Example 1.11

What is the actual weight of a supply of concentrated HCl on which the label reads, specific gravity 1.19 with an assay value of 37%?

$$1.19\text{ g/mL} \times 0.37 = 0.44\text{g/mL of HCl} \qquad \textbf{(Eq. 1.35)}$$

Example 1.12

What is the molarity of this same stock solution? The final units desired are moles per liter (mol/L). The molarity of the solution is

$$\frac{0.44\ \cancel{g}\ \text{HCl}}{\cancel{mL}} \times \frac{1\ \text{mol HCl}}{36.5\ \cancel{g}\ \text{HCl}} \times \frac{1000\ \cancel{mL}}{\text{L}} \qquad \text{(Eq. 1.36)}$$

$$= 12.05\ \text{mol/L or } 12\text{M}$$

Conversions

To convert one unit into another, the same approach of canceling out like units can be applied. In some instances, a chemistry laboratory may report a given analyte using two different concentration units—for example, calcium. The recommended SI unit for calcium is millimoles per liter. The more traditional units are milligrams per deciliter (mg/dL). Again, it is important to understand the relationship between the units given and those needed in the final answer.

Example 1.13

Convert 8.2 mg/dL calcium to millimoles per liter (mmol/L). The gmw of calcium is 40 g. If there are 40 g per mol, then it follows that there are 40 mg per mmol. The final units needed are mmol/L. The equation becomes

$$\frac{8.2\ \cancel{mg}}{\cancel{dL}} \times \frac{10\ \cancel{dL}}{1\ \cancel{L}} \times \frac{1\ \text{mmol}}{40\ \cancel{mg}} = \frac{2.05\ \text{mmol}}{\text{L}} \qquad \text{(Eq. 1.37)}$$

Once again, the systematic stepwise approach of canceling similar units can be used for this conversion problem.

A frequently encountered conversion problem or, more precisely, a dilution problem occurs when a weaker concentration or different volume is needed than the stock substance available, but the *concentration terms* are the same. The following formula is used where V_1 is the volume of the first substance, C_1 is the concentration of the first substance, V_2 is the volume of the second substance, and C_2 is the concentration of the second substance:

$$V_1 \times C_1 = V_2 \times C_2 \qquad \text{(Eq. 1.38)}$$

This formula is useful only if the concentration and volume units between the substances are the *same* and if three of four variables are known.

Example 1.14

What volume is needed to make 500 mL of a 0.1 M solution of Tris buffer from a solution of 2 M Tris buffer?

Identify the known values:

Concentration of initial substance (C_1) = 2 M

Volume of the product (V_2) = 500 mL

Concentration of the product (C_2) = 0.1 M

And the equation becomes:

$$V_1 \times 2\ \text{M} = 0.1\ \text{M} \times 500\ \text{mL}$$
$$(V_1)(2\ \text{M}) = (0.1\ \text{M})(500\ \text{mL});$$
$$(V_1)(2\ \text{M}) = 50\ \text{mL}$$

$$\text{Therefore, } V_1 \frac{50\ \text{mL}}{2} = 25\ \text{mL} \qquad \text{(Eq. 1.39)}$$

It requires 25 mL of the 2 M solution to make up 500 mL of a 0.1 M solution.

This problem differs from the other conversions in that it is actually a dilution of a stock solution. While this approach will provide how much stock is needed when making the solution, the laboratorian must subtract the obtained volume value from the final desired volume to determine the amount of diluent needed, in this case 475 mL. A more involved discussion of dilution problems follows.

Dilutions

A **dilution** represents the part(s) of concentrated material to the total final volume of a solution. It consists of the parts of the substance being diluted in the total numbers of parts of the solution. In contrast, ratio refers to part substance to part substance. The most common dilution uses one part patient serum plus one part saline. This is a 1:1 **ratio** of serum to saline. It is a 1:2 *dilution* which can also be expressed as a fraction (½ dilution). After analysis, the laboratory result is multiplied by the reciprocal of the dilution (2/1) which is known as the *dilution factor*. Dilutions are required when the result is above the linearity of the assay. A dilution is an expression of concentration. Because a dilution is made by adding a more concentrated substance to a diluent, the dilution is always less concentrated than the original substance. There is an inverse relationship between the dilution factor and concentration. As the *dilution factor* increases, the *concentration decreases*. A dilution can be expressed as either a fraction or a ratio.[20]

Example 1.15

What dilution is needed to make a 100 mmol/L sodium solution from a 3000 mmol/L stock solution?

$$\frac{100\ \text{mmol}}{\text{L}} \times \frac{\text{L}}{3000\ \text{mmol}} = \frac{1}{30} \qquad \text{(Eq. 1.40)}$$

The dilution indicates 1 part stock solution is needed to make a *total volume* of 30 mL. To prepare this dilution, 1 mL of stock solution is added to 29 mL of diluent to achieve a total final volume of 30 mL. Note that the *dilution* indicates the *parts per total* amount. In making the dilution, the sum of the amount of the stock material plus the amount of the diluent must equal the *total volume*.

It is important to differentiate when a dilution or a ratio is stated within a procedure. For example, making a "1-**in**-4" dilution means adding one part stock to obtain a *total* of four parts. That is, one part of stock would be added to three parts of diluent. The dilution would be 1/4. Analyses performed on the diluted material would need to be multiplied by 4, the dilution factor, to get the final concentration. The *dilution factor* is the reciprocal of the dilution. This is very different from a procedure indicating preparation of a "1-**to**-4" ratio! In this instance, the dilution would be 1/5, where there is 1 part of serum and 4 parts of diluent. It is important that you fully understand the meaning of these expressions. Dilutions should be made using reagent-grade water, saline, or method-specific diluent. The diluted sample should be thoroughly mixed before analysis.

Example 1.16

Make 150 mL of the 100 mmol/L sodium solution using a 1:30 dilution.

Begin with the dilution and set up a ratio between the desired part (mL) to the total parts (150 mL) to determine the amount of sodium solution needed. The equation becomes

$$\frac{1}{30} = \frac{x\ \text{mL}}{150\ \text{mL}} \qquad \text{(Eq. 1.41)}$$

$$x = 5\ \text{mL}$$

Note that 5/150 reduces to the dilution of 1/30. To make up this solution, 5 mL of sodium solution is added to 145 mL of diluent, making the total volume equal to 150 mL.

Example 1.17

The formula $(V_1)(C_1) = (V_2)(C_2)$ may also be used for simple dilution calculations. This is acceptable, as long as stock volume is subtracted from the total final volume for the correct *diluent* volume.

$$(V_1)(C_1) = (V_2)(C_2)(x)(3000\ \text{mmol/L})$$

$$= (150\ \text{mL})(100\ \text{mmol/L})$$

$$x = 5$$

150 − 5 = 145 mL of diluent shouldbe added to 5 mL of stock (Eq. 1.42)

Simple Dilutions

When making a *simple dilution*, the laboratorian must decide on the total volume desired and the amount of stock to be used.

Example 1.18

A 1/10 dilution of serum can be achieved by using any of the following approaches. A ratio of 1:9—one part serum and nine parts diluent (saline) can be achieved using the following:

- 100 µL of serum added to 900 µL of saline
- 20 µL of serum added to 180 µL of saline
- 1 mL of serum added to 9 mL of saline
- 2 mL of serum added to 18 mL of saline

Note that the sum of the ratio of serum to diluent (1:9) needed to make up each dilution satisfies the dilution factor (1/10) of stock material to total volume. When thinking about the stock to diluent volume, subtract the parts of stock needed from the total volume or parts to get the number of diluent "parts" needed. Once the volume of each part, usually stock, is known, multiply the diluent parts needed to obtain the correct volume.

Example 1.19

You have a 10 g/dL stock of protein standard. You need a 2 g/dL standard. You only have 0.200 mL of 10 g/dL stock to use. The procedure requires 0.100 mL.

Solution:

$$\frac{2\ \text{g/dL}}{10\ \text{g/dL}} = \frac{1}{5} = \text{Dilution} \qquad \text{(Eq. 1.43)}$$

You will need 1 part or volume of stock in a total of 5 parts or volumes. Subtracting 1 from 5 yields that 4 parts or volumes of diluent is needed (**Figure 1.13**). In this instance, you need at least

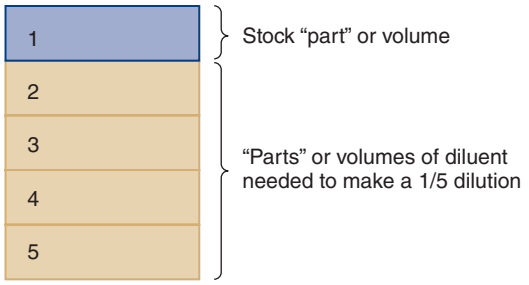


Figure 1.13 Simple dilution. Consider this diagram depicting a substance having a 1/5 dilution factor. The dilution factor represents that 1 part of stock is needed from a total of 5 parts. To prepare this dilution, you would determine the volume of 1 "part," usually the stock or patient sample. The remainder of the "parts" or total would constitute the amount of diluent needed, or four times the volume used for the stock.

0.100 mL for the procedure. You have 0.200 mL of stock. You can make the dilution in various ways, as seen in **Example 1.20**.

Example 1.20

There are several ways to prepare a 1/5 dilution having only 0.200 mL of stock and needing a total minimum volume of 0.100 mL.

- Add 0.050 mL stock (1 part) to 0.200 mL of diluent (4 parts × 0.050 mL).
- Add 0.100 mL of stock (1 part) to 0.400 mL of diluent (4 parts × 0.100 mL).
- Add 0.200 mL of stock (1 part) to 0.800 mL of diluent (4 parts × 0.200 mL).

The dilution factor is also used to determine the final concentration of a dilution by multiplying the original concentration by the inverse of the dilution factor or the dilution factor denominator when it is expressed as a fraction.

Example 1.21

Determine the concentration of a 200 µmol/mL human chorionic gonadotropin (hCG) standard that was diluted 1/50. This value is obtained by multiplying the original concentration, 200 µmol/mL hCG, by the dilution factor, 1/50. The result is 4 µmol/mL hCG. Quite often, the concentration of the original material is needed.

Example 1.22

A 1/2 dilution of serum with saline had a creatinine result of 8.6 mg/dL. Calculate the actual serum creatinine concentration.

Dilution factor: 1/2

Dilution result = 8.6 mg/dL

Because this result represents 1/2 of the concentration, the inverse (or reciprocal) of the dilution is used, and the serum creatinine value must take into consideration this dilution, so the actual value is

$$2 \times 8.6 \text{ mg/dL} = 17.2 \text{ mg/dL} \qquad \textbf{(Eq. 1.44)}$$

Serial Dilutions

A **serial dilution** may be defined as multiple, progressive dilutions that dilutes highly concentrated solutions to produce solutions with lower concentration. Serial dilutions are extremely useful when the volume of concentrate or diluent is in short supply and its use needs to be minimized, or when a number of dilutions are required, such as in determining a titer. For instance, the volume of patient sample available to the laboratory may be small (e.g., pediatric samples), and a serial dilution may be needed to ensure that sufficient sample is available for analysis.

The serial dilution is initially made in the same manner as a simple dilution. Subsequent dilutions will then be made from each preceding dilution. When a serial dilution is made, certain criteria may need to be satisfied. The criteria vary with each situation but usually include such considerations as the total volume desired, the amount of diluent or concentrate available, the dilution factor, the final concentration needed, and the support materials required.

Example 1.23

A three-tube, twofold serial dilution is to be made on a sample. To start, the three tubes must be labeled. It is arbitrarily decided that the total volume for each dilution is to be 1 mL. Into the first tube, 0.5 mL of diluent is added and then 0.5 mL of patient sample. This satisfies the "twofold" or 1/2 dilution for tube 1. In the next tube, 0.5 mL of diluent is again added, along with 0.5 mL of well-mixed liquid from tube 1. This satisfies the 1/2 dilution in tube 2, bringing the total tube dilution to 1/4. For the third tube, 0.5 mL of diluent is added, along with 0.5 mL of well-mixed liquid from tube 2. This satisfies the 1/2 dilution within the tube but brings the total tube dilution to 1/8. The calculation for these values is

$$\frac{1}{2}(\text{Tube 1 Dilution}) \times \frac{1}{2}(\text{Tube 2 Dilution}) \qquad \textbf{(Eq. 1.45)}$$

$$= \frac{1}{4} \text{ total dilution for tube 2}$$

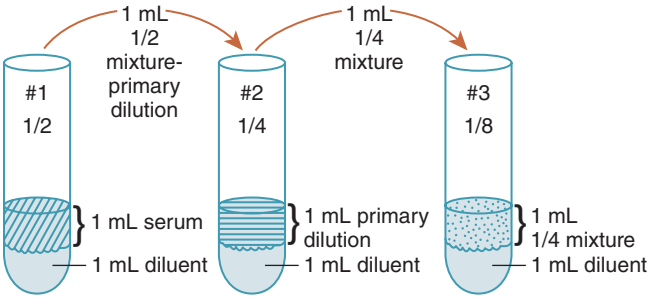


Figure 1.14 Serial dilution.

Making a 1/2 dilution of the 1/4 dilution will result in the next dilution (1/8) in Tube 3. To establish the dilution factor needed for subsequent dilutions, it is helpful to solve the following equation for (*x*):

$$\text{Stock/preceding concentration} \times (x) = (\text{final dilution factor}) \quad \textbf{(Eq. 1.46)}$$

Refer to **Figure 1.14** for an illustration of this serial dilution.

Example 1.24

Another type of dilution combines several dilution factors that are not multiples of one another. In our previous example, 1/2, 1/4, and 1/8 dilutions are all related to one another by a factor of 2. Consider the situation when 1/10, 1/20, 1/100, and 1/200 dilution factors are required. There are several approaches to solving this type of dilution problem. One method is to treat the 1/10 and 1/20 dilutions as one serial dilution problem, the 1/20 and 1/100 dilutions as a second serial dilution, and the 1/100 and 1/200 dilutions as the last serial dilution. Another approach is to consider what dilution factor of the concentrate is needed to yield the final dilution. In this example, the initial dilution is 1/10, with subsequent dilutions of 1/20, 1/100, and 1/200. The first dilution may be accomplished by adding 1 mL of stock to 9 mL of diluent. The total volume of solution is 10 mL. Our initial dilution factor has been satisfied. In making the remaining dilutions, 2 mL of diluent is added to each test tube.

$$\text{Initial/preceding dilution} \times (x) = \text{dilution needed}$$

Solve for (*x*).

Using the dilution factors listed above and solving for (*x*), the equations become

$$1/10 \times (x) = 1/20$$

where $(x) = 2$ (or 1 part stock to 1 part diluent)

$$1/20 \times (x) = 1/100$$

where $(x) = 5$ (or 1 part stock to 4 parts diluent)

$$1/100 \times (x) = 1/200$$

where $(x) = 2$ (or 1 part stock to 1 part diluent) **(Eq. 1.47)**

In practice, the 1/10 dilution must be diluted by a factor of 2 to obtain a subsequent 1/20 dilution. Because the second tube already contains 2 mL of diluent, 2 mL of the 1/10 dilution should be added (1 part stock to 1 part diluent). In preparing the 1/100 dilution from this, a 1/5 dilution factor of the 1/20 mixture is required (1 part stock to 4 parts diluent). Because this tube already contains 2 mL, the volume of diluent in the tube is divided by its parts, which is 4; thus, 500 µL, or 0.500 mL, of stock should be added. The 1/200 dilution is prepared in the same manner using a 1/2 dilution factor (1 part stock to 1 part diluent) and adding 2 mL of the 1/100 to the 2 mL of diluent already in the tube.

Water of Hydration

Some compounds are available in a hydrated form. A reagent protocol often designates the use of an anhydrous form of a chemical; frequently, however, all that is available is a hydrated form. To obtain a correct gmw for these chemicals, the attached water molecule(s) must be included.

Example 1.25

How much $CuSO_4 \cdot 5H_2O$ must be weighed to prepare 1 L of 0.5 *M* $CuSO_4$? When calculating the gmw of this substance, the water weight of 90 g must be considered so that the gmw is 250 g rather than gmw of $CuSO_4$ alone (160 g). Therefore,

$$\frac{250\ \cancel{g}\ CuSO_4\ 5H_2O}{\cancel{mol}} \times \frac{0.5\ \cancel{mol}}{1\ \cancel{g}} = 125\ g/L \quad \textbf{(Eq. 1.48)}$$

Cancel out like terms to obtain the result of 125 g/L of the hydrated form $CuSO_4 \cdot 5H_2O$.

Example 1.26

A procedure requires 0.9 g of $CuSO_4$. All that is available is $CuSO_4 \cdot 5H_2O$. What weight of $CuSO_4 \cdot 5H_2O$ is needed? Calculate the percentage of $CuSO_4$ present in $CuSO_4 \cdot 5H_2O$. Again, using the gmw, the percentage is

$$\frac{160}{250} = 0.64, \text{ or } 64\% \quad \textbf{(Eq. 1.49)}$$

Therefore, 1 g of $CuSO_4 \cdot 5H_2O$ contains 0.64 g of $CuSO_4$, so the equation becomes

$$\frac{0.9 \text{ g } CuSO_4 \text{ needed}}{0.64 \ CuSO_4 \text{ in } CuSO_4 \cdot 5H_2O} = 1.41 \text{ g } CuSO_4 \cdot 5H_2O \text{ required} \quad \textbf{(Eq. 1.50)}$$

Graphing and Beer's Law

The Beer-Lambert law (**Beer's law)** mathematically establishes the relationship between analyte concentration and absorbance of light in many photometric determinations. Beer's law states that the concentration of a substance is directly proportional to the amount of light absorbed or inversely proportional to the logarithm of the light transmitted. Beer's law can be expressed as

$$A = abc \quad \textbf{(Eq. 1.51)}$$

where A is absorbance; a is the absorptivity constant for a particular compound at a given wavelength under specified conditions (such as temperature and pH); b is the length of the light path; and c is the concentration.

If a method follows Beer's law, then absorbance is proportional to concentration as long as the length of the light path and the absorptivity of the absorbing species remain unaltered during the analysis. In practice, however, there are limits to the predictability of a linear response. In automated systems, adherence to Beer's law is often determined by checking the linearity of the test method over a wide concentration range. The *limits of linearity* often represent the reportable range of an assay. This term should not be confused with the reference ranges associated with clinical significance of a test. Assays measuring absorbance generally obtain the concentration results by using a graph of Beer's law, known as a standard graph or curve. This graph is made by plotting absorbance versus the concentration of known standards (**Figure 1.15**). Because most photometric assays set the initial absorbance to zero (0) using a reagent blank, the initial data points are 0,0. Graphs should be labeled properly and the concentration units must be given. The horizontal axis is referred to as the *x*-axis, whereas the vertical line is the *y*-axis. By convention in the clinical laboratory, concentration is usually plotted on the *x*-axis. On a standard graph, only the standard and the associated absorbances are plotted.

In terms of transmitted light, Beer's law is expressed as percent transmission (%T), where T is

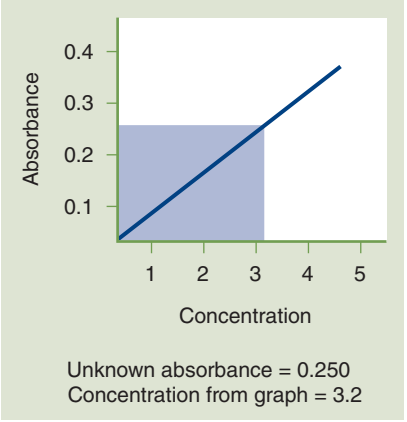


Figure 1.15 Standard curve.

defined as the ratio of the radiant energy transmitted divided by the radiant incident energy of the sample (I). The mathematical derivation of %T can be expressed as

$$\%T = \frac{I}{I_0} \times 100 \quad \textbf{(Eq. 1.52)}$$

Where I_0 is the incident light, and I is the transmitted light. The relationship between transmission and absorbance yields a nonlinear curve, and can be calculated using the following formula:

$$A = -\log\left(I/I_0\right) = \log(100\,\%) - \log\%T = 2 - \log\%T \quad \textbf{(Eq. 1.53)}$$

Once a standard curve has been established, it is permissible to run just one standard, or *calibrator*, as long as the system remains unchanged. A **One-point calibration** or calculation refers to the calculation of the comparison of the known standard or calibrator concentration and its corresponding absorbance to the absorbance of an unknown value according to the following ratio, where C_u and A_u indicate the concentration and absorbance, respectively, for the unknown value:

$$\frac{\text{Concentration of standard } (C_s)}{\text{Absorbance of standard } (A_s)} = \frac{\text{Concentration of standard } (C_u)}{\text{Absorbance of standard } (A_u)} \quad \textbf{(Eq. 1.54)}$$

Solving for the concentration of the unknown, the equation becomes

$$C_u = \frac{(A_u)(C_s)}{A_s} \quad \textbf{(Eq. 1.55)}$$

Example 1.27

The biuret assay method for protein is very stable and follows Beer's law. Rather than make up a standard graph, one protein standard of 6 g/dL concentration was assayed. The measured absorbance of the standard was 0.400, and the measured absorbance of the unknown was 0.350. Determine the value of the unknown in g/dL.

$$C_u = \frac{(0.350)(6\ \text{g/dL})}{(0.400)} = 5.25\ \text{g/dL} \qquad \textbf{(Eq. 1.56)}$$

This method of calculation is acceptable as long as everything in the system, including the instrument and lot number of reagents, remains the same. If anything in the system changes, a new standard graph should be generated. Verification of linearity and/or calibration is required whenever a system changes or becomes unstable. Regulatory agencies often prescribe the condition of verification as well as how frequently the linearity needs to be performed.

Enzyme Calculations

Another application of Beer's law is the calculation of enzyme assay results. When calculating enzyme results, the rate of change in absorbance over time is often monitored continuously during the reaction to give the difference in absorbance, known as the **delta absorbance**, or ΔA. Instead of using a standard graph or a one-point calculation, the molar absorptivity of the product is used. If the absorptivity constant and absorbance, in this case ΔA, are known, Beer's law can be used to calculate the enzyme concentration directly without the need of a standard graph, as follows:

$$A = abC$$
$$C = \frac{A}{ab} \qquad \textbf{(Eq. 1.57)}$$

When the absorptivity constant (a) is given in units of grams per liter (moles) through a 1-centimeter (cm) light path, the term *molar absorptivity* (ε) is used. Substitution of ε for a and ΔA for A produces the following Beer's law formula:

$$C = \frac{\Delta A}{\varepsilon} \qquad \textbf{(Eq. 1.58)}$$

For reporting enzyme activity, the IU, or **international unit**, is defined as the amount of enzyme that will catalyze 1 µmol of substrate per minute per liter. These units were often expressed as units per liter (U/L). The designations IU, U, and IU/L were adopted by many clinical laboratories to represent the IU. Although the reporting unit is the same, unless the analysis conditions are identical, use of the IU does not standardize the actual enzyme activity, and therefore, results between different methods of the same enzyme do not result in equivalent activity of the enzyme. For example, an alkaline phosphatase performed at 37°C will catalyze more substrate than if it is run at lower temperature, such as 25°C, even though the unit of expression, U/L, will be the same. The SI recommended unit is the *katal*, which is expressed as moles per liter per second. Whichever unit is used, calculation of the activity using Beer's law requires inclusion of the dilution and, depending on the reporting unit, possible conversion to the appropriate term (e.g., µmol to mol, mL to L, minute to second, and temperature factors). Beer's law for the IU now becomes

$$C = \frac{(\Delta A)10^{-6}(\text{TV})}{(\varepsilon)(b)(\text{SV})} \qquad \textbf{(Eq. 1.59)}$$

where TV is the total volume of sample plus reagents in mL and SV is the sample volume used in mL. The 10^{-6} converts moles to µmol for the IU. If another unit of activity is used, such as the katal, conversion into liters and seconds would be needed, but the conversions to and from micromoles are excluded.

Example 1.28

The ΔA per minute for an enzyme reaction is 0.250. The product measured has a molar absorptivity of 12.2×10^3 at 425 nm at 30°C. The incubation and reaction temperature are also kept at 30°C. The assay calls for 1 mL of reagent and 0.050 mL of sample. Give the enzyme activity results in international units.

Applying Beer's law and the necessary conversion information, the equation becomes

$$C = \frac{(0.250)(10^{-6})(1.050\ \text{mL})}{(12.2\times10^3)(1)(0.050\ \text{mL})} = 430\ \text{U} \qquad \textbf{(Eq. 1.60)}$$

Note: b is usually given as 1 cm; because it is a constant, it may not be considered in the calculation.

Specimen Collection and Handling

The process of specimen collection, handling, and processing remains one of the primary areas of *preanalytic* errors. Careful attention to each phase of the testing process is necessary to ensure proper subsequent testing and reporting of accurate and reliable results. All accreditation agencies require laboratories to clearly define and delineate the procedures used for proper collection, transport, and processing of patient samples and the steps used to minimize and detect any errors, along with the documentation of the resolution of any errors. The Clinical Laboratory Improvement Amendments Act of 1988 (CLIA 88)[21] specifies that procedures for specimen submission and proper handling be documented, including the disposition of any specimen that does not meet the laboratories' criteria of acceptability.

Types of Samples

Phlebotomy, or *venipuncture*, is the act of obtaining a blood sample from a vein using a needle attached to a collection device or a stoppered *evacuated tube*. The collection tubes are available in different volume sizes: from pediatric sizes (≈150 µL) to larger 10 mL tubes. The most frequent site for venipuncture is the medial antecubital vein of the arm. A tourniquet made of pliable nonlatex rubber flat band or tubing is wrapped around the arm, causing cessation of blood flow and dilation of the veins, making for easier detection. The gauge of the needle is inversely related to the size of the needle; the larger the number, the smaller the needle bore and length. Routine venipuncture uses a 23- or 21-gauge needle. A winged infusion set, sometimes referred to as a butterfly because of the appearance of the setup, may be used whenever the veins are fragile, small, or difficult to detect. The butterfly needle is attached to a piece of tubing, which is then attached to a hub or barrel. Because of potential for needlesticks and cost of the product, this practice may be discouraged. However, newer push-button safety devices are now available.

When selecting an appropriate vein, sites adjacent to IV (intravenous) therapy should be avoided. If both arms are involved in IV therapy and the IV cannot be discontinued for a short time, a site *below* the IV site may be considered. In such cases, the initial sample drawn (5 mL) should be discarded because it is most likely contaminated with IV fluid and only subsequent sample tubes should be used for analytic purposes.

In addition to venipuncture, blood samples can be collected using a capillary puncture technique that involves using either the outer area of the bottom of the foot (a heel stick) for infants or the lateral side of the middle or ring finger for individuals 1 year and older (finger stick). A sharp lancet device is used to pierce the skin, and an appropriate capillary or microtainer tube is used for sample collection.[22]

Analytic testing of blood involves the use of whole blood, serum, or plasma. **Whole blood**, as the name implies, contains the liquid portion of the blood, called *plasma*, and its cellular components (red blood cells, white blood cells, and platelets). The collection of whole blood requires the vacuum tube to contain an *anticoagulant*. Complete mixing of the blood immediately following venipuncture is necessary to ensure the anticoagulant adequately inhibits the specimen from clotting. As the tube of whole blood settles, the cells fall toward the bottom, leaving a clear yellow supernatant on top, which is the plasma.

If a tube does *not* contain an anticoagulant, the blood forms a fibrin clot incorporating the cells; this clot consumes fibrinogen. The remaining liquid is called **serum** rather than plasma (**Figure 1.16**). Most testing in the clinical chemistry laboratory is performed on either plasma or serum. The major difference between plasma and serum is that serum does not contain fibrinogen and some potassium is released from platelets (i.e., potassium levels are slightly higher in serum than in plasma). It is important that serum samples be allowed to completely clot (≈30 minutes) in an upright position at room temperature before being centrifuged. Plasma samples also require centrifugation but do not need time to clot, decreasing the turnaround time for testing and reporting results.

Centrifugation of the sample accelerates the process of separating the liquid portion and cellular portion. Specimens should be centrifuged according to recommendations by the tube manufacturer or test protocol, usually approximately 10 minutes at an RCF of 1000 to 2000 *g*, but should avoid the mechanical destruction of red blood cells that can result in hemoglobin release, which is called **hemolysis**.

Arterial blood samples measure blood gases (partial pressures of oxygen and carbon dioxide) and pH. Syringes containing heparin anticoagulant are used instead of evacuated tubes because of the pressure in an arterial blood vessel. The radial artery is the primary arterial site, while there may be times when the brachial or femoral artery may be considered. Arterial punctures are more difficult to perform

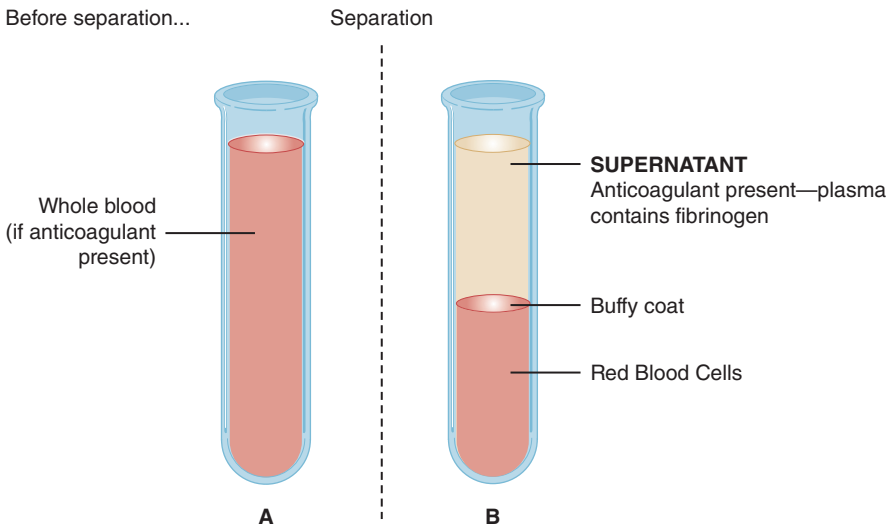


Figure 1.16 Blood sample. **(A)** Whole blood. **(B)** Whole blood after separation.

because of inherent arterial pressure, difficulty in stopping bleeding afterward, and the undesirable development of a hematoma, which cuts off the blood supply to the surrounding tissue.[23]

Continued metabolism may occur if the serum or plasma remains in contact with the cells for any period. Evacuated tubes may incorporate gel-like material that serves as a barrier between the cells and the plasma or serum and seals these compartments from one another during centrifugation. Some gels can interfere with certain analytes, and manufacturer recommendations should be followed.

Proper patient identification is the first step in sample collection. The importance of using the proper collection tube, avoiding prolonged tourniquet application, drawing tubes in the proper order, and proper labeling of tubes cannot be stressed strongly enough. Prolonged tourniquet application causes a stasis of blood flow and an increase in hemoconcentration and anything bound to proteins or the cells. Having patients open and close their hand during phlebotomy is of little value and may cause an increase in potassium or lactic acid and, therefore, should be avoided. IV contamination should be considered if a large increase occurs in the substances being infused, such as glucose, potassium, sodium, and chloride, with a decrease of other analytes such as urea and creatinine. In addition, the proper antiseptic must be used. Isopropyl alcohol wipes, for example, are used for cleaning and disinfecting the collection site; however, isopropyl alcohol is not recommended for disinfecting the site when drawing blood alcohol levels (in such cases, chlorhexidine is used as the disinfectant).

Blood is not the only sample analyzed in the clinical chemistry laboratory. Urine is the next most

CASE STUDY 1.2, PART 3

Recall Mía, the new graduate.

4. Which of the chemistry specimens provides the fastest turn-around time?
5. The whole blood analysis was performed and resulted. When Mía opened the centrifuge, she noticed that the specimens were grossly hemolyzed. What should Mía do?

common fluid for determination. Most quantitative analyses of urine require a timed sample (usually 24 hours); a complete sample (all urine collected within the specified time) can be difficult because many timed samples are collected by the patient in an outpatient situation. Creatinine analysis is often used to assess the completeness of a 24-hour urine sample because creatinine output is relatively free from interference and is stable, with little change in output within individuals. The average adult excretes 1 to 2 g of creatinine per 24 hours. Urine volume differs widely among individuals; however, a 4-L container is adequate (average output is ≈2 L). It should be noted that this analysis differs from the creatinine clearance test used to assess glomerular filtration rate, which compares urine creatinine output with that in the serum or plasma in a specified time interval and urine volume (often correcting for the surface area).

Other body fluids analyzed by the clinical chemistry laboratory include **cerebrospinal fluid (CSF)**, *paracentesis* fluids (pleural, pericardial, and peritoneal), and amniotic fluids. The color and characteristics of the fluid *before* centrifugation should be noted for these samples. Before centrifugation, a laboratorian should also verify that the sample is designated for clinical chemistry analysis *only* because a single fluid sample may be shared among several departments (i.e., hematology or microbiology), and centrifugation could invalidate other laboratory testing in those areas.

CSF is an ultrafiltrate of the plasma and is approximately two-thirds of the plasma glucose value. For glucose and total protein analysis, it is recommended that a blood sample be analyzed concurrently with the analysis of those analytes in the CSF. This will assist in determining the clinical utility of the values obtained on the CSF sample. This is also true for lactate dehydrogenase and protein assays requested on paracentesis fluids. All fluid samples should be handled immediately, without delay between sample procurement, transport, and analysis.

Amniotic fluid may be used to assess fetal lung maturity, congenital diseases, hemolytic diseases, genetic defects, and gestational age. The laboratorian should verify the specific handling of this fluid with the manufacturer of the testing procedure(s).

Sample Processing

When samples arrive in the laboratory, they are first processed. In the clinical chemistry laboratory, this means correctly matching the blood collection tube(s) with the appropriate test requisition and patient identification labels. This is a particularly sensitive area of preanalytic error. Bar code labels (either as 1D linear barcodes, or 2D QR barcodes) or radiofrequency ID chip labeling on primary sample tubes are vital in detecting errors and to minimizing clerical errors. The laboratorian must also ascertain if the sample is acceptable for further processing. The criteria used depends on the test involved but usually include volume considerations (i.e., is there sufficient volume for testing needs?), use of proper anticoagulants or preservatives (i.e., was it collected in the correct evacuated tube?), whether timing is clearly indicated and appropriate for timed testing, and whether the specimen is intact and has been properly transported (e.g., on ice, within a reasonable period, protected from light). Unless a whole blood analysis is being performed, the sample is then centrifuged as previously described and the serum or plasma should be separated from the cells if not analyzed immediately. Today, the use of serum separator tubes and plasma separator tubes is common practice.

Once the sample is processed, the laboratorian should note the presence of any serum or plasma characteristics such as *hemolysis* and **icterus** (increased bilirubin pigment) or the presence of turbidity often associated with **lipemia** (increased lipids). (**See Table 1.5**.) Many analytes are stable at room temperature between 24 to 72 hours. However, if testing is not to be performed within 8 hours, it is recommended that serum and/or plasma be

Table 1.5 Examples of Hemolyzed, Icteric, and Lipemic/Turbid Samples

Hemolyzed (red)

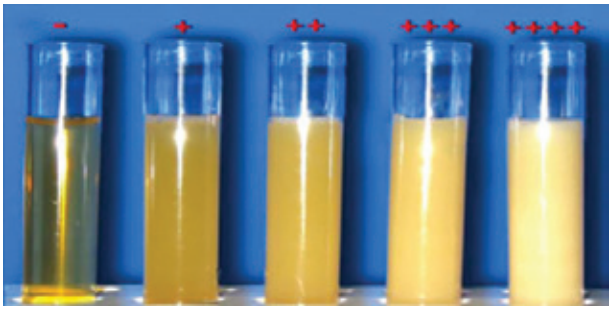
Hemoglobin					
mg/dL	0	50	150	250	525
g/L	0	0.50	1.50	2.50	5.25

Icteric (yellow)

Total Bilirubin

mg/dL	0	1.7	6.6	16	30
µmol/L	0	29.1	112.9	273.6	513

Lipemic (turbid)



Intralipid®

mg/dL	0	125	250	500	1000
mmol/L*	0	1.41	2.83	5.65	11.83

*Concentrations based on dilutions of 20% Intralipid® (or the equivalent).

CLSI C56A_2012 Hemolysis, Icterus, and Lipemia_Turbidity Indices as Indicators of Interference in Clinical Laboratory Analysis, 1st Edition.

All photos used with permission from Siemens Medical Solutions USA, Inc.

refrigerated between 2°C and 8°C.[24] It is important to avoid exposing samples that are light sensitive, such as bilirubin, to artificial or ultraviolet light for extended periods of time.[24] Separated serum and/or plasma may be frozen at –20°C and stored for longer periods without deleterious effects on the results. Repeated cycles of freezing and thawing, like those that occur in so-called frost-free freezers, should be avoided.

Hemolysis can be visually observed in a centrifuged patient sample as a red color due to the release of hemoglobin. There are patient conditions where this may occur in vitro, such as hemolytic anemia, but hemolysis can also be present due to preanalytic collection variables such as inappropriate needle gauge, venipuncture site selection (small veins), and venous trauma or difficulty in specimen collection. Along with the release of hemoglobin, other intracellular components may be released, such as potassium, phosphate, and lactate dehydrogenase, which may impact patient values for these analytes. For analyzers utilizing spectrophotometric or enzymatic detection methods, hemolysis may also cause errors during assay.

Lipemia results when the lipid levels of the patient are elevated and, in turn, is visualized as a creamy or milky appearance to the serum or plasma upon centrifugation. Lipemia can cause a volume displacement for some analytes, such as electrolytes, as well as interference in light-scattering methodologies due to the turbidity present. Icterus is a deep yellow or golden appearance of the serum or plasma due to increased bilirubin levels, and may cause spectral interference on certain analyzers in the chemistry lab. To help determine if interference has occurred, many analyzers are capable of detecting, then estimating, the interferent and the effect on sample values. This assessment of hemolysis, icterus, and lipemia is known as the HIL index. Laboratories have the ability to determine values to establish the alert values of the HIL indices as a means of assessing specimen integrity and acceptability. Table 1.5 illustrates examples of HIL indices in serum or plasma.[25]

Sample Variables

Sample variables (**Table 1.6**) include physiologic considerations, proper patient preparation, and problems in collection, transportation, processing, and storage. Although laboratorians must include mechanisms to minimize the effect of these variables on testing and must document each preanalytic incident, it is often difficult to control the variables that involve individuals outside of the laboratory. The best course of action is to critically assess or anticipate variances, identify potential problems, and implement an action plan that contains policies, procedures, and checkpoints throughout the testing process. Good communication with all personnel involved helps ensure that the right specimen at the right time for the right patient is collected each and every time to meet the needs of the healthcare team including the patient, laboratory, and ordering physician. Most accreditation agencies require that

Table 1.6 Variables Affecting Select Chemistry Analytes

Physiological factors	Diet	Samples requiring fasting (glucose)	
	Medication or herbal supplements	Possible interference with analytical methods (biotin)	
	Circadian rhythm	Analyte changes based on diurnal variation (cortisol)	
		Timing of collection	Serial testing of an analyte (cardiac troponin) or timing based on medication (therapeutic drug monitoring)
	Patient posture	Shift of hemodynamic fluid volume (proteins)	
Patient preparation factors	Fasting	Proper instructions for 8- to 12-hour fasting	
	24-hour urine collection	Proper instructions for collection	
Collection and sample processing factors	Venipuncture technique	Needle selection, site selection to decrease opportunity of hemolysis	
	Tube selection	Appropriate sample tube, inversion following collection, appropriate clotting time	
	Tourniquet use	Prolonged use affects analytes (K^+, lactic acid)	
	Specimen transport and storage	Protection from light (bilirubin) or collect and store on ice (arterial blood gas, ammonia)	

laboratories consider all aspects of preanalytic variation as part of their quality assurance plans, including effective problem solving and documentation.

Physiologic variation refers to changes that occur within the body, such as cyclic changes (diurnal or circadian variation) or those resulting from exercise, diet, stress, gender, age, underlying medical conditions (e.g., fever, asthma, and obesity), drugs, or posture. Samples may be drawn on patients who are fasting (usually overnight for at least 8 hours). When fasting, many patients may drink water to avoid becoming dehydrated, which can lead to falsely elevated electrolyte results. Patient preparation for timed samples or those requiring specific diets or other instructions must be well written and verbally explained to patients. Elderly patients often misunderstand or are overwhelmed by the directions given to them. Collection and processing variations are related to those factors discussed under specimen processing. Clerical errors are the most frequently encountered, followed by other pre-analytical variables including inadequate separation of cells from serum, improper storage, and collection.

Chain of Custody

When laboratory tests are likely linked to a crime or accident, they become forensic in nature. In these cases, documented specimen identification is required at each phase of the process. Each facility has its own forms and protocols; however, the patient, and usually a witness, must identify the sample. The sample should be collected and then sealed with a tamper-proof seal. Any individual in contact with the sample must document receipt of the sample, the condition of the sample at the time of receipt, and the date and time it was received. In some instances, one witness verifies the entire process and cosigns as the sample moves along. Any analytic test could be used as part of legal testimony; therefore, the laboratorian should give each sample—even without the documentation—the same attention given to a forensic sample.

Electronic and Paper Reporting of Results

Electronic transmission of laboratory data and the use of physician order entry, electronic medical record, coding, billing, and other data management systems maintain the integrity of data generated by providing reporting guidelines and safeguards to ensure privacy of the data and records. There are various data management systems in use by healthcare agencies related to accessing laboratory information. For example, the Logical Observation Identifiers Names and Codes (LOINC) system, International Federation of Clinical Chemistry/International Union of Pure and Applied Chemistry (IFCC/IUPAC), ASTM, Health Level Seven International (HL7), and Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT) are databases that use unique coding systems for laboratory observations. There are also additional proprietary systems in use. To standardize these processes and to protect the confidentiality of patient information as required by the Health Insurance Portability and Accountability Act (HIPAA), the Healthcare Common Procedure Coding System Level II (HCPCS) test and services coding system was developed by the Centers for Medicare and Medicaid Services (CMS) to be recognized by all insurers for reimbursement purposes. The International Classification of Diseases (ICD) developed by the World Health Organization (WHO) uses codes identifying patient diseases and conditions. In the United States, ICD-11 is currently in place. The clinical modifications are maintained by the National Center for Health Statistics. Incorporated into the HCPCS system is the Current Procedural Terminology (CPT) codes, developed by the American Medical Association, which identify almost all laboratory tests and procedures. The CPT codes are divided into different subcategories, with tests or services assigned five-digit numbers followed by the name of the test or service. Together, these standard coding systems help patient data and tracking of disease transmission between all stakeholders such as physicians, patients, epidemiologists, and insurers.

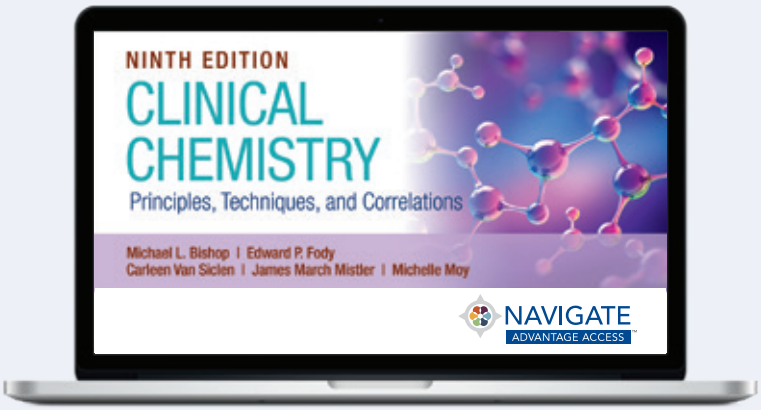
Clinical laboratory procedures are found in CPT Category I with coding numbers falling between 80,000 and 89,999. There can be several codes for a given test based on the reason and type of testing, and there are codes given for common profiles or array of tests that represent each test's separate codes. For example, blood glucose testing includes the codes 82947 (quantitative except for strip reading), 82948 (strip reading), and 82962 (self-monitoring by FDA-cleared device), and the comprehensive metabolic panel (80053) includes albumin, alkaline phosphatase, total bilirubin, blood urea nitrogen, total calcium, carbon dioxide, chloride, creatinine, glucose, potassium, total protein, sodium, and alanine and aspartate transaminases and their associated codes. At a minimum, any laboratory reporting system must include a unique patient identifier, test name, and code that relates back to the HCPCS and ICD databases. For reporting purposes, whether paper or electronic, the report should include the unique patient identifier and test name (including any appropriate abbreviations), the test value with the unit of measure, date and time of collection, sample information, reference ranges, plus any other pertinent information for proper test interpretation. Results that are subject to autoverification should be indicated in the report. **Table 1.7** lists the information that is often required by accreditation agencies.[26]

Table 1.7 Minimum Elements of Paper or Electronic Patient Reports

Name and address of laboratory performing the analysis including any reference laboratories used
Patient name and identification number or unique identifier
Name of physician or person ordering the test
Date and time of specimen collection
Date and time of release of results (or available if needed)
Specimen source or type
Test results and units of measure if applicable
Reference ranges, when available
Comments relating to any sample or testing interferences that may alter interpretation

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Thompson A, Taylor BN. *Guide for the Use of the International System of Units (SI)*. NIST Special Publication 811 (version 3.2); 2008. Gaithersburg, MD: National Institute of Standards and Technology. https://physics.nist.gov/cuu/pdf/sp811.pdf.
2. Armbruster D, Donnelly J. Harmonization of clinical laboratory test results: The role of the IVD industry. *EJIFCC*. 2016;27(1):37–47.
3. American Chemical Society (ACS). *Reagent Chemicals* [online edition]. Washington, DC: ACS Press, 2021. Published online February 28, 2017. doi: 10.1021/acsreagents.
4. Department of Labor and Occupational Safety and Health Administration (OSHA). *Occupational Exposure to Hazardous Chemicals in Laboratories*. Washington, DC: OSHA; 2021. (Federal Register, 29 CFR, Part 1910.1450.). Accessed February 25, 2021.
5. National Institute of Standards and Technology (NIST). *Standard Reference Materials*. Washington, DC: U.S. Department of Commerce; Updated June 2020. https://www.nist.gov/srm-definitions. Accessed March 1, 2021.
6. IUPAC. *Compendium of Chemical Terminology*. 2nd ed. (the "Gold Book"). Compiled by A.D. McNaught and A. Wilkinson. Oxford: Blackwell Scientific Publications; 1997. XML online corrected version: http://goldbook.iupac.org (2006-2014) created by M. Nic, J. Jirat, Kosata; updates compiled by A. Jenkins, ISBN 0-9678550-9-8. doi: 10.1351/goldbook. Last update 2019-07-01; version 3.0.1.
7. International Organization for Standardization, European Committee for Standardization (ISO). *In Vitro Diagnostic Medical Devices—Requirements for establishing metrological traceability of values assigned to calibrators, trueness control materials and human samples*. Geneva, Switzerland: ISO; 2020. (Publication No. ISO EN 17511:2020).
8. Clinical Laboratory Standards Institute/National Committee for Clinical Laboratory Standards. *Preparation and Testing of Reagent Water in the Clinical Laboratory; Approved Guideline*. 4th ed. Wayne, PA: CLSI; 2006. (Publication No. GP40.) Amended June 2018.
9. College of American Pathologists (CAP). *Laboratory General Checklist: Quality of Water and Glassware Washing*. Northfield, IL: June 04, 2020.
10. National Institutes of Health, National Institute of Allergy and Infectious Diseases. *Guidelines for Good Clinical Laboratory Practice Standards*. https://www.niaid.nih.gov/sites/default/files/gclp.pdf. Published July 2013. Content reviewed October 09, 2020. Accessed March 01, 2021.
11. National Institute of Standards and Technology (NIST). *Calibration Uncertainties of Liquid-in-Glass Thermometers Over the Range from –20°C to 400°C*. Gaithersburg, MD: U.S. Department of Commerce; 2003. Updated February 17, 2017. Accessed March 01, 2021.
12. American Society for Testing and Materials (ASTM). *E969-02 (2019), Standards Specification for Volumetric (Transfer) Pipets*. West Conshohocken, PA: ASTM International; 2019. https://www.astm.org
13. American Society for Testing and Materials (ASTM). *E288-10 (2017), Standard Specification for Glass, Volumetric Flasks*. West Conshohocken, PA: ASTM International; 2017. https://www.astm.org.
14. American Society for Testing and Materials (ASTM). *E438-92 (2018), Standard Specification for Glasses in Laboratory Apparatus*. West Conshohocken, PA: ASTM International; 2018, https://www.astm.org.
15. Eppendorf. *Standard Operating Procedure for Manual Dispensing Tool*. Hamburg, Germany: Eppendorf AG; 2019. https://www.eppendorf.com/fileadmin/knowledgebase/asset/CN-zh/155258_SOP.pdf. Accessed June 24, 2021.
16. Gilson. *Guide to Pipetting*. 3rd ed. Middleton, WI: Gilson. Procedure LT 800550/A, October 2015. https://www.gilson.com/pub/media/docs/GuideToPipettingE.pdf. Accessed March 02, 2021.
17. American Society for Testing and Materials (ASTM). *E617-13, Standard Specification for Laboratory Weights and Precision Mass Standards*. West Conshohocken, PA: ASTM International; 2013. https://www.astm.org.

18. Eppendorf. User Guide No. 025, *Calibration and Adjustment of Dispensing Systems in the Laboratory.* Hamburg, Germany: Eppendorf AG; 2017.
19. Mettler Toledo. *Routine Testing of Lab Balances.* Mettler Toledo Group, https://www.mt.com/contacts. Accessed March 2021.
20. Doucette LJ. *Mathematics for the Clinical Laboratory*. 2nd ed. Philadelphia, PA: Saunders; 2010.
21. Clinical Laboratory Standards Institute (CLSI). *Procedures for the Handling and Processing of Blood Specimens*. 4th ed. Wayne, PA: CLSI; 2010. (Publication No. H18-A4, Vol. 30).
22. Centers for Disease Control and Prevention, Division of Laboratory Systems. *Clinical Laboratory Improvement Amendments*. CFR Part 493, Laboratory requirements. http://www.ecfr.gov. Accessed March 2021.
23. Clinical Laboratory Standards Institute (CLSI). *Procedures and Devices for the Collection of Diagnostic Blood Specimens by Skin Puncture*. 6th ed. Wayne, PA: CLSI; 2010. (Publication No. H3-A6, Vol. 27).
24. Clinical Laboratory Standards Institute (CLSI). *Procedures for the Handling and Processing of Blood Specimens for Common Laboratory Tests*. 4th ed. Wayne, PA: CLSI; 2010. (Publication No. GP44-A4, Vol 30).
25. Clinical Laboratory Standards Institute (CLSI). *Hemolysis, Icterus, and Lipemia/Turbidity Indices as Indicators of Interference in Clinical Laboratory Analysis*. Wayne, PA: CLSI; 2012. (Publication No. C56-A).
26. College of American Pathology; Commission on Laboratory Accreditation; Laboratory Accreditation Program Laboratory General Checklist Website. https://webapps.cap.org/apps/docs/laboratory_accreditation/checklists/laboratory_general_sep07.pdf. Accessed November 2015.

CHAPTER 2

Laboratory Safety and Regulations

Brandy Greenhill

CHAPTER OUTLINE

KEY TERMS

Airborne pathogens
Biohazard
Bloodborne pathogens
Carcinogens
Chemical hygiene plan
Corrosive chemicals
Cryogenic material
Exposure control plan
Fire tetrahedron
Hazard Communication Standard
Hazardous materials
High-efficiency particulate air (HEPA) filters
Laboratory standard
Mechanical hazards
Medical waste
National Fire Protection Association (NFPA)
Occupational Safety and Health Act (OSHA)
Radioactive materials
Reactive chemicals
Safety Data Sheets (SDSs)
Standard precautions
Teratogens

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Discuss safety awareness for clinical laboratory personnel.
- List the responsibilities of employer and employee in providing a safe workplace.
- Identify hazards related to handling chemicals, biologic specimens, and radiologic materials.
- Choose appropriate personal protective equipment when working in the clinical laboratory.
- State the classes of fires and the types of fire extinguishers to use for each.
- Describe steps used as precautionary measures when working with electrical equipment, cryogenic materials, and compressed gases and avoiding mechanical hazards associated with laboratory equipment.
- Select the correct means for disposal of waste generated in the clinical laboratory.
- Outline the steps required in documentation of an accident in the workplace.

Laboratory Safety and Regulations

Clinical laboratorians, by the nature of the work they perform, are exposed daily to a variety of real or potential hazards: electric shock, toxic vapors, compressed gases, flammable liquids, radioactive material, corrosive substances, mechanical trauma, poisons, and the inherent risks of handling biologic materials, to name a few.[1] Each clinician should develop an understanding of the risks associated with these hazards and must be "safety conscious" at all times.[2]

Laboratory safety necessitates the effective control of all hazards that exist in the clinical laboratory at any given time. Safety begins with the recognition of hazards and is achieved through the application of common sense, a safety-focused attitude, good personal behavior, good housekeeping in all laboratory work and storage areas, and, above all, the continual practice of good laboratory technique. In most cases, accidents can be traced directly to two primary causes: unsafe acts (not always recognized by personnel) and unsafe environmental conditions.[3] This chapter discusses laboratory safety as it applies to the clinical laboratory.

Occupational Safety and Health Act

Public Law 91-596, better known as the **Occupational Safety and Health Act (OSHA)**, was enacted by the U.S. Congress in 1970. The goal of this federal regulation was to provide all employees (clinical laboratory personnel included) with a safe work environment. Under this legislation, the Occupational Safety and Health Administration (also known as OSHA) is authorized to conduct on-site inspections to determine whether an employer is complying with

CASE STUDY 2.1

Remember Miles and Mía from Chapter 1. Both Miles and Mía volunteered to be on the laboratory safety committee for their health system representing their individual hospitals. Each was tasked with performing an audit of personal protective equipment (PPE) within their department and familiarizing themselves with the laboratory safety policies and procedures in order to become a lab safety resource for their coworkers.

the mandatory standards.[4] Safety is no longer only a moral obligation but also a federal law. In about half of the states, this law is administered by individual state agencies rather than by the federal OSHA. These states still fall within delineated OSHA regions, but otherwise they bear all administrative, consultation, and enforcement responsibilities. The state regulations must be at least as stringent as the federal ones, and many states incorporate large sections of the federal regulations verbatim.

OSHA standards that regulate safety in the laboratory include the Bloodborne Pathogen Standard, Formaldehyde Standard, Laboratory Standard, Hazard Communication Standard, Respiratory Protection Standard, Air Contaminants Standard, and Personal Protective Equipment Standard. Because laws, codes, and ordinances are updated frequently, current reference materials should be reviewed. Assistance can be obtained from local libraries, the Internet, and federal, state, and local regulatory agencies. The primary standards applicable to clinical laboratory safety are summarized next.

Bloodborne Pathogens [29 CFR 1910.1030][4]

This standard applies to exposure to all blood or other potentially infectious materials in any occupational setting. It defines terminology relevant to such exposures and mandates the development of an **exposure control plan**. This plan must cover specific preventative measures including exposure evaluation, engineering controls, work practice controls, and administrative oversight of the program. Standard precautions and personal protective equipment (PPE) are foremost among these infection control measures. The **standard precautions** concept (formerly referred to as universal precautions) is basically an approach to infection control in which all human blood, tissue, and most fluids are handled as if known to be infectious for the human immunodeficiency virus (HIV), hepatitis B virus (HBV), and other **bloodborne pathogens**, including hepatitis C virus (HCV). The standard also provides detailed directions for decontamination and the safe handling of potentially infectious laboratory supplies and equipment, including practices for managing laundry and infectious wastes. Employee information and training are covered regarding recognition of hazards and risk of infection. There is also a requirement for HBV vaccination or formal declination within 10 days of assuming duties that present exposure. In the event of an actual exposure, the standard outlines the procedure for post exposure medical evaluation, counseling, and recommended testing or post exposure prophylaxis.

Hazard Communication [29 CFR 1910.1200][4]

This subpart to OSHA's Toxic and Hazardous Substances regulations is intended to ensure that the hazards of all chemicals used in the workplace have been evaluated and that this hazard information is successfully transmitted to employers and their employees who use the substances. Informally referred to as the OSHA "HazCom Standard," it defines *hazardous substances* and provides guidance for evaluating and communicating identified hazards. The primary means of communication are through proper labeling, the development and use of **Safety Data Sheets (SDSs)**, and employee education.[5]

Occupational Exposure to Hazardous Chemicals in Laboratories [29 CFR 1910.1450][4]

This second subpart to OSHA's Toxic and Hazardous Substances regulations is also known as the "OSHA Lab Standard." It was intended to address the shortcomings of the Hazard Communication Standard regarding its application peculiar to the handling of hazardous chemicals in laboratories, whose multiple small-scale manipulations differ from the industrial volumes and processes targeted by the original HazCom Standard. The Lab Standard requires the appointment of a *chemical hygiene officer* and the development of a **chemical hygiene plan** to reduce or eliminate occupational exposure to hazardous chemicals. This plan is required to describe the

Case Study 2.1, Part 2

1. A coworker asked Miles which bloodborne pathogen causes the most infections in laboratory personnel. How should Miles respond?

laboratory's methods of identifying and controlling physical and health hazards presented by chemical manipulations, containment, and storage. The chemical hygiene plan must detail engineering controls, PPE, safe work practices, and administrative controls, including provisions for medical surveillance and consultation, when necessary.

Other Regulations and Guidelines

There are other federal regulations relating to laboratory safety, such as the Clean Water Act, the Resource Conservation and Recovery Act (RCRA), and the Toxic Substances Control Act. In addition, clinical laboratories are required to comply with applicable local and state laws, such as fire and building codes. The Clinical and Laboratory Standards Institute (CLSI, formerly National Committee for Clinical Laboratory Standards [NCCLS]) provides excellent general laboratory safety and infection control guidelines in their documents GP17-A3 (*Clinical Laboratory Safety; Approved Guideline*, Second Edition) and M29-A4 (*Protection of Laboratory Workers from Occupationally Acquired Infections; Approved Guideline*, Third Edition).[6]

Safety is also an important part of the requirements for initial and continued accreditation of healthcare institutions and laboratories by voluntary accrediting bodies such as The Joint Commission (TJC; formerly the Joint Commission on Accreditation of Health Care Organizations [JCAHO]) and the Commission on Laboratory Accreditation of the College of American Pathologists (CAP). TJC publishes a yearly accreditation manual for hospitals and the *Accreditation Manual for Pathology and Clinical Laboratory Services*, which includes a detailed section on safety requirements. CAP publishes an extensive inspection checklist (*Laboratory General Checklist*) as part of their *Laboratory Accreditation Program*, which includes a section dedicated to laboratory safety.

Over the past decade, several new laws and directives have been instituted regarding enhanced security measures for specific hazardous substances with potential for nefarious use in terrorist activities. These initiatives are typically promulgated by the Department of Homeland Security in cooperation with the respective agency regulating chemical, nuclear, or biological agents of concern. Although most laboratories do not store or use the large volumes of chemicals required to trigger chemical security requirements, many laboratories do surpass the thresholds for radiological and biological agents. Management and employees must be cognizant of security requirements for substances in quantities qualifying them for regulation under enhanced security measures for chemical (Chemical Facilities Anti-Terrorism Standards, 6 CFR 27), radiological (Nuclear Regulatory Commission [NRC] *Security Orders* and *Increased Controls* for licensees holding sources above *Quantities of Concern*), and biological (Select Agents and Toxins, 42 CFR 73) agents. Most security measures involve restriction of access to only approved or authorized individuals, assessment of security vulnerabilities, secure physical containment of the agents, and inventory monitoring and tracking.

Safety Awareness for Clinical Laboratory Personnel

Safety Responsibility

The employer and the employee share safety responsibility. While the individual employee has an obligation to follow safe work practices and be attentive to potential hazards, the employer has the ultimate responsibility for safety and delegates authority for safe operations to laboratory managers and supervisors. To ensure clarity and consistency, safety management in the laboratory should start with a written safety policy. Laboratory supervisors, who reflect the attitudes of management toward safety, are essential members of the safety program.[7]

Employer's Responsibilities

- Establish laboratory work methods and safety policies.
- Provide supervision and guidance to employees.
- Provide safety information, training, PPE, and medical surveillance to employees.
- Provide and maintain equipment and laboratory facilities that are free of recognized hazards and adequate for the tasks required.

The employee also has a responsibility for his or her own safety and the safety of coworkers. Employee conduct in the laboratory is a vital factor in the achievement of a workplace without accidents or injuries.

Employees' Responsibilities

- Know and comply with the established laboratory safe work practices.
- Have a positive attitude toward supervisors, coworkers, facilities, and safety training.

- Be alert and give prompt notification of unsafe conditions or practices to the immediate supervisor and ensure that unsafe conditions and practices are corrected.
- Engage in the conduct of safe work practices and use of PPE.

Signage and Labeling

Appropriate signs to identify hazards are critical, not only to alert laboratory personnel to potential hazards but also to identify specific hazards that may arise because of emergencies such as fire or explosion. The **National Fire Protection Association (NFPA)** developed a standard hazard identification system (diamond-shaped, color-coded symbol), which has been adopted by many clinical laboratories. Immediately, the emergency personnel can assess health hazards (blue quadrant), flammable hazards (red quadrant), reactivity/stability hazards (yellow quadrant), and other special information (white quadrant). In addition, each quadrant shows the magnitude of severity, graded from a low of 0 to a high of 4, of the hazards within the posted area. This is what is depicted on the sides of gasoline tanker trucks.

Manufacturers of laboratory chemicals also provide precautionary labeling information for users. Information indicated on the product label includes statement of the hazard, precautionary measures, specific hazard class, first aid instructions for internal/external contact, the storage code, the safety code, and personal protective gear and equipment needed. This information is in addition to specifications on the actual lot analysis of the chemical constituents and other product notes (Figure 2.1). Over the last two decades, there has been an effort to standardize hazard terminology and classification under an internationally recognized guideline, titled the *Globally Harmonized System of Classification and Labeling of Hazardous Chemicals* (GHS)[8] (**Figure 2.1**). This system incorporates universal definitions and symbols to clearly communicate specific hazards in a single concise label format (**Figure 2.2**). Although GHS is not yet law or codified as a regulatory standard,

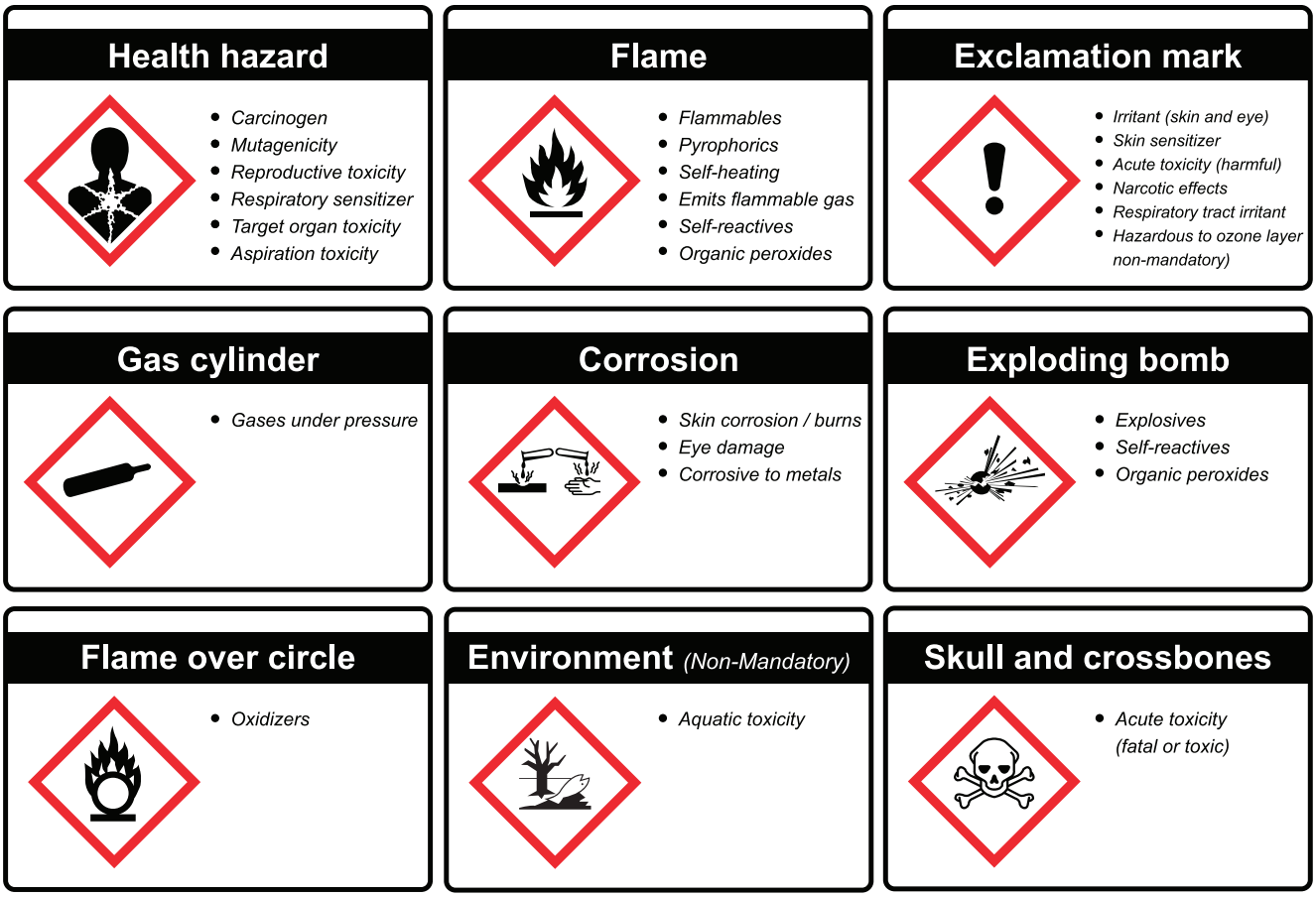


Figure 2.1 Pictographs for general hazards communication.

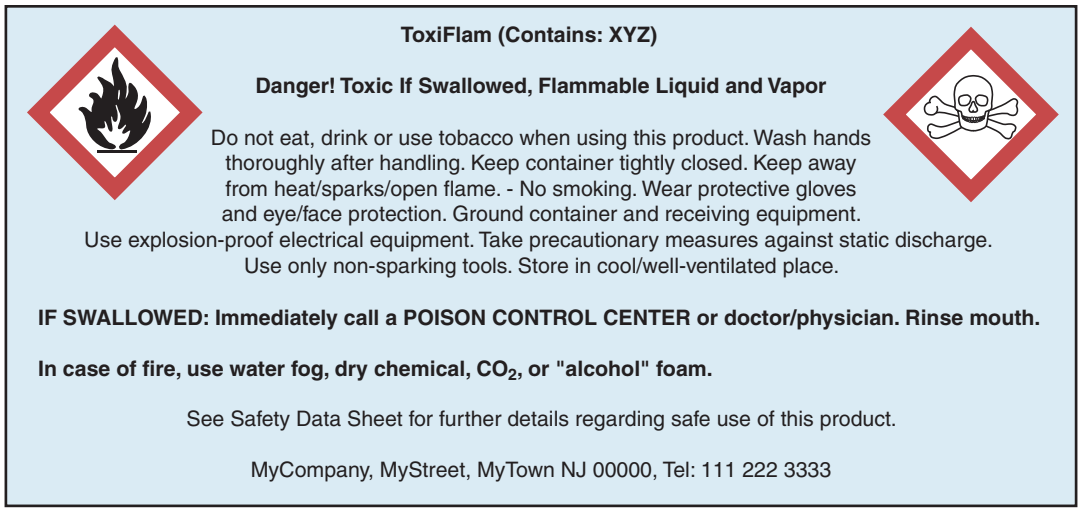


Figure 2.2 Example of a GHS inner container label (e.g., bottle inside a shipping box).

OSHA is presently working to align the existing Hazard Communication Standard with provisions of the GHS and encourages employers to begin adopting the program.

All in-house prepared reagents and solutions should be labeled in a standard manner and include the chemical identity, concentration, hazard warning, special handling, storage conditions, date prepared, expiration date (if applicable), and preparer's initials.

Safety Equipment

Safety equipment has been developed specifically for use in the clinical laboratory. The employer is required by law to have designated safety equipment available, but it is also the responsibility of the employee to comply with all safety rules and to use safety equipment.

All laboratories are required to have safety showers, eyewash stations, and fire extinguishers and to periodically test and inspect the equipment for proper operation. It is recommended that safety showers deliver 30 to 50 gallons of water per minute at 20 to 50 pounds per square inch (psi) and be located in areas where corrosive liquids are stored or used. Eyewash stations must be accessible (i.e., within 100 ft or 10 s travel) in laboratory areas presenting chemical or biological exposure hazards. Other items that must be available for personnel include fire blankets, spill kits, and first aid supplies.

Mechanical pipetting devices must be used for manipulating all types of liquids in the laboratory, including water. Mouth pipetting is strictly prohibited.

Chemical Fume Hoods and Biosafety Cabinets

Fume Hoods

Fume hoods are required to contain and expel noxious and hazardous fumes from chemical reagents. Fume hoods should be visually inspected for blockages. A piece of tissue paper placed at the hood opening will indicate airflow direction. The hood should never be operated with the sash fully opened, and a maximum operating sash height should be established and conspicuously marked. Containers and equipment positioned within hoods should not block airflow. Periodically, ventilation should be evaluated by measuring the face velocity with a calibrated velocity meter. The velocity at the face of the hood (with the sash in normal operating position) should be 100 to 120 ft per minute and fairly uniform across the entire opening. Smoke testing is also recommended to locate no-flow or turbulent areas in the working space. As an added precaution, personal air monitoring should be conducted in accordance with the chemical hygiene plan of the facility.

Biosafety Cabinets

Biological safety cabinets (BSCs) remove particles that may be harmful to the employee who is working with potentially infectious biologic specimens. The Centers for Disease Control and Prevention (CDC) and the National Institutes of Health have described four levels of biosafety, which consist of combinations of

laboratory practices and techniques, safety equipment, and laboratory facilities. The biosafety level of a laboratory is based on the operations performed, the routes of transmission of the infectious agents, and the laboratory function or activity.[6] Accordingly, biosafety cabinets are designed to offer various levels of protection, depending on the biosafety level of the specific laboratory (**Table 2.1**). BSCs should be periodically recertified to ensure continued optimal performance as filter occlusion or rupture can compromise their effectiveness.

Chemical Storage Equipment

Safety equipment is available for the storage and handling of hazardous chemicals and compressed gases. Safety carriers should always be used to transport glass bottles of acids, alkalis, or organic solvents in volumes larger than 500 mL, and approved safety cans should be used for storing, dispensing, or disposing of flammables in volumes greater than 1 quart. Steel safety cabinets with self-closing doors are required for the storage of flammable liquids, and only specially designed, explosion-proof refrigerators may be used to store flammable materials. Only the amount of chemical needed for that day should be available at the bench. Gas cylinder supports or clamps must be used at all times, and larger cylinders should be transported with valve caps on, using handcarts.

PPE and Hygiene

The parts of the body most frequently subject to injury in the clinical laboratory are the eyes, skin, and respiratory and digestive tracts. Hence, the use of PPE and proper hygiene is very important. Safety glasses, goggles, visors, or work shields protect the eyes and face from splashes and impact. Contact lenses do not offer eye protection. Therefore, it is strongly recommended that they not be worn in the clinical chemistry laboratory. However, if worn, protective eyewear is required. If any solution is accidentally splashed into the eye(s), thorough irrigation is required.

Gloves and rubberized sleeves protect the hands and arms when using caustic chemicals. Gloves are required for routine laboratory use. Nitrile gloves, for example, offer a wider range of compatibility with organic solvents than do latex gloves and therefore are routinely used. Laboratory coats, preferably with knit-cuffed sleeves, should be full length, buttoned, and made of liquid-resistant material. When performing manipulations prone to splash hazards,

Table 2.1 Comparison of Biosafety Cabinet Characteristics

BSC Class	Face Velocity	Airflow Pattern	Applications	
			Nonvolatile Toxic Chemicals and Radionuclides	Volatile Toxic Chemicals and Radionuclides
I	75	In at front through HEPA to the outside or into the room through HEPA	Yes	When exhausted outdoors
I, B2	100	No recirculation; total exhaust to the outside through a HEPA filter	Yes	Yes (small amounts)
II, A1	75	70% recirculated to the cabinet work area through HEPA; 30% balance can be exhausted through HEPA back into the room or to outside through a canopy unit	Yes (minute amounts)	No
II, A2	100	Similar to II, A1, but has 100 lfm intake air velocity and plenums are under negative pressure to room; exhaust air can be ducted to the outside through a canopy unit	Yes	When exhausted outdoors (*formally* "B3") (minute amounts)
II, B1	100	30% recirculated, 70% exhausted. Exhaust cabinet air must pass through a dedicated duct to the outside through a HEPA filter	Yes	Yes (minute amounts)

BSC, biological safety cabinet; HEPA, high-efficiency particulate air; lfm, linear feet per minute.

Adapted from Centers for Disease Control and Prevention, National Institutes of Health. *Biosafety in Microbiological and Biomedical Laboratories*. 5th ed. Washington, DC: U.S. Government Printing Office; 2009.

Case Study 2.1, Part 3

2. A coworker asked Mía what the minimum requirement for PPE in the laboratory was. How should Mía respond?
3. Miles reviewed the PPE audit from the previous quarter and noticed it was reported as 100% each month. As he observed his coworkers, many had lab coats on, but the lab coats were unbuttoned. One of his most experienced coworkers was not wearing a lab coat. Instead, it was hanging on the back of the chair. What data should Miles include in his PPE audit report?

© dotshock/Shutterstock.

© Ariel Skelley/DigitalVision/Getty Images.

the laboratory coat should be supplemented with an impermeable apron and/or sleeve garters, constructed of suitable material to guard against the substances. Proper footwear is required. Shoes constructed of porous materials, open-toed shoes, and sandals are considered ineffective against spilled hazardous liquids and should not be worn in the laboratory.

Respirators may be required for various procedures in the clinical laboratory. Whether used for biologic or chemical hazards, the correct type of respirator must be used for the specific hazard. Respirators with **high-efficiency particulate air (HEPA) filters** must be worn when engineering controls are not feasible, such as when working directly with patients with tuberculosis (TB) or when performing procedures that may aerosolize specimens of patients with a suspected or confirmed case of TB. Training, maintenance, and written protocol for use of respirators are required according to the respiratory protection standard.

Each employer must provide (at no charge) fluid-resistant laboratory coats, gloves, or other protective equipment to all employees who may be exposed to biologic or chemical hazards. It is the employer's responsibility to clean and maintain any PPE used by more than one person. All contaminated PPE must be removed and properly cleaned or disposed of before leaving the laboratory.

Hand washing is a crucial component of both infection control and chemical hygiene. After removing gloves, hands should be washed thoroughly with soap and warm water, even if glove breakthrough or contamination is not suspected. The use of antimicrobial soap is not as important as the physical action of washing the hands with water and any mild soap. After any work with highly toxic or carcinogenic chemicals, the face should also be washed.

Biologic Safety

General Considerations

All blood samples and other body fluids should be collected, transported, handled, and processed using standard precautions (i.e., presumed to be infectious). Gloves, gowns, and face protection must be used during manipulations or transfers when splashing or splattering is most likely to occur. Consistent and thorough hand washing is an essential component of infection control. Antiseptic gels and foams may be used at waterless stations between washes, but they should not take the place of an actual hand wash.[6,7]

Centrifugation of biologic specimens produces finely dispersed aerosols that are a high-risk source of infection. Ideally, specimens should remain capped during centrifugation, or several minutes should be allowed to elapse after centrifugation is complete before opening the lid. As a preferred option, the use of a sealed-cup centrifuge is recommended. These sealed vessels can then be brought to a biosafety cabinet to be opened.

Spills

Any blood, body fluid, or other potentially infectious material spill must be promptly cleaned up, and the area or equipment must be disinfected immediately. Safe cleanup includes the following recommendations:

- Alert others in area of the spill.
- Wear appropriate protective equipment.
- Use mechanical devices to pick up broken glass or other sharp objects.
- Absorb the spill with paper towels, gauze pads, or tissue.

- Clean the spill site using a common aqueous detergent.
- Disinfect the spill site using approved disinfectant or 10% bleach, using appropriate contact time.
- Rinse the spill site with water.
- Dispose of all materials in appropriate biohazard containers.

Bloodborne Pathogens

In December 1991, OSHA issued the final rule for occupational exposure to *bloodborne pathogens*. To minimize employee exposure, each employer must have a written *exposure control plan*. The plan must be available to all employees whose duties may result in reasonably anticipated occupational exposure to blood or other potentially infectious materials. The exposure control plan must be discussed with all employees and be available to them while they are working. The employee must be provided with adequate training in all techniques described in the exposure control plan at initial work assignment and annually thereafter. All necessary safety equipment and supplies must be readily available and inspected on a regular basis.[9]

Clinical laboratory personnel are knowingly or unknowingly in frequent contact with potentially biohazardous materials. In recent years, new and serious occupational hazards to personnel have arisen, and this problem has been complicated because of the general lack of understanding of the epidemiology, mechanisms of transmission of the disease, or inactivation of the causative agent. Special precautions must be taken when handling all specimens because of the continual increase in the proportion of infectious samples received in the laboratory. Therefore, in practice, specimens from patients with confirmed or suspected hepatitis, acquired immunodeficiency syndrome (AIDS), or other potentially infectious diseases should be handled no differently from other routine specimens. Adopting a standard precautions policy, which considers blood and other body fluids from all patients as potentially infective, is required.

Airborne Pathogens

Because of a global resurgence of TB, OSHA issued a statement in 1993 that the agency would enforce CDC Guidelines for Preventing the Transmission of Tuberculosis in Health Care Facilities. The purpose of the guidelines is to encourage early detection, isolation, and treatment of active cases. A TB exposure control program must be established, and risks to laboratory workers must be assessed. In 1997, a proposed standard (29 CFR 1910.1035, Tuberculosis)[4] was issued by OSHA only to be withdrawn again when it was determined that existing CDC guidelines could be enforced by OSHA through its "general duty" clause and Respiratory Protection Standard. The CDC guidelines require the development of a *tuberculosis infection control program* by any facility involved in the diagnosis or treatment of cases of confirmed infectious TB. TB isolation areas with specific ventilation controls must be established in healthcare facilities. Those workers in high-risk areas may be required to wear a respirator for protection. All healthcare workers considered to be at risk must be screened for TB infection.

Other specific pathogens, including viruses, bacteria, and fungi, may be considered **airborne pathogens** transmission risks. Protective measures in the clinical laboratory generally involve work practice and engineering controls focused on prevention of aerosolized particles, containment/isolation, and respiratory protection of N-95 (filtration of 95% of particles >0.3 µm) or better.

Shipping

Clinical laboratories routinely ship regulated material. The U.S. Department of Transportation (DOT) and the International Air Transport Association (IATA) have specific requirements for carrying regulated materials. There are two types of specimen classifications. Known or suspect infectious specimens are labeled *infectious substances* if the pathogen can be readily transmitted to humans or animals. *Diagnostic specimens* are those tested as routine screening or for initial diagnosis. Each type of specimen has rules and packaging requirements. The DOT guidelines are found in the *Code of Federal Regulations, Title 49, Subchapter C*; IATA publishes its own manual, *Dangerous Goods Regulations*.

Chemical Safety[10]

Hazard Communication

In the August 1987 issue of the *Federal Register*, OSHA published the new **Hazard Communication Standard** (Right to Know Law, 29 CFR 1910.1200).[4] The Right to Know Law was developed for employees who may be exposed to hazardous chemicals in the workplace. Employees must be informed of the health risks associated with those chemicals. The

intent of the law is to ensure that health hazards are evaluated for all chemicals that are produced and that this information is relayed to employees.

To comply with the regulation, clinical laboratories must:

- Plan and implement a written hazard communication program.
- Obtain SDSs for each hazardous compound present in the workplace and have the SDSs readily accessible to employees.
- Educate all employees annually on how to interpret chemical labels, SDSs, and health hazards of the chemicals and how to work safely with the chemicals.
- Maintain hazard warning labels on containers received or filled on-site.

In 2012, OSHA adopted significant changes to the Hazard Communication Standard to facilitate a standardization of international hazard communication programs. This new initiative was titled the Globally Harmonized System of Classification and Labelling of Chemicals, or GHS.[8] The primary improvements to the program involved more specific criteria for classification of chemicals; a uniform system of chemical labeling, including intuitive pictographs; and replacing the existing Material Safety Data Sheet (MSDS) program with the new SDS format. These changes were phased in over a 3-year period, with the final requirements effective in June of 2016.

Safety Data Sheet

The SDS is a major source of safety information for employees who may use **hazardous materials** in their occupations. Employers are responsible for obtaining the SDS from the chemical manufacturer or developing an SDS for each hazardous agent used in the workplace. The information contained in the SDS must follow a specific format, addressing each of the following 16 items:

- Section 1: Identification
- Section 2: Hazard identification
- Section 3: Ingredients information
- Section 4: First aid procedures
- Section 5: Fire-fighting procedures
- Section 6: Accidental-release measures
- Section 7: Handling and storage
- Section 8: Exposure controls and personal protection
- Section 9: Physical and chemical properties
- Section 10: Stability and reactivity
- Section 11: Toxicological information
- Section 12: Ecological information
- Section 13: Disposal considerations
- Section 14: Transport information
- Section 15: Regulatory information
- Section 16: Other information, including date of preparation or last revision

The SDS must provide the specific compound identity, together with all common names. All information sections must be completed, and the date that the SDS was printed must be indicated. Copies of the SDS must be readily accessible to employees during all shifts.

OSHA Laboratory Standard

Occupational Exposure to Hazardous Chemicals in Laboratories (29 CFR 1910.1450),[4] also known as the **laboratory standard**, was enacted in May 1990 to provide laboratories with specific guidelines for handling hazardous chemicals. This OSHA standard requires each laboratory that uses hazardous chemicals to have a written **chemical hygiene plan**. This plan provides procedures and work practices for regulating and reducing exposure of laboratory personnel to hazardous chemicals. *Hazardous chemicals* are those that pose a physical or health hazard from acute or chronic exposure. Procedures describing how to protect employees against **teratogens** (substances that affect cellular development in a fetus or embryo), carcinogens, and other toxic chemicals must be described in the plan. Training in the use of hazardous chemicals must be provided to all employees and must include recognition of signs and symptoms of exposure, location of SDS, the chemical hygiene plan, and how to protect themselves against hazardous chemicals. A chemical hygiene officer must be designated for any laboratory using hazardous chemicals. The protocol must be reviewed annually and updated when regulations are modified or chemical inventory changes. Remember that practicing consistent and thorough hand washing is an essential component of preventative chemical hygiene.

Toxic Effects from Hazardous Substances

Toxic substances have the potential of producing deleterious effects (local or systemic) by direct chemical action or interference with the function of body systems. They can cause acute or chronic effects related to the duration of exposure (i.e., short-term, or single contact, versus long-term,

or prolonged, repeated contact). Almost any substance, even the most seemingly benign, can pose risk of damage to a worker's lungs, skin, eyes, or mucous membranes following long- or short-term exposure and can be toxic in excess. Moreover, some chemicals are toxic at very low concentrations. Exposure to toxic agents can be through direct contact (absorption), inhalation, ingestion, or inoculation/injection.

In the clinical chemistry laboratory, laboratorians should be particularly aware of toxic vapors from chemical solvents, such as acetone, chloroform, methanol, or carbon tetrachloride, that do not give explicit sensory irritation warnings, as do bromide, ammonia, and formaldehyde. Air sampling or routine monitoring may be necessary to quantify dangerous levels. Mercury is another frequently disregarded source of poisonous vapors. It is highly volatile and toxic and is rapidly absorbed through the skin and respiratory tract. Mercury spill kits should be available in areas where mercury thermometers are used. Most laboratories are phasing out the use of mercury and mercury-containing compounds. Laboratories should have a policy on mercury reduction or elimination and a method for legally disposing of mercury. Several compounds, including formaldehyde and methylene chloride, have substance-specific OSHA standards, which require periodic monitoring of air concentrations. Laboratory engineering controls, PPE, and procedural controls must be adequate to protect employees from these substances.

Storage and Handling of Chemicals

To avoid accidents when handling chemicals, it is important to develop respect for all chemicals and to have a complete knowledge of their properties. This is particularly important when transporting, dispensing, or using chemicals that, when in contact with certain other chemicals, could result in the formation of substances that are toxic, flammable, or explosive. For example, acetic acid is incompatible with other acids such as chromic and nitric acid, carbon tetrachloride is incompatible with sodium, and flammable liquids are incompatible with hydrogen peroxide and nitric acid.

Arrangements for the storage of chemicals will depend on the quantities of chemicals needed and the nature or type of chemicals. Proper storage is essential to prevent and control laboratory fires and accidents. Ideally, the storeroom should be organized so that each class of chemicals is isolated in an area that is not used for routine work. An up-to-date inventory should be kept that indicates location of chemicals, minimum/maximum quantities required, and shelf life. Some chemicals deteriorate over time and become hazardous (e.g., many ethers and tetrahydrofuran form explosive peroxides). Storage should not be based solely on alphabetical order because incompatible chemicals may be stored next to each other and react chemically. They must be separated for storage, as shown in **Table 2.2**.

Table 2.2 Storage Requirements

Substance	Stored Separately
Flammable liquids	Flammable solids
Mineral acids	Organic acids
Caustics	Oxidizers
Perchloric acid	Water-reactive substances
Air-reactive substances	Others
Heat-reactive substances requiring refrigeration	
Unstable substances (shock-sensitive explosives)	

Flammable/Combustible Chemicals

Flammable and combustible liquids, which are used in numerous routine procedures, are among the most hazardous materials in the clinical chemistry laboratory because of possible fire or explosion. They are classified according to flash point, which is the temperature at which sufficient vapor is given off to form an ignitable mixture with air. A flammable liquid has a flash point below 37.8°C (100°F), and combustible liquids, by definition, have a flash point at or above 37.8°C (100°F). Some commonly used flammable and combustible solvents are acetone, benzene, ethanol, heptane, isopropanol, methanol, toluene, and xylene. It is important to remember that flammable or combustible chemicals also include certain gases, such as hydrogen, and solids, such as paraffin.

Corrosive Chemicals

Corrosive chemicals are injurious to the skin or eyes by direct contact or to the tissue of the respiratory and gastrointestinal tracts if inhaled or ingested.

Typical examples include acids (acetic, sulfuric, nitric, and hydrochloric) and bases (ammonium hydroxide, potassium hydroxide, and sodium hydroxide). External exposures to concentrated corrosives can cause severe burns and require immediate flushing with copious amounts of clean water.

Reactive Chemicals

Reactive chemicals are substances that, under certain conditions, can spontaneously explode or ignite or that evolve heat or flammable or explosive gases. Some strong acids or bases react with water to generate heat (exothermic reactions). Hydrogen is liberated if alkali metals (sodium or potassium) are mixed with water or acids, and spontaneous combustion also may occur. The mixture of oxidizing agents, such as peroxides, and reducing agents, such as hydrogen, generates heat and may be explosive.

Carcinogenic Chemicals

Carcinogens are substances that have been determined to be cancer-causing agents. OSHA has issued lists of confirmed and suspected carcinogens and detailed standards for the handling of these substances. Benzidine is a common example of a known carcinogen. If possible, a substitute chemical or different procedure should be used to avoid exposure to carcinogenic agents. For regulatory (OSHA) and institutional safety requirements, the laboratory must maintain an accurate inventory of carcinogens.

Chemical Spills

Strict attention to good laboratory technique can help prevent chemical spills. However, emergency procedures should be established to handle any accidents. If a spill occurs, the first step should be to assist/evacuate personnel, and then confinement and cleanup of the spill can begin. There are several commercial spill kits available for neutralizing and absorbing spilled chemical solutions (**Figure 2.3**). However, no single kit is suitable for all types of spills. Emergency procedures for spills should also include a reporting system.

Figure 2.3 Spill cleanup kit.

Radiation Safety

Environmental Protection

A radiation safety policy should include environmental and personnel protection. All areas where **radioactive materials** are used or stored must be posted with caution signs, and traffic in these areas should be restricted to essential personnel only. Regular and systematic monitoring must be emphasized, and decontamination of laboratory equipment, glassware, and work areas should be scheduled as part of routine procedures. Records must be maintained as to the quantity of radioactive material on hand as well as the quantity that is disposed. A Nuclear Regulatory Commission (NRC) or agreement state license is required if the amount of radioactive material exceeds a certain level. The laboratory safety officer must consult with the institutional safety officer about these requirements.

Personal Protection

It is essential that only properly trained personnel work with radioisotopes. Good work practices must consistently be employed to ensure that contamination and inadvertent internalization are avoided. Users should be monitored to ensure that the maximal permissible dose of radiation is not exceeded. Radiation monitors must be evaluated regularly to detect degree of exposure for the laboratory

Table 2.3 Examples of Nonionizing Radiation in Clinical Laboratories

Type	Approximate Wavelength	Source Equipment Example	Protective Measures
Low frequency	>1 cm	Radiofrequency coil in inductively coupled plasma–mass spectrometer	Engineered shielding and posted pacemaker warning
Microwaves	3 m–3 mm	Energy beam microwave used to accelerate tissue staining in histology prep processes	Engineered shielding
Infrared	750 nm–0.3 cm	Heat lamps, lasers	Containment and appropriate warning labels
Visible spectrum	400–750 nm	General illumination and glare	Filters, diffusers, and nonreflective surfaces
Ultraviolet	4–400 nm	Germicidal lamps used in biologic safety cabinets	Eye and skin protection; UV warning labels

employee. Records must be maintained for the length of employment plus 30 years.

Nonionizing Radiation

Nonionizing forms of radiation are also a concern in the clinical laboratory. Equipment often emits a variety of wavelengths of electromagnetic radiation that staff must be protected against through engineered shielding or use of PPE (**Table 2.3**). These energies have varying biologic effects, depending on wavelength, power intensity, and duration of exposure. Laboratorians must be knowledgeable regarding the hazards presented by their equipment to protect themselves and ancillary personnel.

Fire Safety

The Chemistry of Fire

Fire is basically a chemical reaction that involves the rapid oxidation of a combustible material or fuel, with the subsequent liberation of heat and light. In the clinical chemistry laboratory, all the elements essential for fire to begin are present—fuel, heat or ignition source, and oxygen (air). However, recent research suggests that a fourth factor is present. This factor has been classified as a reaction chain in which burning continues and even accelerates. It is caused by the breakdown and recombination of the molecules from the material burning with the oxygen in the atmosphere.

The fire triangle has been modified into a three-dimensional pyramid known as the **fire tetrahedron** (**Figure 2.4**). This modification does not contradict established procedures in dealing with a fire but does provide additional means by which fires may be prevented or extinguished. A fire will extinguish if any of the three basic elements (heat, air, or fuel) is removed.

Classification of Fires

Fires have been divided into four classes based on the nature of the combustible material and requirements for extinguishment:

Class A: ordinary combustible solid materials, such as paper, wood, plastic, and fabric

Class B: flammable liquids/gases and combustible petroleum products

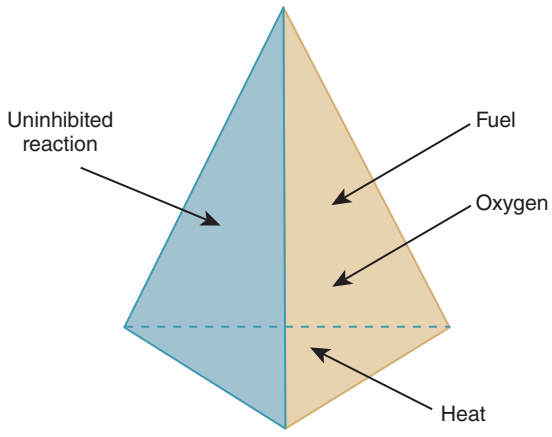


Figure 2.4 Fire tetrahedron.

Class C: energized electrical equipment
Class D: combustible/reactive metals, such as magnesium, sodium, and potassium

Types and Applications of Fire Extinguishers

Just as fires have been divided into classes, fire extinguishers are divided into classes that correspond to the type of fire to be extinguished. Be certain to choose the right type—using the wrong type of extinguisher may be dangerous. For example, do not use water on burning liquids or electrical equipment.

Pressurized water extinguishers, as well as foam and multipurpose dry chemical types, are used for Class A fires. Multipurpose dry-chemical and carbon dioxide extinguishers are used for Class B and C fires. Halogenated hydrocarbon extinguishers are particularly recommended for use with computer equipment. Class D fires present special problems, and extinguishment is left to trained firefighters using special dry chemical extinguishers (**Figure 2.5**). Generally, all that can be done for a Class D fire in the laboratory is to try and isolate the burning metal from combustible surfaces with sand or ceramic barrier material. Personnel should know the location and type of portable fire extinguisher near their work area and know how to use an extinguisher before a fire occurs.

Case Study 2.1, Part 4

4. Help Mía explain the RACE and PASS acronyms to her coworkers.

© Ariel Skelley/DigitalVision/Getty Images.

In the event of a fire, remember the acronym RACE: First *rescue* all personnel, patients, and visitors who are in immediate danger. Next, *activate* the fire alarm and report the fire. It is important to *confine* the fire by closing doors. Lastly, *extinguish* the fire if possible. Personnel should work as a team to carry out emergency procedures. Fire drills must be conducted regularly and with appropriate documentation. Fire extinguishers must be inspected monthly to ensure that they are mounted, visible, accessible, and charged.

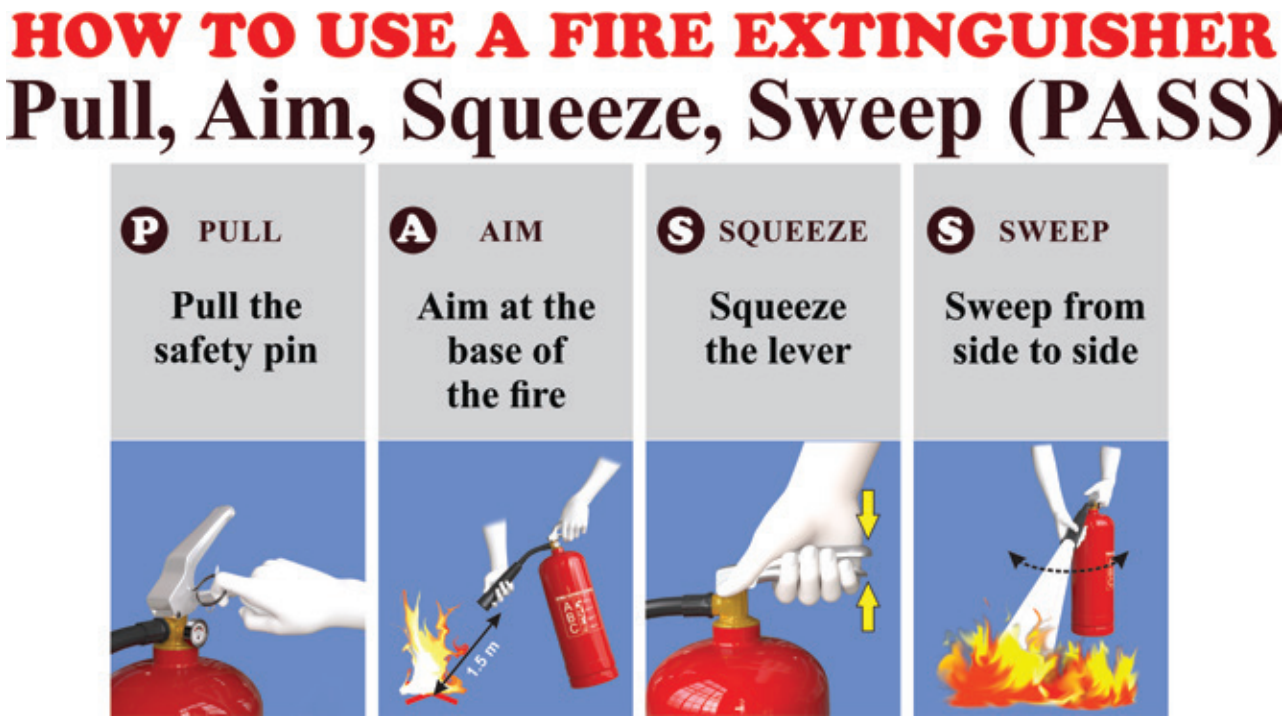


Figure 2.5 Proper use of fire extinguishers.

© Ajamal/Shutterstock.

Control of Other Hazards

Electrical Hazards

Most individuals are aware of the potential hazards associated with the use of electrical appliances and equipment. Direct hazards of electrical energy can result in death, shock, or burns. Indirect hazards can result in fire or explosion. Therefore, there are many precautionary procedures to follow when operating or working around electrical equipment:

- Use only explosion-rated (intrinsically wired) equipment in hazardous atmospheres.
- Be particularly careful when operating high-voltage equipment, such as electrophoresis apparatus.
- Use only properly grounded equipment (three-prong plug).
- Check for frayed electrical cords.
- Promptly report any malfunctions or equipment producing a "tingle" for repair.
- Do not work on "live" electrical equipment.
- Never operate electrical equipment with wet hands.
- Know the exact location of the electrical control panel for the electricity to your work area.
- Use only approved extension cords in temporary applications and do not overload circuits. (Some local regulations prohibit the use of any extension cord.)
- Have ground, polarity, and leakage checks and other periodic preventive maintenance performed on outlets and equipment.

Compressed Gas Hazards

Compressed gases, which serve a number of functions in the laboratory, present a unique combination of hazards in the clinical laboratory: danger of fire, explosion, asphyxiation, or mechanical injuries. There are several general requirements for safely handling compressed gases:

- Know the gas that you will use.
- Store tanks in a vertical position.
- Keep cylinders secured at all times.
- Never store flammable liquids and compressed gases in the same area.
- Use the proper regulator, tubing, and fittings for the type of gas in use.
- Do not attempt to control or shut off gas flow with the pressure relief regulator.
- Keep removable protection caps in place until the cylinder is in use.
- Make certain that acetylene tanks are properly piped (the gas is incompatible with copper tubing).
- Do not force a "frozen" or stuck cylinder valve.
- Use a hand truck to transport large cylinders.
- Always check cylinders on receipt and then periodically for any problems such as leaks.
- Make certain that the cylinder is properly labeled to identify the contents.
- Empty tanks should be marked "empty."

Cryogenic Materials Hazards

Liquid nitrogen is probably one of the most widely used cryogenic fluids (liquefied gases) in the laboratory. There are, however, several hazards associated with the use of any **cryogenic material**: fire or explosion, asphyxiation, pressure buildup, embrittlement of materials, and tissue damage similar to that of thermal burns.

Only containers constructed of materials designed to withstand ultralow temperatures should be used for cryogenic work. In addition to the use of eye/face protection, hand protection to guard against the hazards of touching supercooled surfaces is recommended. The gloves, of impermeable material, should fit loosely so that they can be taken off quickly if liquid spills on or into them. Also, to minimize violent boiling/frothing and splashing, specimens to be frozen should always be inserted into the coolant very slowly. Cryogenic fluids should be stored in well-insulated but loosely stoppered containers that minimize loss of fluid resulting from evaporation by boil-off and that prevent plugging and pressure buildup.

Mechanical Hazards

In addition to physical hazards such as fire and electric shock, laboratory personnel should be aware of the **mechanical hazards** of equipment such as centrifuges, autoclaves, and homogenizers.

Centrifuges, for example, must be balanced to distribute the load equally. The operator should never open the lid until the rotor has come to a complete stop. Safety interlocks on equipment should never be rendered inoperable.

Laboratory glassware itself is another potential hazard. Agents such as glass beads or boiling chips should be added to help eliminate bumping and boiling over when liquids are heated. Tongs or insulated gloves should be used to remove hot glassware from ovens, hot plates, or water baths. Glass pipettes should be handled with extra care, as should sharp

instruments such as needles, scalpel blades, and other tools. A glassware inspection program should be in place to detect signs of wear or fatigue that could contribute to breakage or injury. All infectious *sharps* must be disposed in OSHA-approved containers to reduce the risk of injury and infection.

Ergonomic Hazards

Although increased mechanization and automation have made many tedious and repetitive manual tasks obsolete, laboratory processes often require repeated manipulation of instruments, containers, and equipment. These physical actions can, over time, contribute to repetitive strain disorders such as tenosynovitis, bursitis, and ganglion cysts. The primary contributing factors associated with repetitive strain disorders are position/posture, applied force, and frequency of repetition. Remember to consider the design of hand tools (e.g., ergonomic pipettes), adherence to ergonomically correct technique, and equipment positioning when engaging in any repetitive task. Chronic symptoms of pain, numbness, or tingling in extremities may indicate the onset of repetitive strain disorders. Other hazards include acute musculoskeletal injury. Remember to lift heavy objects properly, keeping the load close to the body and using the muscles of the legs rather than the back. Gradually increase force when pushing or pulling and avoid pounding actions with the extremities.

Disposal of Hazardous Materials

The safe handling and disposal of chemicals and other materials require a thorough knowledge of their properties and hazards. Generators of hazardous wastes have a moral and legal responsibility, as defined in applicable local, state, and federal regulations, to protect both the individual and the environment when disposing off waste. There are four basic waste disposal techniques: flushing down the drain to the sewer system, incineration, landfill burial, and recycling.

Chemical Waste

In some cases, it is permissible to flush water-soluble substances down the drain with copious quantities of water. However, strong acids or bases should be neutralized before disposal. The laboratory must adhere to institutional, local, and state regulations regarding the disposal of strong acids and bases. Foul-smelling chemicals should never be disposed of down the drain. Possible reaction of chemicals in the drain and potential toxicity must be considered when deciding if a particular chemical can be dissolved or diluted and then flushed down the drain. For example, sodium azide, which is used as a preservative in many manufactured reagents, forms explosive salts with metals, such as the copper, in pipes. Some institutions ban the use of sodium azide due to this hazard. Others require special handling and disposal of the chemical. In all cases, check with the local water reclamation district or publicly owned treatment works for specific limitations before utilizing sewer disposal.[11]

Other liquid wastes, including flammable solvents, must be collected in approved containers and segregated into compatible classes. If practical, solvents such as xylene and acetone may be filtered or redistilled for reuse. If recycling is not feasible, disposal arrangements should be made by specifically trained personnel. Flammable material can also be burned in specially designed incinerators with afterburners and scrubbers to remove toxic products of combustion.

Also, before disposal, hazardous substances that are explosive (e.g., peroxides) and carcinogens should be transformed to less hazardous forms whenever feasible. Solid chemical wastes that are unsuitable for incineration may be amenable to other treatments or buried in an approved, permitted landfill. Note that certain chemical wastes are subject to strict "cradle to grave" tracking under the Resource Conservation and Recovery Act (RCRA), and severe penalties are associated with improper storage, transportation, and disposal.[11]

Radioactive Waste

The manner of use and disposal of isotopes is strictly regulated by the NRC or NRC agreement states and depends on the type of waste (soluble or insoluble), its level of radioactivity, and the radiotoxicity and half-life of the isotopes involved. The radiation safety officer should always be consulted about policies dealing with radioactive waste disposal. Many clinical laboratories transfer radioactive materials to a licensed receiver for disposal.

Biohazardous Waste

On November 2, 1988, President Reagan signed into law The Medical Waste Tracking Act of 1988. Its purpose was to (1) charge the Environmental Protection Agency with the responsibility to establish a program

to track medical waste from generation to disposal, (2) define medical waste, (3) establish acceptable techniques for treatment and disposal, and (4) establish a department with jurisdiction to enforce the new laws. Several states have implemented the federal guidelines and incorporated additional requirements. Some entities covered by the rules are any health care–related facility including, but not limited to, ambulatory surgical centers; blood banks and blood drawing centers; clinics, including medical, dental, and veterinary; clinical, diagnostic, pathologic, or biomedical research laboratories; emergency medical services; hospitals; long-term care facilities; minor emergency centers; occupational health clinics and clinical laboratories; and professional offices of physicians and dentists.

Medical waste (regulated waste) is defined as *special waste from health care facilities* and is further defined as solid waste that, if improperly treated or handled, "may transmit infectious diseases." (For additional information, see the digital resources in the Navigate 2 offering that comes with this book. Also see the Joint Commission Web site: https://www.jointcommission.org/standards/standard-faqs/hospital-and-hospital-clinics/environment-of-care-ec/000001237/). It comprises animal waste, bulk blood and blood products, microbiologic waste, pathologic waste, and sharps. The approved methods for treatment and disposition of medical waste are incineration, steam sterilization, burial, thermal inactivation, chemical disinfection, or encapsulation in a solid matrix.[11]

Generators of medical waste must implement the following procedures:

- Employers of healthcare workers must establish and implement an infectious waste program.
- All biomedical waste should be placed in a bag marked with the biohazard symbol and then placed into a leak-proof container that is puncture resistant and equipped with a solid, tight-fitting lid. All containers must be clearly marked with the word **biohazard** or its symbol.
- All sharp instruments, such as needles, blades, and glass objects, should be placed into special puncture-resistant containers before placing them inside the bag and container.
- Needles should not be transported, recapped, bent, or broken by hand.
- All biomedical waste must then be disposed of according to one of the recommended procedures.
- Highly pathogenic waste should undergo preliminary treatment on-site.
- Potentially biohazardous material, such as blood or blood products and contaminated laboratory waste, cannot be directly discarded. Combustible contaminated waste can be incinerated. Noncombustible contaminated waste, such as glassware, should be autoclaved before being discarded. Special attention should be given to the discarding of syringes, needles, and broken glass that could also inflict accidental cuts or punctures. Appropriate containers should be used for discarding these sharp objects.

Accident Documentation and Investigation

Any accidents involving personal injuries, even minor ones, should be reported immediately to a supervisor. Manifestation of occupational illnesses and exposures to hazardous substances should also be reported. Serious injuries and illnesses, including those resulting in hospitalization, disability, or death, must be reported to OSHA or the state-administered program within 8 hours. Under OSHA regulations, employers are required to maintain records of occupational injuries and illnesses for the length of employment plus 30 years. The record-keeping requirements include a first report of injury, an accident investigation report, and an annual summary that is recorded on an OSHA injury and illness log (Form 300).

The first report of injury is used to notify the insurance company and the human resources or safety department that a workplace injury has occurred. The employee and the supervisor usually complete the report, which contains information on the employer and injured person, as well as the time and place, cause, and nature of the injury. The report is signed and dated; then, it is forwarded to the institution's risk manager or insurance representative.

The investigation report should include information on the injured person, a description of what happened, the cause of the accident (environmental or personal), other contributing factors, witnesses, the nature of the injury, and actions to be taken to prevent a recurrence. This report should be signed and dated by the person who conducted the investigation.

Annually, a log and summary of occupational injuries and illnesses should be completed and forwarded to the U.S. Department of Labor, Bureau of Labor Statistics' OSHA injury and illness log

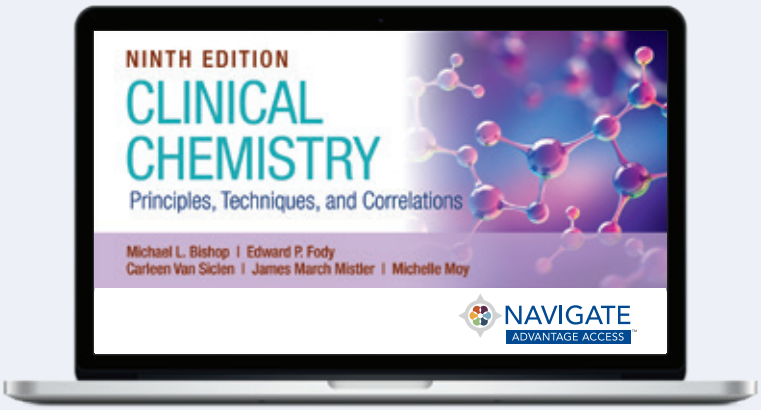
(Form 300). The standardized form requests depersonalized information similar to the first report of injury and the accident investigation report. Information about every occupational death, nonfatal occupational illness, biologic or chemical exposure, and nonfatal occupational injury that involved loss of consciousness, restriction of work or motion, transfer to another job, or medical treatment (other than first aid) must be reported.

Because it is important to determine why and how an accident occurred, an accident investigation should be conducted. Most accidents can be traced to one of two underlying causes: environmental (unsafe conditions) or personal (unsafe acts). Environmental factors include inadequate safeguards, use of improper or defective equipment, hazards associated with the location, or poor housekeeping. Personal factors include improper laboratory attire, lack of skills or knowledge, specific physical or mental conditions, and attitude. The employee's positive motivation is important in all aspects of safety promotion and accident prevention.

It is particularly important that the appropriate authority be notified immediately if any individual sustains a contaminated needle puncture during blood collection or a cut during subsequent specimen processing or handling. For a summary of recommendations for the protection of laboratory workers, refer to *Protection of Laboratory Workers from Occupationally Acquired Infections; Approved Guideline*, Third Edition, M29-A3 (CLSI).[6]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



Bibliography and Suggested Reading

Allocca JA, Levenson HE. *Electrical and Electronic Safety*. Reston, VA: Reston Publishing Company; 1985.

American Chemical Society, Committee on Chemical Safety. Smith GW, ed. *Safety in Academic Chemistry Laboratories*. Washington, DC: American Chemical Society; 1985.

Boyle MP. Hazardous chemical waste disposal management. *Clin Lab Sci.* 1992;5:6.

Brown JW. Tuberculosis alert: an old killer returns. *MLO Med Lab Obs.* 1993;25:5.

Bryan RA. Recommendations for handling specimens from patients with confirmed or suspected Creutzfeldt-Jakob disease. *Lab Med.* 1984;15:50.

Chervinski D. Environmental awareness: it's time to close the loop on waste reduction in the health care industry. *Adv Admin Lab.* 1994;3:4.

Committee on Hazardous Substances in the Laboratory, Assembly of Mathematical and Physical Sciences, National Research Council. *Prudent Practices for Handling Hazardous Chemicals in Laboratories*. Washington, DC: National Academies Press; 1981.

Gile TJ. Hazard-communication program for clinical laboratories. *Clin Lab.* 1988;1:2.

Gile TJ. An update on lab safety regulations. *MLO Med Lab Obs.* 1995;27:3.

Hayes DD. Safety considerations in the physician office laboratory. *Lab Med.* 1994;25:3.

Hazard communication. *Federal Register.* 1994;59:27.

Karcher RE. Is your chemical hygiene plan OSHA proof? *MLO Med Lab Obs.* 1993;25:7.

Le Sueur CL. A three-pronged attack against AIDS infection in the lab. *MLO Med Lab Obs.* 1989;21:37.

Miller SM. Clinical safety: dangers and risk control. *Clin Lab Sci.* 1992;5:6. National Institute for Occupational Safety and Health (NIOSH).

National Institutes of Health, Radiation Safety Branch. *Radiation: The National Institutes of Health Safety Guide*. Washington, DC: U.S. Government Printing Office; 1979.

National Regulatory Committee, Committee on Hazardous Substances in the Laboratory. *Prudent Practices for the Handling of Hazardous Chemicals in Laboratories*. Washington, DC: National Academies Press; 1981.

National Regulatory Committee, Committee on Hazardous Substances in the Laboratory. *Prudent Practices for Disposal of Chemicals from Laboratories*. Washington, DC: National Academies Press; 1983.

National Regulatory Committee, Committee on the Hazardous Biological Substances in the Laboratory. *Biosafety in the Laboratory: Prudent Practices for the Handling and Disposal of Infectious Materials*. Washington, DC: National Academies Press; 1989.

Otto CH. Safety in health care: prevention of bloodborne diseases. *Clin Lab Sci.* 1992;5:6.

Pipitone DA. *Safe Storage of Laboratory Chemicals*. New York, NY: Wiley; 1984.

Rose SL. *Clinical Laboratory Safety*. Philadelphia, PA: JB Lippincott; 1984.

Rudmann SV, Jarus C, Ward KM, et al. Safety in the student laboratory: a national survey of university-based programs. *Lab Med.* 1993;24:5.

Stern A, Ries H, Flynn D, et al. Fire safety in the laboratory: part I. *Lab Med.* 1993;24:5.

Stern A, Ries H, Flynn D, et al. Fire safety in the laboratory: part II. *Lab Med.* 1993;24:6.

Wald PH, Stave GM. *Physical and Biological Hazards of the Workplace*. New York, NY: Van Nostrand Reinhold; 1994.

References

1. National Safety Council (NSC). *Fundamentals of Industrial Hygiene*. 5th ed. Chicago, IL: NSC; 2002.
2. Clinical Laboratory Standards Institute (CLSI). *Clinical Laboratory Safety (Approved Guideline, GP17-A3)*. Wayne, PA: CLSI; 2012.
3. Furr AK. *Handbook of Laboratory Safety*. 5th ed. Boca Raton, FL: CRC Press; 2000.
4. Occupational exposure to bloodborne pathogens; final rule. *Federal Register.* 1991;56:235.
5. Occupational Safety and Health Administration. Subpart Z 29CFR 1910.1000-1450.
6. Clinical Laboratory Standards Institute (CLSI). *Protection of Laboratory Workers from Occupationally Acquired Infection (Approved Guideline, M29-A4)*. Wayne, PA: CLSI; 2014.
7. Gile TJ. *Complete Guide to Laboratory Safety*. Marblehead, MA: HCPro, Inc.; 2004.
8. United Nations. *Globally Harmonized System of Classification and Labelling of Chemicals*. New York, NY; Geneva: United Nations; 2003.
9. Centers for Disease Control and Prevention, National Institutes of Health. *Biosafety in Microbiological and Biomedical Laboratories*. 5th ed. Washington, DC: U.S. Government Printing Office; 2009.
10. *N10SH Pocket Guide to Chemical Hazards*. Atlanta, GA: Center for Disease Control; 2007.
11. Clinical Laboratory Standards Institute (CLSI). *Clinical Laboratory Waste Management (Approved Guideline, GP05-A3)*. Wayne, PA: CLSI; 2011.

Quality Management in the Clinical Laboratory

Yun L. Trull, Michael W. Rogers, Cyndy B. Levtzow, Matthew P. A. Henderson, Christopher R. McCudden, and Monte S. Willis

CHAPTER OUTLINE

KEY TERMS

Accuracy
Analytic sensitivity
Analytic specificity
Analytical measurement range (AMR)
Bias
Clinically reportable range (CRR)
Coefficient of determination (r^2)
Coefficient of variation (CV)
Constant error
Correlation coefficient (r)
Descriptive statistics
Diagnostic sensitivity
Diagnostic specificity
Dispersion
False negative
False positive
Histograms
Lean Six Sigma
Levey-Jennings control chart
Limit of detection (LOD)
Limit of quantification (LOQ)
Linear regression
Mean
Median
Medical decision level
Mode
Negative predictive value
Positive predictive value
Precision
Probability
Proficiency testing (PT)
Proportional error
Quality assurance
Quality control
Random error
Reference range
Sensitivity
Shift

Slope
Specificity
Standard deviation (SD)
Standard deviation index (SDI)
Systematic error
Therapeutic range
Total allowable error
Trend
True negative
True positive
Y-intercept

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define the following terms: quality control, quality assurance, accuracy, precision, descriptive statistics, reference range, random error, analytical sensitivity, analytical specificity, systematic error, and confidence intervals.
- Calculate the following: mean, median, range, variance, and standard deviation.
- Understand why statistics are needed for effective quality management.
- Evaluate a descriptive statistics equation with confidence.
- Understand the types, uses, and requirements for reference ranges.
- Explain the requirement(s) to verify or establish a reference range.
- Evaluate laboratory data using multi-rules for quality control.
- Graph laboratory data and determine significant constant or proportional errors.
- Identify trends and/or shifts given laboratory data.
- List the requirements when performing a method evaluation.
- Interpret a set of laboratory method evaluation data.
- Discuss proficiency testing programs in the clinical laboratory.
- Describe the process of quality improvement.

It is widely accepted that the majority of medical decisions are supported using laboratory data. Therefore, the laboratory results generated must be accurate and provide the appropriate diagnostic information to assist the health care provider in treating patients. Determining and maintaining **accuracy** and **precision** requires considerable effort and cost, entailing the use of a series of approaches depending on the complexity of the test. To begin, one must appreciate what *quality* is and how quality is measured and managed. To this end, it is vital to understand basic statistical concepts that enable the laboratorian to measure quality. The laboratory must ensure accurate and reliable test results, which is achieved by a process known as **quality assurance (QA)** and **quality control (QC)**. Quality assurance is the process that guarantees that the laboratory results are as accurate as possible and includes all three phases of testing: pre-analytic, analytic, and post-analytic. Quality control is the specific process of monitoring and assessing the analytical phase of testing. Before implementing a new test, it is important to determine if the test can perform acceptably by meeting or exceeding defined quality criteria; *method evaluation* is used to verify the acceptability of new methods prior to test implementation and reporting patient results. This chapter describes basic statistical concepts and provides an overview

CASE STUDY 3.1

Sample Labeling

Remember Mía from Chapter 1? Today, Mía is working on the urinalysis bench. She receives a urine specimen with a request for a complete urinalysis. The cup is labeled properly, so Mía begins the testing. She finishes the testing, reviews her results, and reports the results to the nurse on Floor 3 East. Several minutes later, Mía receives a telephone call from the Floor 3 East charge nurse, informing her that the urine was reported on the wrong patient. She is told that the cup was labeled incorrectly before it was brought to the laboratory.

1. What is the problem in this case, and where did it occur?
2. Would Mía's laboratory's QC system be able to detect or prevent this type of error?

of the procedures necessary to implement a new method and ensure its continued accuracy.

Basic Concepts

Each day, high-volume clinical laboratories generate thousands of results. This wealth of clinical laboratory data must be summarized and critically reviewed to monitor test performance. The foundation for monitoring performance (known as QC) is **descriptive statistics**, or numerical values that summarize a given data set.

Descriptive Statistics: Measures of Center, Spread, and Shape

When examined closely, a collection of seemingly similar things always has at least slight differences for any given characteristic (e.g., smoothness, size, color, weight, volume, and potency). Similarly, laboratory data will have at least slight measurement differences. For example, if glucose on a given specimen is measured 100 times in a row, there would be a range of values obtained. Such differences in laboratory values can result from a variety of factors. Although measurements will differ, their values form patterns that can be visualized and analyzed collectively. Laboratorians view and describe these patterns using graphical representations and descriptive statistics (**Figure 3.1**).

When comparing and analyzing collections or sets of laboratory data, patterns can be described by their center, spread, and shape. Although comparing the center of data is most common, comparing the spread can be even more powerful. Assessment of data **dispersion**, or spread, allows laboratorians to assess the predictability (and the lack of) in a laboratory test or measurement.

Measures of Center

The three most frequently used descriptions of the center of a data set (Figure 3.1) are the **mean**, the **median**, and the **mode**. The *mean* is most commonly used and often called the *average*. The *median* is the "middle" point and is often used with skewed data, so its calculation is not significantly affected by outliers. The *mode* is rarely used as a measure of the data's center but is more often used to describe data that seem to have two centers (i.e., bimodal). The mean is calculated by summing the observations and dividing by the number of the critically evaluated observations (**Eq. 3.1**).

The mean equation is as follows:

$\sum$ Add up . . .

$\sum_{i=1}^{n} x_i$ the total number (n) of the data points (x) and . . . (**Eq. 3.1**)

$\sum_{i=1}^{n} x_i/n$ divide by the total number (n) of data points

$\overline{x}\sum_{i=1}^{n} x_i/n$ to find the mean ($\overline{x}$ or "$x$bar")

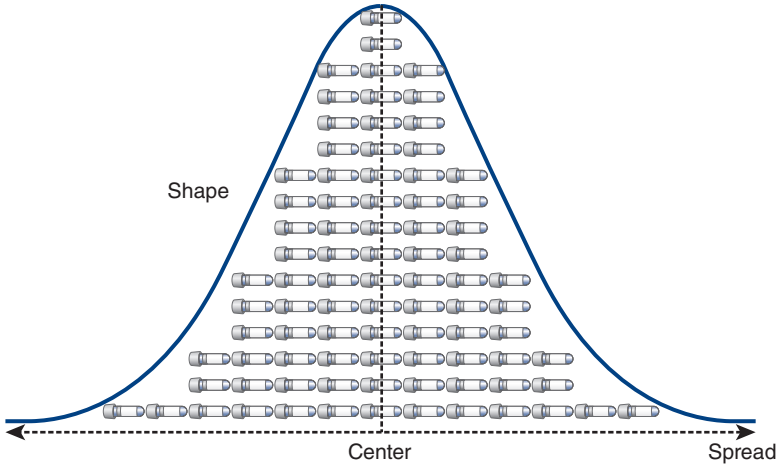


Figure 3.1 Basic measures of data include the center, spread, and shape. The center can be defined by the mean (x-bar character), median, or mode. Spread is defined by the standard deviation and coefficient of variation. Shape is defined by how the distribution of data relates the center. This is an example of data that have "normal" or Gaussian distribution.

The summation sign, Σ, is an abbreviation for $(x_1 + x_2 + x_3 + \cdots + x_n)$ and is used in many statistical formulas. Often, the mean of a specific data set is called $\bar{x}$, or "*x* bar."

The median is the middle of the data after the data have been *rank ordered*. The median is the value that divides the data in half. To determine the median, values are rank ordered from least to greatest and the middle value is selected. For example, given a sample of 5, 4, 6, 5, 3, 7, 5, the rank order of the points is 3, 4, 5, 5, 5, 6, 7. Because there are an odd number of values in the sample, the middle value (median) is 5; the value 5 divides the data in half. Given another sample with an even number of values 5, 4, 6, 8, 9, 7, the rank order of the points is 4, 5, 6, 7, 8, 9. The two "middle" values are 6 and 7. Adding them yields 13 (6 + 7 = 13); division by 2 provides the median (13/2 = 6.5). The median value of 6.5 divides the data in half.

The mode is the most frequently occurring value in a data set. Although it is seldom used to describe data, it is referred to when in reference to the shape of data—a bimodal distribution, for example. In the sample 3, 4, 5, 5, 5, 6, 7, the value that occurs most often is 5. The mode of this set is then 5. The data set 3, 4, 5, 5, 5, 6, 7, 8, 9, 9, 9 has two modes, 5 and 9.

After describing the center of the data set, it is very useful to indicate how the data are distributed (spread). The spread represents the relationship of all the data points to the mean (Figure 3.1). There are four commonly used descriptions of spread: (1) range, (2) standard deviation (SD), (3) coefficient of variation (CV), and (4) standard deviation index (SDI). The easiest measure of spread to understand is the range. The range is simply the largest value in the data minus the smallest value, which represents the extremes of data one might encounter. **Standard deviation** (also called "s," SD, or σ) is the most frequently used measure of variation. SD describes the distribution of all data points around the mean. Although calculating SD can seem somewhat intimidating, the concept is straightforward; in fact, all of the descriptive statistics and even the inferential statistics have a combination of mathematical operations that are by themselves no more complex than a square root. The SD and, more specifically, the variance represent the "average" distance from the center of the data (the mean) and every value in the data set. The **coefficient of variation (CV)** allows a laboratorian to compare SDs with different units and reflects the SDs in percentages. The **standard deviation index (SDI)** refers to the difference between the measured value and the mean expressed as a number of SDs. An SDI = 0 indicates the value is accurate or in 100% agreement, where as an SDI = 3 is 3 SDs away from the target (mean) and indicates inaccuracy. The SDI may be positive or negative. Similar to CV, it is a way to reflect ranges in a relative manner regardless of how low or high the values are.

Range is one description of the spread of data. It is simply the difference between the highest and lowest data points: range = high − low. For the sample 5, 4, 6, 5, 3, 7, 5, the range is 7 − 3 = 4. The range is often a good measure of dispersion for small samples of data. It does have a serious drawback: the range is susceptible to extreme values or outliers (i.e., incorrect quality control material or wrong data entry).

To calculate the SD of a data set, it is easiest to determine the variance (s^2). Variance is similar to the mean in that it is an average. Variance is the average of the squared distances of all values from the mean:

$$s^2 = \sum_{i=1}^{n} (x_i - \bar{x}) / n - 1 \qquad \textbf{(Eq. 3.2)}$$

As a measure of dispersion, variance represents the difference between each value and the average of the data. Given the values 5, 4, 6, 5, 3, 7, 5, variance can be calculated as shown below:

$$\begin{aligned}
&\bar{x} && (5 + 4 + 6 + 5 + 3 + 7 + 5) / 7 = 5 \\
&(x_i - \bar{x}) && (5 - 5) + (4 - 5) + (6 - 5) \\
& && + (5 - 5) + (3 - 5) \\
& && + (7 - 5) + (5 - 5) \\
&(x_i - \bar{x})^2 && (0)^2 + (-1)^2 + (1)^2 + (0)^2 \\
& && + (-2)^2 + (2)^2 + (0)^2 \\
&\sum_{i=1}^{n} (x_i - \bar{x})^2 && 0 + 1 + 1 + 0 + 4 \\
& && + 4 + 0 = 10 \\
&\sum_{i=1}^{n} (x_i - \bar{x})^2 / n - 1 && 10 / (7 - 1) = 10 / 6 \\
&s^2 \sum_{i=1}^{n} (x_i - \bar{x})^2 / n - 1 && 1.67
\end{aligned} \qquad \textbf{(Eq. 3.3)}$$

To calculate the SD (or "*s*"), simply take the square root of the variance:

$$s(\sigma) = \sqrt{s^2} = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 / n - 1} \qquad \textbf{(Eq. 3.4)}$$

Although it is important to understand how these measures are calculated, most analyzers or instruments, laboratory information systems, and laboratory

statistical software packages determine these automatically.

Another way of expressing SD is in terms of the CV. The CV is calculated by dividing the SD by the mean and multiplying by 100 to express it as a percentage:

$$CV(\%) = \frac{100s}{\bar{x}} \qquad \textbf{(Eq. 3.5)}$$

The CV simplifies the comparison of SDs of test results expressed in different units and concentrations. As shown in **Table 3.1**, analytes measured at different concentrations can have a drastically different SD but a comparable CV. The CV is used extensively to summarize QC data. The CV of highly precise analyzers can be lower than 1%. For values at the low end of the **analytical measurement range (AMR)**, or the range of analyte concentrations that can be directly measured without dilution, concentration, or another pretreatment, the acceptable CV can range as high as 50% as allowed by the Clinical Laboratory Improvement Amendments, also known as CLIA.

Another way of expressing SD is in terms of the SDI, where the SDI is calculated by dividing the difference of the measured laboratory value or mean minus the target or group mean, and dividing by the assigned or group SD. The SDI can be a negative or positive value.

$$SDI = \frac{(\text{laboratory mean} - \text{group mean})}{\text{GroupSD}} \qquad \textbf{(Eq. 3.6)}$$

Measures of Shape

Although data sets can exhibit hundreds of different "shapes"—distributions—the most commonly discussed is the Gaussian distribution (also called normal distribution; Figure 3.1). The Gaussian distribution describes many continuous laboratory variables and shares several unique characteristics: the mean, median, and mode are identical; the distribution is symmetric—meaning half the values fall to the left of the mean and the other half fall to the right, with the peak of the curve representing the average of the data. This symmetrical shape is often called a "bell curve."

The total area under the Gaussian curve is 1.0, or 100%. Much of the area—68.3%—under the "normal" curve is between ±1 SD ($\mu \pm 1\sigma$) (**Figure 3.2A**). Most of the area—95.4%—under the "normal" curve is between ±2 SDs ($\mu \pm 2\sigma$; **Figure 3.2B**). And almost all of the area— 99.7%—under the "normal" curve is between ±3 SDs ($\mu \pm 3\sigma$) (**Figure 3.2C**). (Note that μ represents the average of the total population, whereas the mean of a specific data set is $\bar{x}$, or "*x* bar.")

The "68–95–99 Rule" summarizes the above relationships between the area under a Gaussian distribution and the SD. In other words, given any Gaussian distributed data, ≈68% of the data fall between ±1 SD from the mean; ≈95% of the data fall between ±2 SDs from the mean; and ≈99% fall between ±3 SDs from the mean. Likewise, if you selected a value in a data set that is Gaussian distributed, there is a 0.68 chance of it lying between ±1 SD from the mean; there is a 0.95 likelihood of it lying between ±2 SDs; and there is a 0.99 probability of it lying between ±3 SDs. (Note: the terms "chance," "likelihood," and "probability" are synonymous in this example.) **Probability** is the extent to which an event occurs.

As will be discussed in the **reference range** section, most patient data are not normally distributed. These data may be skewed or exhibit multiple centers (bimodal, trimodal, etc.), as shown in **Figure 3.3**. Plotting data in **histograms** (graph of bars of data) as shown in the figure is a useful and

Table 3.1 Comparison of SD and CV for Two Different Analytes

FSH Concentration	SD	CV	βhCG Concentration	SD	CV
1	0.09	9.0	10	0.8	8.0
5	0.25	5.0	100	5.5	5.5
10	0.40	4.0	1000	52.0	5.2
25	1.20	4.8	10,000	500.00	5.0
100	3.80	3.8	100,000	4,897.0	4.9

SD, standard deviation; CV, coefficient of variation; FSH, follicle-stimulating hormone; βhCG, β-human chorionic gonadotropin.

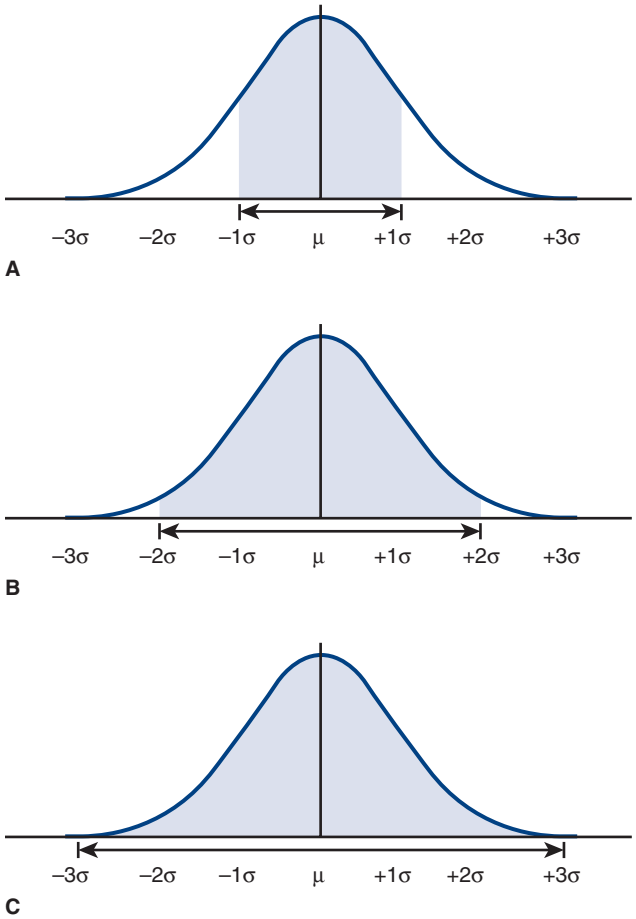


Figure 3.2 A normal distribution contains **(A)** ≈68% of the results within ±1 SD (1s or 1σ), **(B)** 95% of the results within ±2s (2σ), and **(C)** ≈99% of the results within ±3σ.

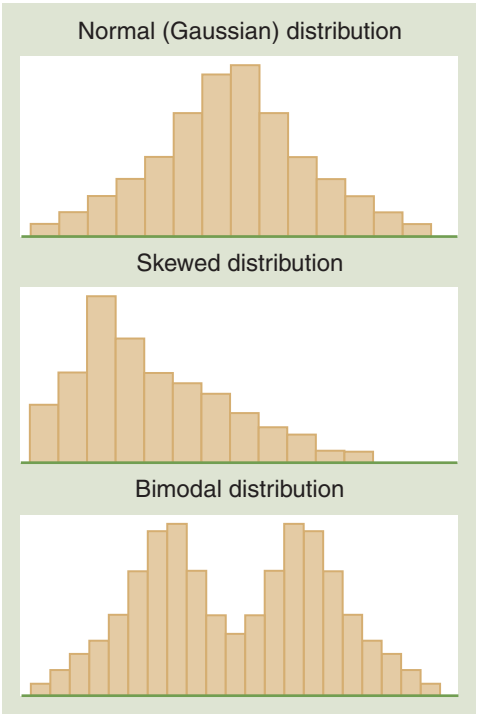


Figure 3.3 Examples of normal (Gaussian), skewed, and bimodal distributions. The type of statistical analysis that is performed to analyze the data depends on the distribution (shape).

easy way to visualize distribution. However, there are also mathematical analyses (e.g., normality tests) that can confirm if data fit a given distribution. The importance of recognizing whether data are normally distributed, skewed, or multimodal is related to the ways that data with different distributions can be statistically analyzed.

Measurement of Precision

Precision estimates the random error associated with the test method and detects any problems affecting its reproducibility. There are two types of precision: within-run (repeatability) and within-laboratory, also referred to as across-run or between-run precision. Within-run precision is defined as the closeness of agreement between results of successive measurements obtained under identical conditions.[1] In within-run precision studies, the same laboratorian performs 20 replicate analyses in a single run on a single day. Similarly, within-laboratory precision is estimated by measuring a sample 20 times over multiple days. It is recommended that this study be performed over a 10- to 20-day period, incorporating one or two analytic runs (runs with patient samples or QC materials) per day, preferred is an AM run and a PM run.[1,2] A common precision study is a 2 × 2 × 10 study (40 total data points), where two controls are run twice a day (AM and PM) for 10 days. The rationale for evaluating precision over several days is logical. Running multiple samples on the same day does a good job of estimating within-run precision within a single day (simple precision). Still, it underestimates long-term changes and testing variables that occur over time. A better estimation of the random error over time is given by running multiple samples on different days. It is essential that more than one concentration be tested in these studies, with materials ideally spanning the clinically appropriate and analytical measurement range of concentrations. For a new glucose method, this might include samples in the hyperglycemic range (i.e., ≥150 mg/dL) and the hypoglycemic range (i.e., ≤70 mg/dL). After these data are collected, the mean, SD, CV, and SDI are calculated.

Precision studies are also required to verify a manufacturer's precision claims for a given test. In this

Box 3.1 Sources of Analytic Variability

Operator technique	Environmental conditions (e.g., temperature, humidity)
Instrument differences	Reagents
Test accessories	Power surges
Contamination	Matrix effects (hemolysis, lipemia, serum proteins)

circumstance, it is recommended to test at least two levels in replicates over five days.

The SD and the CV indicate the random error or imprecision associated with the test procedure. The SD of the controls shows the within-run imprecision analyzed within one run. The total imprecision may be obtained from the SD of control data with the one or two data points accumulated per day over a period. The total imprecision is the most accurate assessment of performance that would affect the values a provider might see and reflects differences in operators, pipettes, and variations in environmental changes such as temperature, reagent stability, etc (**Box 3.1**). In practice, however, within-run imprecision is used more commonly than total imprecision.

Acceptable Performance Criteria: Precision Studies

A laboratory recorded QC data during its first week of operation. For calcium, one level of the data (mean = 8.5 mg/dL) is shown in **Table 3.2A**. The laboratory serves outpatient services and operates Monday through Saturday, 6 AM to 6 PM. QC samples were run twice daily, in the morning and in the afternoon, to bracket daily patient testing. During the initial evaluation of the calcium assay, a simple within-run imprecision study was performed with a replicate of 10. The within-run imprecision study data are shown in **Table 3.2B**. Total imprecision of the week is calculated with the 12 data points, as 0.08 mg/dL; and within-run imprecision is estimated with the 10 repeats as 0.05 mg/dL. Between-run imprecision is calculated based on the equation below as 0.06 mg/dL.

$$S^2_{Total} = S^2_{Within\text{-}run} + S^2_{Between\text{-}run} \qquad \textbf{(Eq. 3.7)}$$

The acceptability of analytic error is based on how the test is used to make clinical interpretations.[3,4] The determination of whether total precision is adequate is based on the total imprecision being less than a pre-determined fraction of the **total allowable error, (TE_a)**. For imprecision, one-third of total error is a rule of thumb; some laboratories may choose one-fourth of the total error for analytes that are very precise and accurate. It is not recommended to use all of the allowable error for imprecision (random error) as it leaves no room for systematic error (bias or inaccuracy). In this case, one-fourth of the total error can be used. Total allowable error for calcium being 1.0 mg/dL, as shown in **Table 3.3**, error budget of imprecision is 0.25 mg/dL (0.25 × 1.0 mg/dL). Therefore, the weekly total imprecision of 0.08 mg/dL is way below the error budget of 0.25 mg/dL, and acceptable. But a closer look at the imprecision data can reveal potential problems. The relatively large within-run imprecision (0.05 mg/dL) compared to the total imprecision (0.08 mg/dL) may indicate a problem in instrument pipetting. If PM QC data are higher than AM QC data, which are influenced by periodic daily changes such as laboratory

Table 3.2 Laboratory Calcium QC Data (Unit: mg/dL). (A) QC samples during a week of operation. (B) Within-run imprecision of initial evaluation.

(a)	Mon	Tues	Wed	Thurs	Fri	Sat				
AM	8.5	8.4	8.5	8.5	8.6	8.6				
PM	8.5	8.5	8.6	8.6	8.6	8.7				
(b)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
	8.5	8.5	8.5	8.4	8.5	8.4	8.4	8.5	8.5	8.5

Table 3.3 Performance Standards for Common Clinical Chemistry Analytes as Defined by CLIA

Calcium, total	Target ±1.0 mg/dL
Chloride	Target ±5%
Cholesterol, total	Target ±10%
Cholesterol, HDL	Target ±20%
Glucose	Target ±8%
Potassium	Target ±0.3 mmol/L
Sodium	Target ±4 mmol/L
Total protein	Target ±8%
Triglycerides	Target ±15%
Urea nitrogen	Greater of target ±2 mg/dL or ±9%
Uric acid	Target ±10%

Adapted from Centers for Medicare and Medicaid Services (CMS), Health and Human Services; Centers for Disease Control and Prevention (CDC), Health and Human Services. Clinical Laboratory Improvement Amendments of 1988 (CLIA) Proficiency Testing Regulations Related to Analytes and Acceptable Performance. *Proposed Rules. Fed Regist.* 2019;84:1536–1567.

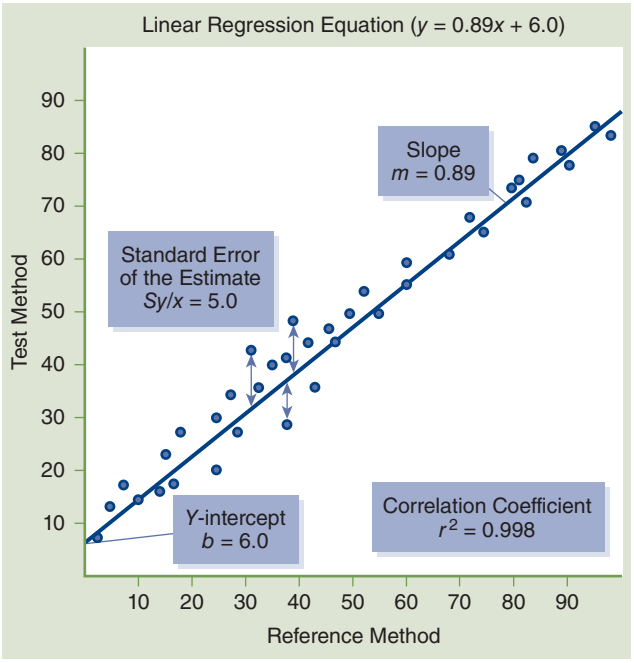


Figure 3.4 A generic example of a linear regression. A linear regression compares two tests and yields important information about systematic and random errors. Systematic error is indicated by changes in the *y*-intercept (constant error) and the slope (proportional error). Random error is indicated by the standard error of the estimate ($S_{y/x}$); $S_{y/x}$ basically represents the distance of each point from the regression line. The correlation coefficient indicates the strength of the relationship between the tests.

temperature and humidity. This lab may need better environmental control. QC data may trend higher toward the end of the week, which may be caused by deteriorating reagents or instruments drifting from calibration set points.

Descriptive Statistics of Groups of Paired Observations

While the use of basic descriptive statistics is satisfactory for examining a single method, laboratorians frequently need to compare two different methods. This is most commonly encountered in comparison of methods experiments. A comparison of methods experiment involves measuring patient specimens by both an existing (reference) method and a new (test) method (described in the Reference Range and Method Evaluation sections later). The data obtained from these comparisons consist of two measurements for each patient specimen. It is easiest to visualize and summarize the paired-method comparison data graphically (**Figure 3.4**). By convention, the values obtained by the reference method are plotted on the *x*-axis, and the values obtained by the test method are plotted on the *y*-axis.

In Figure 3.4, the agreement between the two methods is estimated from the straight line that best fits the points. Whereas visual estimation may be used to draw the line, a statistical technique known as **linear regression** analysis provides an objective measure of the line of best fit for the data. Three factors are generated in a linear regression: the **slope**, the ***y*-intercept**, and the **correlation coefficient (*r*)**. In Figure 3.4, there is a linear relationship between the two methods over the entire range of values. The linear regression is defined by the equation $y = mx + b$. The slope of the line is described by m, and the value of the *y*-intercept (b) is determined by plugging $x = 0$ into the equation and solving for y. The correlation coefficient (r) is a measure of the strength of the relationship between the two methods. The correlation coefficient can have values from −1 to 1, with the sign indicating the direction of relationship between the two variables. A positive r indicates that both variables increase and decrease together, whereas a negative r indicates that as one variable increases, the other decreases. An r value of 0 indicates no relationship, whereas $r = 1.0$ indicates a perfect relationship. Although many equate high positive values of r (0.95 or higher) with excellent agreement between the test and

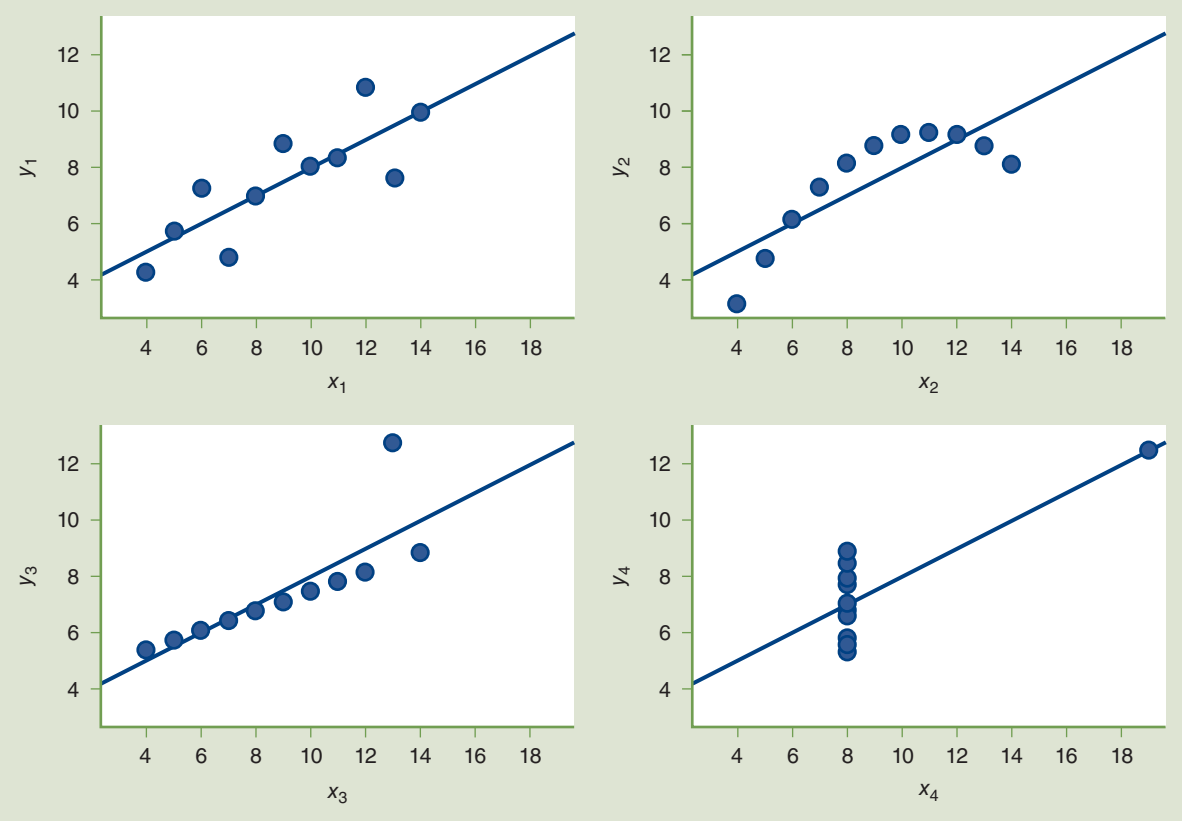


Figure 3.5 Anscombe's quartet demonstrates the need to visually inspect data. In each panel, $y = 0.5\,x + 3$, $r_2 = 0.816$, $S_{y/x} = 4.1$.

comparative methods, most clinical chemistry comparisons should have correlation coefficients greater than 0.98. When r is less than 0.99, the regression formula can underestimate the slope and overestimate the y-intercept. The absolute value of the correlation coefficient can be increased by widening the concentration range of samples being compared. However, if the correlation coefficient remains less than 0.99, then alternate regression statistics or modified comparison value sets should be used to derive more realistic estimates of the regression, slope, and y-intercept, which are beyond the scope of this chapter.[5,6] Visual inspection of data is essential prior to drawing conclusions from the summary statistics as demonstrated by the famous Anscombe quartet (**Figure 3.5**). In this data set, the slope, y-intercept, and correlation coefficients are all identical, yet visual inspection reveals that the underlying data are completely different.

An alternate approach to visualizing paired data is the difference plot, which is also known as the Bland-Altman plot (**Figure 3.6**). A difference plot indicates either the percent or absolute **bias** (or the difference) between the reference and test method values over the average range of values. This approach permits simple comparison of the differences to previously established maximum limits. As is evident in Figure 3.6, it is easier to visualize any concentration-dependent differences than by linear regression analysis. In this example, the percent difference is clearly greatest at lower concentrations, which may not be obvious from a regression plot.

The difference between test and reference method results is called *error*. There are two kinds of error measured in method comparison experiments: random and systematic. **Random error** is present in all measurements and can be either positive or negative; typically, a combination of both positive and negative errors occur on both sides of the assigned target value. Random error can be a result of many factors including instrument, operator, reagent, and environmental variation. Random error is calculated as the SD of the points about the regression line ($S_{y/x}$). $S_{y/x}$ essentially refers to average distance of the data from the regression line (Figure 3.4). The higher the $S_{y/x}$, the wider is the scatter and the higher is the amount of random error. In Figure 3.4, the $S_{y/x}$

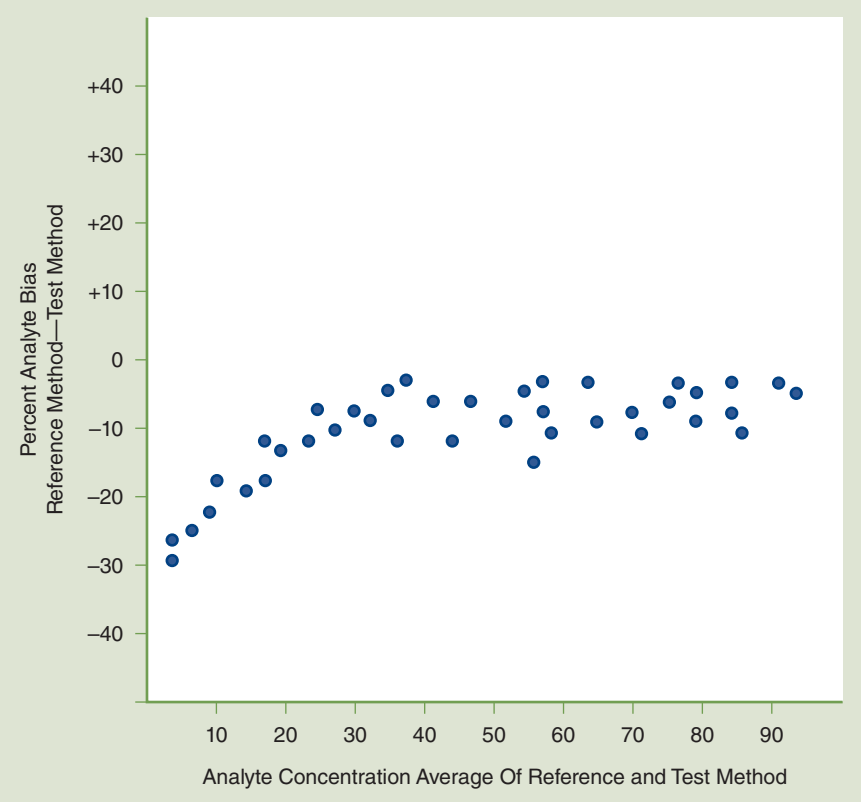


Figure 3.6 An example of a difference (Bland-Altman) plot. Difference plots are a useful tool to visualize concentration-dependent error.

is 5.0. If the points were perfectly in line with the linear regression, the $S_{y/x}$ would equal 0.0, indicating no random error is present. $S_{y/x}$ is also known as the standard error of the estimate S_E (**Box 3.2**).

Systematic error influences observations consistently in one direction (higher or lower). The measures of slope and *y*-intercept provide estimates of the systematic error. Systematic error can be further broken down into **constant error** and **proportional error**. Constant systematic error exists when there is a continual difference between the test method and the comparative method values, regardless of the concentration. In Figure 3.4, there is a constant difference of 6.0 between the test method values and the comparative method values. This constant difference, reflected in the *y*-intercept, is called *constant systematic error*. Proportional error exists when the differences between the test method and the comparative method values are proportional to the analyte concentration. Proportional error is present when the slope is not 1. In the example, the slope of 0.89 represents the proportional error, where samples will be underestimated in a concentration-dependent fashion by the test method compared with the reference method; the error is proportional because it will increase with the analyte concentration.

Box 3.2 Types of Error in Laboratory Testing: A Preview of Things to Come

1. Systematic Error
   - Constant
   - Proportional
2. Random Error

Measurement of Bias

Now let's looks at bias, also known as *accuracy*. Accuracy is the difference between a measured value and its actual value. In order to determine accuracy, a method comparison experiment is usually performed. This examines patient samples by the method being evaluated (test method) with a reference method. It is used primarily to estimate the systematic error in actual patient samples, and it may offer the type of systematic error present (proportional vs. constant).

Ideally, the test method is compared with a standardized reference method (gold standard), a method with acceptable accuracy compared to its imprecision. Many times, reference methods are laborious and time-consuming, as is the case with the isotope dilution mass spectrometry methods used to determine cholesterol. Because most laboratories are understaffed or do not have the technical expertise required to perform reference methods, most test methods are compared with those routinely used. These routine tests have their particular inaccuracies, so it is important to determine what inaccuracies they might have that are documented in the literature. If the new test method replaces the routine method, differences between the two should be well characterized. Also, extensively documented communications with all providers and staff must be performed before implementation.

To compare a test method with a comparative method, it is recommended by Westgard et al.[2] and CLIA[7] that 40 to 100 specimens be run by each method on the same day (preferably within 4 hours), over at least 5 days, with specimens spanning the clinical range and representing a diversity of pathologic conditions. Samples should cover the full analytical measurement range, and it is recommended that 25% be in the lower range, 50% should be in the middle range, and 25% should be in the higher range. Additional samples at the medical decision levels should be a priority during the comparison of methods. As an extra measure of QC, specimens can be analyzed in duplicate. Otherwise, experimental results must be critically reviewed by the laboratory director and evaluation staff, comparing test and comparative method results immediately after analysis. Samples with significant differences should be repeated to rule out technical errors as the source of variation. Daily analysis of two to five patient specimens should be followed for at least 8 days if 20 specimens are compared and for 20 days if 100 specimens are compared in replication studies.[7] A plot of the test method data (y-axis) versus the comparative method (x-axis) helps to visualize the data generated in a comparison-of-methods test (**Figure 3.7A**).[8] As described earlier, if the two methods correlate perfectly, the data pairs plotted as concentration values from the reference method (x) versus the evaluation method (y) will produce a straight line ($y = mx + b$), with a slope of 1.0, a y-intercept of 0, and a correlation coefficient (r) of 1. Data should be plotted daily and inspected for outliers so that original samples can be reanalyzed as needed. Analytical measurement range can also be confirmed in a similar manner quantitatively, with target values plotting on the x-axis and instrument readings on the y-axis.

Statistical Analysis of Bias Studies

While a visual inspection of method comparisons is essential, statistical analysis can be used to make objective decisions on a method's performance. The first and most fundamental statistical analysis for bias studies is the linear regression. Linear regression analysis yields the slope (b), the y-intercept (a), the SD of the points about the regression line ($S_{y/x}$), and the correlation coefficient (r); regression also yields the **coefficient of determination (r^2)**, The r^2 indicates the proportion of variation explained by one variable to predict another and ranges from 0 to 1. An example of these calculations can be found in Figure 3.7, where a comparison of β-human chorionic gonadotropin (βhCG) concentrations on an existing immunoassay system (Reference Method) and to a new system (Test Method). Statistics are calculated to determine a method's types and amounts of error, which is the basis for deciding if the test is still valid to make clinical decisions. Several types of errors can be seen looking at a plot of test method versus comparative method (**Figure 3.8**). When random errors occur (Figure 3.8A), points are randomly distributed around the regression line. Increases in the $S_{y/x}$ statistic reflect random error. Constant error (Figure 3.8B) is seen visually as a shift in the y-intercept. Proportional error (Figure 3.8C) is reflected in alterations in line slope.

Note that a statistical difference does not necessarily indicate that the method is not clinically usable, just that a difference is present. The size and nature (systematic and random error) of the differences determines if a method is clinical useable.[9] In bias studies, the constant systematic error can be determined by the y-intercept and the proportional systematic error can be determined by the slope. Random error can also be determined by the standard error of the estimate ($S_{y/x}$). Importantly, if a nonlinear relationship occurs between the test and comparative methods, linear regression analysis can be used only over the values in the linear range. To make accurate conclusions about the relationship between two tests, it is important to confirm that outliers are true outliers and not the result of technical errors.[10–12]

So far, we have described how we estimate error in test methods in terms of imprecision and inaccuracy. However, tests are performed to answer clinical questions, so to assess how this error might affect clinical judgments, it is assessed in terms of *allowable*

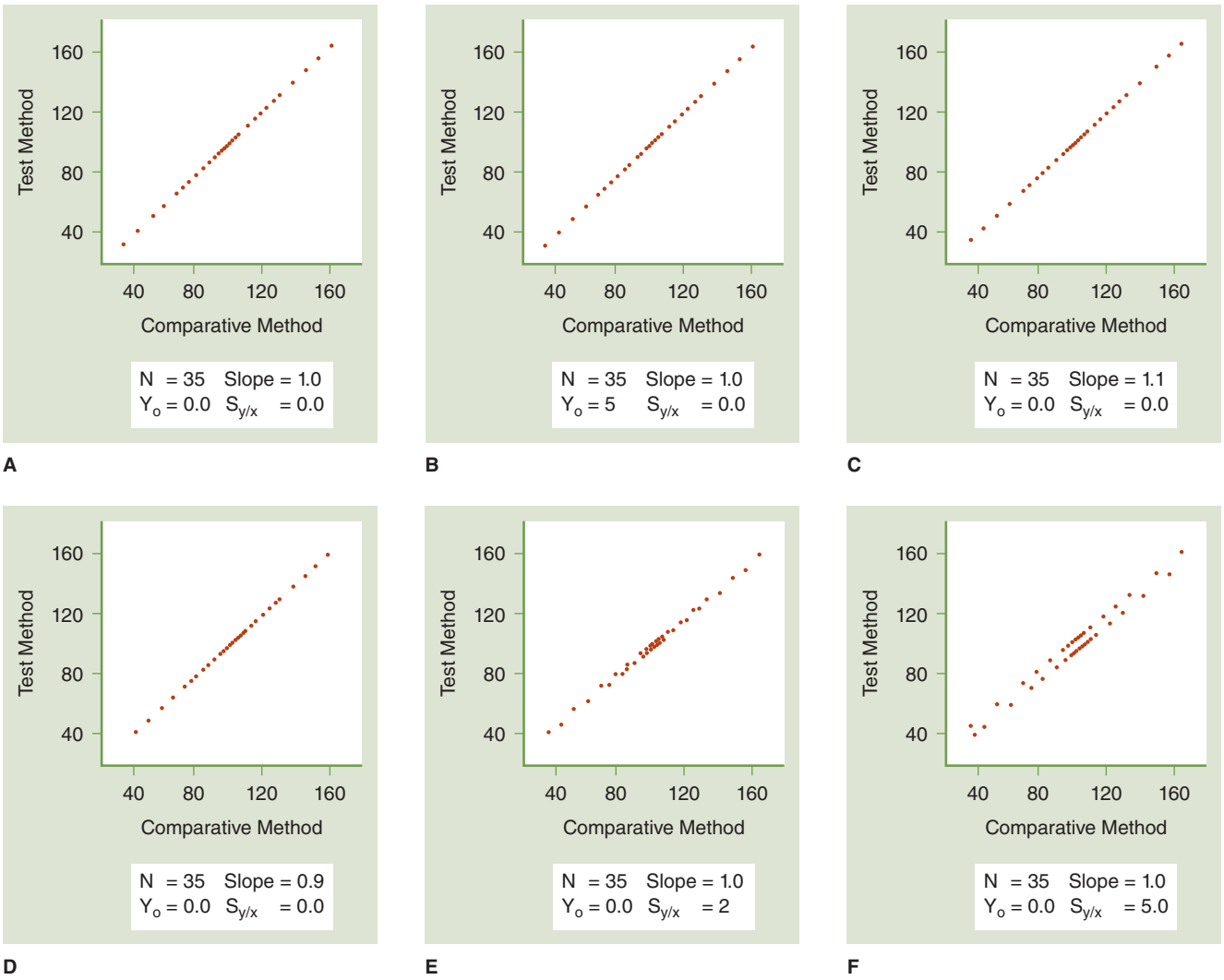


Figure 3.7 Comparison-of-methods experiment. **(A)** shows no error; **(B)** shows constant error; **(C)** and **(D)** show propopotional error; **(E)** and **(F)** show random error.

(analytical) error (E_a).[13] Allowable error is determined for each test based on the amount of error that will not negatively affect clinical decisions. If the combined random and systematic error (total allowable error) is less than E_a, then the performance of the test is considered acceptable. However, if the total error is larger than E_a, corrections (calibration, new reagents, hardware improvements, etc.) must be made to reduce the error or the method is rejected. This process ensures that the laboratory provides accurate, clinically usable information to healthcare providers to manage their patients effectively. To emphasize this point, laboratorians should consider that providers are rarely aware of the imprecision, bias, and performance of a given test that they rely on to make decisions, as this is an assumed understanding of laboratory analysis and quality. It is the responsibility of the laboratory to ensure quality.

Inferential Statistics

The next level of complexity beyond paired descriptive statistics is inferential statistics. Inferential statistics are used to draw conclusions (inferences) regarding the means or SDs of two sets of data. Inferential statistical analyses are most commonly encountered in research studies but also can be used in comparison-of-methods studies.

An important consideration for inferential statistics is the distribution of the data (shape). The distribution of the data determines what kind of inferential statistics can be used to analyze the data. Normally distributed (Gaussian) data are typically analyzed using what are known as "parametric" tests, which include F-test, *t*-test, or analysis of variance (ANOVA). F-test is a method to test whether there is significant difference between variances of two groups of data. Student's *t*-test is used to compare

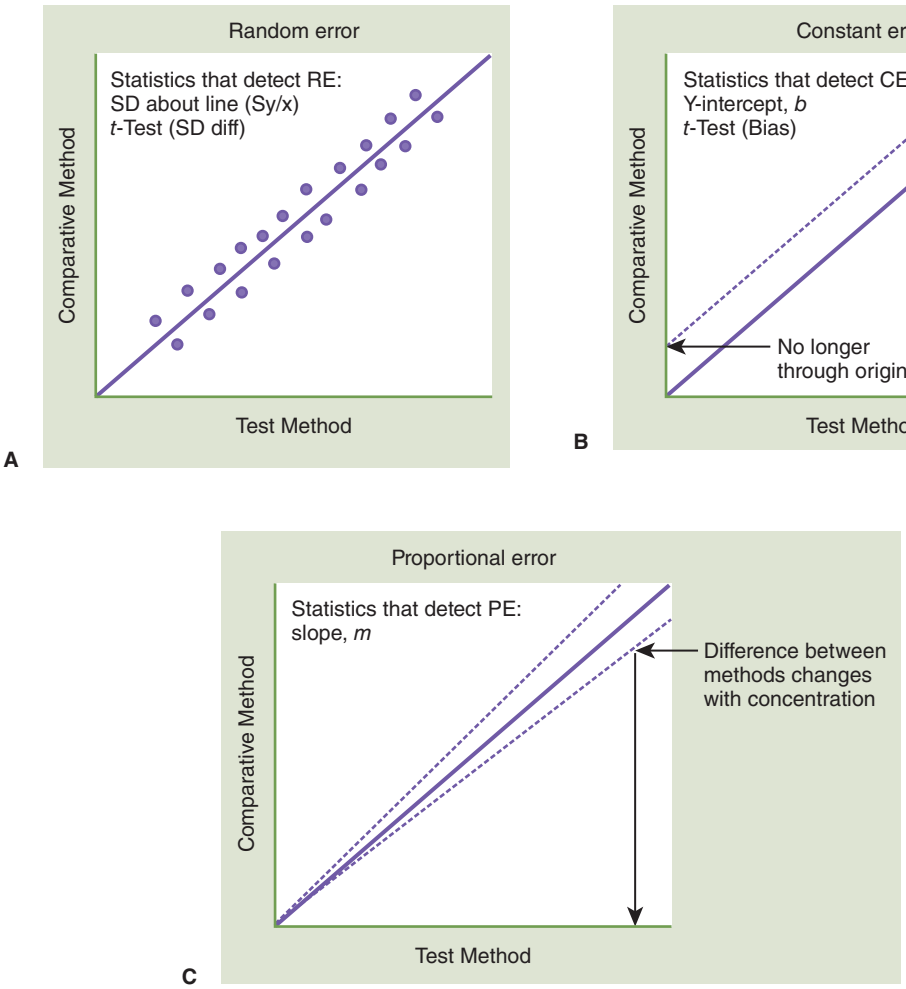


Figure 3.8 Examples of **(A)** random, **(B)** constant, and **(C)** proportional error using linear regression analysis.

Adapted from Westgard JO. *Basic Method Evaluation.* 4th ed. Madison, WI: Westgard Quality Corp.; 2020.

means of two groups of data, assuming the variances are not significantly different. ANOVA tests whether several different groups (three or more) all have the same mean. Data that are not normally distributed require a "nonparametric" analysis. Nonparametric test methods make no specific assumption about the distribution of data and rank the reference data in order of increasing size. Because the majority of analytes are not normally (Gaussian) distributed, nonparametric tests are the recommended analysis for most reference range studies. For reference range studies, nonparametric methods typically rely on ranking or ordering the values to determine upper and lower percentiles, which is a statistical term usually indicating a value below which percentage of data fall. While many software packages are capable of performing either parametric or nonparametric analyses, it is important for the user to understand that the type of data (shape) dictates which statistical test is appropriate for the analysis. An inappropriate analysis of sound data can yield the wrong conclusion and lead to erroneous provider decisions and patient care/impact consequences.

Quality Control and Quality Improvement

QC in the laboratory involves the systematic monitoring of analytic processes to detect analytic errors that occur during analysis and to ultimately prevent the reporting of incorrect patient test results. In the context of what we have discussed so far, QC is part of the performance monitoring that occurs after a test has been established. In general, monitoring of analytic methods is performed by assaying stable control materials and comparing their determined values with their expected values. The expected values are represented by intervals of acceptable values with upper and lower limits, known as *control limits*. When the expected values are within

the control limits, the operator can be reasonably assured that the analytic method is properly reporting values as approved during the method validation. However, when observed values fall outside the control limits, the operator must be aware of possible problems and the need for further investigation before reporting potentially erroneously patient results. The principles of visualizing QC data were initially applied to the clinical laboratory in the 1950s by Levey and Jennings.[14] Many important modifications have been made to these systems since that time, and they are discussed in general in this section.

Specimens analyzed for QC purposes are known as *QC materials,* commonly referred to as controls. These materials must be available in sufficient quantity to last for extended periods determined by QC material stability and manufacturer-determined expiration dates. QC materials should be the same matrix as the patient specimens. For example, a glucose assay performed on serum should have QC materials that are prepared in serum. Variation between vials should be minimal so that differences seen over time can be attributed to the analytic method itself and not variation in the QC material. Control material concentrations should span the clinically important range of the analyte at appropriate decision levels. For example, sodium QC materials might be tested at 130 and 150 mmol/L, representing cutoff values for hyponatremia and hypernatremia, respectively. QC for general chemistry assays generally uses two levels of control, while immunoassays commonly use three levels of control. Today, laboratories more often purchase manufactured control materials for QC instead of preparing the materials themselves. These materials are often lyophilized (dehydrated to powder) for stability and can be reconstituted in specific diluents or matrices representing serum/plasma, urine, whole blood, or cerebrospinal fluid (CSF). Both assayed or unassayed control materials can be purchased. Assayed controls give expected target ranges, often including the mean and SD using common analytic methods. Conversely, unassayed QC materials do not have assigned analyte values provided by the manufacturer. The laboratory, rather than the manufacturer, assigns expected results to unassayed QC material.

While assayed controls are more expensive because of the additional characterization, they allow another external check of method accuracy. Some laboratories will prepare internal quality control materials using human serum (i.e., serum pools) to address concerns about artificial matrixes and analyte values used in commercial materials.

Because some commercially prepared control materials are lyophilized and require reconstitution before use, the diluent should be carefully added and mixed. Incomplete mixing yields a partition of supernatant liquid and underlying sediment and will result in incorrect control values. Frequently, the reconstituted material will be more turbid (cloudy) than the actual patient specimen. Stabilized frozen controls do not require reconstitution but may behave differently from patient specimens in some analytic systems. It is important to carefully evaluate these stabilized controls with any new instrument system. Improper preparation and handling of QC materials is the main reason for QC failures in the laboratory and requires staff education and material management to minimize. For example, if one user mixes the QC material thoroughly after warming it for several hours, while another user runs it cold with minimal mixing, it may be expected that the QC results will vary while the process itself may be perfectly stable. More manufactured liquid QC is becoming available; however, it is usually more expensive, but it does remove any variation due to the preparation of the QC materials.

QC Charts

A common method to assess the determination of control materials over time is by the use of a **Levey-Jennings control chart** (**Figure 3.9**). Control charts graphically represent the observed values of a control material over time in the context of the upper and lower control limits in relation to the target value. When the observed value falls with the control limits, it can be interpreted that the method is performed adequately. Points falling outside the control limits suggest that problems may be developing. Control limits are expressed as the mean ± SD using formulas previously described in this chapter. Control charts can detect errors in accuracy and imprecision over time. Analytic errors that can occur can be separated into random and systematic errors. The underlying rationale for running repeated assays is to detect random errors that affect precision. Random errors may be caused by variations in technique. Systematic errors arise from factors that contribute to constant differences between measurements; these errors may be either positive or negative. Systematic errors may be due to several factors, including poorly made

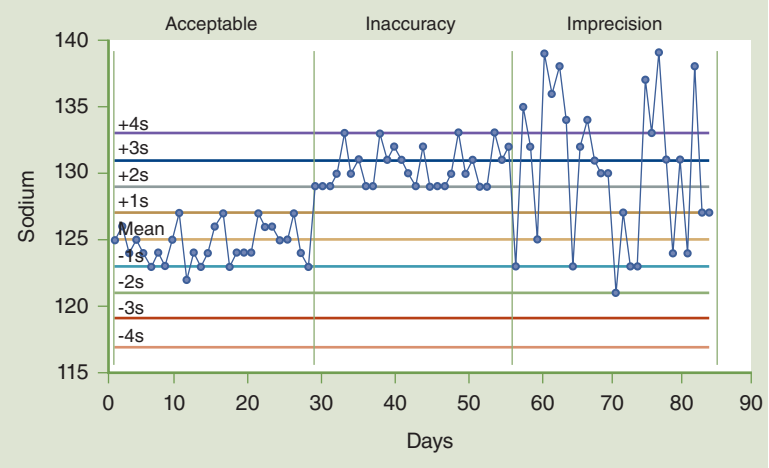


Figure 3.9 Levey-Jennings control chart. Results are plotted over time to identify acceptable quality control and failures due to inaccuracy and imprecision.

standards, reagents, instrumentation problems, poorly written procedures, or inadequate staff training.

Operation of a QC System

The QC system in the clinical laboratory is used to monitor analytic variations that can occur.

The QC program can be thought of as a three-stage process:

1. Establishing or verifying allowable statistical limits of variation for each analytic method
2. Using these limits as criteria for evaluating the QC data generated for each test
3. Taking action to remedy errors real time when indicated
   a. Finding the cause(s) of error
   b. Taking corrective action
   c. Reanalyzing control and patient data

Establishing Statistical Quality Control Limits

Prior to implementing a new lot number of QC material, both levels of QC material must be analyzed for 20 days. Exceptions include assays that are highly precise (CV < 1%), such as arterial blood gases, where 5 days is adequate. Also, if the QC material is a new lot number with similar ranges for an existing lot number to be implemented, a modified precision study can be done with as few as 10 values with monitoring post-implementation. Repeat analysis of the QC material allows for the calculation of the mean and SD. Initial estimates of the mean and SD control limits may be somewhat inaccurate because of the low number of data points. To produce more reliable data, estimates of the QC mean and SDs should be reviewed monthly and updated to include

CASE STUDY 3.2, PART 1

QC Program for POCT

Mía's laboratory oversees the QC program for the Point-of-Care Testing (POCT) glucometers in use at the hospital. Mía notices that the Floor 4 West staff is not following the proper procedure for running QC. She noticed the glucometer QC was re-run three times in a row in order to get results within the control limits. The first two runs were both 3 SD high. The last run did return results that were less than 2 SDs high.

3. Assist Mía in explaining the correct follow-up procedure for dealing with these out-of-control results to the Floor 4 West staff.

cumulative data. When changing to a new lot number of similar material, laboratorians use the newly obtained mean (or manufacturer mean and SD) as the target mean but retain the previous SD. As more data are obtained, all data should be averaged to derive the best estimates of the mean and SD.[15]

The distribution of error is assumed to be symmetrical and bell shaped (Gaussian) as shown in **Figure 3.10**. Examples of a **shift**, an abrupt change in the analytic process, and a **trend**, a gradual change in the analytic process, can be seen. Control limits are set to include most observed values (95% to 99.7%), corresponding to the mean ±2 or 3 SDs. Therefore, observation of values in the distribution tails should be rare (1/20 for 2 SDs; 3/1000 for 3 SDs). Observations outside the control limits suggest changes in the analytic methods. If the process is in control, no more than 0.3% of the points will be outside the 3 SDs (3s) limits. Analytic methods are considered in control if a symmetrical distribution of control values about the mean is seen, and few values outside the 2 SDs (2s) control limits are observed. Some laboratories define a method out of control if a control value exceeds the 2s limits. Other laboratories use the 2s limit as a warning limit and the 3s limit as an error limit. In this case, a control point between 2s and 3s would alert the laboratorian to a potential problem, while a point greater than 3s would require a corrective action. The selection of control rules and numbers should be related to the goals set by the laboratory.[16]

It is essential to recognize the distinction between QC limits and QC goals or specifications (**Figure 3.11**). *QC specifications* are the criteria used to decide that a given method meets the clinical requirements. QC specifications may be derived from the total allowable error, biological variation, or other medical decision criteria. *QC limits* are defined by the process itself based on its natural variation. A good description of the difference is that QC specifications are the "voice of the customer," while QC limits are the "voice of the process."

CASE STUDY 3.3

A Quality Control Decision

Mía's supervisor requested she evaluate the following glucose control data set. The laboratory is very busy, and Mía enlists your assistance performing the task below.

© Ariel Skelley/DigitalVision/Getty Images.

Day	Low	High
1	86	215
2	82	212
3	83	218
4	87	214
5	85	220
6	81	217
7	88	223

4. Calculate the mean and standard deviation for the above data set.

5. Plot these control data by day (one graph for each level, *x*-axis = day, *y*-axis = concentration). Indicate the mean and the upper and lower control limits (mean ±2 standard deviations) with horizontal lines.

6. You are working on the night shift at a community hospital and are the only person in the laboratory. You are running glucose quality control and obtain the following:

Low control value = 90 mg/dL; High control value = 230 mg/dL

a. Plot these controls on the process control chart (Levey-Jennings) you created above.
b. Are these values within the control limits?
c. What do you observe about these control data?
d. What might be a potential problem?
e. What is an appropriate next step?

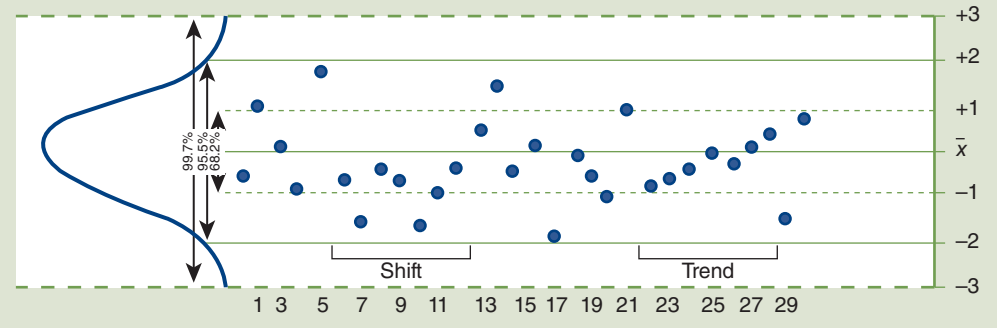


Figure 3.10 Control chart showing the relationship of control limits to the Gaussian distribution. Daily control values are graphed, and they show examples of a shift (an abrupt change in the analytic process) and a trend (a gradual change in the analytic process).

Thus, QC specifications are "the performance the healthcare provider needs," whereas QC limits are the "natural variation of a given test"; more simply, QC specifications are what you want and QC limits are what you have. A process tells the user its limits, and the user decides if those limits meet their specification.

From a practical perspective, QC limits define if a process is in control on a daily basis, and QC specifications define if the overall performance of the process meets the quality goals. Consider that a process may be out of control (e.g., have a 1_{3s} flag, as discussed in the following section) but still meet the QC specification limits. Conversely, a poorly performing method might have natural variation QC limits, which exceed the QC specification. QC limits must be within QC specifications or the lab will not be able to consistently report accurate results.

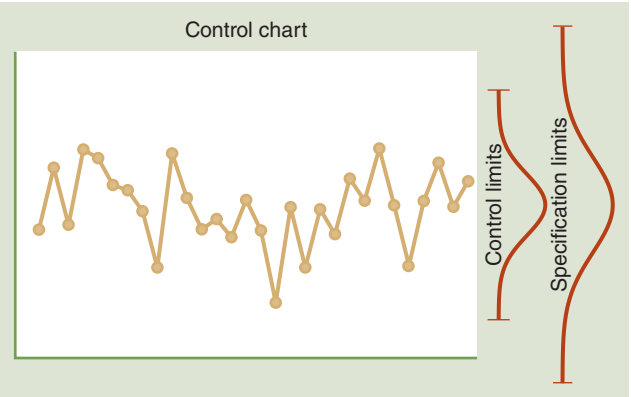


Figure 3.11 QC specification limits are the criteria used to decide that a given method meets the clinical requirements based on the healthcare provider's needs. QC control limits are defined by the process itself based on its natural variation.

Importantly, QC specifications do not have any effect on or causal relationship with QC limits; setting QC specifications to a CV of 0.1% will not make a method more precise any more than widening specifications to a CV of 100% would make the method become imprecise. However, the relationship between QC limits and QC specifications *can* be used to establish which multi-rules to use.

Multi-Rules Rule!

The use of the statistical process control chart (Levey-Jennings) was pioneered by Shewhart in the 1920s. *Multi-rules* were formalized by the Western Electric Company and later applied to the clinical laboratory by Westgard and Groth.[17] Multi-rules establish a criterion for judging whether an analytic process is out of control. To simplify the various control rules, abbreviations are used to refer to the various control rules (**Table 3.4**). Control rules indicate the number of control observations per analytic run, followed by the control amount in subscript.[18] For example, the 1_{3s} rule indicates that one data point cannot exceed 3 SDs (3s). If the 1_{3s} rule is not triggered, the analytic run will be accepted (i.e., results will be reported). If the QC results are more than 3 SDs (i.e., the 1_{3s} rule is violated), the run may be rejected and there will be additional investigation. The type of rule violated indicates what type of error exists. For example, a 1_{3s} rule violation may indicate a loss of precision or "random error" (Table 3.4).

Analogous to overlapping diseased and healthy patient results, it is important to consider that not every rule violation indicates that a process is out of control. The 1_{2s} rule, for example, will be outside the 2s limit in 5% of the runs with normal analytic variation. The 10_{x} rule is violated if 10 consecutive QC

CASE STUDY 3.4

QC Rule Interpretation

Assist Miles, a laboratorian in your facility, in explaining the R_{4s} rule to the clinical pathology resident — including what type of error it detects.

© dotshock/Shutterstock.

values fall consistently on one side of the mean. The more levels of QC material analyzed, the higher the probability of a rule violation even in the absence of true error. When two controls are used, there is an approximately 10% chance that at least one control will be outside the 2s limits; when four controls are used, there is a 17% chance. For this reason, many laboratories use 2s limits either as a warning or not at all when used with other QC rules rather than criteria for run rejection. QC rules should be selected based on the performance of the test relative to the QC specifications. If a test is very precise and well within the QC specification, fewer QC rules are needed. For example, if there are 6 standard deviations between the natural process variation and the QC specification limits, a simple 1_{3S} rule will easily detect that a process is out of control long before it exceeds the specification limits. Conversely, a test that just meets the QC specification would need multiple rules to ensure that error is detected before exceeding the specification. By judiciously selecting QC rules based on method performance and specifications, error detection can be maximized while minimizing false rejections.

A final practical concept related to QC materials is repeats. It is common practice to repeat a control when a QC rule is violated. In the context of imprecision, this practice does not make much sense. A repeat sample is likely to regress to the mean and therefore fall within the limits, even if the process is out of control. If ever, QC repeats only make sense in the context of a shift or bias, where the bias would persist between samples. Collectively, operators should review the Levey-Jennings chart and think about their process and what a given QC violation means before acting. The operators' job is not getting QC to pass, but to report accurate results.

Table 3.4 Westgard QC Multi-Rules

Rule	Explanation	Error Detection
1_{2s}	One control observation exceeding the mean ±2s.	A warning rule that initiates testing of control data by other rules
1_{3s}	One control observation exceeding the mean ±3s.	High sensitivity to random error
2_{2s}	Two control observations consecutively exceeding the same +2s or −2s.	High sensitivity to systematic error
R_{4s}	One control exceeding the +2s and another exceeding the −2s.	Detection of random error
4_{1s}	Four consecutive control observations exceeding +1s or −1s.	Detection of systematic error
10_{x}	Ten consecutive control observations falling on one side or the other of the mean (no requirement for SD size).	Detection of systematic error

SD, standard deviation.

© Jones & Bartlett Learning.

Proficiency Testing

In addition to daily QC practices, laboratories are required to participate in external **proficiency testing (PT)** programs. Proficiency testing is a method used to validate a particular measurement process, where the results are compared with other external laboratories to give an objective indication of test accuracy. Acceptable performance in proficiency testing programs is required by the Clinical Laboratory Improvement Amendments (**CLIA**), the College of American Pathologists (CAP), and The Joint Commission (TJC) to maintain laboratory accreditation. Laboratories are required to enroll in a proficiency testing program approved by the Centers for Medicare and Medicaid Services (CMS) for each regulated analyte performed. When a laboratory performs proficiency testing, there are strict requirements, as follows:

1. The laboratory must incorporate proficiency testing into its routine workflow as much as possible.
2. The test values/samples must not be shared with other laboratories until after the deadline of submission of results to the proficiency provider. Referral of proficiency samples to another lab is

prohibited, and acceptance of proficiency samples from another lab is prohibited.
3. Proficiency samples are tested by bench technical staff who normally conduct patient testing; there can be no unnecessary repeats or actions outside of how a patient sample would be tested and reported.
4. Testing should be completed within the usual time it would take for routine patient testing.
5. Proficiency samples are to be performed and submitted on the primary analyzer when there are multiple analyzers in the laboratory following CLIA guidelines (D2006 493.801). In addition, at least twice per year, other analyzers within the laboratory are to be compared to the primary analyzer.
6. All proficiency failures and significant shifts and trends must be reviewed, investigated, and resolved within 30 days of final receipt of proficiency results. Emphasis must be placed on investigating potential patient impact during the time of proficiency testing.
7. Proficiency testing program must demonstrate a dynamic and real-time review time of all proficiency results by the laboratory director and delegated management personnel.

The bottom line is that the sample should be treated like a patient sample to yield a true indication of test accuracy. Proficiency samples are not to be analyzed more than once unless this is the standard policy for patient testing. After analyzing the PT specimens, the performing laboratorian must sign an affidavit attesting that the PT specimens were tested in the same manner as patient specimens and that the PT specimens were not shared outside the CLIA laboratory. Participation in a proficiency testing is another aspect of quality assurance and an important tool in the ongoing process of monitoring test performance.

Most clinical laboratories subscribe to programs provided by multiple proficiency agencies. The CAP program has been in existence for 50 years, and it is the standard for clinical laboratory proficiency testing. The majority of analytes are monitored with CAP proficiency surveys or alternate surveys. If there is no commercial proficiency testing program available for an analyte, the laboratory is required to implement an alternate proficiency test scheme, which can be comparison of test results using split specimens between two CLIA laboratories twice a year.

For a proficiency test, a series of unknown samples are sent several times per year to the laboratory from the program offering this analysis, such as CAP. The sample challenges are from 2 to 5 in quantity for each analyte for each testing event based on complexity and span a range of target values within the AMR and **medical decision levels** (value for an analyte that represents the boundary between different therapeutic approaches, sometimes referred to as critical values or panic values). These samples will also verify the **clinically reportable range (CRR)**, or the range of analyte that a method can quantitatively report, allowing for dilution, concentration, or another pretreatment used to extend the AMR. The samples are analyzed in the same manner as patient specimens as much as possible, and the results are reported to the proficiency program. The program then compiles the results from all of the laboratories participating in the survey (called *peer laboratories*) and sends a performance report back to each participating laboratory. Each analyte has a defined performance criteria (e.g., ±3 SDs to peer mean), where laboratories using the same method are graded by comparing them with the group. Some PT are not quantitative and are qualitative or semiquantitative (i.e., rapid pregnancy test, RPR titers) compared with other laboratories. Areas of pathology other than clinical chemistry are also subjected to mandatory proficiency qualitative/interpretive testing, including anatomic pathology, clinical microbiology, and clinical microscopy.

An example of a hypothetical survey is shown above. The βhCG survey was the eighth sample sent in that year (βhCG-08). The mean of all the laboratories using the same method was

Indicators of Analytic Performance

Proficiency testing
Internal quality control
Laboratory inspections (accreditation)
Quality assurance monitoring
Clinical utilization

Example of Proficiency Test Results for βhCG

βhCG-08: CAP value = 75.58; SD = 4.80; CV = 6.4%; *n* = 47 peer laboratories

Evaluation criteria: Peer group ±3 SDs; acceptable range 65.7–85.2 mIU/mL

Testing laboratory value = 71.54; SDI = –0.84 acceptable

75.58 mIU/mL. The SD and CV are indicated, as is the number of laboratories that participate in that survey ($n = 47$). The acceptance criteria for this test are established as within ±3 SDs (i.e., between 65.7 and 85.2 mIU/mL). The laboratory's result was 71.54 mIU/mL, which is −0.84 SD from the mean and is within the acceptable limits.

The acceptability criteria for proficiency testing are provided by the PT program. For regulated analytes, these criteria are often the CLIA limits (see Table 3.3). For nonregulated analytes, acceptable criteria are often determined by the scientific community at large. For example, the acceptability criterion for lactate dehydrogenase is ±20% or 3 SDs (whichever is greater) based on peer group data.

Proficiency testing allows each laboratory to compare its test results with those of peer laboratories that use the same or similar instruments and methods. Proficiency testing provides performance data for a given analyte at a specific point in time. Comparison of performance to a robust, statistically valid peer group is essential to identify areas for improvement. Areas of improvement that may be identified in a single PT event or over multiple events include variation from peer group results, imprecision, and/or results that trend above or below the mean consistently or at specific analyte concentrations. Use of these data allows laboratories to continuously monitor and improve their test performance. The PT samples also can serve as valuable troubleshooting aids when investigating problem analytes. In some hospitals, PT samples are also included in the laboratorian annual competency and/or new employee training program. Proficiency tests can also be beneficial in validating the laboratory's measurement method, technical training, and total allowable error limits for new tests.

Proficiency testing programs require thorough investigation of discrepant results for any analyte (i.e., failure). Laboratories may be asked to submit information that could include current and historical PT reports, QC and equipment monitoring, analysis and corrective action of the problem that caused the failure, and the steps taken to ensure the reliability of patient test results (patient impact). If the laboratory cannot resolve analyte testing discrepancies over several testing events, the testing facility may be at risk of losing the authority to perform patient testing for the analyte(s) in question. Also, if the laboratory violates any of the mandatory proficiency requirements referenced earlier, it may be at risk of losing its accreditation and authority to perform patient testing for all tests.

To develop and manage a successful PT program for the clinical laboratory, it is important to understand the documented requirements from the two main accreditation bodies. A large PT program often requires considerable personnel resources and costs for the laboratory, and is an essential factor for providing a quality management system. As an example of the scale and volume of PT, a laboratory (for an 800-bed hospital with outreach clinics) may perform about 10,000 PT in a year for roughly 500 individual analytes. Even a smaller laboratory (with three analyzers and a limited test menu) may perform several hundred proficiencies per year. Besides meeting required accreditation standards, PT allows the laboratory to objectively ensure patient results are accurate.

Lean Six Sigma

In the previous section, we discussed how quality control (QC) continuously monitors processes. In this section, we expand on these ideas to discuss the concept of quality improvement (QI) and quality assurance (QA) (**Figure 3.12**). Quality improvement and quality assurance (QI/QA) goes beyond monitoring, detecting, and preventing errors and eliminates potential adverse patient impact events. QI/QA achieves new performance levels, not otherwise realized through QC, and addresses chronic problems. **Lean Six Sigma** is the combination of Lean principles and Six Sigma methodology. Lean principles work to eliminate the waste, such as streamlining a process to reduce wait times or modifying a process to reduce cost. The Six Sigma business management strategy seeks to improve a process's performance by identifying and eliminating causes of defects and errors, resulting in eliminating variation in the process. In its simplest form, Lean asks the question "Does this process (or step) need to exist?" Six Sigma asks the question, "How can this process be improved?" Together as Lean Six Sigma, sometimes abbreviated as LSS, they are being increasingly used to reduce waste (Lean) and error (Six Sigma) within the healthcare system, including laboratories.

Quality Assessment

Quality assessment (QA) is the continuous process of monitoring the effectiveness of the Quality Control Plan (QCP). Practices, processes, and resources to consider for monitoring effectiveness of a QCP may include, but are not limited to, (1) routine reviews of

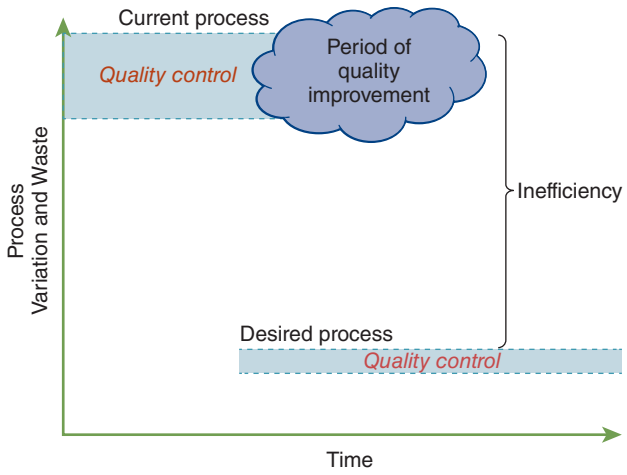


Figure 3.12 QI for a given process. QI seeks to eliminate inefficiencies in a process through reduction of variation and waste. The current process exhibits high variation and waste, and QC measures do not adequately manage the process. After a period of QI, the desired process shows much smaller variation and waste and tighter control of the process using the new QC measures. Lean Six Sigma methodology seeks to identify, measure, and eliminate the large gaps of inefficiency in a process. Examples from the laboratory include data accuracy, turnaround times, and reagent inventory.

Quality Control Performance, (2) proficiency performance reviews, (3) patient chart reviews, (4) specimen rejection logs, (5) turnaround time reports, (6) complaint reports, (7) variance reports, and (8) safety reports. The process of QA involves ongoing monitoring and assessing of performance in a defined way. The laboratory must establish and follow written policies and procedures to monitor and assess, and when indicated, correct problems identified. The monitoring should include, but is not limited to, the following risk assessment components: specimen, test system, reagents, environment, and testing personnel. The QA must also include a review of the effectiveness of corrective actions taken to resolve and monitor problems identified. The laboratory must update the risk assessment and modify the QCP, as necessary, based on the information obtained from the QA. A common confusion when first being introduced to the concepts of QA is understanding how it is different from the quality control assessments that are done in parallel and are closely related. An example of a QC activity may be recording the room temperature on a log sheet; the parallel QA activity may then be reviewing the room temperature log for problems and evidence of corrective actions. Similarly, a QC activity may be documenting personnel training, and a parallel QA activity would

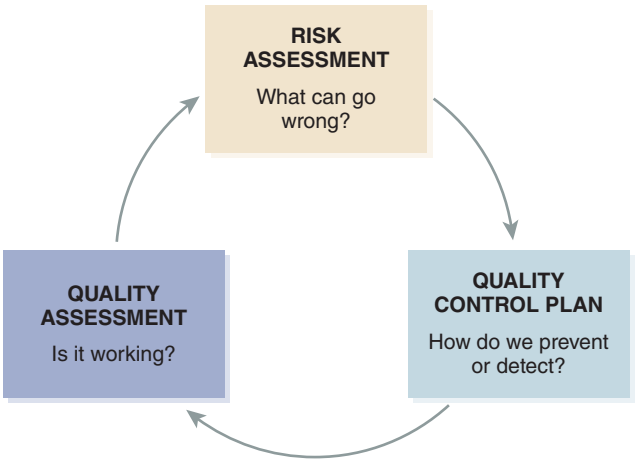


Figure 3.13 Quality assessment is integral to a complete Individual Quality Control Program (IQCP).

be reviewing personnel training records for completion of required training and competency assessments. QA is an integral part of the QCP, receiving input from the QCP and being used to assess risk, as outlined in **Figure 3.13**.

Individual Quality Control Program (IQCP): An Option to Streamline QCP

Under the Clinical Laboratory Improvement Amendments (CLIA), the Centers for Medicare and Medicaid Services (CMS) has implemented IQCP as an acceptable QC option. The new requirements for IQCP provide laboratories with the framework to implement IQCP, when appropriate, and offer flexibility to design a QC plan that meets the needs of the laboratory. These changes will ensure that accredited laboratories remain in compliance with the CLIA and CMS regulations. As one of the leading accrediting organization, the College of American Pathologists (CAP) revised its checklists in the 2015 edition to address the IQCP changes. These adjustments ensure that CAP-accredited laboratories will remain CLIA and CMS compliant. For more information, the CMS has downloadable information regarding implementation. Like the QCP, the IQCP integrates QCP with QA and risk assessment, customized to individual laboratories.

Reference Range Studies

Laboratory test data are used to make medical diagnoses, assess physiologic function, and manage therapy. When interpreting laboratory data, clinicians

compare the measured test result from a patient with a reference range.

Reference ranges include all the data points that define the range of observations; for example, if the interval is 5 to 10, a patient result of 5 would be considered at the lower end but within the range. The upper and lower reference limits are set to define a specified percentage (usually 95%) of the values for a population. This means that a percentage of patients will fall outside the reference range in the absence of any condition or disease (usually 5% total, or 2.5% on either side of the reference range). Reference ranges are sometimes erroneously called "normal ranges." Normal ranges are the range of results between two medical decision levels that typically correspond to the central 95% of results from a healthy patient population. Note: Of the results, 2.5% will be above the upper limit, and 2.5% will be below the lower limit of the normal range, hence, 5% will be outside the "normal range." While all normal ranges are in fact reference ranges, not all reference ranges are normal ranges. This is exemplified by the reference range for therapeutic drug levels, also known as a drug's **therapeutic range** (the reference range applied to a therapeutic drug). In this case, a "normal" individual would not have any drug in their system, whereas a patient on therapy has a defined target range. It is important to note that reference ranges do not imply an absolute maximum and minimum value for a test, only a range of values commonly seen in a healthy population.

The clinical laboratory is required by good laboratory practice and accreditation agencies (i.e., the CAP, TJC) to either verify or establish reference ranges for any new tests or significant changes in methodology (**Box 3.3**).

The core protocols for both establishing and verifying reference ranges are reviewed in this section. Other terms are used for values or ranges that help the clinician determine the relationship of patients' test results to statistically determined values or ranges for the clinical condition under treatment.

The application of reference ranges can be grouped into three main categories: diagnosis of a disease or condition (**Table 3.5**), monitoring of a physiologic condition (**Table 3.6**), or monitoring

Box 3.3 Examples of CAP Checklist Requirements Regarding Reference Intervals for Laboratory Inspection

The laboratory verifies or establishes its reference intervals

NOTE: Reference intervals are important to allow a clinician to assess patient results against an appropriate population. The reference intervals must be established or verified for each analyte and specimen source (eg, blood, urine, cerebrospinal fluid), when appropriate. For example, a reference interval can be verified by testing samples from 20 healthy representative individuals; if no more than two results fall outside the proposed reference interval, that interval can be considered verified for the population studied.

If a formal reference interval study is not possible or practical, then the laboratory should carefully evaluate the use of published data for its own reference intervals, and retain records of this evaluation. For many analytes (eg, therapeutic drugs, cholesterol, and CSF total protein), literature references or a manufacturer's package insert information may be appropriate.

Evidence of Compliance

- Record of reference interval study **OR** records of verification of manufacturer's stated interval when reference interval study is not practical (eg, unavailable normal population) **OR** other methods approved by the laboratory/section director

The laboratory evaluates the appropriateness of its reference intervals and takes corrective action if necessary

NOTE: Criteria for evaluation of reference intervals include:

1. Change of analytic methodology
2. Change in patient population

If it is determined that the range is no longer appropriate for the patient population, corrective action must be taken.

Evidence of Compliance

- Records of evaluation and corrective action, if indicated

Table 3.5 Thyroid-Stimulating Hormone (TSH) Reference Ranges

Patients	Age	TSH (μIU/ML)
Pediatric	0–3 d	1.00–20.00
	3–30 d	0.50–6.50
	31 d–5 m	0.50–6.00
	6 m–18 y	0.50–4.50
Adults, ambulatory, healthy	>18 y	0.60–3.30

Data from Dugaw KA, Jack RM, Rutledge J. Pediatric reference ranges for TSH, free T4, total T4 total T3 and T3 uptake on the vitros ECi analyzer. *Clin Chem.* 2001;47:A108.

therapeutic drugs (**Table 3.7**). These different applications require different approaches for determination of a reference range. Specifically, therapeutic drug targets are not derived from a healthy population, and unique physiologic conditions require the appropriate reference population.

Table 3.6 hCG at Defined Gestational Age

Approximate βhCG Levels at Defined Gestational Age Units MIU/ML (U/L)			
Weeks of Pregnancy	**Mean**	**Range**	**N**
4	1110	40–4480	42
5	8050	270–28,700	52
6	29,700	3700–84,900	67
7	58,800	9700–120,000	62
8	79,500	31,000–184,000	37
9	91,500	61,200–152,000	25
10	71,000	22,000–143,000	12
14	33,100	14,300–75,800	219
15	27,500	12,300–60,300	355
16	21,900	8800–54,500	163
17	18,000	8100–51,300	68
18	18,400	3900–49,400	30
19	20,900	3600–56,600	14

Table 3.7 Therapeutic Management Targets for Digoxin

Normal	0.8–1.8 ng/mL (collected 6 h after dose)
Critical	>2.0 ng/mL

Tables 3.5 to 3.7 also demonstrate the complexity of reference ranges when multiple levels (partitions) of reference ranges are required by the clinician. The framework for verifying or establishing reference ranges is one that can be overwhelming and not feasible for many valid reasons for large and small clinical laboratories. The costs, personnel, and resource requirements mandate that the reference range experiment be feasible, well defined, and structured. The clinical laboratory director may determine that a review of literature references or manufacturer's package inserts is appropriate in assigning reference ranges for an analyte or this additional information may allow for the shorter reference range verification study (i.e., 20 study individuals).

Establishing Reference Ranges

The Clinical and Laboratory Standards Institute (CLSI) has published a preferred guideline/resource for establishing or verification of reference ranges (**Box 3.4**).[21]

Determining Whether to Establish or Transfer and Verify Reference Ranges

Whether to transfer and verify a reference range or establish an entirely new reference range for a new method/analyte depends on several factors, such as the presence of an existing reference range for assay and on the results of a statistical analysis comparing the test method with the reference method. The most basic method of comparison involves plotting a reference method against a test method and fitting a linear regression, described earlier in Figure 3.7. If the correlation coefficient is 1.0, the slope is 1, and intercept is 0, the two methods agree and may not require new reference ranges. In this case, a simple reference range verification study is all that may be required. Conversely, if the two methods differ considerably, then a new reference range needs to be established or external resources reviewed.

CASE STUDY 3.5

Reference Range Study Design

Remember Miles and Mía from Chapter 1? The laboratory supervisor asked Miles and Mía to design a reference range study for a new test. Per the kit package insert, the analyte results are known to be different in men and women and are also affected by the consumption of aspirin, age, and time of day. Mía and Miles want to create a questionnaire to collect the appropriate information needed to perform a reference range study. How should they account for these variables in the data collection?

Data Analysis to Establish a Reference Range

To establish a reference range, it is recommended that the study include at least 120 individuals. This can be challenging and costly, but it may be necessary for esoteric and laboratory-developed tests. The reference range is calculated statistically using methods that depend on the distribution of the data. In the most basic sense, data may be either normally distributed (Gaussian) or skewed (non-Gaussian) (see Figure 3.3). If reference data are normally distributed, the reference range can be determined using a parametric method. A parametric method defines the values as following a normal (Gaussian) distribution by the mean ±1.96 SDs; by centering on the mean, this formula will include the central 95% of values as given in the example in **Figure 3.14A**.

Most analytes do not display a normal (Gaussian) distribution. For example, the distribution of βhCG in pregnant individuals over a range of gestational ages is skewed (**Figure 3.14B**); the selection of a wide range of gestational ages was selected as an example to emphasize the need for nonparametric intervals. Data that are not normally distributed (i.e., non-Gaussian) must be analyzed using nonparametric analyses. Nonparametric determination of the reference range is analyzed using percentiles, which do not depend on the distribution. The reference range is determined by using the central 95% of values; the reference range is therefore defined by the 2.5th to the 97.5th percentiles, as demonstrated in Figure 3.14B. To calculate the interval, values are ranked from lowest to highest and the 2.5th and 97.5th percentiles are then calculated as follows:

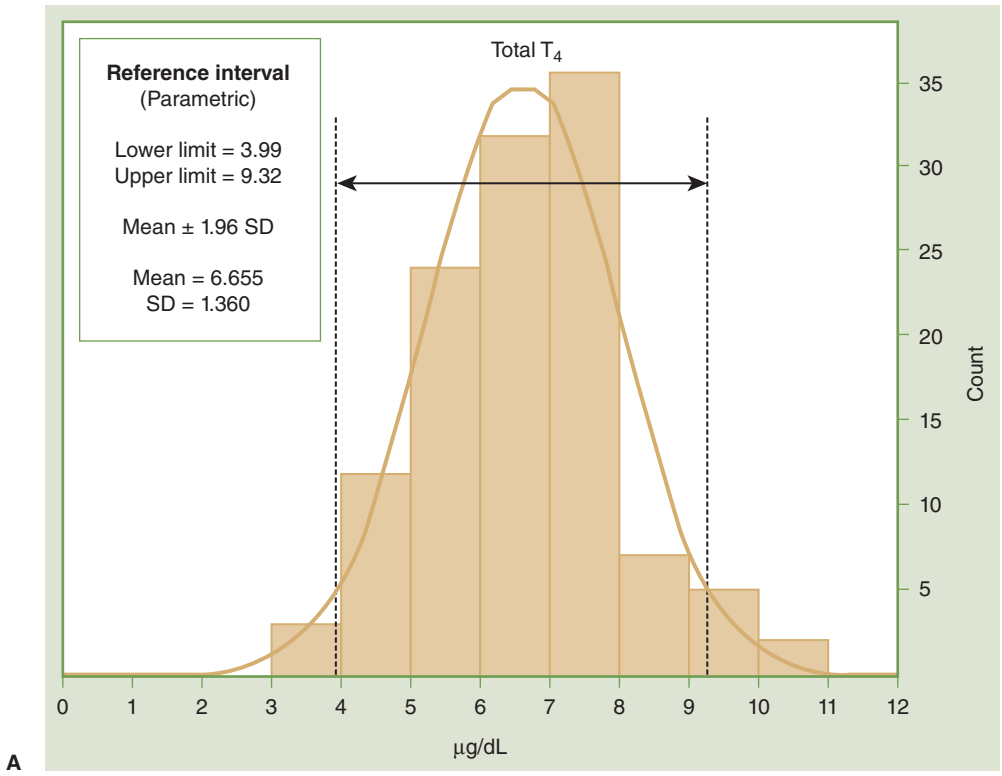
$$n = \text{number of reference specimens}$$
$$2.5\text{th percentile} = 0.025(n + 1) \qquad \textbf{(Eq. 3.8)}$$
$$97.5\text{th percentile} = 0.975(n + 1)$$

Most reference range analyses are determined using nonparametric analysis.

With available statistical software packages, reference ranges are rarely determined manually, although basic reference range verification can be done with in a spreadsheet with minimal effort just reviewing

Box 3.4 To Establish a Reference Range Study

1. Define an appropriate list of biological variations and analytic interferences from medical literature.
2. Choose selection and partition (e.g., age, gender, etc.) criteria.
3. Complete a written consent form and questionnaire to capture selection criteria.
4. Categorize the potential reference individuals based on the questionnaire findings.
5. Exclude individuals from the reference sample group based on exclusion criteria.
6. Define the number of reference individuals in consideration of desired confidence limits and statistical accuracy.
7. Standardize collection and analysis of reference specimens for the measurement of a given analyte consistent with the routine practice of patients.
8. Inspect the reference value data and prepare a histogram to evaluate the distribution of data.
9. Identify possible data errors and/or outliers and then analyze the reference values.
10. Document all of the previously mentioned steps and procedures.



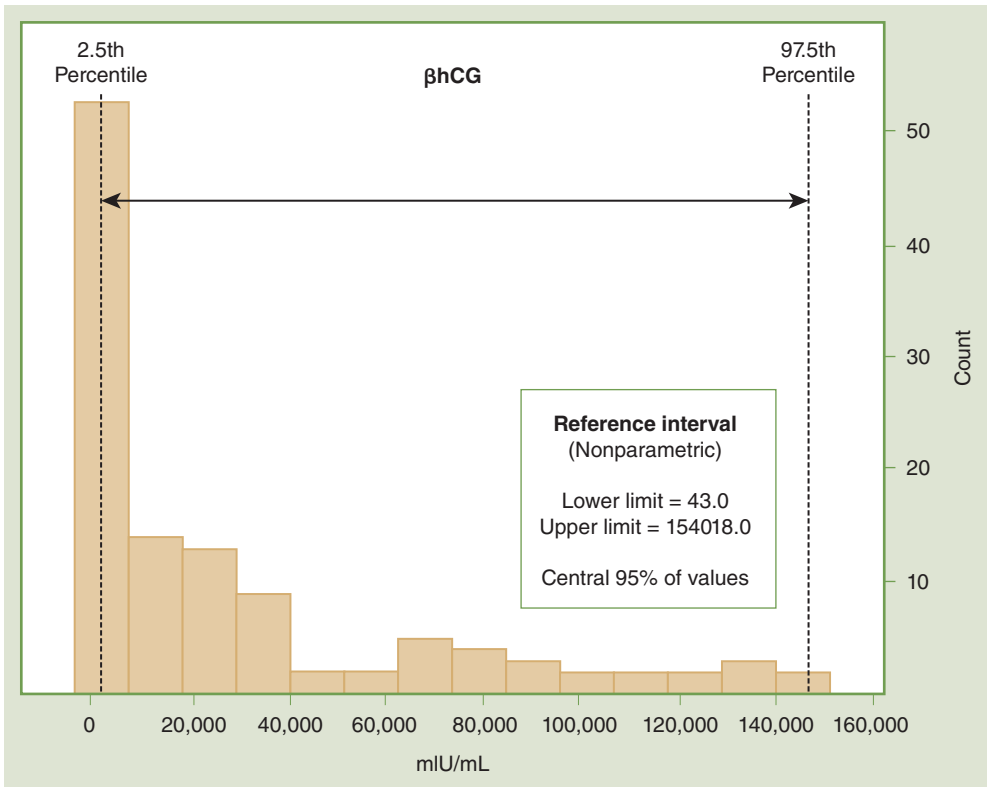


Figure 3.14 **(A)** Histogram of total thyroxine (TT_4) levels in a real population illustrating a shape indicative of a Gaussian distribution, which is analyzed by parametric statistics. The reference range is determined from the mean ±1.96 SDs. **(B)** Histogram of β-human chorionic gonadotropin (βhCG) levels in a population of pregnant women demonstrating non-Gaussian data and nonparametric determination of the reference range. The reference range is determined from percentiles to include the central 95% of values, although the selection of a wide range of gestational ages makes this a poor population for a reference range study, it does demonstrate the application of nonparametric intervals.

patient results (>90% or $n \varepsilon$ 36 out of $n = 40$ total) that fall within the proposed reference range. However they are determined, it is important to understand how basic statistical concepts are used to establish reference ranges.

Data Analysis to Transfer and Verify a Reference Range

When possible, clinical laboratories rely on assay manufacturers or on published primary literature as a source of reference ranges. This avoids the expensive and lengthy process of establishing a reference range on a minimum of 120 healthy people. The CLSI allows less vigorous studies to verify a reference range with as few as 20 subject specimens.[21] The main assumption in using transference studies is that the reference method is of high quality and the subject populations are similar. The manufacturer's reported 95% reference limits may be considered valid if no more than 10% of the tested subjects fall outside the original reported limits. **Figure 3.15** shows an example for verification of the manufacturer's reference range for free thyroxine (fT_4). In this example, fewer than 10% are outside the manufacturer's limits, enabling the reference range to be adopted by the laboratory. **Figure 3.16** demonstrates a simple algorithm to verify reference ranges.

It is considered good laboratory practice to monitor reference ranges regularly and communicate and document to the physicians and providers. Some common problems that occur when determining reference ranges are given in **Table 3.8**. To help identify reference range problems, the clinical laboratorian should be aware of common flags. These flags often come in the form of an event or communication that alerts the laboratory that there is a potential problem with a test. Flags for reference ranges can include vendor notifications, clinician queries of a particular test, and shifts/trends in large-average numbers of patients over time. In addition, all studies need to be evaluated as to patient population sampled and its representation of the desired donor population. Any of these or other related factors may warrant a review of existing reference ranges.

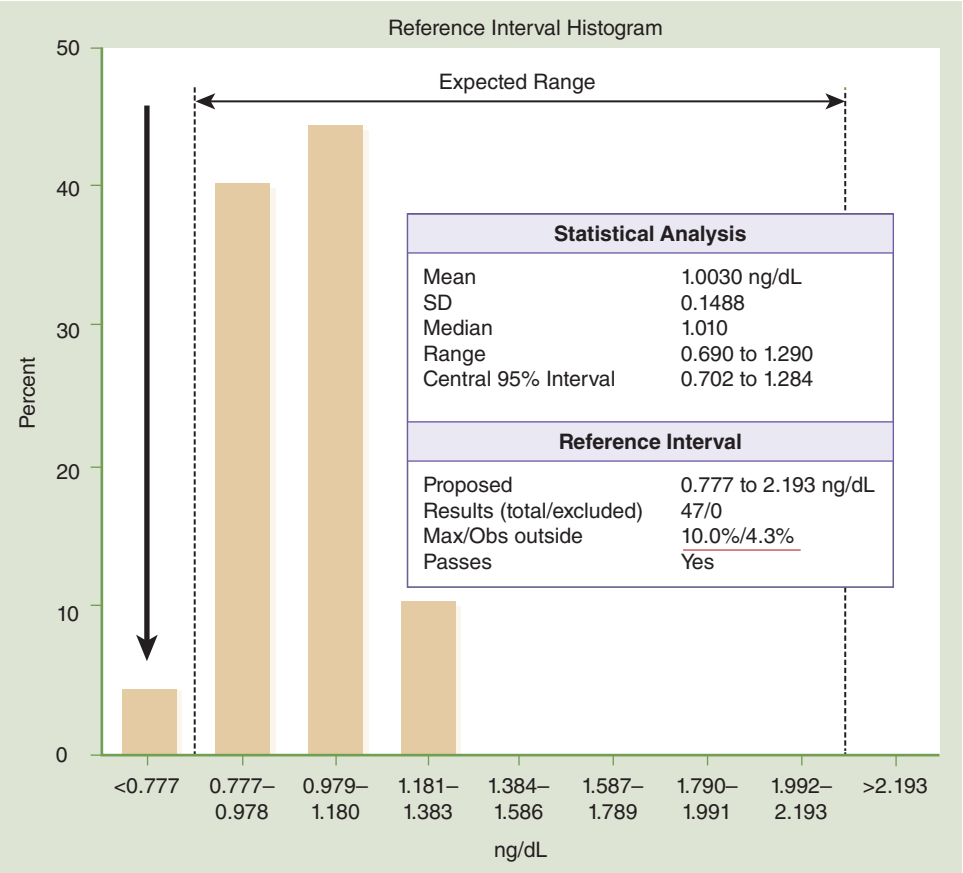


Figure 3.15 Reference verification test for free thyroxine (fT_4). Only 4.3% of the values are outside the expected range (*arrow*). The test passes because this is less than the allowable number of outlying samples (10%) (underlined in *red*).

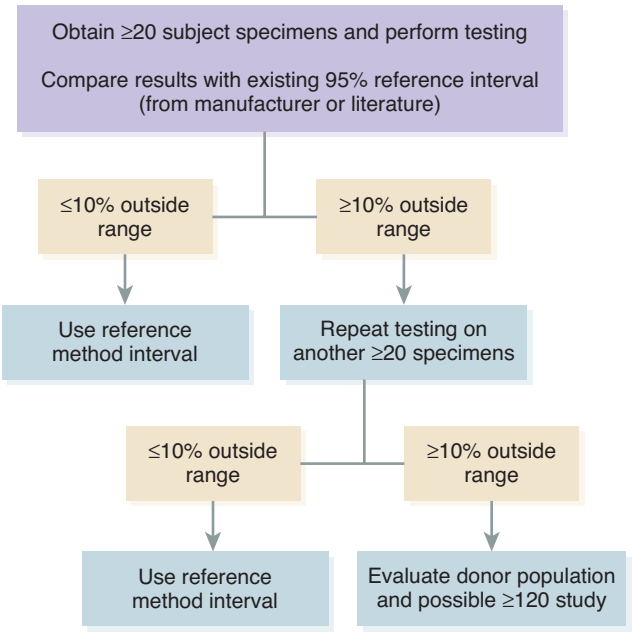


Figure 3.16 Algorithm to test whether a reference range can be verified. A published reference range (from a manufacturer or scientific literature) can be adopted if only a few samples fall outside the range. When possible, laboratory reference ranges are verified because of the time and expense of establishing a new interval.

Not all analytes use population-, gender-, or age-based reference ranges. For example, therapeutic drugs are not found in healthy individuals, such that the target values are based either on minimum effective concentrations or on toxicity limits. The target limits are derived from clinical or pharmacological studies rather than the healthy population studies described above. Cholesterol and hemoglobin A1c (HbA1c) are also used in the context of clinical outcomes rather than reflecting values in a healthy population. The quintessential example of this is vitamin D testing in extreme latitudes or smog-ridden large cities, where even a healthy population will be deficient.

Table 3.8 Common Problems Encountered when Monitoring Reference Ranges

Changes in reagent formulations by the vendor (e.g., new antibody)
Minor changes in reagents due to lot-to-lot variations
Differences between reference range and test populations—selection bias

Method Evaluation

Regulatory Aspects of Method Evaluation

Currently, clinical laboratories more often select and evaluate commercially developed methods instead of developing their own. Most commercially developed tests have U.S. Food and Drug Administration (FDA) approval, only requiring a laboratory to provide a limited but regulatory mandated evaluation of a method to verify the manufacturer's performance claims and to see how well the method works specifically in the laboratory and patient population served.

The Centers for Medicare and Medicaid Services (CMS) and the U.S. Food and Drug Administration (FDA) are the primary government agencies that influence laboratory testing methods in the United States. The FDA regulates laboratory instruments and reagents, and the CMS regulates the CLIA.[22] Most large laboratories in the United States are accredited by the CAP and TJC.

The FDA's Office of In Vitro Diagnostic Device Evaluation and Safety (OIVD) regulates diagnostic tests. Tests are categorized into one of four groups: (1) waived, (2) moderate complexity, (3) high complexity, or (4) emergency use authorization (EUA). The FDA clears waived tests to be so simple that they are most likely accurate but most importantly would pose negligible risk of harm to the patient if not performed correctly. A few waived testing examples include urine dipstick tests, qualitative pregnancy testing, rapid strep, rapid HIV, rapid HCV, and glucose tests. For waived testing, personnel requirements are not as stringent as moderately complex testing or high-complexity testing, but there are still mandatory training, competency, and quality control requirements. Most automated methods are moderate complexity, while manual methods and methods requiring more interpretation are high complexity. The patient management/treatment impact is also a factor in determining whether a test is waived, moderate complexity, or high complexity. The CLIA final rule requires that waived tests simply follow the manufacturer's instructions. Emergency use authorization (EUA) is a way for the FDA to allow use of waived, moderate-, or high-complexity tests to be used in emergency situations without going through the tenuous process of regular authorization. For instance, most of the COVID-19 test (both waived and nonwaived tests) were all EUAs in an effort to quickly and efficiently test patients to ensure the

Table 3.9 General CLIA Regulations of Method Validation

Nonwaived FDA–Approved Tests	1. Demonstrate test performance comparable to that established by the manufacturer. a. Accuracy b. Precision c. Reportable range 2. Verify reference (normal) values appropriate for patient population.
Nonwaived FDA–Approved Tests Modified or Laboratory Developed Tests (LDTs)	1. Determine a. Accuracy b. Precision c. Analytic sensitivity d. Analytic specificity (including interfering substances) e. Reportable range of test results f. Reference/normal ranges g. Other performance characteristics, such as specimen and reagent stability h. Calibration and quality control procedures

Data from Clinical Laboratory Improvement Amendments of 1988; final rule. *Fed Regist.* 7164 [42 CFR 493 1253]: Department of Health and Human Services, Centers for Medicare and Medicaid Services; 1992.

safety of the public health during the Coronavirus pandemic of 2020 and 2021. Both moderate- and high-complexity tests require in-laboratory validation. However, FDA-approved nonwaived tests may undergo a more basic validation process (**Table 3.9**), whereas a more extensive validation is required for tests developed by laboratories (Table 3.9).

Allowable Analytical Error

An important aspect of method evaluation is to determine if the random and systematic errors (total error) are less than the total allowable analytical error (TE_a). In the past, there have been several methodologies that have been used to establish TE_a, including physiologic variation, multiples of the reference range, and pathologist judgment.[23–26] The Clinical Laboratory Improvement Amendments of 1988 (CLIA 88) published total allowable error (TE_a) for an array of clinical tests.[27] The TE_a limits published by the CLIA specify the maximum error allowable by federally mandated proficiency testing (see examples in Table 3.3). These performance standards are now being used to determine the acceptability of clinical chemistry analyzer performance.[28,29] If the test does not meet the allowable error criteria, it must be modified to reduce error or be rejected.

Comprehensive bias studies are demanding on personnel, time, and budgets. This has led to the description of abbreviated experiments that could be undertaken to estimate imprecision and inaccuracy.[30] CLIA has published guidelines for such an abbreviated application, which can be used by a laboratory to confirm that the precision and accuracy performance is consistent with the manufacturer's reported claims. These studies can be completed in 5 working days, making it likely laboratories will use these guidelines to set up new methodologies. The CLSI also has several documents that describe best practices for method evaluation with worked examples.

Method Evaluation Acceptance Criteria

Method evaluations are based on *precision* and *accuracy*, which should be compared with the maximum allowable error budgets based on medical criteria and regulatory guidelines. Remember, precision is the dispersion of repeated measurements around a mean (true level), as shown in **Figure 3.17A** with the mean represented as the bull's-eye. Precision is estimated from studies in which multiple aliquots of the same specimen (with a constant concentration) are analyzed repetitively. Accuracy, or the difference between a measured value and its actual value (also known a *bias*), is due to the presence of a systematic error, as represented in **Figure 3.17B**. Systematic error can be due to constant or proportional error and is estimated from three types of study: (1) recovery, (2) interference, and (3) a comparison of methods study.

Collectively, the data gathered in precision studies and accuracy studies, the analytical measurement range, method comparison, and reference range studies are used to guide test implementation decisions. That is, the data do not define if a test method

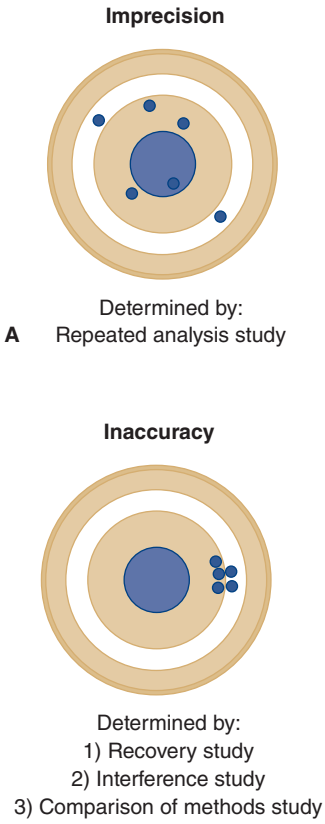


Figure 3.17 Graphic representation of **(A)** imprecision and **(B)** inaccuracy on a dartboard configuration with bull's-eye in the center.

is acceptable by itself. Clinical judgment is required to determine if the analytical performance is acceptable for clinical use with consideration for the nature and application of the analyte. For example, imprecision for a pregnancy test around the cutoff value of 5 or 10 mIU/mL is more of a concern than it would be at 15,000 mIU/mL. Likewise, proportional bias for troponin at a high concentration is of less concern than it is near the clinical decision point (troponin is an important analyte used to determine myocardial infarction). Thus, method evaluation studies and statistical analysis are necessary but not sufficient to determine if a test is valid. The validation and statistics serve as the basis on which to make the decision.

Diagnostic Efficiency

Ideally, healthy patients would have completely distinct laboratory values from patients with disease. However, the reality is that laboratory values usually overlap significantly between these populations. To determine how good a given test is at detecting and predicting the presence of disease (or a physiologic condition), there are a number of different parameters that are used. These parameters are broadly defined as *diagnostic efficiency*, which can be broken down into sensitivity, specificity, and predictive values.

Measures of Diagnostic Efficiency

Parameters of diagnostic efficiency are intended to quantify how useful a test is for a given disease or condition. For example, βhCG is used as a test to diagnose pregnancy. While βhCG is excellent for this purpose, there are instances where βhCG may be increased because of other causes, such as cancer (trophoblastic tumors), or below the cutoff, as is the case very early in pregnancy.

It is important to recognize that there is both diagnostic (clinical) and analytic sensitivity. **Diagnostic sensitivity**, also referred to simply as **sensitivity**, is the ability of a test to detect a given disease or condition, while **analytic sensitivity** refers to the ability to detect small quantities or changes in an analyte. The analytic sensitivity also determines the lower **limit of detection (LOD)** for a given analyte (the lowest amount that can be reliably detected), whereas clinical sensitivity refers to the proportion of individuals with that disease who test positively with the test. The LOD is different from the **limit of quantification (LOQ)**, which is the lowest amount of analyte that can be reported while achieving a precision target. *Sensitivity* can be calculated from simple ratios (**Figure 3.18A**). Patients with a condition who are correctly classified by a test to have the condition are called **true positives** (TPs). Patients with the condition who are classified by the test as not having the condition are called **false negatives** (FNs). Using the βhCG test as an example, sensitivity can be calculated as follows:

$$\text{Diagnostic sensitivity (\%)} = \frac{\text{(No. of pregnant with positivee test)}}{\text{(No. of pregnant individuals tested)}}$$

$$= \text{(TP)} / \text{(TP + FN)} \quad \textbf{(Eq. 3.9)}$$

Another measure of clinical performance is **diagnostic specificity**. Diagnostic specificity (also known as **specificity**) is defined as the proportion of individuals without a condition who have a negative test for that condition (**Figure 3.18B**). Note that there is also an **analytic specificity**, which refers to

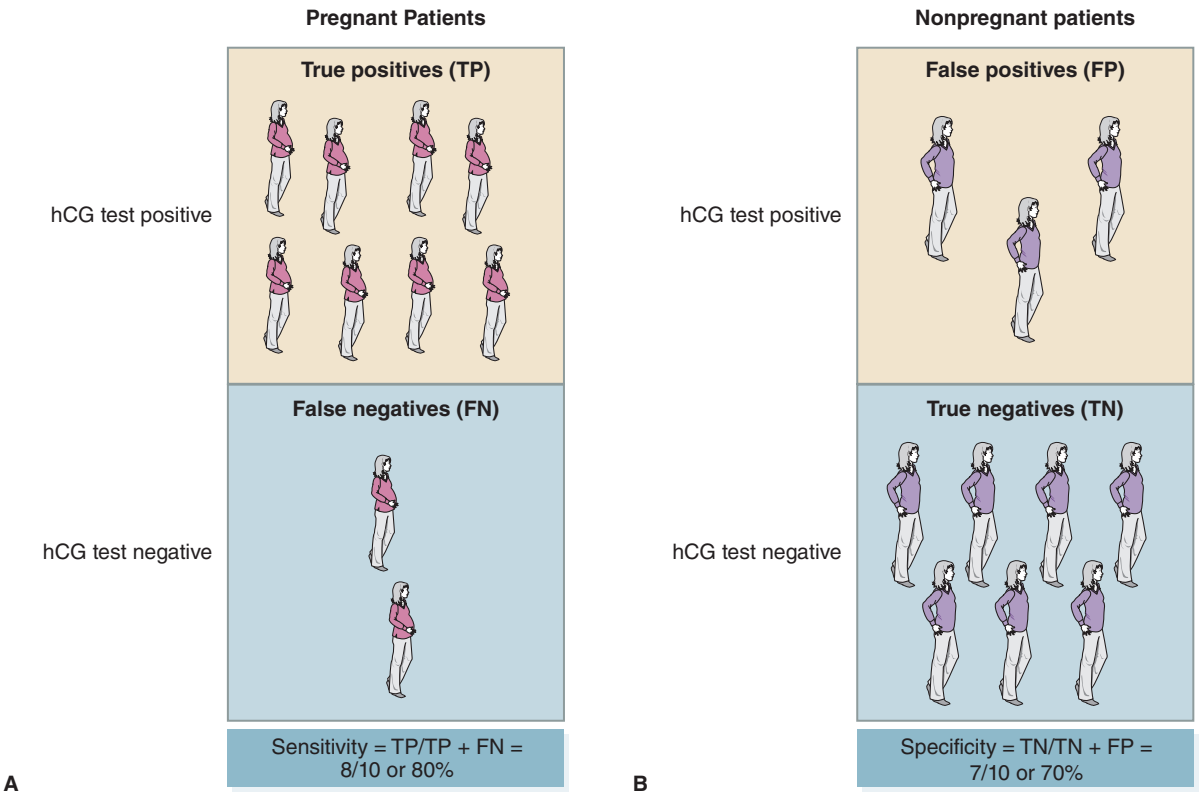


Figure 3.18 The sensitivity and specificity of β-human chorionic gonadotropin (βhCG) for pregnancy. **(A)** Sensitivity refers to the ability to detect pregnancy. **(B)** Specificity refers to the ability of the test to correctly classify nonpregnant women. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

cross-reactivity with other substances (i.e., the ability to only detect the desired analyte). Continuing with βhCG as an example, diagnostic specificity refers to the percentage of nonpregnant individuals that have a negative test compared with the number of nonpregnant individuals tested. Patients who are not pregnant and have a negative βhCG test are called **true negatives** (TNs), whereas those who are incorrectly classified as pregnant by the test are called **false positives** (FPs). Clinical specificity can be calculated as follows:

$$\text{Specificity (\%)} = \frac{\text{(No. of nonpregnant with negative test)}}{\text{(No. of nonpregnant individuals tested)}} = \text{(TN) / (TN + FP)} \quad \textbf{(Eq. 3.10)}$$

For example, a sensitivity of 100% plus a specificity of 100% means that the test detects every patient with disease and that the test is negative for every patient without the disease. Because of the overlap in laboratory values between people with and without disease, this is, of course, almost never the case.

There are other measures of diagnostic efficiency such as predictive values. There are predictive values for both positive and negative test results. The **positive predictive value** (PPV) of a test refers to the probability of an individual having the disease if the result is abnormal ("positive" for the condition). Conversely, the **negative predictive value (NPV)** of a test refers to the probability that a patient does not have a disease if a result is within the reference range (test is negative for the disease) (**Figure 3.19**). Predictive values are also calculated using ratios of TPs, TNs, FPs, and FNs as follows:

$$\text{PPV} = \frac{\text{(No. of pregnant with positive test)}}{\text{(No. with positive test)}} \quad \textbf{(Eq. 3.11)}$$
$$\text{(TP) / (TP + FP)}$$

$$\text{NPV} = \frac{\text{(No. of nonpregnant with negative test)}}{\text{(No. with negaative test)}}$$
$$\text{(TN) / (TN + FN)} \quad \textbf{(Eq. 3.12)}$$

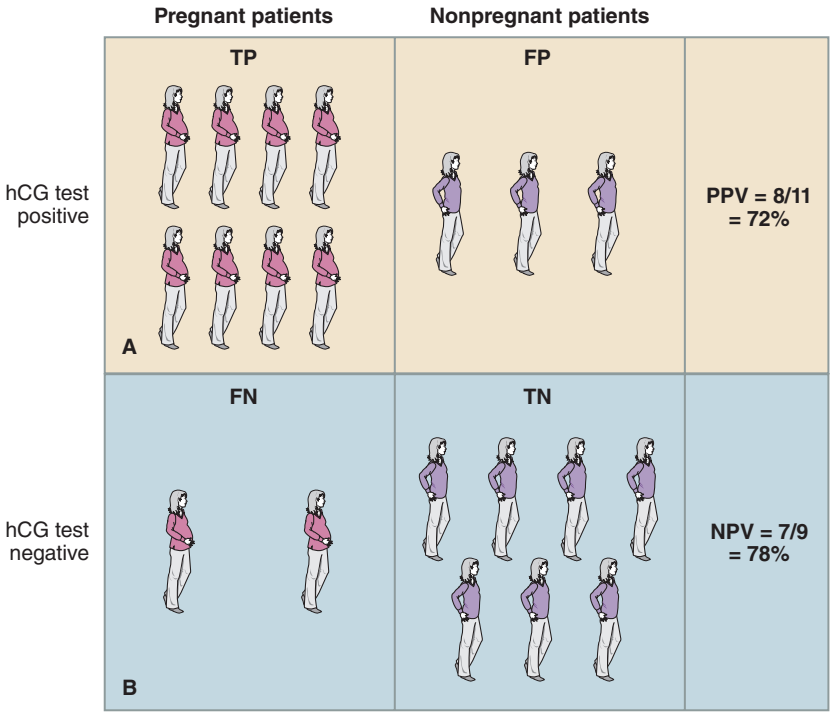
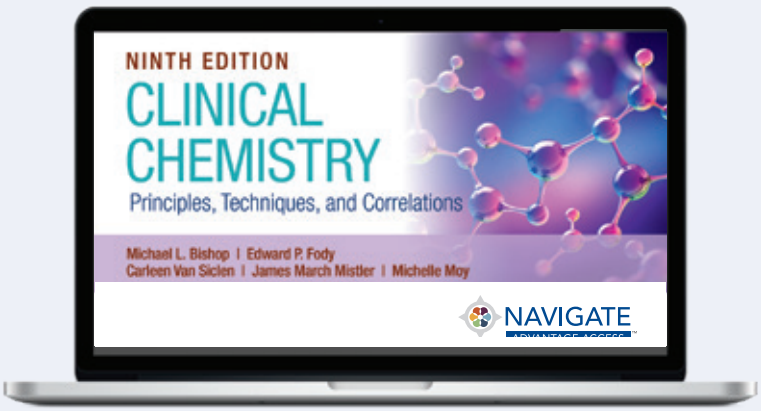


Figure 3.19 Positive and negative predictive values using β-human chorionic gonadotropin (βhCG) as a test for pregnancy. **(A)** Predictive value of a positive test (PPV) indicates the probability of being pregnant if the test is positive. **(B)** Predictive value of a negative test (NPV) refers to the probability of being nonpregnant if the test is negative. FN, false negative; FP, false positive; TN, true negative; TP, true positive.

Using the data from Figure 3.19, if the βhCG test is "positive," there is a 72% chance the patient is pregnant; if the test is negative, then there is a 78% chance the patient is not pregnant. It is important to understand that unlike sensitivity and specificity, predictive values depend on the prevalence of the condition in the population studied. Prevalence refers to the proportion of individuals within a given population who have a particular condition.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



Online Resources

AACC	http://aacc.org
Centers for Disease Control and Prevention	http://www.cdc.gov
CLIA	http://www.cms.hhs.gov/clia
CLSI	http://www.clsi.org
College of American Pathologists	http://CAP.org
FDA	http://www.FDA.gov
Lab Tests Online	http://www.labtestsonline.org
Westgard QC	http://www.westgard.com

References

1. Clinical and Laboratory Standards Institute (CLSI). *Evaluation of Precision of Quantitative Measurement Procedures.* Wayne, PA: CLSI 2014 (Document no. EP05-A3).
2. Westgard JO, de Vos DJ, Hunt MR, et al. Concepts and practices in the evaluation of clinical chemistry methods. II. Experimental procedures. *Am J Med Technol.* 1978;44: 420–430.
3. Westgard JO, Burnett RW. Precision requirements for cost-effective operation of analytic processes. *Clin Chem.* 1990;36:1629–1632.
4. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. *Clin Chem.* 1974;20:825–833.
5. Westgard JO, de Vos DJ, Hunt MR, et al. Concepts and practices in the evaluation of clinical chemistry methods. V. Applications. *Am J Med Technol.* 1978;44:803–813.
6. Wakkers PJ, Hellendoorn HB, Op de Weegh GJ, et al. Applications of statistics in clinical chemistry. A critical evaluation of regression lines. *Clin Chim Acta.* 1975;64: 173–184.
7. Clinical and Laboratory Standards Institute (CLSI). *Measurement Procedure Comparison and Bias Estimation Using Patient Samples.* Wayne, PA: CLSI 2018 (Document no. EP09c).
8. Shahzad K, Kim DH, Kang MJ. Analytic evaluation of the beta-human chorionic gonadotropin assay on the Abbott IMx and Elecsys 2010 for its use in doping control. *Clin Biochem.* 2007;40:1259–1265.
9. Westgard JO, Hunt MR. Use and interpretation of common statistical tests in method-comparison studies. *Clin Chem.* 1973;19:49–57.
10. Westgard JO. Precision and accuracy: concepts and assessment by method evaluation testing. *Crit Rev Clin Lab Sci.* 1981;13:283–330.
11. Cornbleet PJ, Gochman N. Incorrect least-squares regression coefficients in method-comparison analysis. *Clin Chem.* 1979;25:432–438.
12. Feldmann U, Schneider B, Klinkers H, et al. A multivariate approach for the biometric comparison of analytical methods in clinical chemistry. *J Clin Chem Clin Biochem.* 1981;19:121–137.
13. Westgard JO, de Vos DJ, Hunt MR, et al. Concepts and practices in the evaluation of clinical chemistry methods: IV. Decisions of acceptability. *Am J Med Technol.* 1978;44: 727–742.
14. Levey S, Jennings ER. The use of control charts in the clinical laboratory. *Am J Clin Pathol.* 1950;20:1059–1066.
15. Westgard JO, Barry PL, Hunt MR, et al. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem.* 1981;27:493–501.
16. Clinical and Laboratory Standards Institute (CLSI). *Statistical Quality Control for Quantitative Measurements: Principles and Definitions*. Wayne, PA: CLSI; 2006.
17. Westgard JO, Groth T. Power functions for statistical control rules. *Clin Chem.* 1979;25:863–869.
18. Westgard JO, Groth T, Aronsson T, et al. Performance characteristics of rules for internal quality control: probabilities for false rejection and error detection. *Clin Chem.* 1977;23:1857–1867.
19. Clinical Laboratory Improvement Amendments of 1988; final rule. *Fed Regist*. 7164 [42 CFR 493.1213]: Department of Health and Human Services, Centers for Medicare and Medicaid Services; 1992.
20. van der Meulen EA, Boogaard PJ, van Sittert NJ. Use of small-sample-based reference limits on a group basis. *Clin Chem.* 1994;40:1698–1702.
21. Clinical and Laboratory Standards Institute (CLSI). *Defining, Establishing, and Verifying Reference Intervals in the Clinical Laboratory; Approved Guideline.* 3rd Edition. Wayne, PA: CLSI; 2010 (Document no. EP28-A3c).
22. Centers for Disease Control and Prevention (CDC), Centers for Medicare and Medicaid Services (CMS), Department of Health and Human Services (HHS). Medicare, Medicaid, and CLIA programs; laboratory requirements relating to quality systems and certain personnel qualifications. Final rule. *Fed Regist*. 2003;68:3639–3714.
23. Barnett RN. Medical significance of laboratory results. *Am J Clin Pathol.* 1968;50:671–676.
24. Fraser CG. Data on biological variation: essential prerequisites for introducing new procedures? *Clin Chem.* 1994;40: 1671–1673.
25. Fraser CG. *Biological Variation: From Principles to Practice*. Washington, DC: AACC Press; 2001.
26. Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. *Clin Chem.* 1963;9:217–233.
27. Centers for Medicare and Medicaid Services (CMS), Health and Human Services; Centers for Disease Control and Prevention (CDC), Department of Health and Human Services (HHS). Clinical Laboratory Improvement

Amendments of 1988 (CLIA) Proficiency Testing Regulations Related to Analytes and Acceptable Performance. Proposed rule. *Fed Regist*. 2019;84:1536–1567.

28. Ehrmeyer SS, Laessig RH, Leinweber JE, et al. 1990 Medicare/CLIA final rules for proficiency testing: minimum intralaboratory performance characteristics (CV and bias) needed to pass. *Clin Chem*. 1990;36: 1736–1740.
29. Westgard JO, Seehafer JJ, Barry PL. European specifications for imprecision and inaccuracy compared with operating specifications that assure the quality required by US CLIA proficiency-testing criteria. *Clin Chem*. 1994;40(7 pt 1): 1228–1232.
30. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307–310.

Analytic Techniques

Karen K. Apolloni

CHAPTER OUTLINE

KEY TERMS

Atomic absorption spectrophotometry
Chemiluminescence
Chromatography
Electrochemistry
Electrophoresis
Fluorescence
Fluorometry
Gas chromatography
High-performance liquid chromatography
Ion-selective electrodes
Mass spectrometry
Osmometry
Spectrophotometry

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to:

- Explain the general principles of each analytic method.
- Discuss the limitations of each analytic technique.
- Compare and contrast the various analytic techniques.
- State existing clinical applications for each analytic technique.
- Describe the operation and component parts of the following instruments: spectrophotometer, atomic absorption spectrometer, fluorometer, ion-selective electrode, pH electrode, osmometer, gas chromatograph, and mass spectrometer.
- Outline spectrophotometer quality assurance procedures.

A variety of analytic techniques are incorporated into instrumentation and are in widespread use in the modern clinical chemistry laboratory. The majority of analytic techniques fall into one of four basic disciplines within the field of analytic chemistry: spectrometry (including **spectrophotometry**, **atomic absorption spectrometry**, and **mass spectrometry** [MS]); luminescence (including **fluorescence** and **chemiluminescence**); electroanalytic methods (including **electrophoresis**, potentiometry, and amperometry); and **chromatography** (including gas, liquid, and thin layer).

Spectrophotometry

Instruments that measure electromagnetic radiation have several concepts and components in common. Shared instrumental components are discussed in some detail in a later section. Photometric instruments measure light intensity without consideration of wavelength. Most instruments today use filters (photometers), prisms, or gratings (spectrometers) to select (isolate) a narrow range of the incident wavelength. Radiant energy that passes through an object will be partially reflected, absorbed, and transmitted.

Electromagnetic radiation is described as photons of energy traveling in waves. The relationship between wavelength and energy E is described by Planck's formula:

$$E = h\nu \qquad \text{(Eq. 4.1)}$$

where h is a constant (6.62×10^{-27} erg sec), known as Planck constant, and ν is frequency. Because the frequency of a wave is inversely proportional to the wavelength, it follows that the energy of

CASE STUDY 4.1, PART 1

Remember Miles and Mía from Chapter 1? The laboratory is placing a spectrophotometer back in service after being in storage for 6 months. The instrument manuals are no longer available for this model. Miles and Mía, who manage quality control for the laboratory, are tasked with getting it ready for use.

1. What procedures should Miles and Mía develop to validate that the instrument is working properly for clinical use?

© dotshock/Shutterstock.

© Ariel Skelley/DigitalVision/Getty Images.

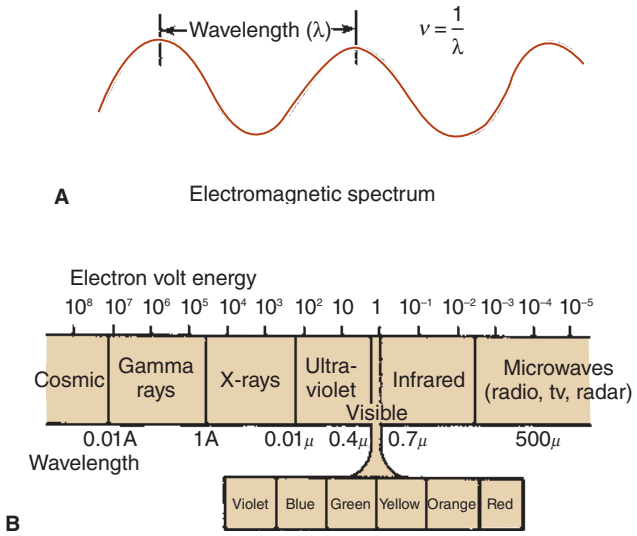


Figure 4.1 Electromagnetic radiation—relationship of energy and wavelength.

electromagnetic radiation is inversely proportional to wavelength. **Figure 4.1A** shows this relationship. Electromagnetic radiation includes a spectrum of energy from short-wavelength, highly energetic gamma rays and x-rays on the left in **Figure 4.1B** to long-wavelength radiofrequencies on the right. Visible light falls in between, with the color violet at 400 nm and red at 700 nm wavelengths being the approximate limits of the visible spectrum.

The instruments discussed in this section measure either absorption or emission of radiant energy to determine the concentration of atoms or molecules. The two phenomena, absorption and emission, are closely related. For a ray of electromagnetic radiation to be absorbed, it must have the same frequency as a rotational or vibrational frequency in the atom or molecule that it strikes. Levels of energy that are absorbed move in discrete steps, and any particular type of molecule or atom will absorb only certain energies and not others. When energy is absorbed, valence electrons move to an orbital with a higher energy level. Following energy absorption, the excited electron will fall back to the ground state by emitting a discrete amount of energy in the form of a characteristic wavelength of radiant energy.

Absorption or emission of energy by atoms results in a line spectrum. Because of the relative complexity of molecules, they absorb or emit a band of energy over a large region. Light emitted by incandescent solids (tungsten or deuterium) is in a continuum. The three types of spectra are shown in **Figure 4.2**.[1–3]

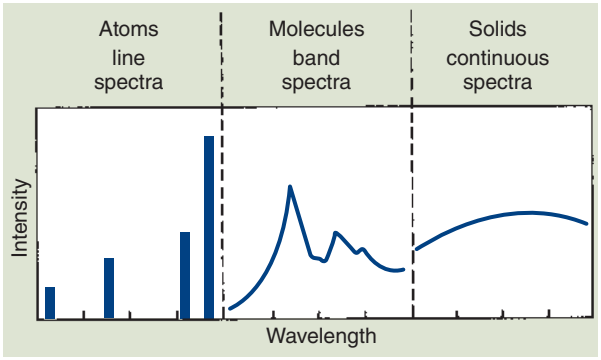


Figure 4.2 Characteristic absorption or emission spectra.

Data from Coiner D. *Basic Concepts in Laboratory Instrumentation.* Bethesda, MD: ASMT Education and Research Fund; 1975-1979.

Beer's Law

The relationship between absorption of light by a solution and the concentration of that solution has been described by Beer and others. Beer's law states that the concentration of a substance is directly proportional to the amount of light absorbed or inversely proportional to the logarithm of the transmitted light. Percent transmittance (%*T*) and absorbance (*A*) are related photometric terms that are explained in this section.

Figure 4.3A shows a beam of monochromatic light entering a solution. Some of the light is absorbed. The remainder passes through, strikes a light detector, and is converted to an electric signal. Percent transmittance is the ratio of the radiant energy transmitted (*T*) divided by the radiant energy incident on the sample (*I*). If all light is absorbed or blocked, %*T* is equal to zero. A level of 100% *T* is obtained if no light is absorbed. In practice, the solvent without the constituent of interest is placed in the light path, as in **Figure 4.3B**. Most of the light is transmitted, but a small amount is absorbed by the solvent and cuvette or is reflected away from the detector. The electrical readout of the instrument is set arbitrarily at 100% *T*, while the light is passing through a "blank" or reference. The sample containing absorbing molecules to be measured is then placed in the light path. The difference in amount of light transmitted by the blank and that transmitted by the sample is due only to the presence of the compound being measured. The %*T* measured by commercial spectrophotometers is the ratio of the sample transmitted beam divided by the blank transmitted beam, multiplied by 100.

Equal thicknesses of an absorbing material will absorb a constant fraction of the energy incident upon the layers. For example, in a tube containing layers of solution (**Figure 4.4A**), the first layer

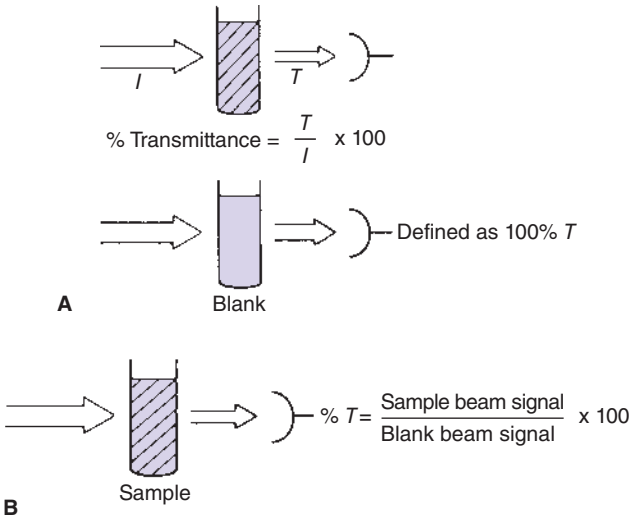


Figure 4.3 Percent transmittance (%*T*) defined.

transmits 70% of the light incident upon it. The second layer will, in turn, transmit 70% of the light incident upon it. Thus, 70% of 70% (49%) is transmitted by the second layer. The third layer transmits 70% of 49%, or 34% of the original light. Continuing on, successive layers transmit 24% and 17%, respectively. The %*T* values, when plotted on linear graph paper, yield the curve shown in **Figure 4.4B**. Considering each equal layer as many monomolecular layers, we can translate layers of material to concentration. If semi log graph paper is used to plot the same figures, a straight line is obtained (**Figure 4.4C**), indicating that, as concentration increases, %*T* decreases in a logarithmic manner.

Absorbance *A* is the amount of light absorbed. It cannot be measured directly by a spectrophotometer but rather is mathematically derived from %*T* as follows:

$$\%T = \frac{I}{I_0} \times 100 \qquad \textbf{(Eq. 4.2)}$$

where I_0 is the incident light and I is the transmitted light.

Absorbance is defined as follows:

$$A = -\log (I/I_0) = \log (100\%) - \log \%T$$
$$= 2 - \log \%T \qquad \textbf{(Eq. 4.3)}$$

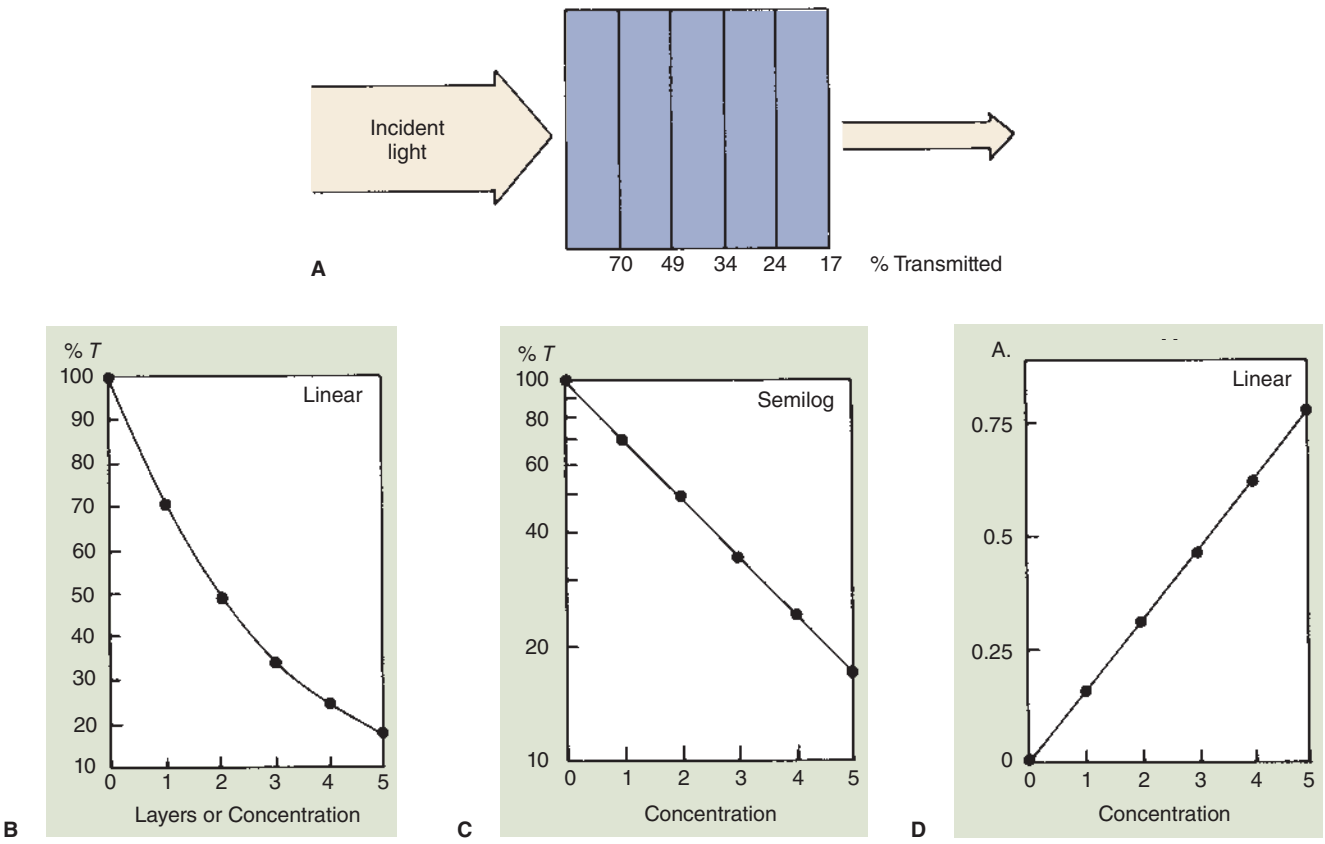


Figure 4.4 **(A)** Percent of original incident light transmitted by equal layers of light-absorbing solution. **(B)** Percent *T* versus concentration on linear graph paper. **(C)** Percent *T* versus concentration on semi log graph paper. **(D)** *A* versus concentration on linear graph paper.

According to Beer's law, absorbance is directly proportional to concentration (**Figure 4.4D**):

$$A = \varepsilon \times b \times c \qquad \text{(Eq. 4.4)}$$

where ε = molar absorptivity, the fraction of a specific wavelength of light absorbed by a given type of molecule; b is the length of light path through the solution; and c is the concentration of absorbing molecules.

Absorptivity depends on the molecular structure and the way in which the absorbing molecules react with different energies. For any particular molecular type, absorptivity changes as wavelength of radiation changes. The amount of light absorbed at a particular wavelength depends on the molecular and ion types present and may vary with concentration, pH, or temperature.

Because the path length and molar absorptivity are constant for a given wavelength, absorbance is directly proportional to concentration.

$$A \sim c$$

Unknown concentrations are determined from a calibration curve that plots absorbance at a specific wavelength versus concentration for standards of known concentration. For calibration curves that are linear and have a zero y-intercept, unknown concentrations can be determined from a single calibrator. Not all calibration curves result in straight lines. Deviations from linearity are typically observed at high absorbances. The stray light within an instrument will ultimately limit the maximum absorbance that a spectrophotometer can achieve, typically 2.0 absorbance units.

Spectrophotometric Instruments

A spectrophotometer is used to measure the light transmitted by a solution to determine the concentration of the light-absorbing substance in the solution. **Figure 4.5** illustrates the basic components of a single-beam spectrophotometer, which are described in subsequent sections.

Components of a Spectrophotometer

Light Source

The most common source of light for work in the visible and near-infrared regions is the incandescent tungsten or tungsten-iodide lamp. Only about 15% of radiant energy emitted falls in the visible region, with most emitted as near-infrared.[1–3] Often, a heat-absorbing filter is inserted between the lamp and the sample to absorb the infrared radiation.

The lamps usually used for ultraviolet (UV) work are the deuterium discharge lamp and the mercury arc lamp. Deuterium provides continuous emission down to 165 nm. Low-pressure mercury lamps emit a sharp line spectrum, with both UV and visible lines. Medium- and high-pressure mercury lamps emit a continuum from UV to the mid-visible region. The most important factors for a light source are range, spectral distribution within the range, the source of radiant production, stability of the radiant energy, and temperature.

Monochromators

Isolation of individual wavelengths of light is an important and necessary function of a monochromator. The degree of wavelength isolation is a function of the type of device used and the width of entrance and exit slits. The band-pass of a monochromator defines the range of wavelengths transmitted and is calculated as width at half the maximum transmittance (**Figure 4.6**).

Numerous devices are used for obtaining monochromatic light. The least expensive are colored glass filters. These filters usually pass a relatively wide band of radiant energy and have a low transmittance of the selected wavelength. Although not precise, they are simple and inexpensive.

Interference filters produce monochromatic light based on the principle of constructive interference of waves. Two pieces of glass, each mirrored on one side, are separated by a transparent spacer that is precisely one-half the desired wavelength. Light waves

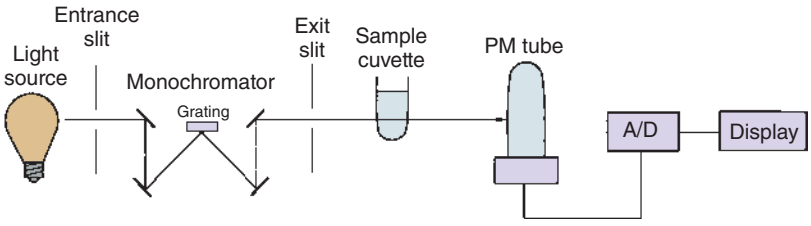


Figure 4.5 Single-beam spectrophotometer.

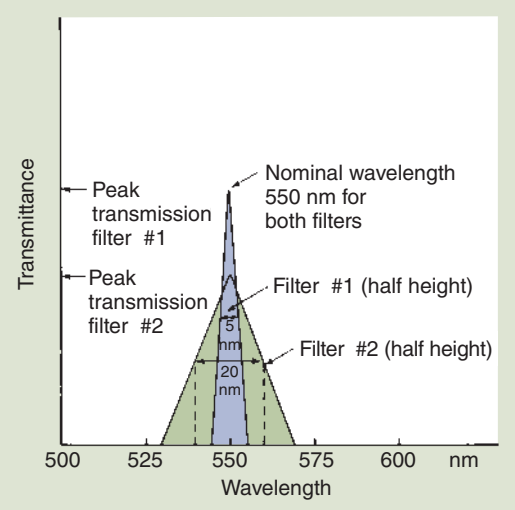


Figure 4.6 Spectral transmittance of two monochromators with band pass at half height of 5 and 20 nm.

enter one side of the filter and are reflected at the second surface. Wavelengths that are twice the space between the two glass surfaces will reflect back and forth, reinforcing others of the same wavelengths and finally passing through. Other wavelengths will cancel out because of phase differences (destructive interference). Because interference filters also transmit multiples of the desired wavelengths, they require accessory filters to eliminate these harmonic wavelengths. Interference filters can be constructed to pass a very narrow range of wavelengths with good efficiency.

The simple glass prism is another type of monochromator. A narrow beam of light focused on a prism is refracted as it enters the denser glass. Short wavelengths are refracted more than long wavelengths, resulting in dispersion of white light into a continuous spectrum. The prism can be rotated, allowing only the desired wavelength to pass through an exit slit.

Diffraction gratings are commonly used as monochromators. A diffraction grating consists of many parallel grooves (15,000 or 30,000 per inch) etched onto a polished surface. Diffraction, the separation of light into component wavelengths, is based on the principle that wavelengths bend as they pass a sharp corner. The degree of bending depends on the wavelength. As the wavelengths move past the corners, wave fronts are formed. Those that are in phase reinforce one another, whereas those not in phase cancel out and disappear. This results in complete spectra. Gratings with very fine line rulings produce a widely dispersed spectrum. They produce linear spectra, called orders, in both directions from the entrance slit. Because the multiple spectra have a tendency to cause stray light problems, accessory filters are used.

Sample Cell

The next component of the basic spectrophotometer is the sample cell or cuvette, which typically has a flat surface. The light path must be kept constant to have absorbance proportional to concentration. This is easily checked by preparing a colored solution to read midscale when using the wavelength of maximum absorption. Each cuvette to be tested is filled, readings are taken, and results are compared against an acceptable tolerance (e.g., ±0.25% *T*). Cuvettes are sold in matched sets. Square cuvettes have plane-parallel optical surfaces and a constant light path. Cuvettes with scratched optical surfaces scatter light and should be discarded. Inexpensive glass cuvettes can be used for applications in the visible range, but they absorb light in the UV region. Quartz cuvettes enable transmission of light and are used when substances absorb in this region (e.g., NADH at 340 nm).

Photodetectors

The purpose of the detector is to convert the transmitted radiant energy into an equivalent amount of electrical energy. The least expensive of the devices is known as a barrier-layer cell, or photocell. The photocell is composed of a film of light-sensitive material, frequently selenium, on a plate of iron. A thin, transparent layer of silver overlays the light-sensitive material. When exposed to light, electrons in the light-sensitive material are excited and released to flow to the highly conductive silver. In comparison with the silver, a moderate resistance opposes the electron flow toward the iron, forming a hypothetical barrier to flow in that direction. Consequently, this cell generates its own electromotive force, which can be measured. The current produced is proportional to the incident radiation. Photocells require no external voltage source but rely on internal electron transfer to produce a current in an external circuit. Because of their low internal resistance, the output of electrical energy is not easily amplified. Consequently, this type of detector is used mainly in filter photometers with a wide bandpass, producing a fairly high level of illumination so that there is no need to amplify the signal. The photocell is inexpensive and durable; however, it is temperature sensitive and nonlinear at very low and very high levels of illumination.

A phototube (**Figure 4.7**) is similar to a photocell in that it has photosensitive material that gives off electrons when light energy strikes it. It differs

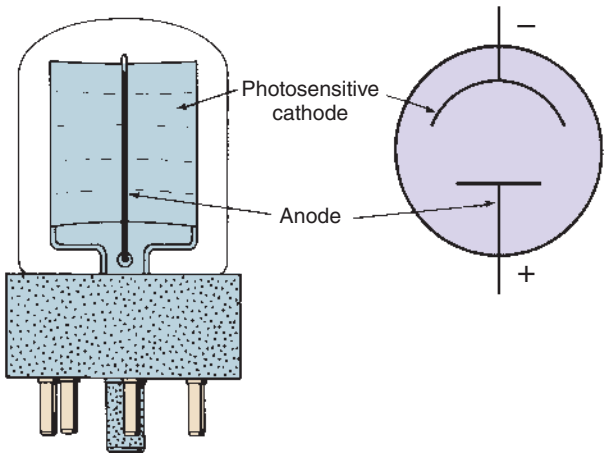


Figure 4.7 Phototube drawing and schematic.

in that an outside voltage is required for operation. Phototubes contain a negatively charged cathode and a positively charged anode enclosed in a glass case. The cathode is composed of a material (e.g., rubidium or lithium) that acts as a resistor in the dark but emits electrons when exposed to light. The emitted electrons jump over to the positively charged anode, where they are collected and return through an external, measurable circuit. The cathode usually has a large surface area. Varying the cathode material changes the wavelength at which the phototube gives its highest response. The photocurrent is linear with the intensity of the light striking the cathode as long as the voltage between the cathode and the anode remains constant. A vacuum within the tubes avoids scattering of the photoelectrons by collision with gas molecules.

The third major type of light detector is the photomultiplier (PM) tube, which detects and amplifies radiant energy. As shown in **Figure 4.8**, incident light strikes the coated cathode, emitting electrons. The electrons are attracted to a series of anodes, known as dynodes, each having a successively higher positive voltage. These dynodes are made of a material that gives off many secondary electrons when hit by single electrons. Initial electron emission at the cathode triggers a multiple cascade of electrons within the PM tube itself. Because of this amplification, the PM tube is 200 times more sensitive than the phototube. PM tubes are used in instruments designed to be extremely sensitive to very low light levels and light flashes of very short duration. The accumulation of electrons striking the anode produces a current signal, measured in amperes, that is proportional to the initial intensity of the light. The analog signal is converted first to a voltage and then to a digital signal through the use of an analog-to-digital converter. Digital signals are processed electronically to produce absorbance readings.

In a photodiode, absorption of radiant energy by a reverse-biased PN junction diode (PN: positive–negative) produces a photocurrent that is proportional to the incident radiant power. Although photodiodes are not as sensitive as PM tubes because of the lack of internal amplification, their excellent linearity, speed, and small size make them useful in applications where light levels are adequate.[4] Photodiode array (PDA) detectors are available in integrated circuits containing 256 to 2048 photodiodes in a linear arrangement. A linear array is shown in **Figure 4.9**. Each photodiode responds to a specific

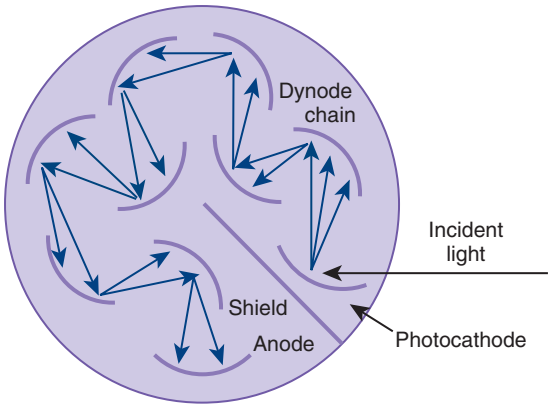


Figure 4.8 Dynode chain in a photomultiplier.

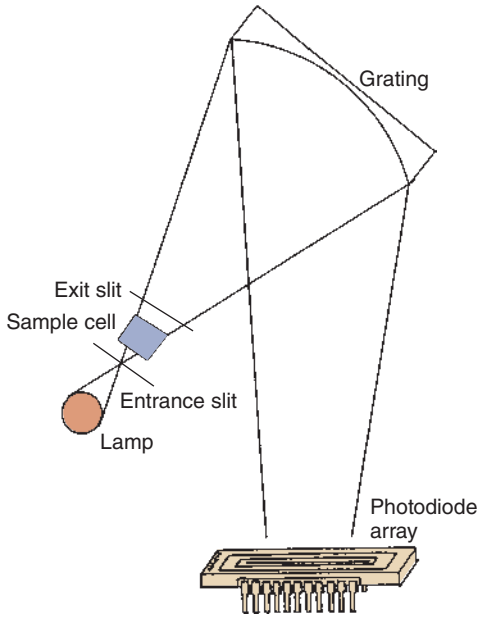


Figure 4.9 Photodiode array spectrophotometer illustrating the placement of the sample cuvette before the monochromator.

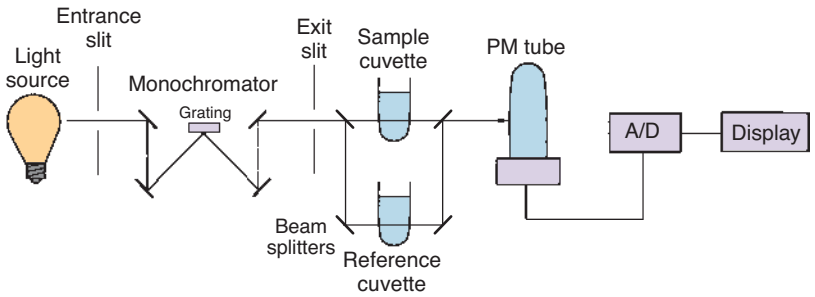


Figure 4.10 Double-beam spectrophotometer.

wavelength, and as a result, a complete UV/visible spectrum can be obtained in less than 1 second. Resolution is 1 to 2 nm and depends on the number of discrete elements. In spectrophotometers using PDA detectors, the grating is positioned after the sample cuvette and disperses the transmitted radiation onto the PDA detector (Figure 4.9).

For single-beam spectrophotometers, the absorbance reading from the sample must be blanked using an appropriate reference solution that does not contain the compound of interest. Double-beam spectrophotometers permit automatic correction of sample and reference absorbance, as shown in **Figure 4.10**. Because the intensities of light sources vary as a function of wavelength, double-beam spectrophotometers are necessary when the absorption spectrum for a sample is to be obtained. Computerized, continuous zeroing, single-beam spectrophotometers have replaced most double-beam spectrophotometers.

Spectrophotometer Quality Assurance

Performing at least the following checks should validate instrument function: wavelength accuracy, stray light, and linearity. Wavelength accuracy means that the wavelength indicated on the control dial is the actual wavelength of light passed by the monochromator. It is most commonly checked using standard absorbing solutions or filters with absorbance maxima of known wavelength. Didymium or holmium oxide in glass is stable and frequently used as filters. The filter is placed in the light path, and the wavelength control is set at the wavelength at which maximal absorbance is expected. The wavelength control is then rotated in either direction to locate the actual wavelength that has maximal absorbance. If these two wavelengths do not match, the optics must be adjusted to calibrate the monochromator correctly.

Some instruments with narrow bandpass use a mercury vapor lamp to verify wavelength accuracy. The mercury lamp is substituted for the usual light source, and the spectrum is scanned to locate mercury emission lines. The wavelength indicated on the control is compared with known mercury emission peaks to determine the accuracy of the wavelength indicator control.

Stray light refers to any wavelengths outside the band transmitted by the monochromator. **Figure 4.11** shows the performance of a spectrophotometer to measure high absorbance in the presence of stray light. The most common causes of stray light are reflection of light from scratches on optical surfaces or from dust particles anywhere in the light path and higher order spectra produced by diffraction gratings. The major effect is absorbance error, especially in the high absorbance range. Stray light is detected by using cutoff filters, which eliminate all radiation at wavelengths beyond the one of interest. To check for

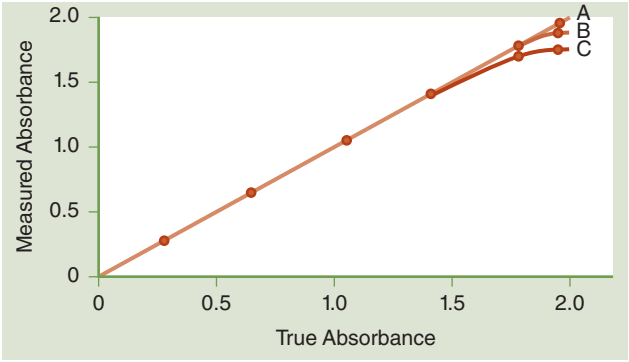


Figure 4.11 Spectrophotometer's ability to measure high absorbance with stray light. (**A**) No stray light, with no deviation from the actual absorbance. (**B**) Some stray light within the instrument showing deviations from the actual at high absorbance. (**C**) A higher degree of stray light showing further deviation from the actual absorbance.

stray light in the near-UV region, for example, a filter that does not transmit in the region of 200 to 400 nm is inserted. If the instrument reading is greater than 0% *T*, stray light is present. Certain liquids, such as $NiSO_4$, $NaNO_2$, and acetone, absorb strongly at short wavelengths and can be used to detect stray light in the UV range.

Linearity is demonstrated when a change in concentration results in a straight-line calibration curve, as discussed under Beer's law. Colored solutions may be carefully diluted and used to check linearity, using the wavelength of maximal absorbance for that color. Sealed sets of different colors and concentrations are available commercially. They should be labeled with expected absorbance for a given bandpass instrument. Less than expected absorbance is an indication of stray light or of a bandpass that is wider than specified. Sets of neutral-density filters to check linearity over a range of wavelengths are also commercially available.

Atomic Absorption Spectrophotometry

The atomic absorption spectrophotometer is used to measure concentration by detecting the absorption of electromagnetic radiation by atoms rather than by molecules. The basic components are shown in **Figure 4.12**. The usual light source, known as a hollow-cathode lamp, consists of an evacuated gas-tight chamber containing an anode, a cylindrical cathode, and an inert gas, such as helium or argon. When voltage is applied, the filler gas is ionized. Ions attracted to the cathode collide with the metal, knock atoms off, and cause the metal atoms to be excited. When they return to the ground state, light energy is emitted that is characteristic of the metal in the cathode. Generally, a separate lamp is required for each metal (e.g., a copper hollow cathode lamp is used to measure this metal).

Electrodeless discharge lamps are a relatively new light source for atomic absorption spectrophotometers. A bulb is filled with argon and the element to be tested. A radiofrequency generator around the bulb supplies the energy to excite the element, causing a characteristic emission spectrum of the element.

The analyzed sample must contain the reduced metal in the atomic vaporized state. Commonly, this is done by using the heat of a flame to break the chemical bonds and form free, unexcited atoms. The flame serves as the sample cell in this instrument, instead of a cuvette. There are various designs; however, the most common burner is the premix long-path burner. The sample, in solution, is aspirated as a spray into a chamber, where it is mixed with air and fuel. This mixture passes through baffles, where large drops fall and are drained off. Only fine droplets reach the flame. The burner is a long, narrow slit, to permit a longer path length for absorption of incident radiation. Light from the hollow-cathode lamp passes through the sample of ground state atoms in the flame. The amount of light absorbed is proportional to the concentration. When a ground state atom absorbs light energy, an excited atom is produced. The excited atom then returns to the

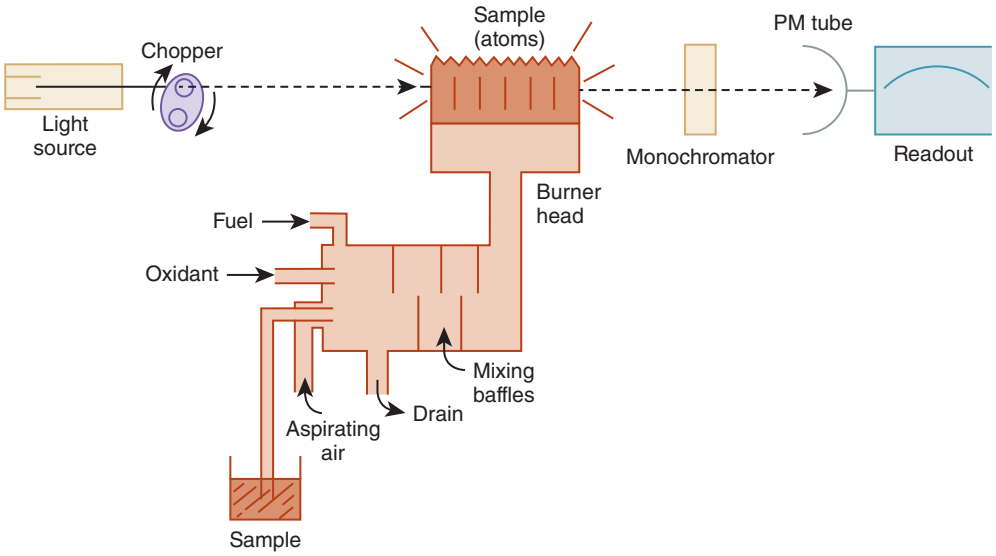


Figure 4.12 Single-beam atomic absorption spectrophotometer—basic components.

ground state, emitting light of the same energy as it absorbed. The flame sample thus contains a dynamic population of ground state and excited atoms, both absorbing and emitting radiant energy. The emitted energy from the flame will go in all directions, and it will be a steady emission. Because the purpose of the instrument is to measure the amount of light absorbed, the light detector must be able to distinguish between the light beam emitted by the hollow-cathode lamp and that emitted by excited atoms in the flame. To do this, the hollow cathode light beam is modulated by inserting a mechanical rotating chopper between the light and the flame or by pulsing the electric supply to the lamp. Because the light beam being absorbed enters the sample in pulses, the transmitted light will also be in pulses. There will be less light in the transmitted pulses because part of it will be absorbed. There are, therefore, two light signals from the flame—an alternating signal from the hollow-cathode lamp and a direct signal from the flame emission. The measuring circuit is tuned to the modulated frequency. Interference from the constant flame emission is electronically eliminated by accepting only the pulsed signal from the hollow cathode.

The monochromator is used to isolate the desired emission line from other lamp emission lines. In addition, it serves to protect the photodetector from excessive light emanating from flame emissions. A PM tube is the usual light detector.

Flameless atomic absorption requires an instrument modification that uses an electric furnace to break chemical bonds (electrothermal atomization). A tiny graphite cylinder holds the sample, either liquid or solid. An electric current passes through the cylinder walls, evaporates the solvent, ashes (heats at a high temperature to leave an ash residue for analysis) the sample, and, finally, heats the unit to incandescence to atomize the sample. This instrument, like the spectrophotometer, is used to determine the amount of light absorbed. Again, Beer's law is used for calculating concentration. A major problem is that background correction is much more necessary and critical for electrothermal techniques than for flame-based atomic absorption methods. Currently, the most common approach uses a deuterium lamp as a secondary source and measures the difference between the two absorbance signals. However, there has also been extensive development of background correction techniques based on the Zeeman effect.[1] The presence of an intense static magnetic field will cause the wavelength of the emitted radiation to split into several components. This shift in wavelength is the Zeeman effect.

Atomic absorption spectrophotometry is sensitive and precise. It is routinely used to measure concentration of trace metals that are not easily excited. It is accurate, precise, and specific. One disadvantage, however, is the inability of the flame to dissociate samples into free atoms. For example, phosphate may interfere with calcium analysis by formation of calcium phosphate. This may be overcome by adding cations that compete with calcium for phosphate. Routinely, lanthanum or strontium is added to samples to form stable complexes with phosphate. Another possible problem is the ionization of atoms following dissociation by the flame, which can be decreased by reducing the flame temperature. Matrix interference, due to the enhancement of light absorption by atoms in organic solvents or formation of solid droplets as the solvent evaporates in the flame, can be another source of error. This interference may be overcome by pretreatment of the sample by extraction.

Recently, inductively coupled plasma (ICP) has been used to increase sensitivity for atomic emission. The torch, an argon plasma maintained by the interaction of a radiofrequency field and an ionized argon gas, is reported to have used temperatures between 5500 and 8000 K. Complete atomization of elements is thought to occur at these temperatures. Use of ICP as a source is recommended for determinations involving refractory elements such as uranium, zirconium, and boron. ICP with MS detection is the most sensitive and specific assay technique for all elements on the periodic chart. Atomic absorption spectrophotometry is used less frequently because of this newer technology.

Fluorometry

As seen with the spectrophotometer, light entering a solution may pass mainly through or may be absorbed partly or entirely, depending on the concentration and the wavelength entering that particular solution. Whenever absorption occurs, there is a transfer of energy to the medium. Each molecular type possesses a series of electronic energy levels and can pass from a lower energy level to a higher energy level only by absorbing an integral unit (quantum) of light that is equal in energy to the difference between the two energy states. There are additional energy levels owing to rotation or vibration of molecular parts. The excited state lasts about 10^{-5} seconds before the electron loses energy and returns to the ground state. Energy is lost by collision, heat loss, transfer to other molecules, and emission of radiant energy. Because the molecules

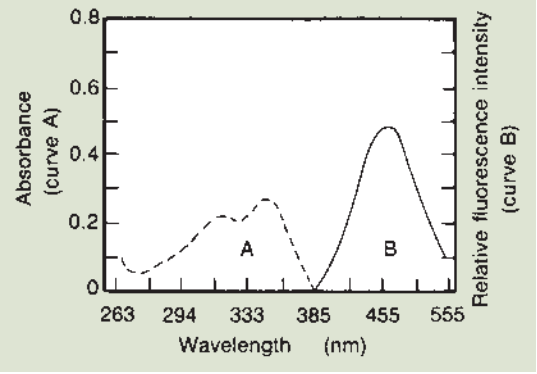


Figure 4.13 Absorption and fluorescence spectra of quinine in 0.1 N sulfuric acid.

Data from Coiner D. *Basic Concepts in Laboratory Instrumentation.* Bethesda, MD: ASMT Education and Research Fund; 1975–1979.

are excited by absorption of radiant energy and lose energy by multiple interactions, the radiant energy emitted is less than the absorbed energy. The difference between the maximum wavelengths, excitation, and emitted fluorescence is called Stokes shift. Both excitation (absorption) and fluorescence (emission) energies are characteristic for a given molecular type; for example, **Figure 4.13** shows the absorption and fluorescence spectra of quinine in sulfuric acid. The dashed line on the left shows the short-wavelength excitation energy that is maximally absorbed, whereas the solid line on the right is the longer wavelength (less energy) fluorescent spectrum.

Fluorometry Instrumentation

Filter fluorometers measure the concentrations of solutions that contain fluorescing molecules. A basic instrument is shown in **Figure 4.14**. The source emits short-wavelength high-energy excitation light. A mechanical attenuator controls light intensity. The primary filter, placed between the radiation source and the sample, selects the wavelength that is best absorbed by the solution to be measured. The fluorescing sample in the cuvette emits radiant energy in all directions. The detector (placed at right angles to the sample cell) and a secondary filter that passes the longer wavelengths of fluorescent light prevent incident light from striking the photodetector. The electrical output of the photodetector is proportional to the intensity of fluorescent energy. In spectrofluorometers, the filters are replaced by a grating monochromator.

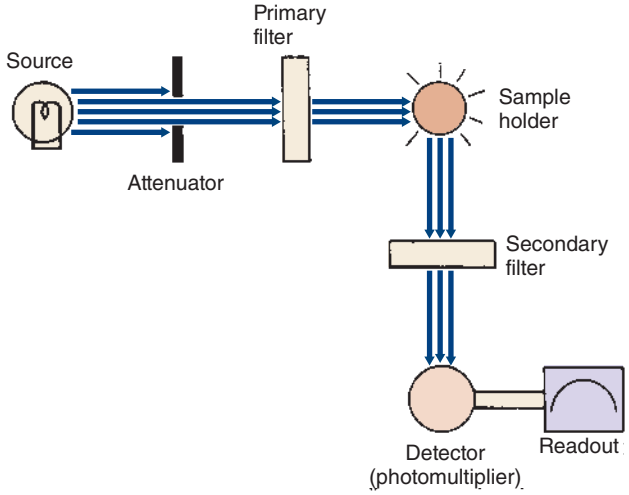


Figure 4.14 Basic filter fluorometer.

Data from Coiner D. *Basic Concepts in Laboratory Instrumentation.* Bethesda, MD: ASMT Education and Research Fund; 1975–1979.

Gas discharge lamps (mercury and xenon arc) are the most frequently used sources of excitation radiant energy. Incandescent tungsten lamps are seldom used because they release little energy in the UV region. Mercury vapor lamps are commonly used in filter fluorometers. Mercury emits a characteristic line spectrum. Resonance lines at 365 to 366 nm are commonly used. Energy at wavelengths other than the resonance lines is provided by coating the inner surface of the lamp with a material that absorbs the 254-nm mercury radiation and emits a broad band of longer wavelengths. Most spectrofluorometers use a high-pressure xenon lamp. These lamps produce a nearly continuous spectrum of wavelengths.

Monochromator fluorometers use gratings for isolation of incident radiation. Light detectors are almost exclusively PM tubes because of their higher sensitivity to low light intensities. Double-beam instruments are used to compensate for instability due to electric power fluctuation.

Fluorescence concentration measurements are related to molar absorptivity of the compound, intensity of the incident radiation, quantum efficiency of the energy emitted per quantum absorbed, and length of the light path. In dilute solutions with instrument parameters held constant, fluorescence is directly proportional to concentration. Generally, a linear response will be obtained until the concentration of the fluorescent species is so high that the sample begins to absorb significant amounts of excitation light. A curve demonstrating nonlinearity as concentration increases is shown in **Figure 4.15**. The solution must absorb less than 5% of the exciting radiation for a linear response to occur.[5] As with all quantitative measurements, a standard curve must be prepared to demonstrate that the concentration used falls in a linear range.

Fluorescence Polarization

In fluorescence polarization, radiant energy is polarized in a single plane. When the sample fluorophore is excited, it emits polarized light along the same plane

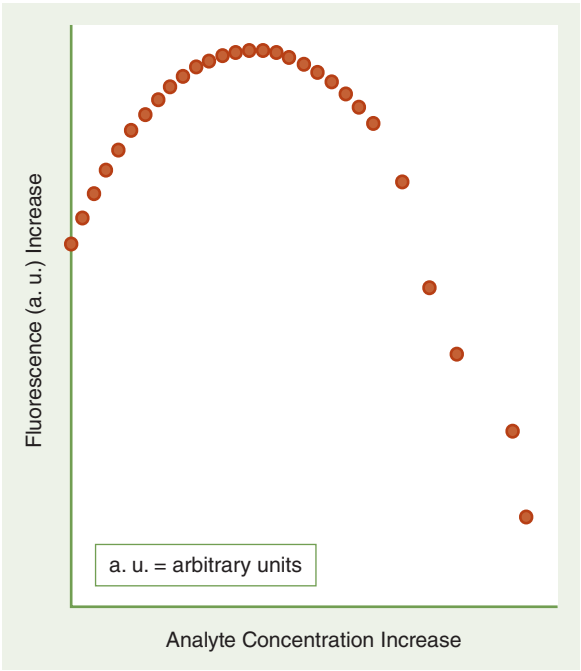


Figure 4.15 Dependence of fluorescence on the concentration of fluorophore.

Data from Guilbault GG. *Practical Fluorescence, Theory, Methods and Techniques.* New York, NY: Marcel Dekker; 1973.

as the incident light if the fluorophore does not rotate in solution (i.e., if it is attached to a large molecule). In contrast, a small molecule emits depolarized light because it will rotate out of the plane of polarization during its excitation lifetime. This technique is widely used for the detection of therapeutic and abused drugs. In the procedure, the sample analyte is allowed to compete with a fluorophore-labeled analyte for a limited antibody to the analyte. The lower the concentration of the sample analyte, the higher the concentration of macromolecular antibody–analyte–fluorophore formed and the lower the depolarization of the radiant light.

Advantages and Disadvantages of Fluorometry

Fluorometry has two advantages over conventional spectrophotometry: specificity and sensitivity. Fluorometry increases specificity by selecting the optimal wavelength for both absorption and fluorescence, rather than just the absorption wavelength seen with spectrophotometry.

Fluorometry is approximately 1000 times more sensitive than most spectrophotometric methods.[5] One reason is because the emitted radiation is measured directly; it can be increased simply by increasing the intensity of the exciting radiant energy. In addition, fluorescence measures the amount of light intensity present over a zero background. In absorbance, however, the quantity of the absorbed light is measured indirectly as the difference between the transmitted beams. At low concentrations, the small difference between 100% *T* and the transmitted beam is difficult to measure accurately and precisely, limiting the sensitivity.

The biggest disadvantage is that fluorescence is very sensitive to environmental changes. Changes in pH affect availability of electrons, and temperature changes the probability of loss of energy by collision rather than fluorescence. Contaminating chemicals or a change of solvents may change the structure. UV light used for excitation can cause photochemical changes. Any decrease in fluorescence resulting from any of these possibilities is known as quenching. Because so many factors may change the intensity or spectra of fluorescence, extreme care is required in analytic technique and instrument maintenance.

Chemiluminescence

In *chemiluminescence* reactions, part of the chemical energy generated produces excited intermediates that decay to a ground state with the emission of photons.[6] The emitted radiation is measured with a PM tube, and the signal is related to analyte concentration. Chemiluminescence is different from fluorescence in that no excitation radiation is required and no monochromators are needed because the chemiluminescence arises from one species. Most importantly, chemiluminescence reactions are oxidation reactions of luminol, acridinium esters, and dioxetanes characterized by a rapid increase in intensity of emitted light followed by a gradual decay. Usually, the signal is taken as the integral of the entire peak. Enhanced chemiluminescence techniques increase the chemiluminescence efficiency by including an enhancer system in the reaction of a chemiluminescent agent with an enzyme. The time course for the light intensity is much longer (60 minutes) than that for conventional chemiluminescent reactions, which last for about 30 seconds (**Figure 4.16**).

Advantages of chemiluminescence assays include subpicomolar detection limits, speed (with flash-type reactions, light is only measured for 10 seconds), ease of use (most assays are one-step procedures), and simple instrumentation.[6] The main disadvantage is that impurities can cause a background signal that degrades the sensitivity and specificity.

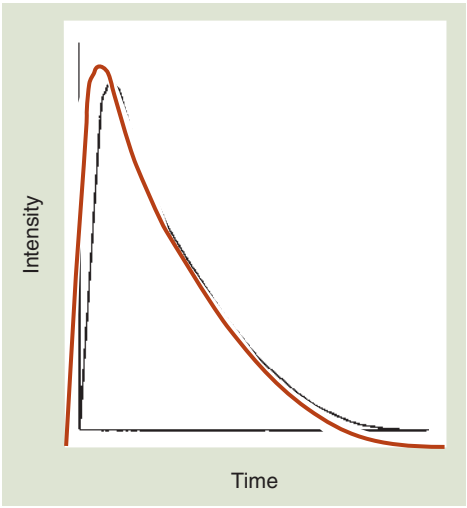


Figure 4.16 Representative intensity-versus-time curve for a transient chemiluminescence signal.

Turbidimetry

Turbidimetric measurements are made with a spectrophotometer to determine the concentration of particulate matter in a sample. The decrease in amount of light transmitted is measured. The amount of light blocked by a suspension of particles at 180° depends not only on concentration but also on particle size. Because particles tend to aggregate and settle out of suspension, sample handling is critical for accurate measurement. Instrument operation is the same as for any spectrophotometer.

Nephelometry

Nephelometry is similar to turbidometry, except that light scattered by the small particles is measured at an angle to the beam incident on the cuvette, instead of at 180°. The amount of scattered light is proportional to the concentration of the analyte. **Figure 4.17** demonstrates two possible optical arrangements for a nephelometer. Light scattering depends on wavelength and particle size. For macromolecules with a size close to or larger than the wavelength of incident light, sensitivity is increased by measuring the forward light scatter.[7] Instruments are available with detectors placed at various forward angles, as well as at 90° to the incident light. Monochromatic light obtains uniform scatter and minimizes sample heating. Certain instruments use lasers as the source of monochromatic light; however, any monochromator may be used. The Siemens BN II analyzer is an example of one instrument that utilizes the principle of nephelometry.

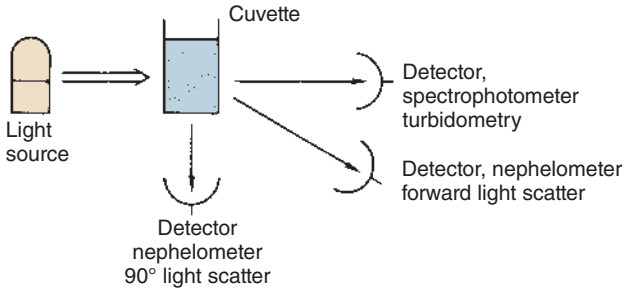


Figure 4.17 Nephelometer versus spectrophotometer—optical arrangements.

Measuring light scatter at an angle other than 180° minimizes error from colored solutions and increases sensitivity. Because both methods depend on particle size, some instruments quantitate initial change in light scatter rather than total scatter. Reagents must be free of any particles, and cuvettes must be free of scratches.

Laser Applications

Laser (light amplification by stimulated emission of radiation) is based on the interaction of radiant energy with suitably excited atoms or molecules. The wavelength, direction of propagation, phase, and plane of polarization of the emitted light are the same as those of the incident radiation. Laser light is polarized and coherent and has narrow spectral width and small cross-sectional area with low divergence. The radiant emission can be very powerful and either continuous or pulsating.

Laser light can serve as the source of incident energy in a spectrometer or nephelometer. Some lasers produce bandwidths of a few kilohertz in both the visible and infrared regions, making these applications about three to six orders more sensitive than conventional spectrometers.[8]

Laser spectrometry can also be used for the determination of structure and identification of samples, as well as for diagnosis. Quantitation of samples depends on the spectrometer used. An example of the clinical application of laser is the hematology and flow cytometer analyzers for the differential analysis of white blood cells.[9]

Electrochemistry

Electrochemistry is the basis for many types of analyses used in the clinical laboratory, including potentiometry, amperometry, coulometry, and polarography. The two basic types of electrochemical

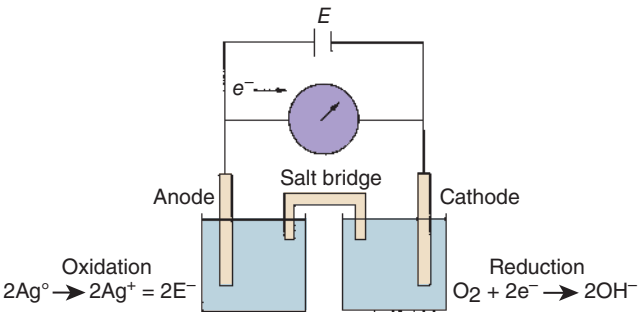


Figure 4.18 Electrochemical cell.

cells involved in these analyses are galvanic and electrolytic cells.

Galvanic and Electrolytic Cells

An electrochemical cell consists of two half-cells and a salt bridge, which can be a liquid or a piece of filter paper saturated with electrolytes, as shown in **Figure 4.18**. Each half-cell contains one electrode, either an anode or a cathode. Instead of two beakers as shown, the electrodes can be immersed in a single, large beaker containing a salt solution. In such a setup, the solution serves as the salt bridge.

In a galvanic cell, as the electrodes are connected, there is spontaneous flow of electrons from the electrode with the lower electron affinity (oxidation). These electrons pass through the external meter to the cathode (reduction), where OH^- ions are liberated. This reaction continues until one of the chemical components is depleted, at which point, the cell is "dead" and cannot produce electrical energy to the external meter.

Current may be forced to flow through the dead cell only by applying an external electromotive force *E*. This is called an electrolytic cell. In short, a galvanic cell can be built from an electrolytic cell. When the external *E* is turned off, accumulated products at the electrodes will spontaneously produce current in the opposite direction of the electrolytic cell.

Half-Cells

It is impossible to measure the electrochemical activity of one half-cell; two reactions must be coupled and one reaction compared with the other. To rate half-cell reactions, a specific electrode reaction is arbitrarily assigned 0.00 V. Every other reaction coupled with this arbitrary zero reaction is either positive or negative, depending on the relative affinity for electrons. The electrode defined as 0.00 V is the standard hydrogen electrode: H_2 gas at 1 atmosphere (atm).

Table 4.1 Standard Reduction Potentials

	Potential, *V*
$Zn^{2+} + 2e \leftrightarrow Z$	−0.7628
$Cr^{2+} + 2e \leftrightarrow Cr$	−0.913
$Ni^{2+} + 2e \leftrightarrow Ni$	−0.257
$2H^+ + 2e \leftrightarrow H_2$	0.000
$Cu^{2+} + 2e \leftrightarrow Cu$	0.3419
$Ag^+ + e \leftrightarrow Ag$	0.7996

Data presented are examples from Lide DR. *CRC Handbook of Chemistry and Physics*. 93rd ed. Boca Raton, FL: CRC Press; 2012–2013.

The hydrogen gas in contact with H^+ in solution develops a potential. The hydrogen electrode coupled with a zinc half-cell is cathodic, with the reaction $2H^+ + 2e^- \rightarrow H_2$, because H_2 has a greater affinity than does Zn for electrons. Cu, however, has a greater affinity than H_2 for electrons, and thus the anodic reaction $H_2 \rightarrow 2H^+ + 2e^-$ occurs when coupled to the Cu-electrode half-cell.

The potential generated by the hydrogen-gas electrode is used to rate the electrode potential of metals in 1 mol/L solution. Reduction potentials for certain metals are shown in **Table 4.1**.[10] A hydrogen electrode is used to determine the accuracy of reference and indicator electrodes, the stability of standard solutions, and the potentials of liquid junctions.

Ion-Selective Electrodes

Potentiometric methods of analysis involve the direct measurement of electrical potential due to the activity of free ions. **Ion-selective electrodes** (ISEs) are designed to be sensitive toward individual ions.

pH Electrodes

An ISE universally used in the clinical laboratory is the pH electrode. The basic components of a pH meter are shown in **Figure 4.19**.

Indicator Electrode

The pH electrode consists of a silver wire coated with AgCl, immersed into an internal solution of 0.1 mmol/L HCl, and placed into a tube containing a special glass membrane tip. This membrane is only sensitive to hydrogen ions (H^+). Glass membranes that are selectively sensitive to H^+ consist of specific quantities of lithium, cesium, lanthanum, barium,

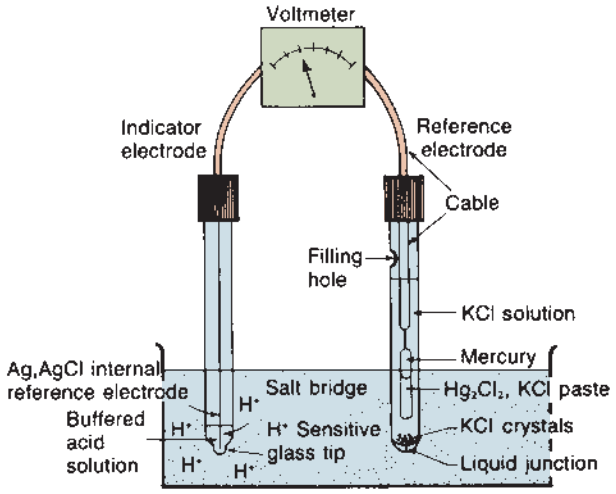


Figure 4.19 Necessary components of a pH meter.

or aluminum oxides in silicate. When the pH electrode is placed into the test solution, movement of H^+ near the tip of the electrode produces a potential difference between the internal solution and the test solution, which is measured as pH and read by a voltmeter. The combination pH electrode also contains a built-in reference electrode, either Ag/AgCl or calomel (Hg/Hg_2Cl_2) immersed in a solution of saturated KCl.

The specially formulated glass continually dissolves from the surface. The present concept of the selective mechanism that causes the formation of electromotive force at the glass surface is that an ion-exchange process is involved. Cationic exchange occurs only in the gel layer. There is no penetration of H^+ through the glass. Although the glass is constantly dissolving, the process is slow, and the glass tip generally lasts for several years. pH electrodes are highly selective for H^+; however, other cations in high concentration interfere, the most common of which is sodium. Electrode manufacturers should list the concentration of interfering cations that may cause error in pH determinations.

Reference Electrode

The reference electrode commonly used is the calomel electrode. Calomel, a paste of predominantly mercurous chloride, is in direct contact with metallic mercury in an electrolyte solution of potassium chloride. As long as the electrolyte concentration and the temperature remain constant, a stable voltage is generated at the interface of the mercury and its salt (mercurous chloride). A cable connected to the mercury leads to the voltmeter. The filling hole is needed for adding potassium chloride solution. A tiny opening at the bottom is required for completion of electric contact between the reference and indicator electrodes. The liquid junction consists of a fiber or ceramic plug that allows a small flow of electrolyte filling solution.

Construction varies, but all reference electrodes must generate a stable electrical potential. Reference electrodes generally consist of a metal and its salt in contact with a solution containing the same anion. Mercury/mercurous chloride, as in this example, is a frequently used reference electrode; the disadvantage is that it is slow to reach a new stable voltage following temperature change and it is unstable above 80°C.[1,2] Ag/AgCl is another common reference electrode. It can be used at high temperatures, up to 275°C, and the AgCl-coated Ag wire makes a more compact electrode than that of mercury. In measurements in which chloride contamination must be avoided, a mercury sulfate and potassium sulfate reference electrode may be used.

Liquid Junctions

Electrical connection between the indicator and reference electrodes is achieved by allowing a slow flow of electrolyte from the tip of the reference electrode. A junction potential is always set up at the boundary between two dissimilar solutions because of positive and negative ions diffusing across the boundary at unequal rates. The resultant junction potential may increase or decrease the potential of the reference electrode. Therefore, it is important that the junction potential be kept to a minimum reproducible value when the reference electrode is in solution.

KCl is a commonly used filling solution because K^+ and Cl^- have nearly the same mobilities. When KCl is used as the filling solution for Ag/AgCl electrodes, the addition of AgCl is required to prevent dissolution of the AgCl salt. One way of producing a lower junction potential is to mix K^+, Na^+, NO^-, and Cl^- in appropriate ratios.

Readout Meter

Electromotive force produced by the reference and indicator electrodes is in the millivolt range. Zero potential for the cell indicates that each electrode half-cell is generating the same voltage, assuming there is no liquid junction potential. The isopotential is that potential at which a temperature change has no effect on the response of the electrical cell. Manufacturers generally achieve this by making the midscale (pH 7.0) correspond to 0 V at all temperatures. They use

an internal buffer whose pH changes due to temperature compensate for the changes in the internal and external reference electrodes.

Nernst Equation

The electromotive force generated because of H^+ at the glass tip is described by Nernst equation, which is shown in a simplified form:

$$\varepsilon = \Delta pH \times \frac{RT \ln 10}{F} = \Delta pH \times 0.059\ V \quad (\textbf{Eq. 4.6})$$

where ε is the electromotive force of the cell, F is the Faraday constant (96,500 C/mol), R is the molar gas constant, and T is temperature, in Kelvin.

As the temperature increases, H^+ activity increases and the potential generated increases. Most pH meters have a temperature compensation knob that amplifies the millivolt response when the meter is on pH function. pH units on the meter scale are usually printed for use at room temperature. On the voltmeter, 59.16 is read as 1 pH unit change. The temperature compensation changes the millivolt response to compensate for changes due to temperature from 54.2 at 0°C to 66.10 at 60°C. However, most pH meters are manufactured for greatest accuracy in the 10°C to 60°C range.

Calibration

The steps necessary to standardize a pH meter are fairly straightforward. First, balance the system with the electrodes in a buffer with a 7.0 pH (**Zone A**). The balance or intercept control shifts the entire slope, as shown in **Figure 4.20**. Next, replace the buffer with one of a different pH (**Zone B**). If the meter does not register the correct pH, amplification of the response changes the slope to match that predicted by Nernst equation. If the instrument does not have a slope control, the temperature compensator performs the same function.

pH Combination Electrode

The most commonly used pH electrode has both the indicator and reference electrodes combined in one small probe, which is convenient when small samples are tested. It consists of an Ag/AgCl internal reference electrode sealed in a narrow glass cylinder with a pH-sensitive glass tip. The reference electrode is an Ag/AgCl wire wrapped around the indicator electrode. The outer glass envelope is filled with KCl and has a tiny pore near the tip of the liquid junction. The solution to be measured must completely cover the glass tip. Examples of other ISEs are shown in **Figure 4.21**. The reference electrode, electrometer, and calibration system described for pH measurements are applicable to all ISEs.

There are three major ISE types: inert metal electrodes in contact with a redox couple, metal electrodes that participate in a redox reaction, and membrane electrodes. The membrane can be solid material (e.g., glass), liquid (e.g., ion-exchange electrodes), or special membrane (e.g., compound electrodes), such as gas-sensing and enzyme electrodes.

The standard hydrogen electrode is an example of an inert metal electrode. The Ag/AgCl electrode is an example of the second type. The electrode process $AgCl + e^- \rightarrow Ag^+ + Cl^-$ produces an electrical potential proportional to chloride ion (Cl^-) activity. When Cl^- is held constant, the electrode is used as a reference electrode. The electrode in contact with varying Cl^- concentrations is used as an indicator electrode to measure Cl^- concentration.

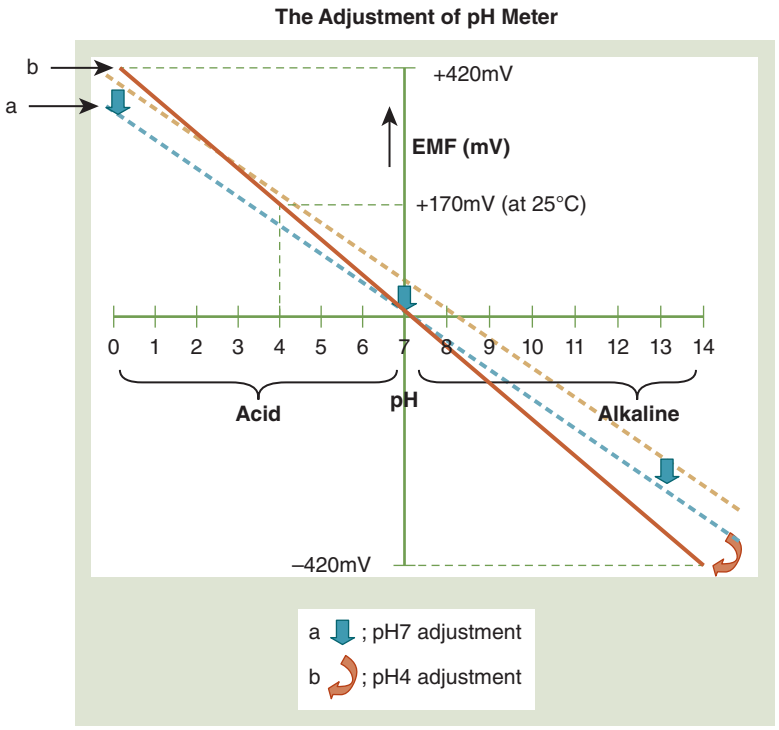
The H^+-sensitive gel layer of the glass pH electrode is considered a membrane. A change in the glass formulation makes the membrane more sensitive to sodium ions (Na^+) than to H^+, creating a sodium ISE. Other solid-state membranes consist of either a single crystal or fine crystals immobilized in an inert matrix such as silicone rubber. Conduction depends on a vacancy defect mechanism, and the crystals are formulated to be selective for a particular size, shape, and charge. Examples include F^--selective electrodes of LaF_3, Cl^--sensitive electrodes with AgCl crystals, and AgBr electrodes for the detection of Br^-.

The calcium ISE is a liquid membrane electrode. An ion-selective carrier, such as dioctylphenyl phosphonate dissolved in an inert water-insoluble solvent, diffuses through a porous membrane. Because the solvent is insoluble in water, the test sample cannot cross the membrane, but calcium ions (Ca^{2+}) are exchanged. The Ag/AgCl internal reference in a filling solution of $CaCl_2$ is in contact with the carrier by means of the membrane.

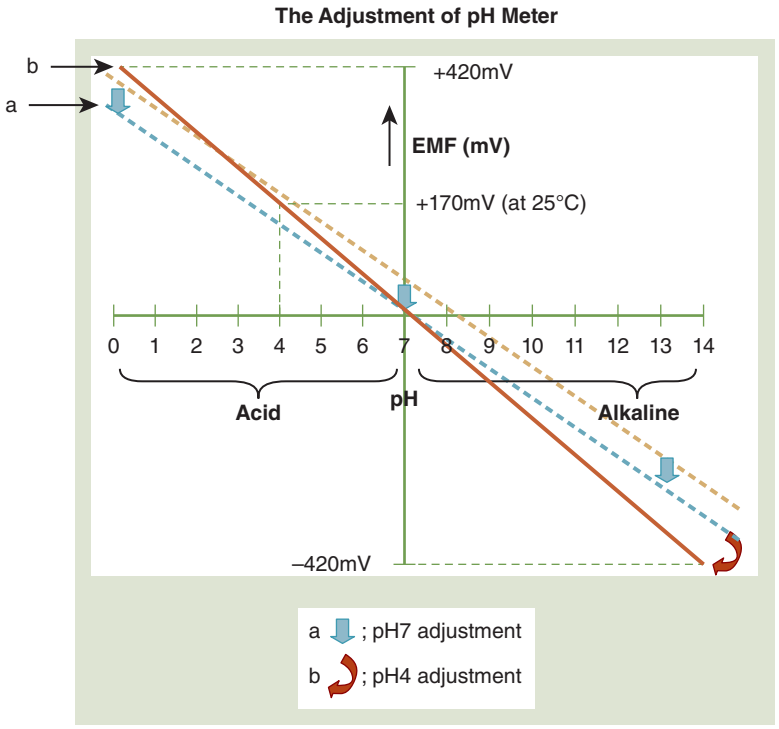
Potassium-selective liquid membranes use the antibiotic valinomycin as the ion-selective carrier. Valinomycin membranes show great selectivity for K^+. Liquid membrane electrodes are recharged every few months to replace the liquid ion exchanger membrane and the porous membrane.

Gas-Sensing Electrodes

Gas electrodes are similar to pH glass electrodes but are designed to detect specific gases (e.g., CO_2 and NH_3) in solutions and are usually separated from



A *EMF: Electromotive force



B *EMF: Electromotive force

Figure 4.20 pH meter calibration.

Data from Willard HH, Merritt LL, Dean JA, et al. *Instrumental Methods of Analysis.* Belmont, CA: Wadsworth; 1981.

the solution by a thin, gas-permeable hydrophobic membrane. **Figure 4.22** shows a schematic illustration of the pCO_2 electrode. The membrane in contact with the solution is permeable only to CO_2, which diffuses into a thin film of sodium bicarbonate solution. The pH of the bicarbonate solution is changed as follows:

$$CO_2 + H_2O \leftrightarrow H^+ + HC_3^- \qquad \textbf{(Eq. 4.7)}$$

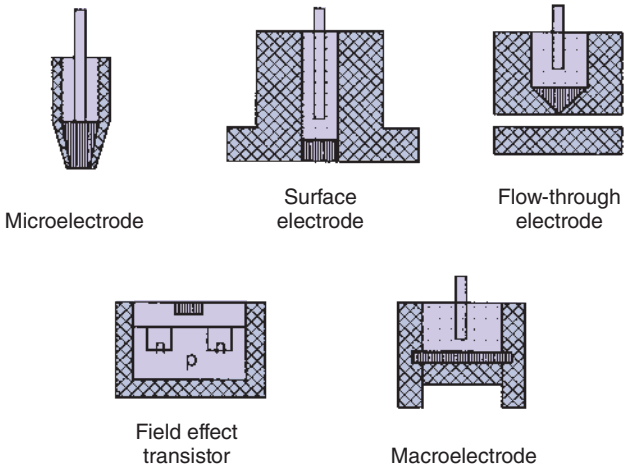


Figure 4.21 Other examples of ion-selective electrodes.

The change in pH of the HCO_3^- is detected by a pH electrode. The pCO_2 electrode is widely used in clinical laboratories as a component of instruments for measuring serum electrolytes and blood gases.

In the NH_3 gas electrode, the bicarbonate solution is replaced by ammonium chloride solution, and the membrane is permeable only to NH_3 gas. As in the pCO_2 electrode, NH_3 changes the pH of NH_4Cl as follows:

$$NH_3 + H_2O \leftrightarrow NH_4^+ + OH^- \qquad \text{(Eq. 4.8)}$$

The amount of OH^- produced varies linearly with the log of the partial pressure of NH_3 in the sample.

Other gas-sensing electrodes function on the basis of an amperometric principle—that is, measurement of the current flowing through an electrochemical cell at a constant applied electrical potential to the electrodes. Examples are the determination of pO_2, glucose, and peroxidase.

The chemical reactions of the pO_2 electrode (Clark electrode), an electrochemical cell with a platinum cathode and an Ag/AgCl anode, are illustrated in Figure 4.18. The electrical potential at the cathode is set to −0.65 V and will not conduct current without oxygen in the sample. The membrane is permeable to oxygen, which diffuses through to the platinum cathode. Current passes through the cell and is proportional to the pO_2 in the test sample.

Glucose determination is based on the reduction in pO_2 during glucose oxidase reaction with glucose and oxygen. Unlike the pCO_2 electrode, the peroxidase electrode has a polarized platinum anode and its potential is set to +0.6 V. Current flows through the system when peroxide is oxidized at the anode as follows:

$$H_2O_2 \rightarrow 2H^+ + 2e^- + O_2 \qquad \text{(Eq. 4.9)}$$

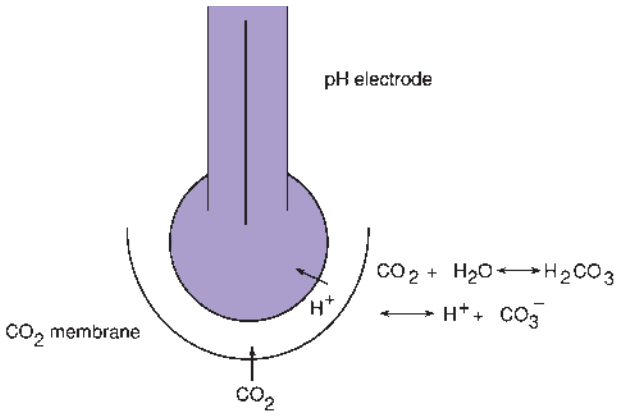


Figure 4.22 The pCO_2 electrode.

Enzyme Electrodes

The various ISEs may be covered by immobilized enzymes that can catalyze a specific chemical reaction. Selection of the ISE is determined by the reaction product of the immobilized enzyme. Examples include urease, which is used for the detection of urea, and glucose oxidase, which is used for glucose detection. A urea electrode must have an ISE that is selective for NH_4^+ or NH_3, whereas glucose oxidase is used in combination with a pH electrode.

Coulometric Titration

In coulometric titration a constant current is applied and the potential of a working electrode is monitored. When all of the analyte has changed state, the change in potential is registered. Coulometric titration is used clinically for sweat chloride determination.

Anodic Stripping Voltammetry

In anodic stripping voltammetry, the analyte is first concentrated onto the surface of an electrode at a constant potential and then goes back into solution as the voltage is changed. Anodic stripping voltammetry is used for the analysis of lead in point-of-care and laboratory settings, although lead testing in the laboratory is currently more commonly performed by electrothermal (graphite furnace) atomic absorption spectroscopy or, preferably, inductively coupled plasma-mass spectrometry (ICP-MS).

Electrophoresis

Electrophoresis is the migration of charged solutes or particles in an electrical field. Iontophoresis refers to the migration of small ions, whereas

zone electrophoresis is the migration of charged macromolecules in a porous support medium such as paper, cellulose acetate, or agarose gel film. An electrophoretogram is the result of zone electrophoresis and consists of sharply separated zones of a macromolecule. In a clinical laboratory, the macromolecules of interest are proteins in serum, urine, cerebrospinal fluid (CSF), other biologic body fluids, erythrocytes, and tissue.

Electrophoresis consists of five components: the driving force (electrical power), the support medium, the buffer, the sample, and the detecting system. A typical electrophoretic apparatus is illustrated in **Figure 4.23**.

Charged particles migrate toward the opposite charged electrode. The velocity of migration is controlled by the net charge of the particle, the size and shape of the particle, the strength of the electric field, chemical and physical properties of the supporting medium, and the electrophoretic temperature. The rate of mobility[11] of the molecule (μ) is given by

$$\mu = Q/6\pi r\eta \qquad \text{(Eq. 4.10)}$$

where Q is net charge of the particle, r is the ionic radius of the particle, and η is the viscosity of the buffer.

From the equation, the rate of migration is directly proportional to the net charge of the particle and inversely proportional to its size and the viscosity of the buffer.

Procedure

The sample is soaked in hydrated support for approximately 5 minutes. The support is put into the electrophoresis chamber, which was previously filled with the buffer. Sufficient buffer must be added to the chamber to maintain contact with the support. Electrophoresis is carried out by applying a constant voltage or constant current for a specific time. The support is then removed and placed in a fixative or rapidly dried to prevent diffusion of the sample. This is followed by staining the zones with an appropriate dye. The uptake of dye by the sample is proportional to sample concentration. After excess dye is washed away, the supporting medium may need to be placed in a clearing agent. Otherwise, it is completely dried.

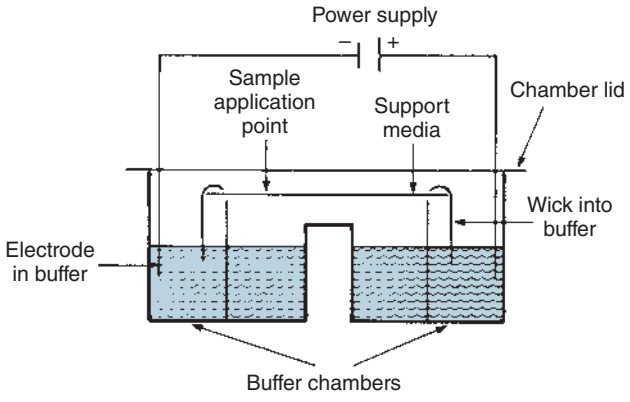


Figure 4.23 Electrophoresis apparatus—basic components.

Power Supply

Power supplies operating at either constant current or constant voltage are available commercially. In electrophoresis, heat is produced when current flows through a medium that has resistance, resulting in an increase in thermal agitation of the dissolved solute (ions) and leading to a decrease in resistance and an increase in current. The increase leads to increases in heat and evaporation of water from the buffer, increasing the ionic concentration of the buffer and subsequent further increases in the current. The migration rate can be kept constant by using a power supply with constant current. This is true because, as electrophoresis progresses, a decrease in resistance as a result of heat produced also decreases the voltage.

Buffers

Two buffer properties that affect the charge of ampholytes are pH and ionic strength. The ions carry the applied electric current and allow the buffer to maintain constant pH during electrophoresis. An ampholyte is a molecule, such as a protein, for which the net charge can be either positive or negative. If the buffer is more acidic than the isoelectric point (pI) of the ampholyte, it binds H^+, becomes positively charged, and migrates toward the cathode. If the buffer is more basic than the pI, the ampholyte loses H^+, becomes negatively charged, and migrates toward the anode. A particle without a net charge will not migrate, remaining at the point of application. During electrophoresis, ions cluster around a migrating particle. The higher the ionic concentration, the greater the size of the ionic cloud and the lower the mobility of the particle. Greater ionic strength produces sharper protein-band separation but leads to increased heat production, which may cause denaturation of heat-labile proteins. Consequently, the optimal buffer concentration should be determined for any electrophoretic system. Generally, the most widely used buffers are made of

monovalent ions because their ionic strength and molality are equal.

Support Materials

Cellulose Acetate

Paper electrophoresis use has been replaced by cellulose acetate or agarose gel in clinical laboratories. Cellulose is acetylated to form cellulose acetate by treating it with acetic anhydride. Cellulose acetate, a dry, brittle film composed of about 80% air space, is produced commercially. When the film is soaked in buffer, the air spaces fill with electrolyte and the film becomes pliable. After electrophoresis and staining, cellulose acetate can be made transparent for densitometer quantitation. The dried transparent film can be stored for long periods. Cellulose acetate prepared to reduce electroendosmosis is available commercially. Cellulose acetate is also used in isoelectric focusing.

Agarose Gel

Agarose gel is another widely used supporting medium. Used as a purified fraction of agar, it is neutral and, therefore, does not produce electroendosmosis. After electrophoresis and staining, it is destained (cleared), dried, and scanned with a densitometer. The dried gel can be stored indefinitely. Agarose gel electrophoresis requires small amounts of sample (~2 mL); it does not bind protein and, therefore, migration is not affected.

Polyacrylamide Gel

Polyacrylamide gel electrophoresis involves separation of protein on the basis of charge and molecular size. Layers of gel with different pore sizes are used. The gel is prepared before electrophoresis in a tube-shaped electrophoresis cell. The small-pore separation gel is at the bottom, followed by a large-pore spacer gel and, finally, another large-pore gel containing the sample. Each layer of gel is allowed to form a gelatin before the next gel is poured over it. At the start of electrophoresis, the protein molecules move freely through the spacer gel to its boundary with the separation gel, which slows their movement. This allows for concentration of the sample before separation by the small-pore gel. Polyacrylamide gel electrophoresis separates serum proteins into 20 or more fractions rather than the usual 6 fractions separated by cellulose acetate or agarose. It is widely used to study individual proteins (e.g., isoenzymes).

Starch Gel

Starch gel electrophoresis separates proteins on the basis of surface charge and molecular size, as does polyacrylamide gel. The procedure is not widely used because of technical difficulty in preparing the gel.

Treatment and Application of Sample

Serum contains a high concentration of protein, especially albumin, and therefore, serum specimens are routinely diluted with buffer before electrophoresis. In contrast, urine and CSF are usually concentrated. Hemoglobin hemolysate is used without further concentration. Generally, sample preparation is performed according to the manufacturer's instructions of the electrophoretic supplies.

Cellulose acetate and agarose gel electrophoresis require approximately 2 to 5 µL of sample. These are the most common routine electrophoreses performed in clinical laboratories. Because most commercially manufactured plates come with a thin plastic template that has small slots through which samples are applied, overloading of agarose gel with sample is not a frequent problem. After serum is allowed to diffuse into the gel for approximately 5 minutes, the template is blotted to remove excess serum before being removed from the gel surface. Sample is applied to cellulose acetate with a twin-wire applicator designed to transfer a small amount.

Detection and Quantitation

Separated protein fractions are stained to reveal their locations. Different stains come with different plates from different manufacturers. The simplest way to accomplish detection is visualization under UV light, whereas densitometry is the most common and reliable way to quantitate the protein bands. Most densitometers will automatically integrate the area under a peak, and the result is printed as percentage of the total. A schematic illustration of a densitometer is shown in **Figure 4.24**.

Electroendosmosis

The movement of buffer ions and solvent relative to the fixed support is called endosmosis or electroendosmosis. Support media, such as paper, cellulose acetate, and agar gel, take on a negative charge from adsorption of hydroxyl ions. When current is applied to the electrophoresis system, the hydroxyl ions remain fixed while the free positive ions move toward the cathode. The ions are highly hydrated, resulting in net cathodic

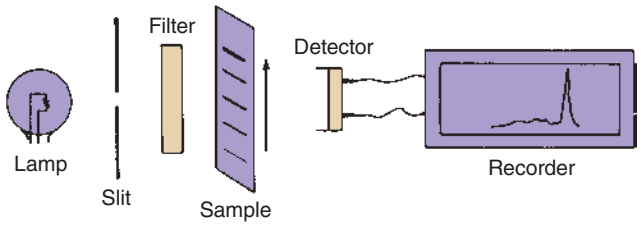


Figure 4.24 Densitometer—basic components.

movement of solvent. Molecules that are nearly neutral are swept toward the cathode with the solvent. Support media such as agarose and acrylamide gel are essentially neutral, eliminating electroendosmosis. The position of proteins in any electrophoresis separation depends not only on the nature of the protein but also on all other technical variables.

Isoelectric Focusing

Isoelectric focusing is a modification of electrophoresis. Charged proteins migrate through a support medium that has a continuous pH gradient. Individual proteins move in the electric field until they reach a pH equal to their isoelectric point, at which point they have no charge and cease to move.

Immunofixation Electrophoresis

Immunofixation electrophoresis (IFE) is used in the clinical laboratory to characterize monoclonal proteins in serum, urine, or cerebrospinal fluid (CSF). A serum, urine, or CSF sample is placed in all six lanes of an agarose gel and electrophoresed to separate the proteins. Cellulose acetate (or some other porous material) is saturated with an Ab reagent and then applied to one lane of the separated protein. If the Ab reagent recognizes the protein, an insoluble complex is formed. After staining and drying of the agarose film, interpretation is based on the migration and appearance of bands. Monoclonal proteins present will appear as a discrete band (with both a heavy and a light chain monospecific antiserum occurring at the same position). Polyclonal proteins will appear as a diffuse band. The concentration of patient sample may need adjustment to ensure the reaction is in the zone of equivalence. See Chapter 6, *Amino Acids and Proteins*, for figures and disease state discussion.

Capillary Electrophoresis

In capillary electrophoresis (CE), separation is performed in narrow-bore, fused silica capillaries (inner diameter 25 to 75 μm). Usually, the capillaries are

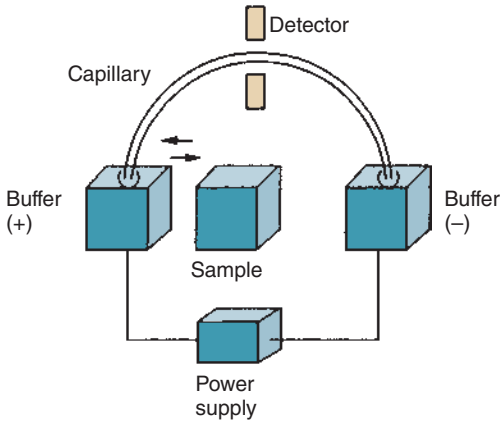


Figure 4.25 Schematic of capillary electrophoresis instrumentation. Sample is loaded on the capillary by replacing the anode buffer reservoir with the sample reservoir.

only filled with buffer, although gel media can also be used. A CE instrumentation schematic is shown in **Figure 4.25**. Initially, the capillary is filled with buffer and then the sample is loaded; applying an electric field performs the separation. Detection can be made near the other end of the capillary directly through the capillary wall.[12]

A fundamental CE concept is the electroosmotic flow (EOF). EOF is the bulk flow of liquid toward the cathode upon application of an electric field, and it is superimposed on electrophoretic migration. EOF controls the amount of time solutes remain in the capillary. Cations migrate fastest because both EOF and electrophoretic attraction are toward the cathode; neutral molecules are all carried by the EOF but are not separated from each other; and anions move slowest because, although they are carried to the cathode by the EOF, they are attracted to the anode and repelled by the cathode (**Figure 4.26**). Widely used for monitoring separated analytes, UV-visible detection is performed directly on the capillary; however, sensitivity is poor because of the small dimensions of the capillary, resulting in a short path length. Fluorescence, laser-induced fluorescence, and chemiluminescence detection can be used for higher sensitivity.

CE has been used for the separation, quantitation, and determination of molecular weights of proteins and peptides; for the analysis of polymerase chain reaction products; and for the analysis of inorganic ions, organic acids, pharmaceuticals, optical isomers, and drugs of abuse in serum and urine.[13] While traditionally serum protein electrophoresis for the diagnosis of plasma cell dyscrasias has been performed using polyacrylamide gel

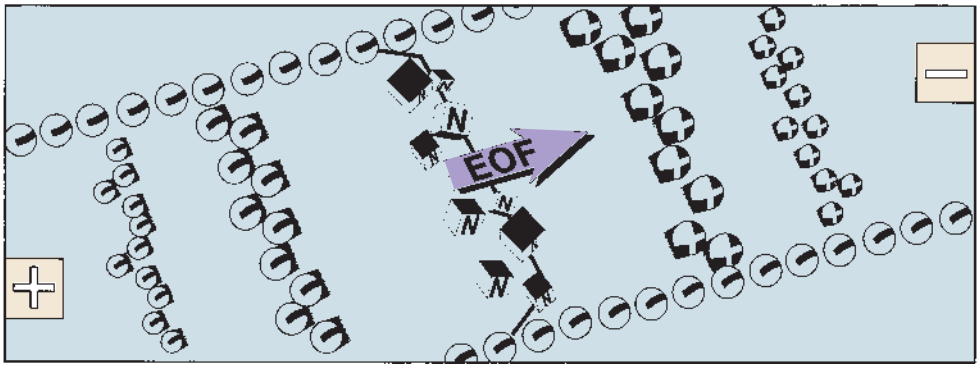


Figure 4.26 Differential solute migration superimposed on electro-osmotic flow in capillary zone electrophoresis.

electrophoresis, CE has now become widely used for this analysis due to its faster run time and its relative automation.

Two-Dimensional Electrophoresis

This electrophoresis assay combines two different electrophoresis dimensions to separate proteins from complex matrices such as serum or tissue. In the first dimension, proteins are resolved according to their isoelectric points (pIs), using immobilized pH gradients. Commercial gradients are available in a variety of pH ranges. In the second dimension, proteins are separated according to their relative size (molecular weight), using sodium dodecyl sulfate–polyacrylamide gel electrophoresis. A schematic of this is shown in **Figure 4.27**. Gels can be run under denaturing or nondenaturing conditions (e.g., for the maintenance of enzyme activity) and visualized by a variety of techniques, including the use of colorimetric dyes (e.g., Coomassie blue or silver stain) and radiographic, fluorometric, or chemiluminescence of appropriately labeled polypeptides. These latter techniques are considerably more sensitive than the colorimetric dyes.

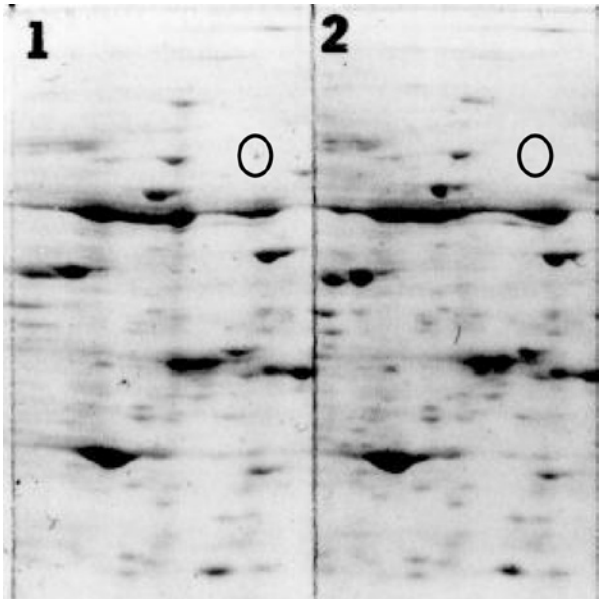


Figure 4.27 Hypothetical example of a two-dimensional electrophoretogram from a patient with a disease (*panel 1*) compared with a normal subject (*panel 2*). The patient exhibits a protein (*oval*) that is not expressed in the normal subject. This protein might be a potential marker for this disease.

Gels courtesy of Kendrick Laboratories, Madison, WI.

Osmometry

Osmometry is the principle of measuring the concentration of solute particles in a solution using one of the four colligative properties discussed below. An osmometer is used to perform this measurement. The mathematic definition of osmometry is

$$\text{Osmolality} = \phi \times n \times C \qquad \text{(Eq. 4.11)}$$

where φ is the osmotic coefficient, n is the number of dissociable particles (ions) per molecule in the solution, and C is the concentration in moles per kilogram of solvent.

The osmotic coefficient is an experimentally derived factor to correct for the fact that some of the molecules, even in a highly dissociated compound, exist as molecules rather than as ions.

The four physical properties (also known as colligative properties) of a solution that change with variations in the number of dissolved particles in the solvent are osmotic pressure, vapor pressure, boiling point, and freezing point. Osmometers measure osmolality indirectly by measuring one of these colligative properties, which change proportionally

with osmotic pressure. Osmometers in clinical use measure either freezing point depression or vapor pressure depression; results are expressed in milliosmolal per kilogram (mOsm/kg) units. Only freezing point osmometry will be discussed in this section as it is the most commonly used in the clinical laboratory.

Freezing Point Osmometer

Figure 4.28 illustrates the basic components of a freezing point osmometer. The sample in a small tube is lowered into a chamber with cold refrigerant circulating from a cooling unit. A thermistor is immersed in the sample. To measure temperature, a wire is used to gently stir the sample until it is cooled to several degrees below its freezing point. It is possible to cool water to as low as −40°C and still have liquid water, provided no crystals or particulate matter is present. This is referred to as a *supercooled solution*. Vigorous agitation when the sample is supercooled results in rapid freezing. Freezing can also be started by "seeding" a supercooled solution with crystals. When the supercooled solution starts to freeze as a result of the rapid stirring, a slush is formed and the solution actually warms to its freezing point temperature. The slush, an equilibrium of liquid and ice crystals, will remain at the freezing point temperature until the sample freezes solid and drops below its freezing point.

Impurities in a solvent will lower the temperature at which freezing or melting occurs by reducing the bonding forces between solvent molecules so that the molecules break away from each other and exist as a fluid at a lower temperature. The decrease in the freezing point temperature is proportional to the number of dissolved particles present.

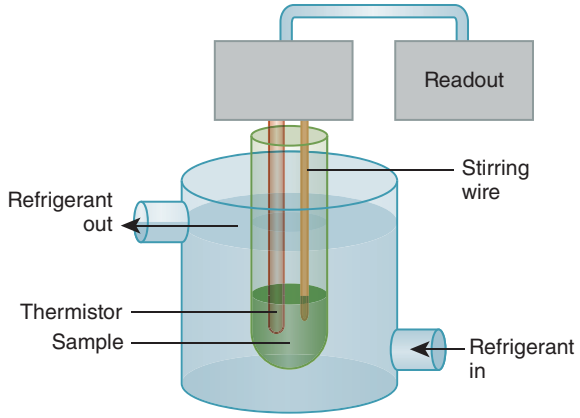


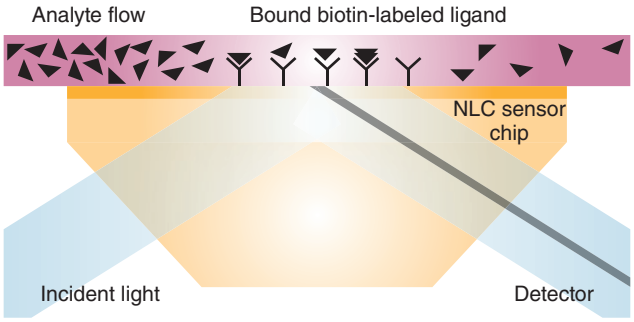
Figure 4.28 Freezing point osmometer.

Data from Coiner D. *Basic Concepts in Laboratory Instrumentation*. Bethesda, MD: ASMT Education and Research Fund; 1975–1979.

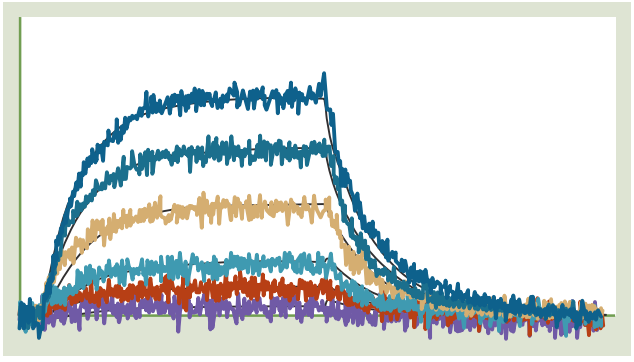
The thermistor is a material that has less resistance when the temperature increases. The readout uses a Wheatstone bridge circuit that detects temperature change as proportional to change in thermistor resistance. Freezing point depression is proportional to the number of solute particles. Standards of known concentration are used to calibrate the instruments in mOsm/kg.

Newer Optical Techniques

Surface plasmon resonance (SPR) and biolayer interferometry (bli) are newer optical techniques currently used mainly in research laboratories. SPR enables the study of the binding of ligands to surface receptors such as membrane proteins in real time. **Figure 4.29A** illustrates a schematic diagram of SPR using the "Kretschmann" geometry, and **Figure 4.29B** illustrates a typical plot of binding affinity as reflected by a change in refractive index versus time. Polarized light passes through a prism and strikes the back of a sensor chip containing a metal such as gold and reflects off the back of the chip onto a detector. At the proper angle of light, the



A



B

Figure 4.29 Schematic of surface plasmon resonance spectroscopy. **(A)** Block diagram of the SPR sensor. **(B)** Output demonstrating the kinetics of analyte binding to the sensor.

Figures courtesy of Bio-Rad Laboratories, Hercules, CA. See text for details.

light is absorbed by the electrons in the metal causing a resonance. This decreases the amount of light reaching the detector. SPR can be used to measure interactions between immobilized antibodies and freely circulating analytes (or vice versa). In a clinical research setting, SPR is being used for the analysis of low molecular weight compounds, proteins and biomarkers, hormones, nucleic acids, circulating antibodies, and infectious microorganisms from biological matrices.[14] The analytical sensitivity of SPR is similar to that of conventional immunoassays. The costs for developing and using SPR may be less expensive since there are no labels involved. Moreover, sensors can be placed into an array for multiplexing purposes.

Biolayer interferometry is a similar label-free technique that uses sensor tips instead of a flat metal surface.[15]

Chromatography

Chromatography refers to the group of techniques used to separate complex mixtures on the basis of different physical interactions between the individual compounds and the stationary phase of the system. The basic components in any chromatographic technique are the mobile phase (gas or liquid), which carries the complex mixture (sample); the stationary phase (solid or liquid), through which the mobile phase flows; the column holding the stationary phase; and the separated components (eluate).

Modes of Separation

Adsorption

Adsorption chromatography, also known as liquid–solid chromatography, is based on the competition between the sample and the mobile phase for adsorption sites on the solid stationary phase. There is an equilibrium of solute molecules being adsorbed to the solid surface and desorbed and dissolved in the mobile phase. The molecules that are most soluble in the mobile phase move fastest; the least soluble move slowest. Thus, a mixture is typically separated into classes according to polar functional groups. The stationary phase can be acidic polar (e.g., silica gel), basic polar (e.g., alumina), or nonpolar (e.g., charcoal). The mobile phase can be a single solvent or a mixture of two or more solvents, depending on the analytes to be desorbed. Liquid–solid chromatography is not widely used in clinical laboratories because of technical problems with the preparation of a stationary phase that has homogeneous distribution of adsorption sites.

Partition

Partition chromatography is also referred to as liquid–liquid chromatography. Separation of solute is based on relative solubility in an organic (nonpolar) solvent and an aqueous (polar) solvent. In its simplest form, partition (extraction) is performed in a separatory funnel. Molecules containing polar and nonpolar groups in an aqueous solution are added to an immiscible organic solvent. After vigorous shaking, the two phases are allowed to separate. Polar molecules remain in the aqueous solvent; nonpolar molecules are extracted in the organic solvent. This results in the partitioning of the solute molecules into two separate phases.

The ratio of the concentration of the solute in the two liquids is known as the partition coefficient:

$$K = \frac{\text{solute in stationary phase}}{\text{solute in mobile phase}} \qquad \textbf{(Eq. 4.12)}$$

Modern partition chromatography uses pseudoliquid stationary phases that are chemically bonded to the support or high-molecular-weight polymers that are insoluble in the mobile phase.[16] Partition systems are considered normal phase when the mobile solvent is less polar than the stationary solvent and reverse phase when the mobile solvent is more polar.

Partition chromatography is applicable to any substance that may be distributed between two liquid phases. Because ionic compounds are generally soluble only in water, partition chromatography works best with nonionic compounds.

Steric Exclusion

Steric exclusion, a variation of liquid–solid chromatography, is used to separate solute molecules on the basis of size and shape. The chromatographic column is packed with porous material, as shown in **Figure 4.30**. A sample containing different-sized molecules moves down the column dissolved in the mobile solvent. Small molecules enter the pores in the packing and are momentarily trapped. Large molecules are excluded from the small pores and so move quickly between the particles. Intermediate-sized molecules are partially restricted from entering the pores and, therefore, move through the column at an intermediate rate that is between those of the large and small molecules.

Early methods used hydrophilic beads of cross-linked dextran, polyacrylamide, or agarose, which

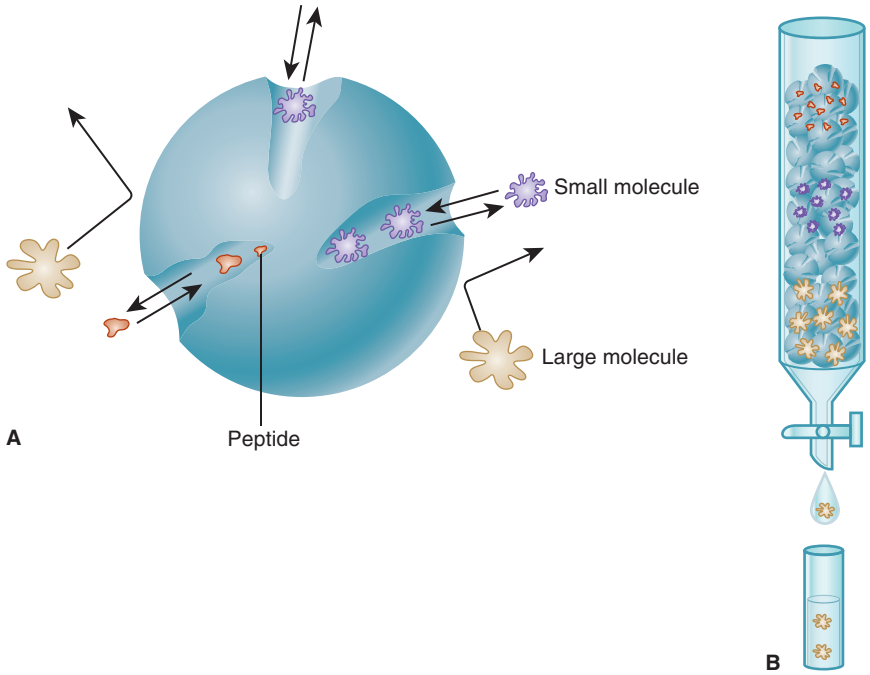


Figure 4.30 Pictorial concept of steric exclusion chromatography. Separation of sample components by their ability to permeate pore structure of column-packing material. Smaller molecules **(A)** permeating the interstitial pores; large, excluded molecules **(B)**.

formed a gel when soaked in water. This method was termed *gel filtration*. A similar separation process using hydrophobic gel beads of polystyrene with a nonaqueous mobile phase was called *gel permeation chromatography*. Current porous packing uses rigid inorganic materials such as silica or glass. The term *steric exclusion* includes all these variations. Pore size is controlled by the manufacturer, and packing materials can be purchased with different pore sizes, depending on the size of the molecules being separated.

Ion-Exchange Chromatography

In ion-exchange chromatography, solute mixtures are separated by virtue of the magnitude and charge of ionic species. The stationary phase is a resin, consisting of large polymers of substituted benzene, silicates, or cellulose derivatives, with charged functional groups. The resin is insoluble in water, and the functional groups are immobilized as side chains on resin beads that are used to fill the chromatographic column. **Figure 4.31A** shows a resin with sulfonate functional groups. Hydrogen (H^+) ions are loosely held and free to react. This is an example of a cation-exchange resin. When a cation such as Na^+ comes in contact with these functional groups, an equilibrium is formed, following the law of mass action. Because there are many sulfonate groups, Na^+ is effectively and completely removed from solution. The Na^+ concentrated on the resin column can be eluted from the resin by pouring acid through the column, driving the equilibrium to the left.

Anion-exchange resins are made with exchangeable hydroxyl ions such as the diethylamine functional group illustrated in **Figure 4.31B**. They are used like cation-exchange resins, except that hydroxyl ions are exchanged for anions. The example shows Cl^- in sample solution exchanged for OH^- from the resin functional group. Anion and cation resins mixed together (mixed-bed resin) are used to deionize water. The displaced protons and hydroxyl ions combine to form water. Ionic functional groups other than the illustrated examples are used for specific analytic

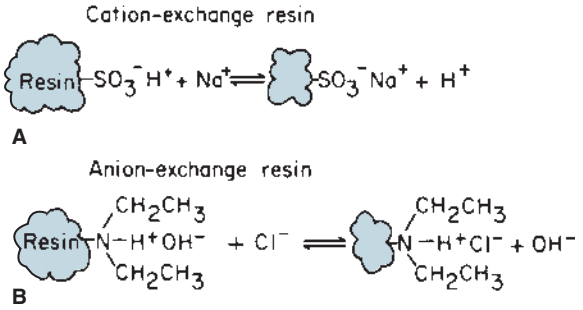


Figure 4.31 Chemical equilibrium of ion-exchange resins. **(A)** Cation-exchange resin. **(B)** Anion-exchange resin.

applications. Ion-exchange chromatography is used to remove interfering substances from a solution, to concentrate dilute ion solutions, and to separate mixtures of charged molecules, such as amino acids. Changing pH and ionic concentration of the mobile phase allows separation of mixtures of organic and inorganic ions.

Chromatographic Procedures

Thin-Layer Chromatography

Thin-layer chromatography (TLC) is a variant of column chromatography. A thin layer of sorbent, such as alumina, silica gel, cellulose, or cross-linked dextran, is uniformly coated on a glass or plastic plate. Each sample to be analyzed is applied as a spot near one edge of the plate, as shown in **Figure 4.32**. The mobile phase (solvent) is usually placed in a closed container until the atmosphere is saturated with solvent vapor. One edge of the plate is placed in the solvent, as shown. The solvent migrates up the thin layer by capillary action, dissolving and carrying sample molecules. Separation can be achieved by any of the four processes previously described, depending on the sorbent (thin layer) and solvent chosen. After the solvent reaches a predetermined height, the plate is removed and dried. Sample components are identified by comparison with standards on the same plate. The distance a component migrates, compared with the distance the solvent front moves, is called the retention factor, R_f:

$$R_f = \frac{\text{distance leading edge of componenet moves}}{\text{total distance solvent front moves}} \quad \textbf{(Eq. 4.13)}$$

Each sample component R_f is compared with the R_f of standards. Using Figure 4.32 as an example, standard A has an R_f value of 0.4, standard B has an R_f value of 0.6, and standard C has an R_f value of 0.8.

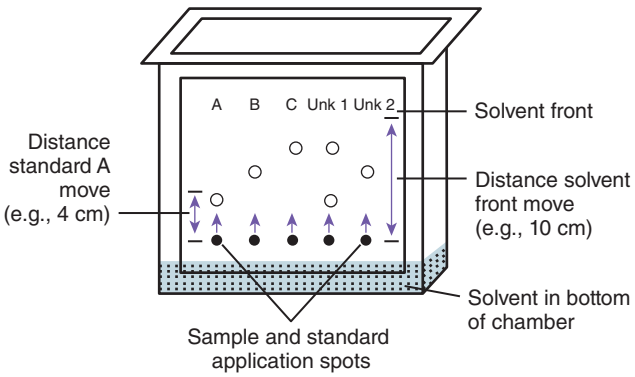


Figure 4.32 Thin-layer chromatography plate in chromatographic chamber.

The first unknown contains A and C, because the R_f values are the same. This ratio is valid only for separations run under identical conditions. Because R_f values may overlap for some components, further identifying information is obtained by spraying different stains on the dried plate and comparing colors of the standards.

TLC is most commonly used as a semiquantitative screening test. Technique refinement has resulted in the development of semiautomated equipment and the ability to quantitate separated compounds. For example, sample applicators apply precise amounts of sample extracts in concise areas. Plates prepared with uniform sorbent thickness, finer particles, and new solvent systems have resulted in the technique of high-performance thin-layer chromatography (HPTLC).[17] Absorbance of each developed spot is measured using a densitometer, and the concentration is calculated by comparison with a reference standard chromatographed under identical conditions.

High-Performance Liquid Chromatography

Modern **high-performance liquid chromatography** (HPLC) uses pressure for fast separations, controlled temperature, inline detectors, and gradient elution techniques.[18,19] **Figure 4.33** illustrates the basic components.

Pumps

A pump forces the mobile phase through the column at a much greater velocity than that accomplished by gravity flow columns and includes pneumatic, syringe, reciprocating, or hydraulic amplifier pumps. The most widely used pump today is the mechanical reciprocating pump, which is used as a multihead pump with two or more reciprocating pistons. During pumping, the pistons operate out of phase (180° for two heads, 120° for three heads) to provide constant flow. Pneumatic pumps are used for preoperative purposes; hydraulic amplifier pumps are no longer commonly used.

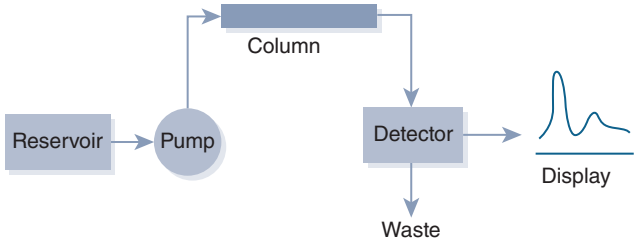


Figure 4.33 High-performance liquid chromatography basic components.

Columns

The stationary phase is packed into long stainless-steel columns. HPLC is usually run at ambient temperatures, although columns can be put in an oven and heated to enhance the rate of partition. Fine, uniform column packing results in narrower chromatogram peaks but requires pressure to force the mobile phase through. The packing can also be pellicular (an inert core with a porous layer), inert and small particles, or macroporous particles. The most common material used for column packing is silica gel. It is very stable and can be used in different ways. It can be used as solid packing in liquid–solid chromatography or coated with a solvent, which serves as the stationary phase (liquid–liquid). As a result of the short lifetime of coated particles, molecules of the mobile-phase liquid are now bonded to the surface of silica particles.

Reversed-phase HPLC is now popular; the stationary phase is nonpolar molecules (e.g., octadecyl C-18 hydrocarbon) bonded to silica gel particles. For this type of column packing, the mobile phase commonly used is acetonitrile, methanol, water, or any combination of solvents. A reversed-phase column can be used to separate ionic, nonionic, and ionizable samples. A buffer is used to produce the desired ionic characteristics and pH for separation of the analyte. Column packings vary in size (3 to 20 mm), using smaller particles mostly for analytic separations and larger ones for preparative separations.

Sample Injectors

A small syringe can be used to introduce the sample into the path of the mobile phase that carries it into the column (Figure 4.33). The best and most widely used method, however, is the loop injector. The sample is introduced into a fixed-volume loop. When the loop is switched, the sample is placed in the path of the flowing mobile phase and flushed onto the column. Loop injectors have high reproducibility and are used at high pressures. Many HPLC instruments have loop injectors that can be programmed for automatic injection of samples. When the sample size is less than the volume of the loop, the syringe containing the sample is often filled with the mobile phase to the volume of the loop before filling the loop. This prevents the possibility of air being forced through the column because such a practice may reduce the lifetime of the column-packing material.

Detectors

Modern HPLC detectors monitor the eluate as it leaves the column and, ideally, produce an electronic signal proportional to the concentration of each separated component. Spectrophotometers that detect absorbances of visible or UV light are most commonly used. Photodiode array (PDA) and other rapid scanning detectors are also used for spectral comparisons and compound identification and purity. These detectors have been used for drug analyses in urine. Obtaining a UV scan of a compound as it elutes from a column can provide important information as to its identity. Unknowns can be compared against library spectra in a similar manner to MS. Unlike gas chromatography/MS, which requires volatilization of targeted compounds, liquid chromatography (LC)/PDA enables direct injection of aqueous urine samples.

Because many biologic substances fluoresce strongly, fluorescence detectors are also used, involving the same principles discussed in the section on spectrophotometric measurements. Another common HPLC detector is the amperometric or electrochemical detector, which measures current produced when the analyte of interest is either oxidized or reduced at some fixed potential set between a pair of electrodes.

A mass spectrometer can also be used as a detector, as described later.

Recorders

The recorder produces a graph, called a chromatogram, that shows detector response versus the time it takes for the mobile phase to pass through the instrument, starting from the time of sample injection (**Figure 4.34**). The retention time is used to identify compounds when compared with standard retention times run under identical conditions. Peak area is proportional to concentration of the compounds that produced the peaks.

When the elution strength of the mobile phase is constant throughout the separation, it is called isocratic elution. For samples containing compounds of widely differing relative compositions, the choice of solvent is a compromise. Early eluting compounds may have retention times close to zero, producing a poor separation (resolution), as shown in Figure 4.34A. Basic compounds often have low retention times because C-18 columns cannot tolerate high pH mobile phases. The addition of cation-pairing reagents to the mobile phase (e.g., octane sulfonic acid) can result in better retention of negatively charged compounds onto the column.

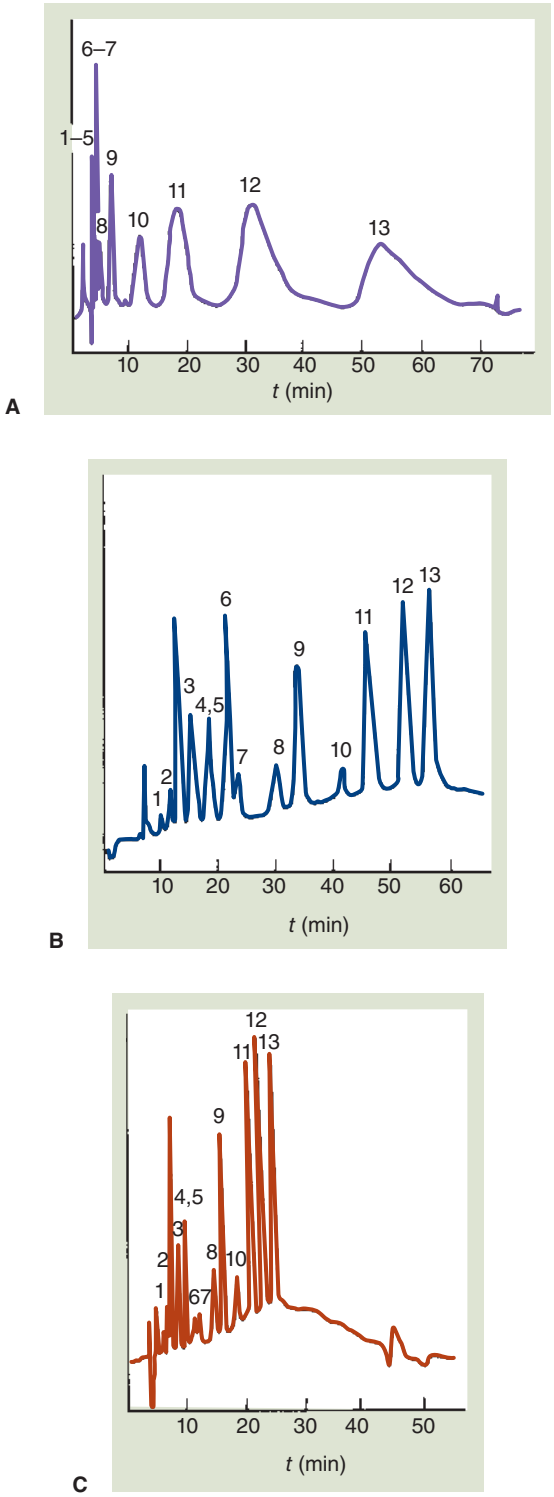


Figure 4.34 Chromatograms. **(A)** Isocratic ion-exchange separation mobile phase contains 0.055 mol/L $NaNO_3$. **(B)** Gradient elution mobile phase gradient from 0.01 to 0.1 mol/L $NaNO_3$ at 2% per minute. **(C)** Gradient elution—5% per minute.

Adapted from Horváth C. *High Performance Liquid Chromatography, Advances and Perspectives.* New York, NY: Academic Press; 1980.

The late-eluting compounds may have long retention times, producing broad peaks (bands) resulting in decreased sensitivity. In some cases, certain components of a sample may have such a great affinity for the stationary phase that they do not elute at all. Gradient elution is an HPLC technique that can be used to overcome this problem. The composition of the mobile phase is varied to provide a continual increase in the solvent strength of the mobile phase entering the column (Figure 4.34B). The same gradient elution can be performed with a faster change in concentration of the mobile phase (Figure 4.34C).

Gas Chromatography

Gas chromatography (GC) is used to separate mixtures of compounds that are volatile or can be made volatile.[20] GC may be gas–solid chromatography, with a solid stationary phase, or gas–liquid chromatography (GLC), with a nonvolatile liquid stationary phase. GLC is commonly used in clinical laboratories. **Figure 4.35** illustrates the basic components of a GC system. The setup is similar to HPLC, except that the mobile phase is a gas and samples are partitioned between a gaseous mobile phase and a liquid stationary phase. The carrier gas can be nitrogen, helium, or argon. The selection of a carrier gas is determined by the detector used in the instrument. The instrument can be operated at a constant temperature or programmed to run at different temperatures if a sample has components with different volatilities. This is analogous to gradient elution described for HPLC.

The sample, which is injected through a septum, must be injected as a gas or the temperature of the injection port must be above the boiling point of the components so that they vaporize upon injection. Sample vapor is swept through the column partially as a gas and partially dissolved in the liquid phase. Volatile compounds that are present mainly in the gas phase will have a low partition coefficient and will move quickly through the column. Compounds with higher boiling points will move slowly through the column. The effluent passes through a detector that produces an electric signal proportional to the concentration of the volatile components. As in HPLC, the chromatogram is used both to identify the compounds by the retention time and to determine their concentration by the area under the peak.

Columns

GLC columns are generally made of glass or stainless steel and are available in a variety of coil configurations and sizes. Packed columns are filled with inert particles such as diatomaceous earth or porous polymer or glass beads coated with a nonvolatile liquid

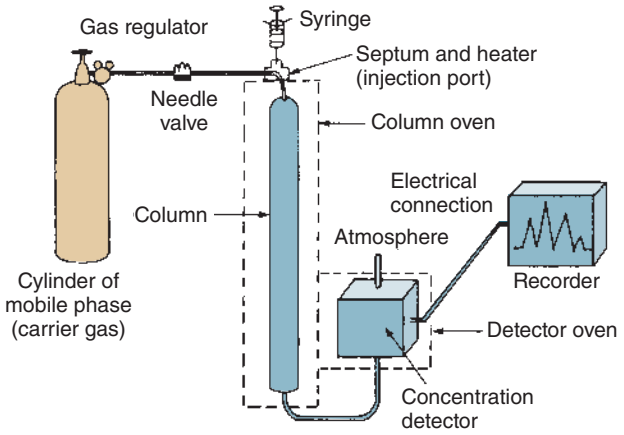


Figure 4.35 Gas–liquid chromatography basic components.

(stationary) phase. These columns are usually 1/8 to 1/4 inch wide and 3 to 12 ft long. Capillary wall-coated open tubular columns have inside diameters in the range of 0.25 to 0.50 mm and are up to 60 m long. The liquid layer is coated on the walls of the column. A solid support coated with a liquid stationary phase may in turn be coated on column walls. The liquid stationary phase must be nonvolatile at the temperatures used, must be thermally stable, and must not react chemically with the solutes to be separated. The stationary phase is termed nonselective when separation is primarily based on relative volatility of the compounds. Selective liquid phases are used to separate polar compounds based on relative polarity (as in liquid–liquid chromatography).

Detectors

Although there are many types of detectors, only thermal conductivity (TC) and flame ionization detectors are discussed because they are the most stable (**Figure 4.36**). TC detectors contain wires (filaments) that change electrical resistance with change in temperature. The filaments form opposite arms of a Wheatstone bridge and are heated electrically to raise their temperature. Helium, which has a high TC, is usually the carrier gas. Carrier gas from the reference column flows steadily across one filament, cooling it slightly. Carrier gas and separated compounds from the sample column flow across the other filament. The sample components usually have a lower TC, increasing the temperature and resistance of the sample filament. The change in resistance results in an unbalanced bridge circuit. The electrical change is amplified and fed to the recorder. The electrical change is proportional to the concentration of the analyte. Flame ionization detectors are widely used

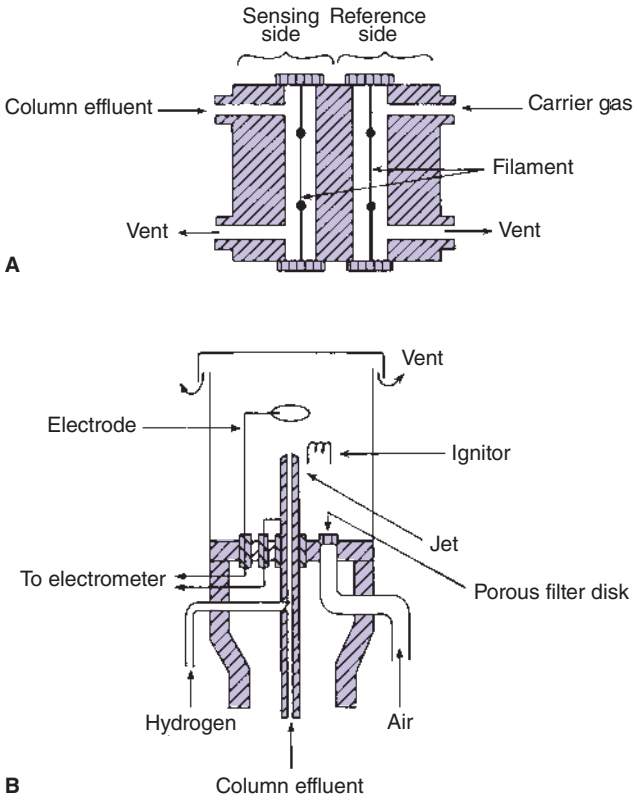


Figure 4.36 **(A)** Schematic diagram of a thermal conductivity detector. **(B)** Schematic diagram of a flame ionization detector.

Adapted from Tietz NW, ed. *Fundamentals of Clinical Chemistry*. Philadelphia, PA: WB Saunders; 1987.

in the clinical laboratory. They are more sensitive than TC detectors. The column effluent is fed into a small hydrogen flame burning in excess air or atmospheric oxygen. The flame jet and a collector electrode around the flame have opposite potentials. As the sample burns, ions form and move to the charged collector. Thus, a current proportional to the concentration of the ions is formed and fed to the recorder.

Mass Spectrometry

Definitive identification of samples eluting from GC or HPLC columns is possible when a mass spectrophotometer (MS) is used as a detector.[21] The coupled techniques, GC/MS and LC/MS, have powerful analytic capabilities with widespread clinical applications. The sample in an MS is first volatilized and then ionized to form charged molecular ions and fragments that are separated according to their mass-to-charge (m/z) ratio; the sample is then measured by a detector, which gives the intensity of the ion current for each species. These steps take place in the four basic components that are standard in all MSs: the sample inlet, ionization source, mass analyzer, and ion detector (**Figure 4.37**). Ultimately,

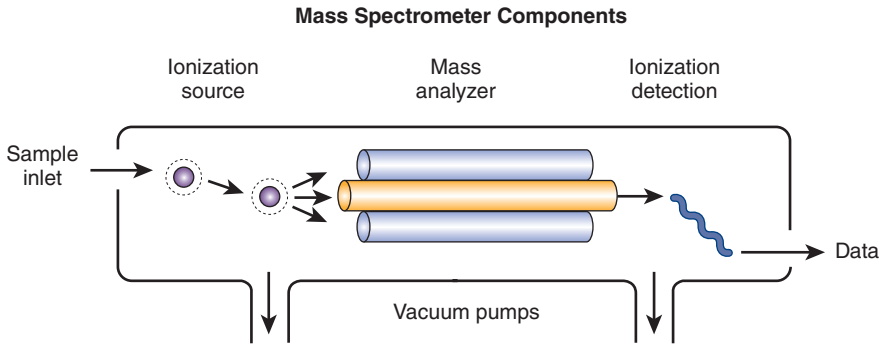


Figure 4.37 The components of a mass spectrometer. In this case, the ionization source pictured is electrospray ionization and the mass analyzer is a quadrupole.

molecule identification is based on the formation of characteristic fragments. **Figure 4.38** illustrates the mass spectrum of Δ9-carboxytetrahydrocannabinol, a metabolite of marijuana.

Sample Introduction and Ionization

Direct infusion is commonly used to interface a GC or LC with an MS; however, the challenge of introducing a liquid sample from an LC column into an MS was a significant barrier until recent technological advances in ionization techniques.

Electron Ionization

The most common form of ionization used in GC/MS is electron ionization (EI). This method requires a source of electrons in the form of a filament to which an electric potential is applied, typically at 70 eV.[22] The molecules in the source are bombarded with high-energy electrons, resulting in the formation of charged molecular ions and fragments. Molecules break down into characteristic fragments according to their molecular structure (**Figure 4.39**). The ions formed and their relative proportions are reproducible and can be used for qualitative identification of the compound. Since most instruments use the same 70 eV potential, the fragmentation of molecules on different days and different instruments is remarkably similar, allowing the comparison of unknown spectra to spectra in a published reference library.[22]

Atmospheric Pressure Ionization

Unlike EI in GC/MS, most LC/MS ionization techniques are conducted at atmospheric pressure. As such, the ion source of this type of instrument is not included in the high-vacuum region of the instrument. Three types of ionization for LC/MS will be discussed here: electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and matrix-assisted laser desorption ionization (MALDI). Many LC/MS

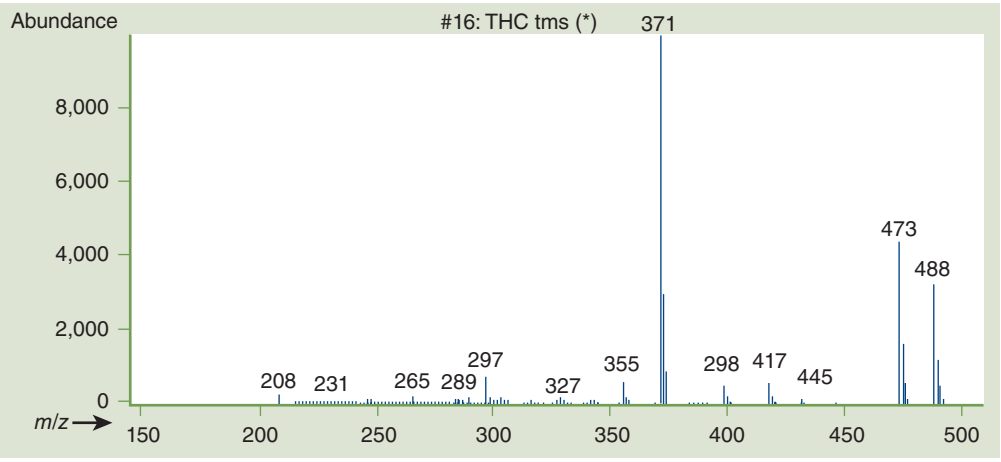


Figure 4.38 Mass spectrum of the trimethylsilane derivative of Δ9-carboxytetrahydrocannabinol (marijuana metabolite).

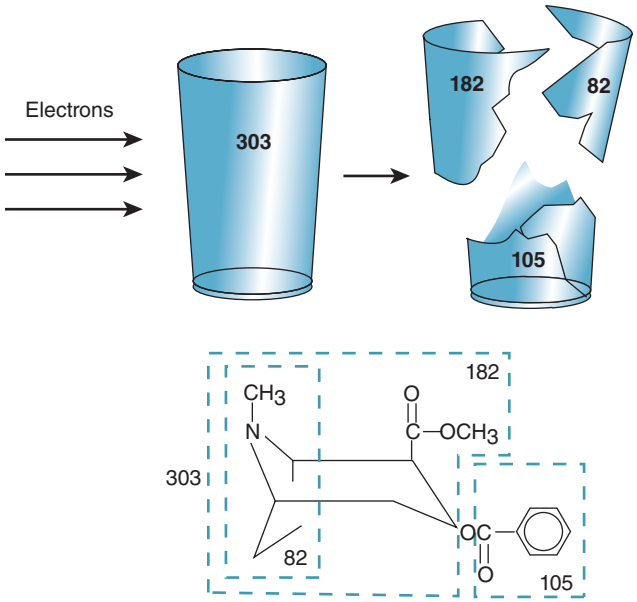


Figure 4.39 Electron bombardment breaks cocaine into fragments, with number and size quantified. Unlike the illustrative glass tumbler, the result of mass fragmentation of cocaine or other chemical compounds is both predictable and reproducible, especially with electron ionization.

techniques employ technologies after the source, in the mass analyzer, to fragment molecules and generate the daughter or fragment ions used in identification. However, ionization techniques used in LC/MS produce fragments and therefore mass spectra that are somewhat less reproducible between instruments than EI used in GC/MS. This may prove to limit the utility of reference library spectra produced in other instruments.

Electrospray Ionization

Thanks to its wide mass range and high sensitivity, ESI can be applied to a wide range of biological macromolecules in addition to small molecules and has become the most common ionization source for LC/MS. ESI involves passing the LC effluent through a capillary to which a voltage has been applied. The energy is transferred to the solvent droplets, which become charged.[22] Evaporation of the solvent through heat and gas causes the droplets to decrease in size, which increases the charge density on the surface. Eventually, the Coulombic repulsion of like charges leads to the ejection of ions from the droplet (**Figure 4.40**).[23] The individually charged molecules are drawn into the MS for mass analysis. ESI is adept at forming singly charged small molecules, but larger molecules can also be ionized using this method. Larger molecules such as proteins become multiply charged in ESI, and since MSs measure the *m/z*, even these large molecules can be observed in an instrument with a relatively small mass range (**Figure 4.41**).[23]

Atmospheric Pressure Chemical Ionization

Another important ionization source is APCI, which is similar to ESI in that the liquid from LC is introduced directly into the ionization source. However, the droplets are not charged and the source contains a heated vaporizer to allow rapid desolvation of the drops.[23] A high voltage is applied to a corona discharge needle, which emits a cloud of electrons to ionize compounds after they are converted to the gas phase.

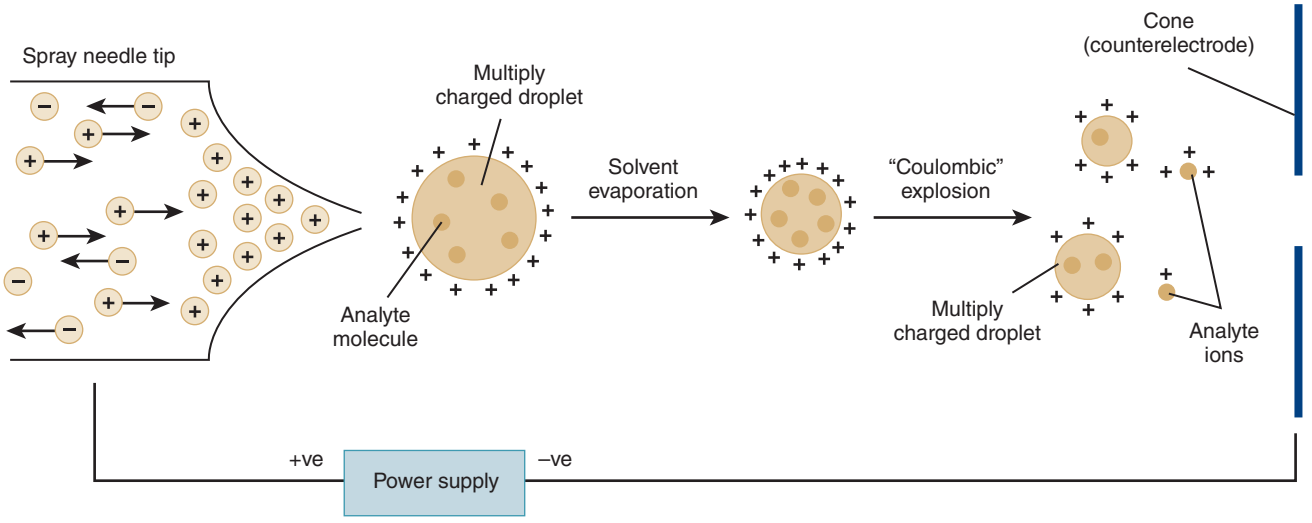


Figure 4.40 Diagram of electrospray ionization, the most common ionization source for liquid chromatography/mass spectrometry.

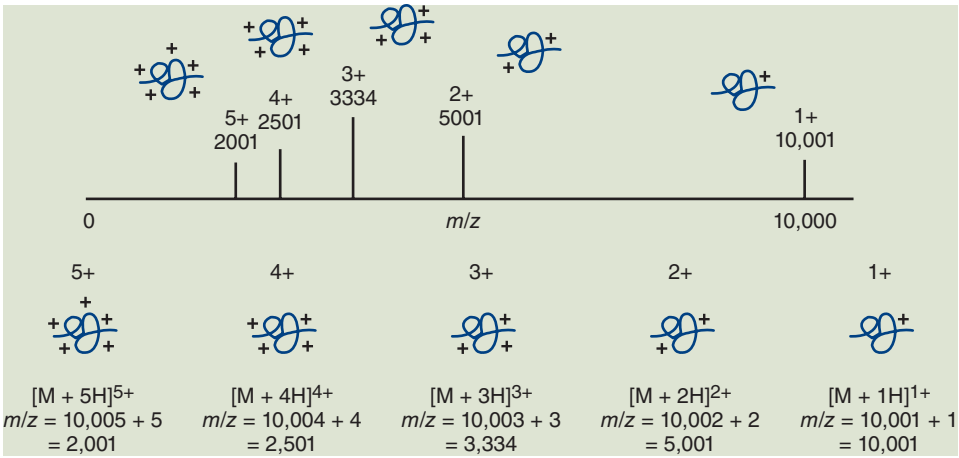


Figure 4.41 A theoretical protein with a molecular weight of 10,000 can be multiply charged, which will generate numerous peaks. A mass spectrometer with a relatively small mass range can still detect the multiply charged ions since the *m/z* is reduced.

Matrix-Assisted Laser Desorption Ionization

Matrix-Assisted Laser Desorption Ionization - Time of Flight (MALDI-TOF) ionization is used for the analysis of biomolecules such as peptides and proteins. Protein samples are mixed with an appropriate matrix solvent and spotted onto a stainless-steel plate. The solvent is dried and the plate is introduced into the vacuum system of the MALDI-TOF analyzer. As shown in **Figure 4.42**, a laser pulse irradiates the sample, causing desorption and ionization of both the matrix and the sample. Because the monitored mass spectral range is high (>500 Da), the ionization of the low-molecular-weight matrix can be readily distinguished from high-molecular-weight peptides and proteins and does not interfere with the assay of the protein. Ions from the sample are focused into the mass spectrometer. The time required for an ion with a given mass to reach the detector is a nonlinear function of the mass, with larger ions requiring more time than smaller ions. The molecular weight of the proteins acquired by mass spectrum is used to determine the identity of the sample. For very large proteins, samples can be pretreated with trypsin, which cleaves peptide bonds after lysine and arginine, to produce lower-molecular-weight fragments that can then be measured.

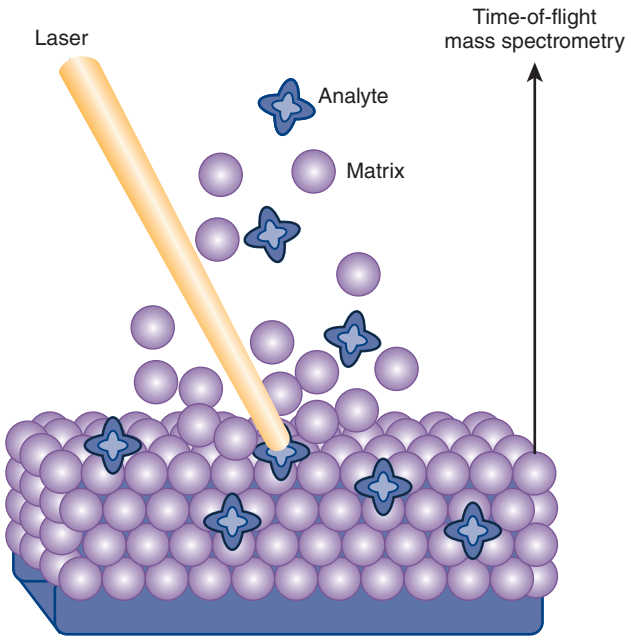


Figure 4.42 Sample desorption process prior to matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) analysis.

Diagram courtesy of Stanford Research Systems, Sunnyvale, CA.

Mass Spectrometer Analyzer

The actual measuring of the *m/z* occurs when the gas phase ions pass into the mass analyzer. Mass analyzers generate electric fields that can manipulate the charged molecules to sort them according to their *m/z*.

Quadrupole

A diagram of a quadrupole MS is shown in **Figure 4.43**. The quadrupole is the most common mass analyzer in use today. The electric field on the two sets of diagonally opposed rods allows only ions of a single selected *m/z* value to pass through the analyzer to the detector. All other ions are deflected into the rods. The rods can be scanned from low to high mass to allow ions of increasing mass to form stable sinusoidal orbits and traverse the filtering sector. This technique

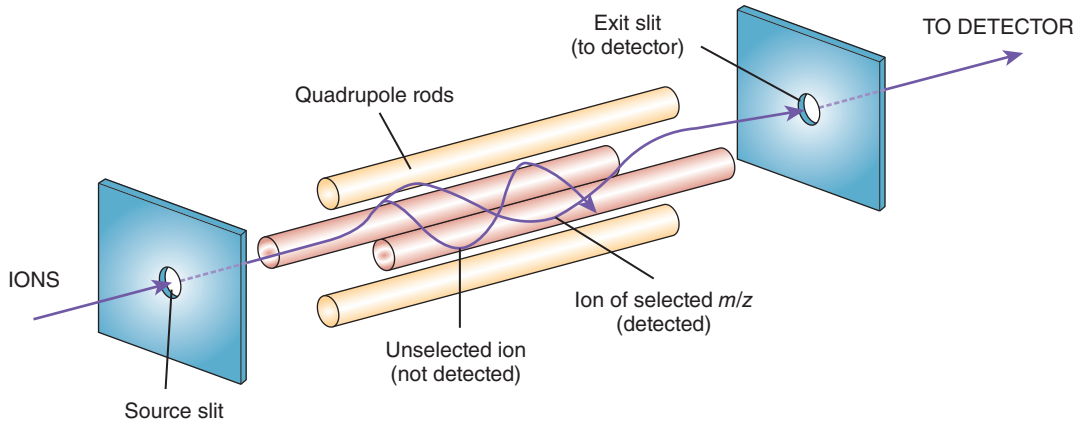


Figure 4.43 Single quadrupole mass spectrometer.

will generate a full-scan mass spectrum. Alternatively, specific masses can be selected to monitor a few target analytes. This technique is called selected ion monitoring (SIM) and it allows for a longer dwell time (time spent monitoring a single ion) and therefore higher sensitivity.[22] A full scan provides more information than does SIM since ions not specifically selected in SIM are not detected. Therefore, a full scan would be preferable for general unknown screening, while SIM analysis is more suitable for target compound analysis.

Ion Trap

The ion trap can be thought of as a modified quadrupole. A linear ion trap employs a stopping potential on the end electrodes to confine ions along the two-dimensional axis of the quadrupoles. In a three-dimensional ion trap, the four rods, instead of being arranged parallel to each other, form a three-dimensional sphere in which ions are "trapped." In all ion traps, after a period of accumulation, the electric field adjusts to selectively destabilize the trapped ions, which are mass-selectively ejected from the cavity to the detector based on their *m/z*.[22] The unique feature of ion trap MSs is that they trap and store ions generated over time, effectively concentrating the ions of interest and yielding a greater sensitivity.

Tandem Mass Spectrometry

Tandem MS (GC/MS/MS and LC/MS/MS) can be used for greater selectivity and lower detection limits. A common form of MS/MS is to link three quadrupoles in series; such an instrument is referred to as a triple quad (**Figure 4.44**). Generally, each quadrupole has a separate function.[23] Following an appropriate ionization method, the first quadrupole (Q1) is used to scan across a preset *m/z* range and select an ion of interest. The second quadrupole (Q2) functions as a collision cell. In a process called collision induced dissociation, the ions are accelerated to high kinetic energy and allowed to collide with neutral gas molecules (usually nitrogen, helium, or argon) to fragment the ions. The single ion that passed through the first analyzer is called the precursor (or parent) ion while the ions formed during fragmentation of the precursor ions are called product (or daughter) ions. The third quadrupole (Q3) serves to analyze the product ions generated in Q2. This last quadrupole can be set to scan all of the product ions to produce a full product ion scan or to selectively allow one or more of these product ions through to the detector in a process called selected reaction monitoring. Various scanning modes commonly used in a triple quad are shown in **Figure 4.45**. In some triple quad instruments, the third quadrupole can also function as a linear ion trap to add further sensitivity to MS/MS.

High-Resolution MS

Newer technologies utilizing high-resolution mass spectrometers based on time-of-flight (TOF) or Orbitrap (Thermo Fisher) technologies have gained popularity in recent years. These instruments can measure large numbers of analytes simultaneously in complex biological matrices and have been particularly useful for drug screening applications.[24] Compared with traditional or "nominal resolution" mass spectrometers that determine masses to approximately 0.5 Da, high-resolution instruments such as TOF and Orbitrap mass spectrometers operate at resolutions that allow the exact mass of an unknown compound to be calculated to approximately 0.001 to 0.0001 Da. The resolution of a mass spectrometer is defined as the mass of a given compound divided by the width of the corresponding peak and is commonly

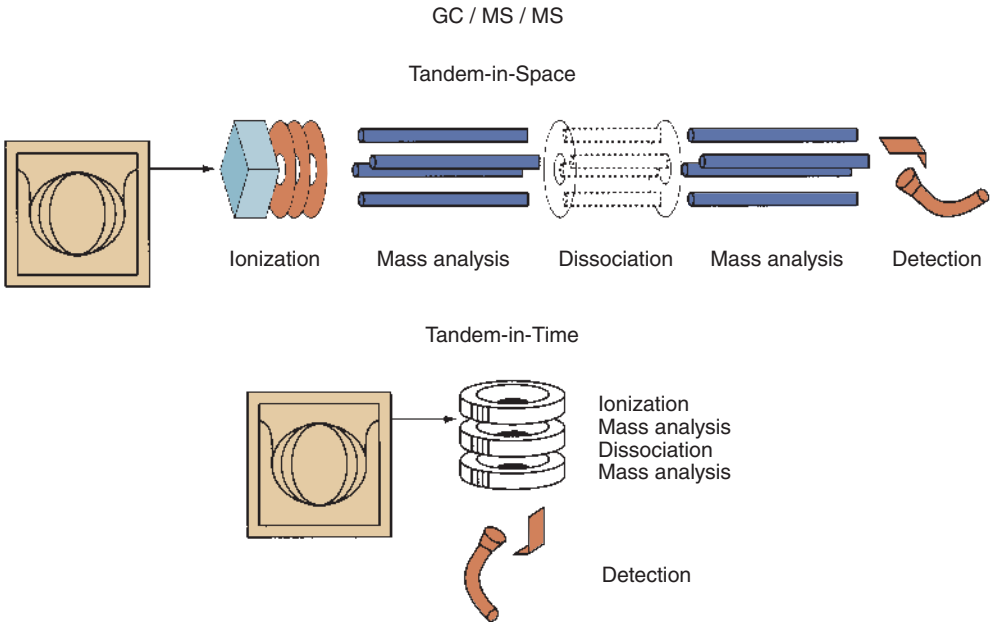


Figure 4.44 Triple quadrupole mass spectrometer.

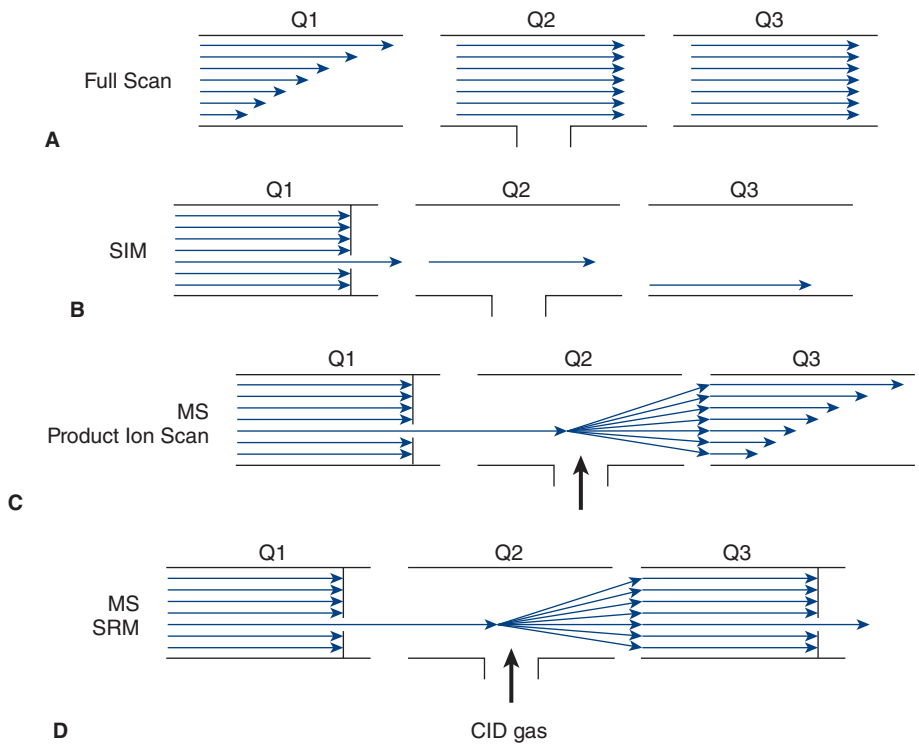


Figure 4.45 Scanning modes used in a triple quadrupole mass spectrometer. **(A)** Full-scan mass spectrometry (MS) detects all ions. **(B)** Selected ion monitoring (SIM) detects ions of one selected *m/z*. **(C)** Product ion scans select ions of one *m/z* in Q1 to pass onto Q2, the collision cell, where the ion is fragmented. All ion fragments are allowed to pass through to the detector. **(D)** Selected reaction monitoring (SRM) is similar to the product ion scan, but only fragments of one selected *m/z* are allowed to pass onto the detector. Both **(C)** and **(D)** are examples of tandem mass spectrometry (MS/MS). CID, collision-induced dissociation.

designated by the term full width at half maximum (FWHM) (**Figure 4.46A**). TOF mass spectrometers achieve resolutions of 10,000 to 50,000 FWHM utilizing the principle that given the same kinetic energy, lighter ions travel faster than heavier ions. By measuring the time taken for an ion to traverse the

flight tube and hit the detector, the *m/z* ratio can be calculated (**Figure 4.46B**). Orbitrap mass spectrometers operate on a different principle. With Orbitrap instruments, ions are injected tangentially to the electric field between the outer barrellike electrode and the inner spindle-like electrode, and the stable orbit achieved is proportional to the *m/z* value (**Figure 4.46C**). Orbitrap mass spectrometers can achieve 100,000 to 250,000 FWHM.

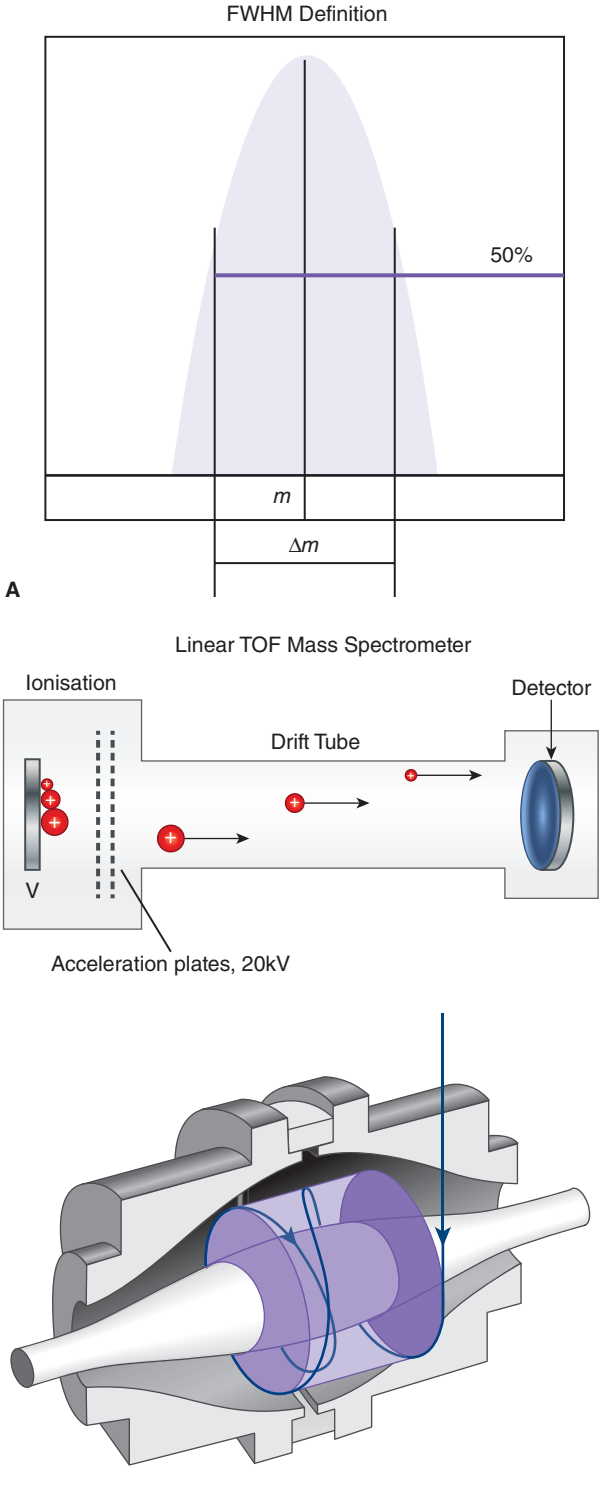


Figure 4.46 **(A)** Pictorial representation of the full width at half maximum (FWHM) definition of mass resolution. **(B)** Diagram of the principle of time-of-flight (TOF) mass spectrometers. **(C)** Diagram of Orbitrap mass spectrometry.

(A) Reprinted with permission from Thermo Fisher Scientific. (B) © Jones & Bartlett Learning, (C) Reprinted with permission from Thermo Fisher Scientific.

Detector

The most common means of detecting ions employs an electron multiplier. In this detector, a series of dynodes with increasing potentials are linked. When ions strike the first dynode surface, electrons are emitted. These electrons are attracted to the next dynode where more secondary electrons are emitted due to the higher potential of subsequent dynodes. A cascade of electrons is formed by the end of the chain of dynodes, resulting in overall signal amplification on the order of 1 million or greater.[23]

Applications of MS in the Clinical Laboratory

Small Molecule Analysis

Mass spectrometers coupled to GC or LC can be used not only for the identification and quantitation of compounds but also for structural information and molecular weight determination (high-resolution MS).[25] GC/MS systems are widely used for measuring drugs of abuse in urine toxicology confirmations. Drugs and metabolites must be extracted from body fluids and typically reacted with derivatizing reagents to form compounds that are more volatile for the GC process. Computerized libraries and matching algorithms are available within the instrument to compare mass spectral results of an unknown substance obtained from a sample to the reference library.

Increasingly, LC/MS (including LC/MS/MS) technology is taking its place alongside GC/MS in clinical laboratories. LC offers a number of advantages over GC. Typically, LC requires less extensive extraction procedures, and derivatization is rarely used, saving time and expense. Solid-phase extraction columns can be incorporated directly into the injector for online purifications (**Figure 4.47**). In addition, polar and heat-labile compounds fare better in LC.[22] However; the chromatography itself in LC can be somewhat less robust than in GC, resulting in wider peaks and more variable retention times and potentially requiring maintenance that is more frequent. Another disadvantage of LC/ MS is the less reproducible mass spectra, as mentioned earlier. One drawback to

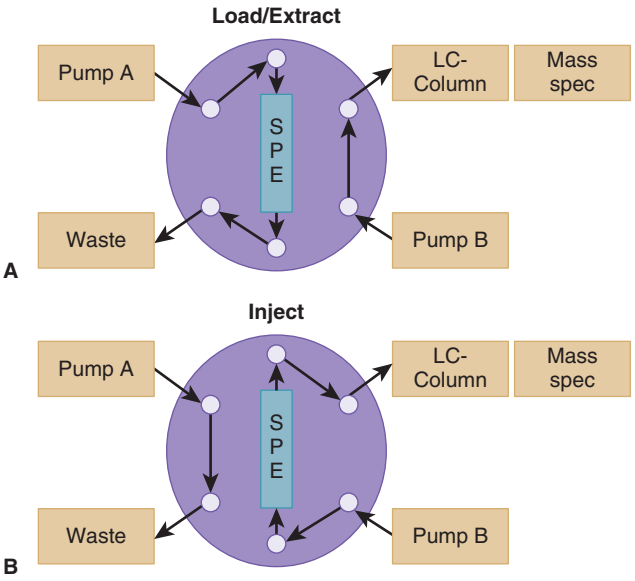


Figure 4.47 Diagram of an online extraction method for LC-MS/MS. An LC injector is depicted. **(A)** In the "load/extract" position, samples are loaded using pump A onto the extraction column while the LC mobile phase is directed to the main analytical column using pump B. **(B)** In the "inject" position, the LC mobile is directed using pump B through the extraction column eluting any compounds directly to the main analytical column. SPE = solid phase extraction column.

implementing LC or GC/MS in clinical laboratories is the long run times associated with the chromatographic separation. For low-volume testing, this is not an issue; however, for medium- or high-volume testing, it remains a significant challenge. To overcome this, laboratories can utilize technologies such as ultra-performance liquid chromatography (UPLC) or multiplexing. With UPLC, both the column and the pumps are robust enough to handle very high pressures (>600 bar), allowing faster flow rates, which in turn dramatically reduces the run time.

For very high-volume testing, several HPLC (or UPLC) pumps can be connected to a single mass spectrometer, which is referred to as multiplexing. This technology works because analyte peaks are only typically approximately 5 to 30 seconds in width, and in theory, that is the only time the LC flow needs to be in-line with the mass spectrometer. LC separations typically require much longer gradients for good resolution, and additionally, the column must be equilibrated before the next run, which takes several minutes. In multiplexing, the LC pumps are staggered so that once the peak of interest has been analyzed by the mass spectrometer, the flow is diverted to waste. At that point, the flow from the next LC pump is placed in-line with the mass spectrometer and the cycle repeats. Multiplexing can reduce the run time by a factor of the number of LC pumps and is well suited for high-volume targeted analysis but not for untargeted screening. For example, if the total run time with a single LC pump takes 8 minutes, then a system with four multiplexed LC pumps would take only 2 minutes to measure the same analyte (see **Figure 4.48**). Utilizing technologies such as UPLC and multiplexing has allowed clinical laboratories to measure very high-volume analytes using MS.

High-resolution mass spectrometry has applications in unknown toxicology analysis whereby the drug or metabolite is unknown to the testing laboratory. With soft ionization (i.e., minimal fragmentation of the parent ion), high-resolution MS enables putative identification of drugs by identifying the compound's molecular weight to three or four decimal points. From this information, the exact molecular formula can be deduced and a search of the literature can uncover compounds that have that molecular formula. By definition, these compounds are stereo or structure isomers of each other. **Figure 4.49** shows the chemical formula of three compounds that have the same *nominal* molecular weight of *m/z* 285 but different *exact* molecular weights. Using high-resolution mass spectrometry, the three compounds can be differentiated from each other. This feature is particularly useful for identification of "designer drugs." These are drugs that have similar pharmacologic activity but differ from well-known drugs with the substitution of a functional group. "Bath salts" (synthetic cathinones) are a group of designer amines that are difficult to detect using conventional targeted mass spectrometric techniques. Designer fentanyl compounds have been produced for many years.

Besides its use in toxicology, LC/MS also has great potential for measuring low-level and mixed-polarity analytes such as vitamin D, testosterone, and immunosuppressant drugs due to its superior sensitivity and specificity over immunoassays. In addition, LC/MS has the advantage of being able to detect multiple analytes (such as a panel of drugs or a series of metabolites) in one run. LC/MS is free from the antibody interferences seen in immunoassays, although LC/MS has its own type of interference in the form of ion suppression. This effect is seen when a co-eluting chemical in the sample prevents a compound of interest from being ionized, thereby reducing or eliminating its signal. LC/MS also requires highly skilled operators and is not nearly as automated as immunoassay instruments.

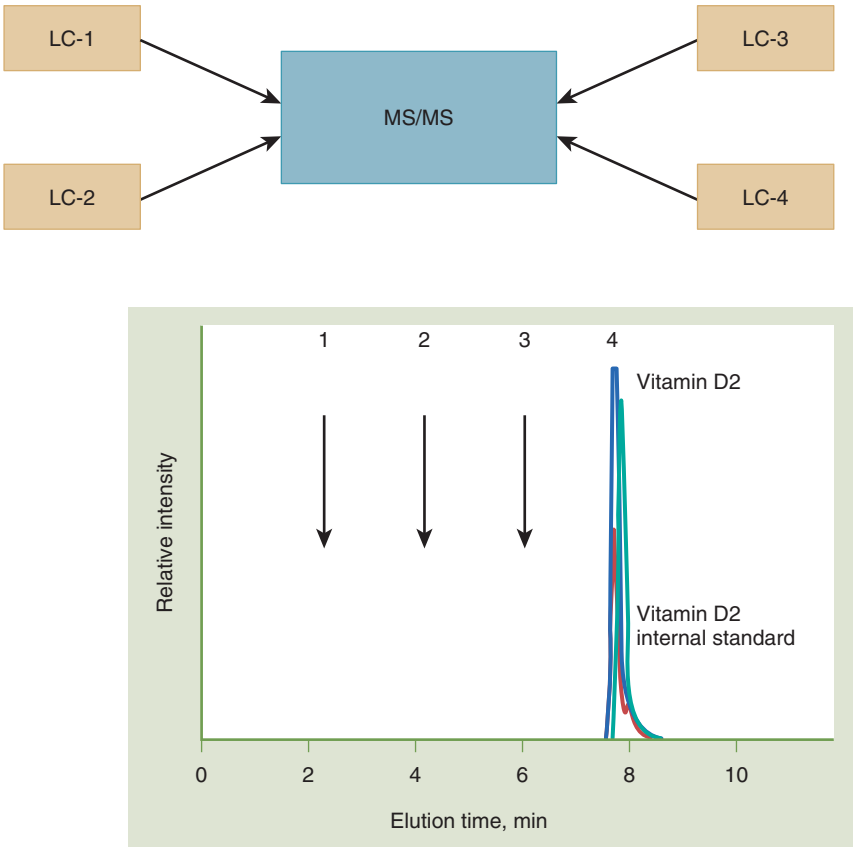


Figure 4.48 Multiplexing of four liquid chromatograms into one mass spectrometer. See text for details.

Mass Spectrometry in Proteomics and Pathogen Identification

The next generation of biomarkers for human diseases will be discovered using techniques found within the research fields of genomics and proteomics. Genomics uses the known sequences of the entire human genome for determining the role of genetics in certain human diseases. Proteomics is the investigation of the protein products encoded by these genes. Protein expression is equal to and, in many cases, more important for disease detection than genomics because these products determine what is currently occurring within a cell, rather than the genes, which indicate what a cell might be capable of performing. Moreover, many (posttranslational) changes can occur to the protein, as influenced by other proteins and enzymes that cannot be easily predicted by knowledge at the genomic level.

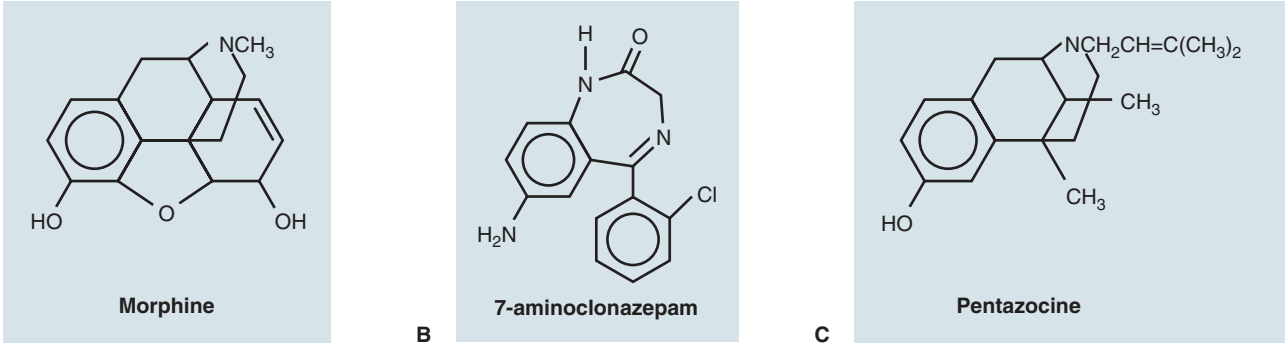


Figure 4.49 Chemical structure of **(A)** morphine, $C_{17}H_{19}NO_3$, *m/z* 285.3377; **(B)** 7-aminoclonazepam (clonazepam metabolite), $C_{15}H_{12}ClNO_3$, *m/z* 285.7283; and **(C)** pentazocine, $C_{19}H_{27}NO$, *m/z* 285.2093.

A "shotgun" approach is often used in the discovery of new biochemical markers. The proteins from samples (e.g., serum, urine, and tissue extract) from normal individuals are compared with those derived from patients with the disease being studied. Before being analyzed on the mass spectrometer, the complex mixture of proteins must first be separated using chromatographic techniques. Proteins can be trypsin digested and separated by HPLC, or techniques such as two-dimensional electrophoresis can be used to separate proteins into individual spots or bands.

Proteins that only appear in either the normal or diseased specimens are further studied. For proteomic analysis, computer programs are available that digitally compare gels to determine spots or areas that are different. When candidate proteins have been found, the spots can be isolated and subjected to traditional trypsin digestion followed by HPLC-MS/MS or to MALDI-TOF MS to identify the protein and possibly any posttranslational modifications that may have occurred.

Pathogen Identification

MALDI-TOF MS is increasingly being used for the identification of pathogens in modern microbiology laboratories.[26] Isolated bacterial or fungal colonies can be directly spotted onto the MALDI plate and ionized, which results in a protein "fingerprint" of the species. This protein "fingerprint" is composed of mainly ribosomal proteins and can be compared with a digital database of species (**Figure 4.50**). Because MALDI-TOF MS requires inexpensive reagents and only takes 5 to 10 minutes per run, it is significantly cheaper and faster than traditional automated biochemical identification techniques.

Mass Spectrometry at the Point of Care

Analytical instrumentation for mass spectrometry is advancing to the point that point-of-care testing applications may be available in the near future. One

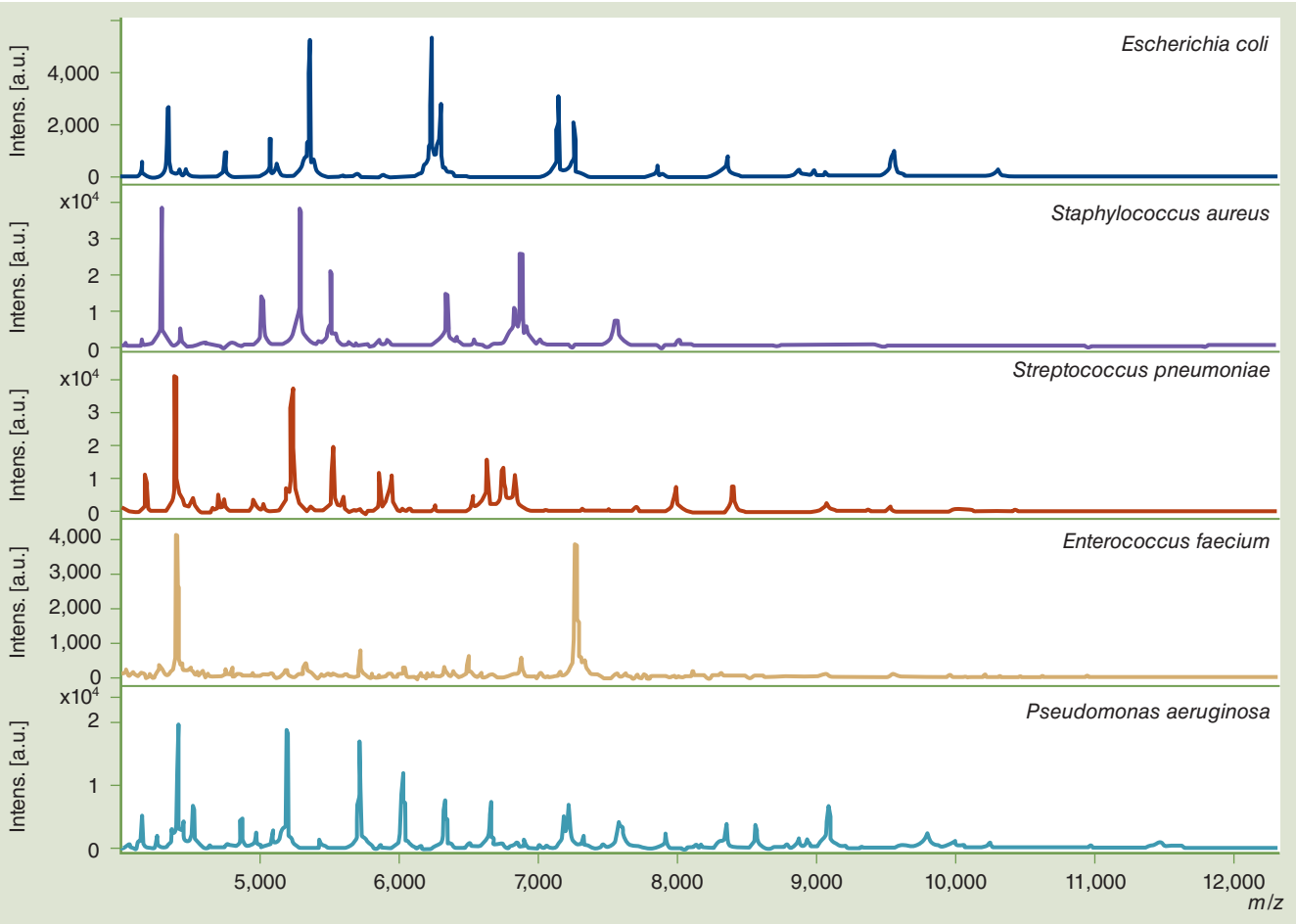


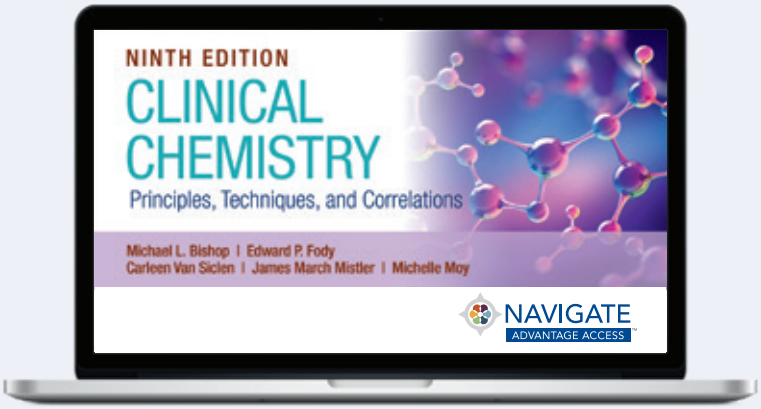
Figure 4.50 Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) spectral fingerprints of five different bacterial species.

Reproduced from Carbonnelle E, Mesquita C, Bille E, et al. *Clin Biochem.* 2011;44:104–109.

area that has received attention is the use of portable mass spectrometry to analyze biopsy specimens intraoperatively.[27] Results of such testing may help to determine if a surgeon has removed all of the tumorous tissue and can terminate the surgery or if the surgeon has to continue the procedure to search for more tumor. For this application to be useful, testing must be rapid and the equipment must be small and simple enough to be useful in the operating room itself.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Christian GD, O'Reilly JE. *Instrumental Analysis*. 2nd ed. Boston, MA: Allyn and Bacon; 1986.
2. Willard HH, Merritt LL, Dean JA, et al. *Instrumental Methods of Analysis*. Belmont, CA: Wadsworth; 1981.
3. Coiner D. *Basic Concepts in Laboratory Instrumentation*. Bethesda, MD: ASMT Education and Research Fund; 1975–1979.
4. Ingle JD, Crouch SR. *Spectrochemical Analysis*. Upper Saddle River, NJ: Prentice-Hall; 1988.
5. Guilbault GG. *Practical Fluorescence, Theory, Methods and Techniques*. New York, NY: Marcel Dekker; 1973.
6. Kricka LJ. Chemiluminescent and bioluminescent techniques. *Clin Chem*. 1991;37:1472.
7. Wild D. *The Immunoassay Handbook*. London, UK: Macmillan Press; 1994.
8. Svelto O. *Principles of Lasers*. 2nd ed. New York, NY: Plenum Press; 1982.
9. *Coulter Hematology Analyzer: Multidimensional Leukocyte Differential Analysis*. Vol. 11(1). Miami, FL: Beckman Coulter; 1989.
10. Lide DR. *CRC Handbook of Chemistry and Physics*. 93rd ed. Boca Raton, FL: CRC Press; 2012–2013.
11. Rifai N. *Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics*. 8th ed. Saunders; 2018.
12. Heiger DN. *High-Performance Capillary Electrophoresis: An Introduction*. 2nd ed. Waldbronn, Germany: Hewlett Packard; 1992.
13. Monnig CA, Kennedy RT. Capillary electrophoresis. *Anal Chem*. 1994;66:280R.
14. Mariani S, Minunni M. Surface plasmon resonance applications in clinical analysis. *Anal Bioanal Chem*. 2014;406:2303–2323.
15. Wilson JL, Scott IM, McMurry JL. Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. *Biochemistry and molecular biology education*. 2010;38:400–407.
16. Parris NA. *Instrumental Liquid Chromatography: A Practical Manual on High Performance Liquid Chromatographic Methods*. New York, NY: Elsevier; 1976.
17. Jurk H. *Thin-Layer Chromatography. Reagents and Detection Methods*. Vol. 1a. Weinheim, Germany: Verlagsgesellschaft; 1990.
18. Bender GT. *Chemical Instrumentation: A Laboratory Manual Based on Clinical Chemistry*. Philadelphia, PA: WB Saunders; 1972.
19. Horváth C. *High Performance Liquid Chromatography, Advances and Perspectives*. New York, NY: Academic Press; 1980.
20. Constantin E, Schnell A. *Mass Spectrometry*. New York, NY: Ellis Horwood; 1990.
21. Kebarle E, Liang T. From ions in solution to ions in the gas phase. *Anal Chem*. 1993;65:972A.
22. Karasek FW, Clement RE. *Basic Gas Chromatography: Mass Spectrometry*. New York, NY: Elsevier; 1988.
23. Levine B. *Principles of Forensic Toxicology*. Washington, DC: AACC Press; 2006.
24. Jiwan JL, Wallemacq P, Hérent MF. HPLC—high resolution mass spectrometry in clinical laboratory? *Clin Biochem*. 2011;44:136–147.
25. Siuzdak, G. *The Expanding Role of Mass Spectrometry in Biotechnology*. San Diego, CA: MCC Press; 2006.
26. Carbonnelle E, Mesquita C, Bille E, et al. MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory. *Clin Biochem*. 2011;44:104–109.
27. Banerjee S. Empowering clinical diagnostics with mass spectrometry. *ACS Omega*. 2020;5(5):2041–2048. https://pubs.acs.org/doi/10.1021/acsomega.9b03764

CHAPTER 5

Clinical Chemistry Automation

Nikola A. Baumann and Brooke M. Katzman

CHAPTER OUTLINE

KEY TERMS

Affinity
Analytic
Antibody
Antigen
Automation
Avidity
Barcode
Bidirectional interface
Centrifugal analysis
Closed tube sampling
Competitive immunoassay
Consumables
Continuous flow
Discrete analysis
Dry chemistry slide
Endpoint
Epitope
Hapten
Heterogeneous
Homogeneous
Hook effect
Kinetic
Middleware
Modular analyzer
Monoclonal
Nephelometry
Noncompetitive immunoassay
Polyclonal
Post-analytic
Pre-analytic
Probe
Random access
Robotics
Rotor
Solid phase
Total laboratory automation
Turbidimetry

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- State the components of total laboratory automation.
- Define continuous flow, discrete analysis, random access, and throughput.
- Name three basic approaches to sample analysis used by automated analyzers.
- Explain the major steps in automated analysis.

- Provide examples of commercially available chemistry analyzers and modular systems.
- Contrast the different automated analysis platforms used by instrument manufacturers.
- Differentiate between turbidimetry and nephelometry.
- Compare and contrast the types of labels used in immunoassays.
- Diagram competitive immunoassays and noncompetitive immunoassays.
- Discuss the benefit of using paramagnetic particles in automated immunoassay systems.
- List examples of the interfering substances in immunoassays.

The modern clinical chemistry laboratory uses a high degree of **automation**. Clinical laboratory automation is characterized by automating, using **robotics** and instrumentation to do tasks that were traditionally manually performed by humans. Many steps in the total testing process can now be performed automatically, permitting the laboratorians to focus on manual or technical processes that increase both efficiency and capacity. The total testing process can be divided into three major phases: **pre-analytic**, **analytic**, and **post-analytic**. These phases correspond to sample processing/preparation, analyte measurement, and data management/sample storage, respectively. Substantial innovation and improvements have occurred in all three areas during the past decade. This chapter focuses on the foundation of automated clinical laboratories beginning with the principles and components of automated chemistry and immunoassay analyzers. Considerations specific to automated immunoassay methods are discussed.

Automated chemistry and immunoassay analyzers allow testing to occur with minimal operator intervention and can be stand-alone instruments or integrated into modules. Steps in the pre-analytic phase can be automated by either stand-alone systems or systems that connect to the analytical modules with tracks that transport the specimens. Post-analytic refrigerated storage and retrieval systems can also be integrated with pre-analytic and analytic automation. These comprehensive automated systems are referred to as **total laboratory automation** (TLA) systems. The chapter will provide an overview of TLA and its functionality.

History and Evolution of Automated Analyzers

Following the introduction of the first automated analyzer by Technicon in 1957, automated instruments proliferated from many manufacturers.[1] This first "AutoAnalyzer" (AA) was a continuous flow, single-channel, sequential batch analyzer capable of providing a single test result on approximately 40 samples per hour. In **continuous flow**, liquids (reagents, diluents, and samples) are pumped through a system of continuous tubing. Samples are introduced in a sequential manner, following each other through the same network and reaction path. A series of air bubbles at regular intervals serve to both separate and clean the tubing. Continuous flow also assists the laboratory that needs to run many samples requiring the same procedure. The more sophisticated continuous flow analyzers used parallel single channels to run multiple tests on each sample. The major drawbacks that contributed to the eventual demise of traditional

CASE STUDY 5.1, PART 1

While Mía was training on the Roche Cobas, the following results were obtained:

K	Na	Cl	CO_2	TBIL	DBILI	AST
5.6H	131L	94	14L	5.2H	3.2H	1078H
>I.H					>I.L	>Abs

© Ariel Skelley/DigitalVision/Getty Images.

continuous flow analyzers (i.e., AA) in the marketplace were significant carryover problems and wasteful use of continuously flowing reagents. Technicon's (now Siemens) answer to these problems was a noncontinuous flow discrete analyzer (the RA1000), using random access fluid to reduce surface tension between samples/reagents and their tubing, thereby reducing carryover. Later, the Chem 1 was developed by Technicon to use Teflon tubing and Teflon oil, virtually eliminating carryover problems. The Chem 1 was a continuous flow analyzer but only remotely comparable to the original continuous flow principle.

The next generation of Technicon instruments to be developed was the Simultaneous Multiple Analyzer (SMA) series. SMA-6 and SMA-12 were analyzers with multiple channels (for different tests), working synchronously to produce 6 or 12 test results simultaneously at the rate of 360 or 720 tests per hour. It was not until the mid-1960s that these continuous flow analyzers had any significant competition in the marketplace.

In 1970, the first commercial centrifugal analyzer was introduced as a spin-off technology from NASA space research. Dr. Norman Anderson developed a prototype in 1967 at the Oak Ridge National Laboratory as an alternative to continuous flow technology, which as noted earlier had significant carryover problems and costly reagent waste. He wanted to perform analyses in parallel and also take advantage of advances in computer technology. The second generation of these instruments (1975) was more successful because of miniaturization of computers and advances in the polymer industry for high-grade, optical plastic cuvettes. **Centrifugal analysis** uses the force generated by centrifugation to transfer and then contain liquids in separate cuvettes for measurement at the perimeter of a spinning **rotor**. Centrifugal analyzers are capable of running multiple samples, one test at a time, in a batch. Batch analysis is their major advantage because reactions in all cuvettes are read virtually simultaneously, so that it takes the same amount of time to run a full rotor of about 30 samples as it would take to run just a few. Laboratories with a high-volume workload of individual tests for routine batch analysis may use these instruments. Again, each cuvette must be uniformly matched to each other to maintain quality handling of each sample. The Cobas-Bio (Roche Diagnostics), with a xenon flash lamp and longitudinal cuvettes,[2] and the Monarch (Fortress Diagnostics), with a fully integrated walk-away design, were two of the most widely used centrifugal analyzers.

Another major development that revolutionized clinical chemistry instrumentation occurred in 1970 with the introduction of the Automatic Clinical Analyzer (ACA, DuPont [now Siemens]). It was the first noncontinuous flow, discrete analyzer, as well as the first instrument to have **random access** capabilities, whereby stat specimens could be analyzed out of sequence on an as-needed basis. Plastic test packs, positive patient identification, and infrequent calibration were among the unique features of the ACA. **Discrete analysis** is the separation of each sample and accompanying reagents in a separate container (i.e., reaction chamber, cuvette, well). Discrete analyzers have the capability of running multiple tests from one sample at a time or multiple samples one test at a time. They are the most popular and versatile analyzers and have almost completely replaced continuous flow and centrifugal analyzers. However, because each sample is in a separate reaction container, uniformity of quality must be maintained in each cuvette so that a particular sample's quality is not affected. The high-volume chemistry and immunoassay analyzers listed in **Table 5.1** are examples of contemporary discrete analyzers with random access capabilities.

Other major milestones were the introduction of thin film analysis technology in 1976 and the production of the Kodak Ektachem (now VITROS) Analyzer (Ortho Clinical Diagnostics) in 1978. This instrument was the first to use microsample volumes and reagents on slides for *dry chemistry analysis* and to incorporate computer technology extensively into its design and use. This dry slide technology is still in use today on the VITROS analyzer and offers several unique advantages that will be discussed below.

Since 1980, several primarily discrete analyzers have been developed that incorporate such characteristics as ion-selective electrodes (ISEs), fiberoptics, polychromatic analysis, continually more sophisticated computer hardware and software for data handling, and larger test menus. The differences among the manufacturers' instruments, operating principles, and technologies are less distinct now than they were in the beginning years of laboratory automation.

Driving Forces and Benefits of Automation

The pace of changes in current routine chemistry analyzers and the introduction of new ones has slowed considerably. Certainly, analyzers are faster and easier to use as a result of continuous reengineering and electronic refinements. Methods are more precise,

Table 5.1 Summary of Features for Selected High Volume Chemistry and Immunoassay Analyzers

	Abbott Diagnostics		Beckman Coulter, Inc.		Ortho Clinical Diagnostics		Roche Diagnostics	Siemens Healthineers
Name of Instrument	**Architect c16000**	**Architect i4000SR**	**AU5800**	**UniCel DxI 8000**	**VITROS 4600**	**VITROS 3600**	**Cobas 8000 Modular**	**Dimension Vista 1500 Intelligent Lab System**
Type of Instrument	chemistry	immunoassay	chemistry	immunoassay	chemistry	immunoassay	combination chemistry/ immunoassay	combination chemistry/ immunoassay
First year sold in the United States	2007	2007	2011	2003	2004	2008	2010	2006
Throughput (tests per hour, depends on test mix)	1800	400	2000–9800	400	600–700	189	9800	2000
No. of assays onboard simultaneously	68	50	54–216	50	125	31	283	>100
No. of open channels	220	0	76	0	20	0	10	10
No. of ion-selective electrode channels	3	N/A	3	N/A	3	N/A	3	3
Detection/ Separation Methodology	photometric, potentiometric, turbidimetric/ N/A	chemiluminescence/ magnetic particle	photometry, potentiometry/ N/A	chemiluminescence/ magnetic particle	photometry, potentiometry, colorimetric, turbidimetric/ N/A	chemiluminescence, enzyme immunoassay, direct enhanced chemiluminescence/ coated microwell	photometry, potentiometry (chemistry), electrochemiluminescence (immunoassay)/magnetic particle	chemiluminescence, enzyme immunoassay, ACMIA, EMIT, LOCI, PETINIA, NEPH/ none

Minimum sample volume aspirated (µL)	2	10	1	5	2	2	1	2
Dedicated pediatric sample cup/ dead volume (µL)	Yes/50	No/50	Yes/50	Yes/100	No/35	Yes/35	Yes/50	No/10
Short sample/clot/ interference detection	Yes/Yes/Yes	Yes/Yes/No	Yes/Yes/Yes	Yes/Yes/No	Yes/Yes/Yes	Yes/Yes/Yes	Yes/Yes/Yes	Yes/Yes/Yes
Onboard test automatic inventory	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Remote troubleshooting by modem	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Data obtained from Dabkowski B. Vendors toil and tinker to refine chemistry analyzers. *CAP Today*. 2011;July:56–71 and directly from vendors.

sensitive, and specific, although some of the same principles are found in today's instruments as in earlier models. Manufacturers have worked successfully toward automation with "walk-away" capabilities and minimal operator intervention.[3] Manufacturers have also responded to the physicians' desire to bring laboratory testing closer to the patient. The introduction of small, portable, easy-to-operate benchtop analyzers in physician office laboratories, as well as in surgical and critical care units that demand immediate laboratory results, has resulted in a hugely successful domain of point-of-care (POC) analyzers.[4] Another specialty area with a rapidly developing arsenal of analyzers is immunochemistry. Immunologic techniques for assaying drugs, specific proteins, tumor markers, and hormones have evolved to an increased level of automation. Instruments that use techniques such as fluorescence polarization immunoassay, **nephelometry**, and competitive and **noncompetitive immunoassays** with chemiluminescent detection have become popular in laboratories.

The most recent milestone in chemistry analyzer development has been the combination of chemistry and immunoassay into a single **modular analyzer**. Modular analyzers combining chemistry and immunoassay capabilities are now available from several vendors that meet the needs of mid- and high-volume laboratories (**Figure 5.1**).

Other forces are also driving the market toward more focused automation. Higher volumes of testing and faster turnaround times have resulted in fewer and more centralized core laboratories performing more comprehensive testing.[5] The use of laboratory panels or profiles has declined, with more diagnostically directed individual tests as dictated by recent policy changes from Medicare and Medicaid. Researchers have known for many years that chemistry panels only occasionally lead to new diagnoses in patients who appear healthy.[6] The expectation of quality results with higher accuracy and precision is ever present with the regulatory standards set by the Clinical Laboratory Improvement Amendments (CLIA), The Joint Commission (TCJ), the College of American Pathologists (CAP), and others. Intense competition among instrument manufacturers has driven automation into more sophisticated analyzers with creative technologies and unique features. Furthermore, escalating costs have spurred health care reform.

There are many advantages to automating procedures. One benefit is to increase the number of tests performed by one laboratorian in a shift. Labor is an expensive commodity, and staffing shortages in clinical laboratories are not uncommon. Through mechanization, the labor component devoted to any single test is minimized, and this effectively lowers the cost per test. A second benefit is minimizing the variation in results from one laboratorian to another. By standardizing the procedure, the coefficient of variation is lowered, and reproducibility is increased. Accuracy is then not dependent on the skill or workload of a particular operator on a particular day. This allows better comparison of results from day to day and week to week. Automation, however, cannot correct for deficiencies inherent in methodology. A third advantage is gained because automation eliminates the potential errors of manual analyses such as volumetric pipetting steps, calculation of results, and transcription of results. A fourth advantage accrues because instruments can use very small amounts of samples and reagents. This allows less blood to be drawn from each patient, and the use of small amounts of reagents decreases the cost

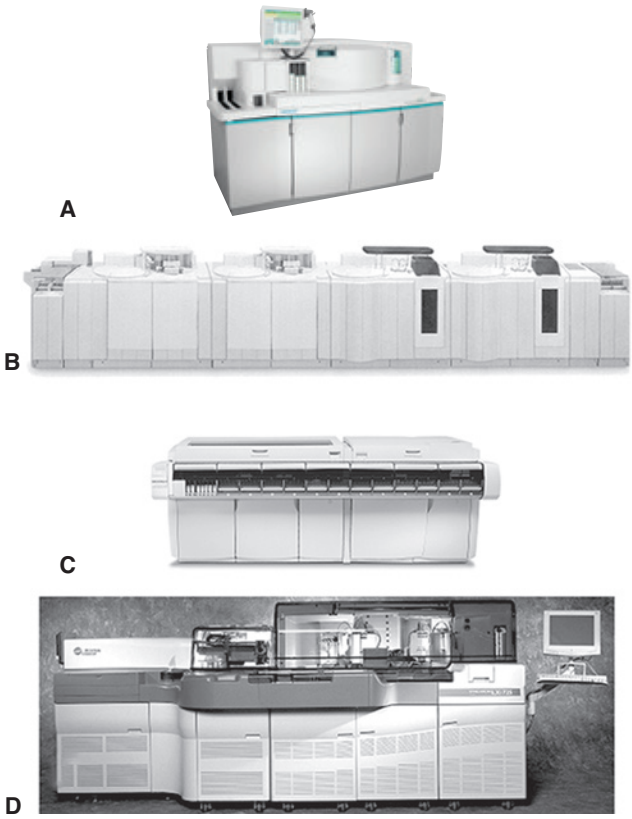


Figure 5.1 Modular chemistry/immunoassay analyzers. **(A)** Siemens Dimension Vista 500. **(B)** Roche Cobas modular analyzers. **(C)** Abbott ARCHITECT ci8200. **(D)** Beckman Coulter Synchron Lxi 725.

(A) Courtesy of Siemens Medical Solutions USA, Inc.; (B) Photograph courtesy of Roche Diagnostics; (C) ARCHITECT is trademark of Abbott or its related companies. Reproduced with permission of Abbott, © 2021. All rights reserved; (D) Photograph courtesy of Beckman Coulter, Inc.

of **consumables**. Consumables are components of a test that are "consumed" or used during analysis and then discarded, such as disposable cuvettes, pipette tips, and reagents. In addition, automation facilitates better space utilization through consolidation of analyzers.

Steps in Automated Analysis

The major processes performed by an automated analyzer can be divided into specimen identification and preparation, chemical reaction, and data collection and analysis. An overview of these operations is provided in **Table 5.2**.

Each step of automated analysis is explained in this section, and several different applications are discussed. Several instruments have been chosen because they have components that represent either common features used in chemistry instrumentation or a unique method of automating a step in a procedure. None of the representative instruments are completely described; rather, the important components are described in the text as examples.

Specimen Preparation and Identification

Most major automated chemistry and immunoassay analyzers can use the original labeled specimen collection tube (also known as the primary tube), after plasma or serum separation; the test tube itself can be used as the sample cup. Samples may also be assayed in labeled aliquot tubes or sample cups in scenarios where the specimen must be manipulated prior to analysis (i.e., small sample volume, filtration, dilution, concentration).

The sample must be properly identified, and its location in the analyzer must be monitored throughout the system. The approach that is commonly used today employs a **barcode** label affixed to the primary collection or aliquot tube. This barcode label contains patient demographics and also includes specific physician-ordered test requests for that patient sample. Specimen identification is traceable throughout the automated analytic process. Most automated analyzers read the linear 1D barcode, whereas the 2D data matrix barcode (also known as a QR code) is more commonly used for positive patient identification at the bedside.

Table 5.2 Summary of Chemistry Analyzer Operations

Identification and Preparation	
1. Sample identification	The analyzer will scan/read the bar code on the labeled primary specimen tube or an aliquot tube. This information can also be entered manually.
2. Determine test(s) to perform	Upon bar code scanning, test order information is retrieved from the LIS and automatically sent to the analyzer via an interface.
Chemical Reaction	
3. Reagent systems and delivery	One or more reagents can be dispensed into the reaction cuvette.
4. Specimen measurement and delivery	A small aliquot of the patient sample is introduced into the reaction cuvette.
5. Chemical reaction phase	The patient sample and reagents are mixed and incubated.
6. Measurement phase	Optical readings may be initiated before or after all reagents have been added.
7. Signal processing and data handling	The analyte concentration (result) is estimated from a calibration curve that is stored in the analyzer.
8. Send result(s) to Middleware/LIS	The analyzer sends results for the ordered tests via an interface to the Middleware/LIS and subsequently to the electronic medical record.

Operations generally occur in the order listed from 1 to 8. However, there may be slight variations in the order. Some steps may be deleted or duplicated. Most analyzers have the capability to dilute the sample and repeat the testing process if the analyte concentration exceeds the linear range of the assay. LIS, laboratory information system/software

Specimen Measurement and Delivery

Most instruments use either circular carousels or rectangular racks to hold primary/aliquot sample tubes or disposable sample cups in the loading or pipetting zone of the analyzer. The slots in the trays or racks are usually numbered to aid in sample identification. The trays or racks move automatically in one-position steps at preselected speeds. The speed determines the number of specimens to be analyzed per hour. As a convenience, the instrument can determine the slot number containing the last sample and terminate the analysis after that sample. The instrument's microprocessor holds the number of samples in memory and aspirates only in positions containing samples.

On the VITROS analyzer, sample cup trays are quadrants that hold 10 samples per quadrant in cups with conical bottoms. The four quadrants fit on a tray carrier (**Figure 5.2**). Although the tray carrier accommodates only 40 samples, more trays of samples can be programmed and then loaded in place of completed trays while tests on other trays are in progress. Roche Cobas analyzers can use five-position racks to hold samples (**Figure 5.3**). A modular analyzer can accommodate as many as 60 of these racks at one time.

A limitation of contemporary systems is that samples are uncapped and exposure of the sample to air can lead to sample evaporation, produce errors in analysis, as well expose the laboratorian to biohazards during the uncapping step. Evaporation of the sample may be significant and may cause the concentration of the constituents being analyzed to rise 50% in 4 hours.[7] For instruments measuring electrolytes, the carbon dioxide present in the samples will be lost to the atmosphere, resulting in low carbon dioxide values. Manufacturers have devised a variety of mechanisms to minimize this effect, such as, lid covers for trays and individual caps that can be pierced, which includes **closed tube sampling** from primary collection tubes.[8]

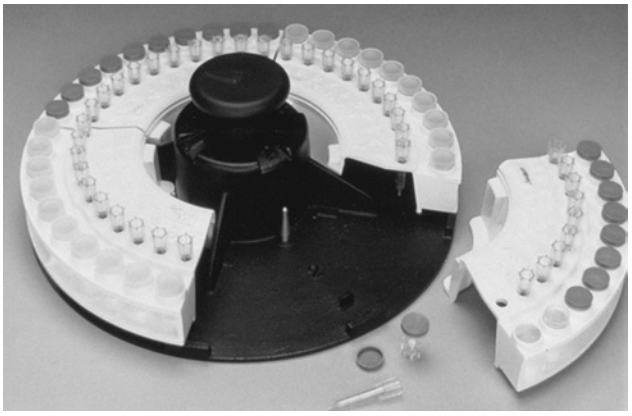


Figure 5.2 VITROS. The four quadrant trays, each holding 10 samples, fit on a tray carrier.

Photograph courtesy of Ortho-Clinical Diagnostics.

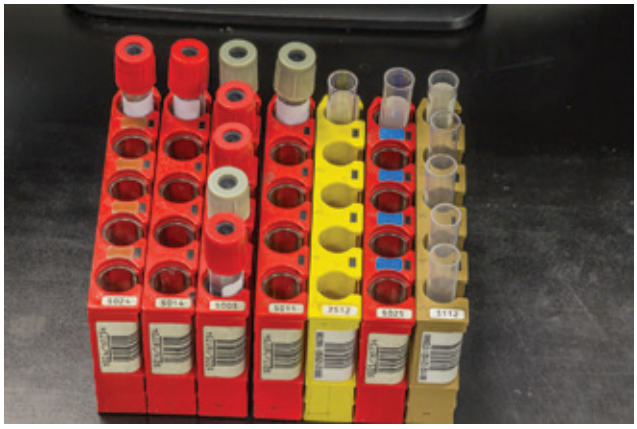


Figure 5.3 Roche five-position rack.

Aliquots of the patient sample are aspirated from the sample into a **probe**. When the instrument is in operation, the probe automatically dips into each sample cup and aspirates a portion of the liquid. The sample probe then delivers the sample to a discrete reaction chamber or cuvette on the analyzer. Sampling probes on instruments using specific sampling cups are programmed or adjusted to reach a prescribed depth in those cups to maximize the use of available sample. Those analyzers capable of aspirating sample from primary collection tubes usually have a parallel liquid-level sensing probe that will control entry of the sampling probe to a minimal depth below the surface of the serum, allowing full aliquot aspiration while avoiding clogging of the probe with serum separator gel or clot (**Figure 5.4**).

Certain pipettors use a disposable pipette tip and an air displacement syringe to measure and deliver both the patient sample and necessary reagents. When this is used, the pipettor may be reprogrammed to measure sample and reagent for batches of different tests comparatively easily. Besides eliminating the effort of priming the reagent delivery system with the new solution, no reagent is wasted or contaminated because nothing but a disposable pipette tip contacts it.

The cleaning of the probe and tubing after each dispensing to minimize the carryover of one sample into the next is a concern for many instruments. In some systems, the reagent or diluent is also dispersed into the cuvette through the same tubing and probe.

Figure 5.4 Dual sample probes of a chemistry analyzer. Note the liquid-level sensor to the left of probes.

Photograph courtesy of Roche Diagnostics.

Deionized water may be dispensed into the cuvette after the sample to produce a specified dilution of the sample and also to rinse the dispensing system.

If a separate probe or tip is used for each sample and discarded after use, carryover is not an issue. VITROS has a unique sample dispensing system. A proboscis presses into a tip on the sample tray, picks it up, and moves over the specimen to aspirate the volume required for the tests programmed for that sample. The tip is then moved over to the slide metering block. When a slide is in position to receive an aliquot, the proboscis is lowered so that a dispensed 10-µL drop touches the slide, where it is absorbed from the nonwetting tip. A stepper motor-driven piston controls aspiration and drop formation. The precision of dispensing is specified at ±5%.

In several discrete systems, the probe is attached by means of non-wettable tubing to precision syringes. The syringes draw a specified amount of sample into the probe and tubing. Then the probe is positioned over a cuvette and the sample is dispensed. The Roche/Hitachi chemistry analyzer used two sample probes to simultaneously aspirate a double volume of sample and deliver it into four individual test channels, all in one operational step (**Figure 5.5**). The loaded probes pass through a fine mist shower bath before delivery to wash off any sample residue adhering to the outer surface of the probes. After delivery, the probes move to a rinse bath station for cleaning the inside and outside surfaces of the probes.

Many chemistry analyzers use computer-controlled stepping motors to drive both the sampling and washout syringes. Every few seconds, the sampling probe enters a specimen container, withdraws the required volume, moves to the cuvette, and dispenses the aliquot with a volume of water to wash the probe. The washout volume is adjusted to yield the final reaction volume. If a procedure's range of linearity is exceeded, the system will retrieve the original sample tube, repeat the test using a portion of the original sample volume for the repeat test, and calculate a new result, taking the dilution into consideration.

Economy of sample size is a major consideration in developing automated procedures, but methodologies have limitations to maintain proper levels of sensitivity and specificity. The factors governing sample and reagent measurement are interdependent. Generally, if sample size is reduced, then either the size of the reaction cuvette and final reaction volume must be decreased or the reagent concentration must be increased to ensure sufficient color development for accurate photometric readings.

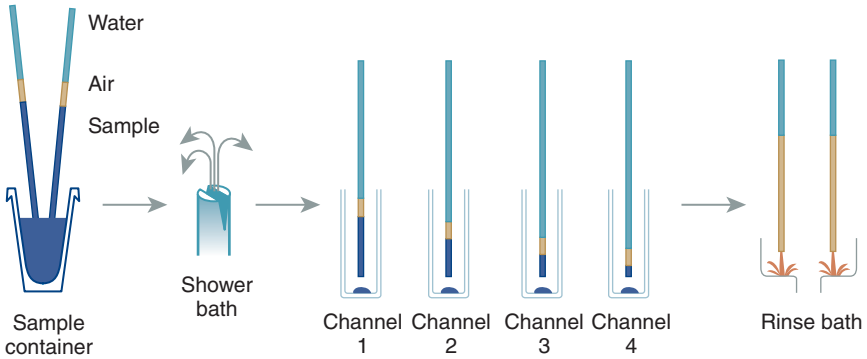


Figure 5.5 Sampling operation of the Hitachi 736 analyzer.

Courtesy of Roche Diagnostics.

Reagent Systems and Delivery

Reagents may be classified as liquid or dry systems for use with automated analyzers. Liquid reagents may be purchased in bulk volume containers or in unit dose packaging as a convenience for stat testing on some analyzers. Dry reagents are packaged in various forms. They may be bottled as lyophilized powder, which requires reconstitution with water or a buffer. Unless the manufacturer provides the diluent, the water quality available in the laboratory is important. A second and unique type of dry reagent is the multilayered **dry chemistry slide** for the VITROS analyzer (rebranded in 2001 as the VITROS MicroSlide technology). These slides have microscopically thin layers of dry reagents mounted on a plastic support. The slides are approximately the size and thickness of a postage stamp.

Reagent handling varies according to instrument capabilities and methodologies. Many test procedures use sensitive, short-lived working reagents; so contemporary analyzers use a variety of techniques to preserve them. One technique is to keep all reagents refrigerated until the moment of need and then quickly preincubate them to reaction temperature or store them in a refrigerated compartment on the analyzer that feeds directly to the dispensing area. Another means of preservation is to provide reagents in a dried, tablet form and reconstitute them when the test is to be run. A third is to manufacture the reagent in two stable components that will be combined at the moment of reaction. If this approach is used, the first component also may be used as a diluent for the sample. The various manufacturers often use combinations of these reagent-handling techniques.

Reagents also must be dispensed and measured accurately. Many instruments use bulk reagents to decrease the preparation and changing of reagents. Instruments that do not use bulk reagents have unique reagent packaging.

To deliver reagents, many discrete analyzers use techniques like those used to measure and deliver the samples. Syringes, driven by a stepping motor, pipette the reagents into reaction containers. Piston-driven pumps, connected by tubing, may also dispense reagents. Another technique for delivering reagents to reaction containers uses pressurized reagent bottles connected by tubing to dispensing valves. The computer controls the opening and closing of the valves. The fill volume of reagent into the reaction container is determined by the precise amount of time the valve remains open.

The VITROS analyzers use slides to contain their entire reagent chemistry system. Multiple layers on the slide are backed by a clear polyester support. The coating itself is sandwiched in a plastic mount. There are three or more layers: (1) a spreading layer, which accepts the sample; (2) one or more central layers, which can alter the aliquot; and (3) an indicator layer, where the analyte of interest may be quantified (**Figure 5.6**). The number of layers varies depending on the assay to be performed. The color developed in the indicator layer varies with the concentration of the analyte in the sample. Physical or chemical reactions can occur in one layer, with the product of these reactions proceeding to another layer, where subsequent reactions can occur. Each layer may offer a unique environment and the possibility to carry out a reaction comparable to that offered in a chemistry assay, or it may promote an entirely different activity that does not occur in the liquid phase. The ability to create multiple reaction sites allows the possibility of manipulating and detecting compounds in ways not possible in solution chemistries. Interfering materials can be left behind or altered in upper layers.

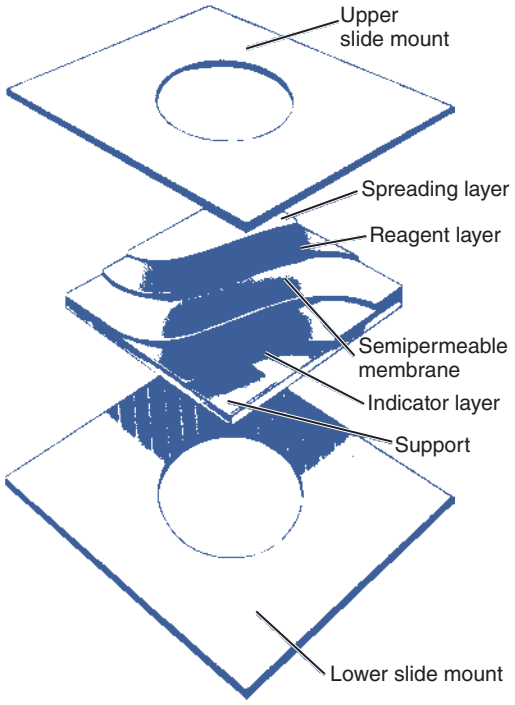


Figure 5.6 VITROS slide with multiple layers contains the entire reagent chemistry system.

Photograph courtesy of Ortho Clinical Diagnostics.

Chemical Reaction Phase

This phase consists of mixing, separation, incubation, and reaction time. In most discrete analyzers, the chemical reactants are held in individual moving containers that are either disposable or reusable.

Content removed due to copyright restrictions

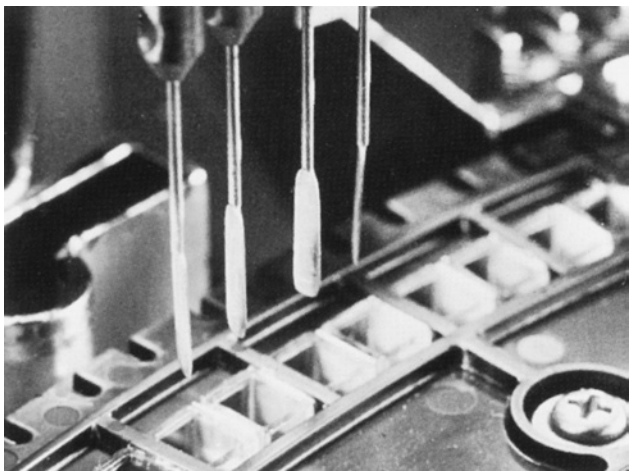


Figure 5.8 Stirring paddles on a chemistry analyzer.

Photograph courtesy of Roche Diagnostics.

These reaction containers also function as the cuvettes for optical analysis. If the cuvettes are reusable, then wash stations are set up immediately after the read stations to clean and dry these containers (**Figure 5.7**). This arrangement allows the analyzer to operate continuously without replacing cuvettes. Examples of this approach include ADVIA Centaur (Siemens), ARCHITECT (Abbott Diagnostics), Cobas (Roche Diagnostics), and UniCel DxC Synchron (Beckman Coulter) analyzers. Alternatively, the reactants may be placed in a stationary reaction chamber in which a flow-through process of the reaction mixture occurs before and after the optical reading.

Mixing

A vital component of each procedure is the adequate mixing of the reagents and sample. Instrument manufacturers go to great lengths to ensure complete mixing. Nonuniform mixtures can result in imprecision in discrete analysis.

Most automated wet-chemistry analyzers use stirring paddles that dip into the reaction container for a few seconds to stir sample and reagents, after which they return to a wash reservoir (**Figure 5.8**). Other instruments use forceful dispensing to accomplish mixing.

Separation

In chemical reactions, undesirable constituents that will interfere with an analysis may need to be separated from the sample before the other reagents are introduced into the system. Protein causes major interference in many analyses. One approach without separating protein is to use a very high reagent-to-sample ratio (the sample is highly diluted) so that turbidity caused by precipitated protein is not detected by the spectrophotometer. Another approach is to shorten the reaction time to eliminate slower-reacting interferents.

In the VITROS MicroSlide technology, the spreading layer of the slide not only traps cells, crystals, and other small particulate matter but also retains large molecules, such as protein. In essence, what passes through the spreading layer is a protein-free filtrate.

Most contemporary discrete analyzers have no automated methodology by which to separate interfering substances from the reaction mixture. Therefore, methods have been chosen that have few interferences or that have known interferences that can be compensated for by the instrument (i.e., using correction formulas).

Incubation

A heating bath in discrete analysis systems maintains the required temperature of the reaction mixture and provides the delay necessary to allow complete color

development. The principal components of the heating bath are the heat transfer medium (i.e., water or air), the heating element, and the thermoregulator. A thermometer is located in the heating compartment of an analyzer and is monitored by the system's computer. On many discrete analyzer systems, the multicuvettes incubate in a water bath maintained at a constant temperature of usually 37°C.

Slide technology incubates colorimetric slides at 37°C. There is a precondition station to bring the temperature of each slide close to 37°C before it enters the incubator. The incubator moves the slides at 12-second intervals in such a manner that each slide is at the incubator exit four times during the 5-minute incubation time. This feature is used for two-point rate methods and enables the first point reading to be taken partway through the incubation time. Potentiometric slides are held at 25°C. The slides are kept at this temperature for 3 minutes to ensure stability before reading.

Reaction Time

Before the optical reading by the spectrophotometer, the reaction time may depend on the rate of transport through the system to the "read" station, timed reagent additions with moving or stationary reaction chambers, or a combination of both processes. An environment conducive to the completion of the reaction must be maintained for a sufficient length of time before spectrophotometric analysis of the product is made. Time is a definite limitation. To sustain the advantage of speedy multiple analyses, the instrument must produce results as quickly as possible.

It is possible to monitor not only completion of a reaction but also the rate at which the reaction is proceeding. The instrument may delay the measurement for a predetermined time or may present the reaction mixtures for measurement at constant intervals of time. Use of rate reactions may have two advantages: the total analysis time is shortened, and interfering chromogens that react slowly may be negated. Reaction rate is controlled by temperature; therefore, the reagent, timing, and spectrophotometric functions must be coordinated to work in harmony with the chosen temperature. The environment of the cuvettes is maintained at a constant temperature by a liquid bath, containing water or some other fluid with good heat transfer properties, in which the cuvettes move.

Measurement Phase

After the reaction is completed, the formed end products must be quantified. Almost all available systems for measurement have been used, such as ultraviolet, fluorescent, and flame photometry; ion-specific electrodes; gamma counters; and luminometers. Still, the most common are visible and ultraviolet light spectrophotometry, although adaptations of traditional fluorescence measurement, such as fluorescence polarization, chemiluminescence, and bioluminescence, have become popular.

Analyzers that measure light require a monochromator to achieve the desired component wavelength. Traditionally, analyzers have used filters or filter wheels to separate light. The old AAs used filters that were manually placed in position in the light path. Many instruments still use rotating filter wheels that are microprocessor controlled so that the appropriate filter is positioned in the light path. However, newer and more sophisticated systems offer the higher resolution afforded by diffraction gratings to achieve light separation into its component colors. Many instruments now use such monochromators with either a mechanically rotating grating or a fixed grating that spreads its component wavelengths onto a fixed array of photo diodes—for example, Roche analyzers (**Figure 5.9**). This latter grating arrangement, as well as rotating filter wheels, easily accommodates polychromatic light analysis, which offers improved sensitivity and specificity over monochromatic measurement. By recording optical readings at different wavelengths, the instrument's computer can then use these data to correct for reaction mixture interferences that may occur at adjacent, as well as desired, wavelengths.

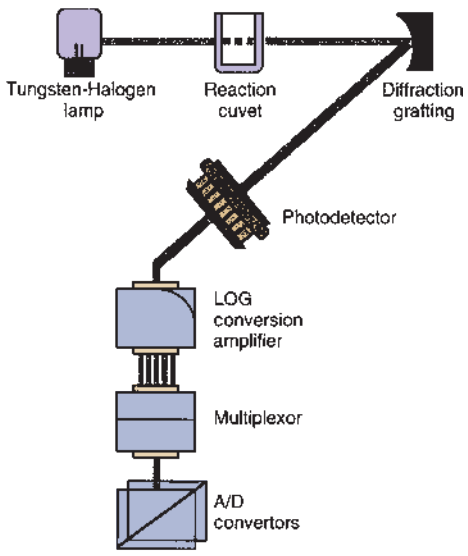


Figure 5.9 Photometer for a chemistry analyzer. Fixed diffraction grating separates light into specific wavelengths and reflects them onto a fixed array of 11 specific photodetectors. Photometer has no moving parts.

Courtesy of Roche Diagnostics.

Many newer instruments use fiberoptics as a medium to transport light signals from remote read stations back to a central monochromator detector box for analysis of these signals. The fiberoptic cables, or "light pipes" as they are sometimes called, are attached from multiple remote stations where the reaction mixtures reside to a centralized monochromator/detector unit that, in conjunction with the computer, sequences and analyzes a large volume of light signals from multiple reactions.

The containers holding the reaction mixture also play a vital role in the measurement phase. In most discrete wet-chemistry analyzers, the cuvette used for analysis is also the reaction vessel in which the entire procedure has occurred. The reagent volume and, therefore, sample size, speed of analysis, and sensitivity of measurement are some aspects influenced by the method of analysis. A beam of light is focused through the container holding the reaction mixture. The amount of light that exits from the container is dictated primarily by the absorbance of light by the reaction mixture. The exiting light strikes a photodetector, which converts the light into electrical energy. Filters and light-focusing components permit the desired light wavelength to reach the photodetector. The photometer continuously senses the sample photodetector output voltage and, as is the process in most analyzers, compares it with a reference output voltage. The electrical impulses are sent to a readout device, such as a printer or computer, for storage and retrieval.

Slide technology depends on reflectance spectrophotometry, as opposed to traditional transmittance photometry, to provide a quantitative result. The amount of chromogen in the indicator layer is read after light passes through the indicator layer, which is reflected from the bottom of a pigment-containing layer (usually the spreading layer) and is returned through the indicator layer to a light detector. For colorimetric determinations, the light source is a tungsten–halogen lamp. The beam focuses on a filter wheel holding up to eight interference filters, which are separated by a dark space. The beam is focused at a 45° angle to the bottom surface of the slide, and a silicon photodiode detects the portion of the beam that reflects down. Three readings are taken for the computer to derive reflectance density. The three recorded signals taken are (1) the filter wheel blocking the beam, (2) reflectance of a reference white surface with the programmed filter in the beam, and (3) reflectance of the slide with the selected filter in the beam.

After a slide is read, it is shuttled back in the direction from which it came, where a trap door allows it to drop into a waste bin. If the reading was the first for a two-point rate test, the trap door remains closed, and the slide reenters the incubator.

The principles of automated immunoassays are discussed below, and the similarities between automated chemistry analyzers and automated immunoassay analyzers are worth noting. There are many fully automated, random-access immunoassay systems, which use chemiluminescence or electrochemiluminescence technology for reaction analysis. In chemiluminescence assays, quantification of an analyte is based on emission of light resulting from a chemical reaction.[9] The principles of chemiluminescent immunoassays are similar to those of radioimmunoassay and immunoradiometric assay, except that an acridinium ester is used as the tracer and paramagnetic particles are used as the solid phase. Sample, tracer, and paramagnetic particle reagent are added and incubated in disposable plastic cuvettes, depending on the assay protocol. After incubation, magnetic separation and washing of the particles are performed automatically. The cuvettes are then transported into a light-sealed luminometer chamber, where appropriate reagents are added to initiate the chemiluminescent reaction. On injection of the reagents into the sample cuvette, the system luminometer detects the chemiluminescent signal. Luminometers are like gamma counters in that they use a photomultiplier tube detector; however, unlike gamma counters, luminometers do not require a crystal to convert gamma rays to light photons. Light photons from the sample are detected directly, converted to electrical pulses, and then counted.

Signal Processing and Data Handling

Before results can be transmitted, ensuring accurate calibration of the test system is essential to obtaining accurate information. There are many variables that may enter into the use of calibration standards, and the matrices of the standards and unknowns may be different. Depending on the methodology, this may or may not present problems. Primary or secondary standards may be used for calibration purposes. A primary standard is a highly purified chemical that can be measured directly to produce an *exact* known concentration and purity. A secondary standard is made from a primary standard. If secondary standards are used to calibrate an instrument, the methods used to derive the standard's constituent values

should be known. Standards containing more than one analyte per vial may cause interference problems. Because there are no primary standards available for enzymes, either secondary standards or calibration factors based on the molar extinction coefficients of the products of the reactions may be used.

The advantage of calibrating an automated instrument is the long-term stability of the standard curve, which only requires monitoring with controls on a daily basis. Some analyzers use low- and high-concentration standards at the beginning of each run and then use the absorbances of the reactions produced by the standards to produce a standard curve electronically for each run. Other instruments are self-calibrating after analyzing standard solutions.

Slide technology requires more sophisticated calculations to produce results. The calibration materials require a protein-based matrix because of the necessity for the calibrators to behave as serum when reacting with the various layers of the slides. Calibrator fluids are bovine serum–based, and the concentration of each analyte is determined by reference methods. **Endpoint** reaction tests and enzymatic methods require three calibrators, while tests requiring a blank need four calibrators. Colorimetric tests use spline fits, which is a data analysis technique used to interpolate data when there are sudden slope changes in a curve, to produce the standardization. In enzyme analysis, a curve-fitting algorithm estimates the change in reflection density per unit time. This is converted to either absorbance or transmission-density change per unit time. Then, a quadratic equation converts the change in transmission density to volume activity (U/L) for each assay.

All advanced automated instruments have some method of reporting results with a link to sample identification. In sophisticated systems, the demographic sample information is entered in the instrument's computer together with the tests required. Then the sample identification is printed with the test results. Most automated instruments report results to the Laboratory Information System (LIS) using a **bidirectional interface**, meaning that the analyzer can read data from the patient sample barcode and also transmit laboratory results back to the LIS in electronic formats. Most laboratories use barcode labels that are generated from the LIS to identify samples. Barcode-labeled samples can be loaded directly on the analyzer without the need to enter identifying information, tests, or other information manually. Microprocessors control the tests, reagents, and timing, while verifying the barcode for each sample. This is the link between the results reported and the specimen identification. Even the simplest of systems sequentially number the test results to provide a connection with the samples.

On most modern automated analyzers, computerized monitoring is available and flagged for the operator for such parameters as specimen integrity, assay linearity, quality control data with various options for statistical display and interpretation, short sample sensing, abnormal patient results, clot detection, reaction vessel or test chamber temperature, and reagent inventories. The monitors can also display patients' results as well as various data flags previously mentioned to assist the operator in troubleshooting prior to reporting the patient results. Data flags should be investigated prior to releasing patient results into the LIS. Most instrument manufacturers offer computer software for preventive maintenance schedules and algorithms for troubleshooting. Some manufacturers also install phone modems on the analyzer for a direct communication link between the instrument and their service center for instant troubleshooting and technical service.

CASE STUDY 5.1, PART 2

Remember Mía, who is training on the Roche Cobas?

1. Can Mía verify the patient results? Why or why not?
2. What action should Mía take?

K	Na	Cl	CO_2	TBIL	DBILI	AST
5.6H	131L	94	14L	5.2H	3.2H	1078H
>I.H					>I.L	>Abs

CASE STUDY 5.2, PART 1

Miles is performing a noncompetitive immunoassay for IgG subclasses. The results are below:

IgG	1490 mg/dL
G1	713 mg/dL
G2	101 mg/dL
G3	60.7 mg/dL
G4	30.7 mg/dL

1. What is the calculated sum of the IgG subclasses?
2. Comparing the calculated sum of the IgG subclasses to the total measured IgG, what action should be taken?

Additional Considerations for Automated Immunoassays

Immunoassays were first developed by Dr. Rosalyn Yalow and colleagues in 1959, who developed a radioimmunoassay (RIA) for the measurement of insulin.[10] Dr. Yalow went on to share the Nobel Prize in physiology or medicine for this discovery in 1977. Today, immunoassays are one of the essential analytical techniques used in clinical chemistry and can be highly automated. The basis of all immunoassays is the binding of **antibody** (Ab) to **antigen** (Ag) for the specific and sensitive detection of an analyte. The design, label, and detection system combine to create many different assays, which enable the measurement of analytes including proteins, hormones, metabolites, therapeutic drugs, and drugs of abuse.

Immunoassay Basics

In an immunoassay, an Ab molecule recognizes and binds to an Ag. Antibodies are immunoglobulin (Ig) molecules with a functional domain known as F(ab) that specifically binds to an antigenic determinant or **epitope** of the antigen (i.e., the site on the antigen). This binding is related to the concentration of each reactant, the specificity of the Ab for the Ag, the **affinity**, **avidity**, and the environmental conditions.

The degree of binding is an important consideration in any immunoassay. Affinity refers to the strength of binding between a single binding site on the Ab and its epitope. Under standard conditions, the affinity of an Ab is measured using a **hapten** (Hp) because the Hp is a low-molecular-weight Ag considered to have only one epitope. The affinity for the Hp is related to the likelihood to bind or to the degree of complementary nature of each. The reversible reaction is summarized in **Equation 5.1**:

$$\text{Hp} + \text{Ab} \rightarrow \text{HpAb complex} \qquad \textbf{(Eq. 5.1)}$$

The binding between an Hp and the Ab obeys the law of mass action and is expressed mathematically in **Equation 5.2**:

$$K_a = \frac{k_1}{k_2} = \frac{[\text{Hp}-\text{Ab}]}{[\text{Hp}][\text{Ab}]} \qquad \textbf{(Eq. 5.2)}$$

K_a is the affinity or equilibrium constant and represents the reciprocal of the concentration of free Hp when 50% of the binding sites are occupied. The greater the affinity of the Hp for the Ab, the smaller is the concentration of Hp needed to saturate 50% of the binding sites of the Ab. For example, if the affinity constant of a **monoclonal** antibody (Mab, one specific antibody) is 3×10^{11} L/mol, then it means that an Hp concentration of 3×10^{-11} mol/L is needed to occupy half of the binding sites. Typically, the affinity constant of Abs used in immunoassay procedures ranges from 10^9 to 10^{11} L/mol, whereas the affinity constant for transport proteins ranges from 10^7 to 10^8 L/mol and the affinity for receptors ranges from 10^8 to 10^{11} L/mol.

As with all chemical reactions, the initial concentrations of the reactants and the products affect the extent of immune complex binding. In immunoassays, the reaction moves forward (to the right) (Eq. 5.1) when the concentration of reactants (Ag and Ab) exceeds the concentration of the product (Ag–Ab complex) and when there is a favorable affinity constant.

The forces that bring an antigenic determinant and an Ab together are noncovalent, reversible bonds that result from the cumulative effects of hydrophobic, hydrophilic, hydrogen bonds, and van der Waals

forces. The most important factor that affects the cumulative strength of bonding is the closeness of fit between the Ab and the Ag. The strength of most of these interactive forces is inversely related to the distance between the interactive sites. The closer the Ab and Ag can physically approach one another, the greater are the attractive forces.

After the Ag–Ab complex is formed, the likelihood of separation (which is inversely related to the tightness of bonding) is referred to as *avidity*. The avidity refers to the cumulative strength of binding of all Ab–epitope pairs and exceeds the sum of single Ab–epitope binding. For example, IgG has only two epitope binding sites while an IgM pentamer has 10 epitope binding sites and therefore, a higher avidity. In general, affinity is a property of the Ag and avidity is a property of the Ab.

The specificity of an Ab is most often described by the Ag that induced the Ab production, the homologous Ag. Ideally, this Ab would react only with that specific Ag. However, an Ab can react with an Ag that is structurally like the homologous Ag. This is referred to as *cross-reactivity*. Considering that an antigenic determinant can be five or six amino acids or one immunodominant sugar, it is not surprising that Ag similarity is common. The greater the similarity between the cross-reacting Ag and the homologous Ag, the stronger is the bond with the Ab.[11]

Reagent Ab production is achieved by polyclonal or monoclonal techniques. In **polyclonal** Ab production, the stimulating Ag is injected in an animal responsive to the Ag; the animal detects this foreign Ag and mounts an immune response to eliminate the Ag. If part of this immune response includes strong Ab production, then blood is collected and Ab is harvested, characterized, and purified to yield the commercial antiserum reagent. This polyclonal Ab reagent is a mixture of different Ab specificities. Some Abs react with the stimulating epitopes and some are endogenous to the host. Multiple Abs directed against the multiple epitopes on the Ag are present and can cross-link the multivalent Ag. Polyclonal Abs are often used as "capture" Abs in sandwich or indirect immunoassays.

In contrast, an immortal cell line produces monoclonal Abs (mAbs); each line produces one specific Ab. This method developed as an extension of the hybridoma work published by Kohler and Milstein in 1975.[12] The process begins by selecting cells with the qualities that will allow the synthesis of a homogeneous Ab. First, a host (commonly, a mouse) is immunized with an Ag (the one to which an Ab is desired); later, the sensitized lymphocytes of the spleen are harvested. Second, an immortal cell line (usually a nonsecretory mouse myeloma cell line that is hypoxanthine guanine phosphoribosyltransferase deficient) is required to ensure that continuous propagation in vitro is viable. These cells are then mixed in the presence of a fusion agent, such as polyethylene glycol, which promotes the fusion of two cells to form a hybridoma. In a selective growth medium, only the hybrid cells will survive. B cells have a limited natural life span in vitro and cannot survive, and the unfused myeloma cells cannot survive due to their enzyme deficiency. If the viable fused cells synthesize Ab, then the specificity and isotype of the Ab are evaluated. Commercial mAb reagent is produced by growing the hybridoma in tissue culture or in compatible animals. An important feature of mAb reagent is that the Ab is homogeneous (a single Ab, not a mixture of Abs). Therefore, it recognizes only one epitope on a multivalent Ag and cannot cross-link a multivalent Ag.

Turbidimetry and Nephelometry

Turbidimetry (immunoturbidimetry) and nephelometry (immunonepholometry) are two related automated methods used to quantitate Ag–Ab complexes (see Chapter 4, Analytic Techniques). In general terms, turbidimetry measures the amount of light that can pass through a sample. As a sample becomes more turbid, more light is blocked by the particles in the sample, and less light passes through. Nephelometry is similar conceptually; however, the scattered light is measured by placing a detector at a defined angle (e.g., 90°) from the incident light. When Ag and Ab combine, immune complexes are formed that act as particles in suspension and thus can scatter light. The size of the particles determines the type of scatter that will dominate when the solution interacts with nearly monochromatic light.[13] When the particle, such as albumin or IgG, is relatively small compared with the wavelength of incident light, the particle will scatter light symmetrically, both forward and backward. A minimum of scattered light is detectable at 90° from the incident light. Larger molecules and Ag–Ab complexes have diameters that approach the wavelength of incident light and scatter light with a greater intensity in the forward direction. The wavelength of light is selected based on its ability to be scattered in a forward direction and the ability of the Ag–Ab complexes to absorb the wavelength of light. To recap, turbidimetry measures the light transmitted and nephelometry measures the light scattered.

Turbidimeters (spectrophotometers or colorimeters) are designed to measure the light passing through a solution, and the photodetector is placed at an angle of 180° from the incident light. If light scattering is insignificant, turbidity can be expressed as "optical density," which is directly related to the concentration of suspended particles and path length. Nephelometers measure light at an angle other than 180° from the incident light; most measure forward light scattered at 90° or less because the sensitivity is increased.

Both methods can be performed in an endpoint or a kinetic mode. In the **endpoint** assay, a measurement is taken at the beginning of the reaction (the background signal) and one is taken at a set time later in the reaction (plateau or endpoint signal). The concentration is determined using a calibration curve. In **kinetic** assays, the rate of complex formation is continuously monitored, and the peak rate is determined. The peak rate is directly related to the concentration of the Ag, although this is not necessarily linear. Thus, a calibration curve is required to determine concentration in unknown samples.

In general, turbidimetry is less sensitive than nephelometry. In turbidimetry, the decrease in the amount of light transmitted due to scattering of particles is measured relative to the intensity of the reference blank. Therefore, when there is no sample, there is 100% transmittance. If the sample concentration of the analyte is high, then a large precipitate is formed resulting in a significant decrease in the amount of light transmitted that can easily be measured by turbidimetry. If, however, the amount of precipitate is small due to a low sample concentration, the amount of light transmitted is decreased minimally and the instrument must measure the difference between two high intensity light signals. Because in nephelometry the amount of light is measured at an angle of 90°, the reference blank produces no signal (no light scattering). When a microprecipitate is formed, the light-scattered signal is more analytically discernible when measured against a dark reference baseline.

Labeled Immunoassays

General Considerations

In all labeled immunoassays, a reagent (Ag or Ab) is usually labeled by attaching a particle or molecule that will better detect lower concentrations of Ag–Ab complexes. Therefore, the label improves analytic sensitivity. All assays have a binding reagent, which can bind to the Ag or ligand. If the binding reagent is an Ab, the assay is an immunoassay.

Immunoassays may be described based on the label, which reactant is labeled, the relative concentration and source of the Ab, the method used to separate free from bound labeled reagents, the signal that is measured, and the method used to assign the concentration of the analyte in the sample. Immunoassay design has many variables to consider, leading to diverse assays.

Labels

The simplest way to identify an assay is by the label used. **Table 5.3** lists the commonly used labels and the methods used to detect the label.

Radioactive Labels

As discussed previously, radioimmunoassay was the first immunoassay developed, and the developers won the Nobel Prize in Medicine in 1977 for their innovation. These original immunoassays used radioactive isotopes (^{125}I, ^{131}I or ^{3}H) to label the analyte. The emitted gamma rays were measured using a scintillation counter. Due to safety concerns over of radioactive reagents and waste, it has been replaced by safer methods in the clinical laboratory.

Enzyme Labels

Enzymes are commonly used to label the Ag/Hp or Ab.[14,15] Horseradish peroxidase (HRP), alkaline phosphatase (ALP), and glucose-6-phosphate dehydrogenase are used most often. Enzymes are biologic catalysts that increase the rate of conversion of substrate to product and are not consumed by the reaction. As such, an enzyme can catalyze many substrate molecules, amplifying the amount of product generated. The enzyme activity may be monitored directly by measuring the product formed or by measuring the effect of the product on a coupled reaction. Depending on the substrate used, the product can be photometric, fluorometric, or chemiluminescent. For example, a typical photometric reaction using HRP-labeled Ab (Ab-HRP) and the substrate (a peroxide) generates a product (oxygen). The oxygen can then oxidize a reduced chromogen (such as reduced orthophenylenediamine [OPD]) to produce a colored compound (oxidized OPD), which is measured using a photometer:

$$\text{Ab} - \text{HRP} + \text{peroxide} \rightarrow \text{Ab} = \text{HRP} + O_2$$
$$O_2 + \text{reduced OPD} \rightarrow \text{oxidized} + \text{OPD} + H_2O \quad \text{(Eq. 5.3)}$$

Table 5.3 Labels and Detection Methods

Immunoassay	Common Label	Detection Method
RIA	^{3}H	Liquid scintillation counter
	^{125}I	Gamma counter
EIA	Horseradish peroxidase	Photometer, fluorometer, luminometer
	Alkaline phosphatase	Photometer, fluorometer, luminometer
	β-D-Galactosidase	Fluorometer, luminometer
	Glucose-6-phosphate dehydrogenase	Photometer, luminometer
CLIA	Isoluminol derivative	Luminometer
	Acridinium esters	Luminometer
FIA	Fluorescein	Fluorometer
	Europium	Fluorometer
	Phycobiliproteins	Fluorometer
	Rhodamine B	Fluorometer
	Umbelliferone	Fluorometer

RIA, Radioimmunoassay; EIA, enzyme immunoassay; CLIA, chemiluminescent immunoassay; FIA, fluorescent immunoassay

Fluorescent Labels

Fluorescent labels (fluorochromes or fluorophores) are compounds that absorb radiant energy of one wavelength and emit radiant energy of a longer wavelength in less than 10^{-4} seconds. Generally, the emitted light is detected at an angle of 90° from the path of excitation light using a fluorometer or a modified spectrophotometer. The difference between the excitation wavelength and emission wavelength (Stokes shift) usually ranges between 20 and 80 nm for most fluorochromes. Some fluorescence immunoassays simply substitute a fluorescent label (such as fluorescein) for an enzyme label and quantitate the fluorescence.[16] Another approach, time-resolved fluorescence immunoassay, uses a highly efficient fluorescent label such as europium chelate,[17] which fluoresces approximately 1000 times slower than the natural background fluorescence and has a wide Stokes shift. The delay allows the fluorescent label to be detected with minimal interference from background fluorescence. The long Stokes shift facilitates measurement of emission radiation while excluding the excitation radiation. This assay shows high sensitivity with minimized background fluorescence.

Luminescent Labels

Luminescent labels emit a photon of light as the result of an electrical, biochemical, or chemical reaction.[18,19] Some organic compounds become excited when oxidized and emit light as they revert to the ground state. Oxidants include hydrogen peroxide, hypochlorite, or oxygen. Sometimes, a catalyst is needed, such as peroxidase, ALP, or metal ions.

Luminol, the first chemiluminescent label used in immunoassays, is a cyclic diacylhydrazide that emits light energy under alkaline conditions in the presence of peroxide and peroxidase. Because peroxidase can serve as the catalyst, assays may use this enzyme as the label; the chemiluminogenic substrate, luminol, will produce light that is directly proportional to the amount of peroxidase present (**Eq. 5.4**):

$$\text{Luminol} + 2H_2O_2 + OH^- \xrightarrow{\text{Peroxidase}} \text{3-aminophthalate} + \text{light (425 nm)} \quad \textbf{(Eq. 5.4)}$$

A popular chemiluminescent label, acridinium esters, is a triple-ringed organic molecule linked by an ester bond to an organic chain. In the presence of hydrogen peroxide and under alkaline conditions, the ester bond is broken and an unstable molecule

(*N*-methylacridone) remains. Light is emitted as the unstable molecule reverts to its more stableground state:

$$\text{Acridinium ester} + 2H_2O_2 + OH^- \rightarrow N\text{-methylacridone} + CO_2 + H_2O + \text{light (430 nm)} \quad \textbf{(Eq. 5.5)}$$

ALP commonly conjugated to an Ab has been used in automated immunoassay analyzers to produce some of the most sensitive chemiluminescent assays. ALP catalyzes adamantyl 1,2-dioxetane aryl phosphate substrates to release light at 477 nm. The detection limit approaches 1 µmol, or approximately 602 enzyme molecules.[20,21]

Assay Design

Competitive Immunoassays. In a **competitive immunoassay**, labeled Ag (Ag*) in the reagent competes with Ag in the patient sample for a limited number of Ab binding sites (**Figure 5.10**). In the competitive assay, the Ag* concentration is constant and limited. As the concentration of patient Ag increases it competes with binding to the Ab resulting in less Ag* bound. Thus, the signal is inversely proportional to analyte concentration.

The Ag–Ab reaction can be accomplished in one step when labeled antigen (Ag*), unlabeled antigen (Ag), and reagent antibody (Ab) are simultaneously incubated together to yield bound, labeled complex (Ag*Ab), bound, unlabeled complex (AgAb), and free labeled Ag (Ag*), as shown in Figure 5.10A and **Equation 5.6**:

$$\text{Ag* (fixed reagent)} + \text{Ag} + \text{Ab (limited reagent)} \rightarrow \text{Ag* Ab} + \text{AgAb} + \text{Ag*} \quad \textbf{(Eq. 5.6)}$$

Alternatively, the competitive assay may be accomplished in sequential steps. First, patient sample containing the Ag to be measured is incubated with the reagent Ab and then labeled Ag is added. After a longer incubation time and a separation step, the bound, labeled Ag is measured. This approach increases the analytic sensitivity of the assay.

Consider the example in **Table 5.4**. A relatively small, yet constant, number of Ab combining sites is available to combine with a relatively large, constant amount of Ag* (tracer) and calibrators with known Ag concentrations. Because the amount of tracer and Ab is constant, the only variable in the test system is the amount of unlabeled Ag. As the concentration of unlabeled Ag increases, the concentration (or percentage) of free tracer increases.

By using multiple calibrators, a dose–response curve is established. As the concentration of unlabeled Ag increases, the concentration of tracer that binds to the Ab decreases. In the example presented in Table 5.4, if the amount of unlabeled Ag is zero, maximum tracer will combine with the Ab. When no unlabeled Ag is present, maximum binding by the tracer is possible; this is referred to as B_0, B_{max}, maximum binding, or the zero standard. When the amount of unlabeled Ag is the same as the tracer, each will bind

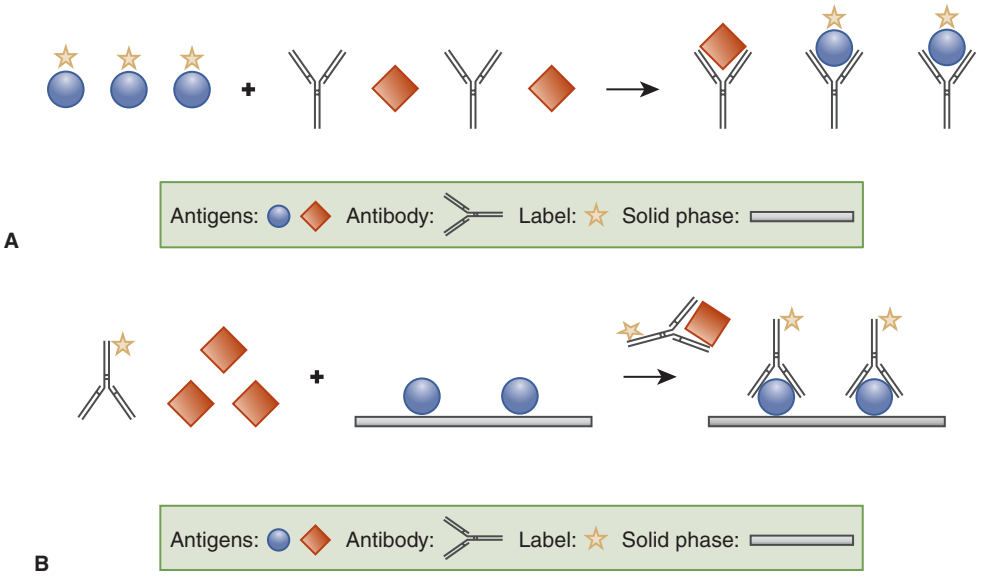


Figure 5.10 Examples of competitive immunoassays. (**A**) Labeled antigen and antigen in patient sample compete for binding to antibody. (**B**) Immobilized Ag and Ag in the patient sample compete for binding to labeled antibody.

Table 5.4 Competitive Binding Assay Example

AG	+	AG*	+	AB	→	AGAB	+	AG*AB	+	AG*
CONCENTRATION OF REACTANTS						CONCENTRATION OF PRODUCTS				
AG		**AG***		**AB**		**AGAB**		**AG*AB**		**AG***
0		200		100		0		100		100
50		200		100		20		80		120
100		200		100		34		66		134
200		200		100		50		50		150
400		200		100		66		34		166

SAMPLE CALCULATIONS		
Dose of [Ag]	**% B**	**B/F**
0	$\frac{100}{200} = 50$	$\frac{100}{100} = 1$
50	$\frac{80}{200} = 40$	$\frac{80}{200} = 67$
100	$\frac{66}{200} = 33$	$\frac{66}{134} = 33$
200	$\frac{50}{200} = 25$	$\frac{50}{150} = 33$
400	$\frac{34}{200} = 17$	$\frac{34}{166} = 20$

AG, unlabeled antigen; AG*, labeled antigen; AB, antibody; AGAB, antigen-antibody complex; AG*AB, labeled antigen-antibody complex.

equally to the Ab. As the concentration of Ag increases in a competitive assay, the amount of tracer that complexes with the binding reagent decreases. If the tracer is of low molecular weight, free tracer is often measured. If the tracer is of high molecular weight, the bound tracer is measured. The data may be plotted in one of three ways: bound/free versus the arithmetic dose of unlabeled Ag, percentage bound versus the log dose of unlabeled Ag, and logit bound/B_0 versus the log dose of the unlabeled Ag (**Figure 5.11**).

The bound fraction can be expressed in several different formats. Bound/free is counts per minute (CPM) of the bound fraction compared with the CPM of the free fraction. Percent bound (% B) is the CPM of the bound fraction compared with the CPM of maximum binding of the tracer (B_0) multiplied by 100. Logit B/B_0 transformation is the natural log of $(B/B_0)/(1 - B/B_0)$. When B/B_0 is plotted on the ordinate and the log dose of the unlabeled Ag is plotted on the abscissa, a straight line with a negative slope is produced using linear regression.

It is important to remember that the best type of curve-fitting technique is determined by experiment and that there is no assurance that a logit-log plot of the data will always generate a straight line. To determine the best method, several different methods of data plotting should be tried when a new assay is introduced. Every time the assay is performed, a dose–response curve should be prepared to check the performance of the assay. The relative error for all RIA dose–response curves is minimal when $B/B_0 = 0.5$ and increases at both high and low concentrations of the plot. As shown in the plot of B/B_0 versus log of the Ag concentration (Figure 5.11), a relatively large change in the concentration at either end of the curve produces little change in the B/B_0 value. Patient

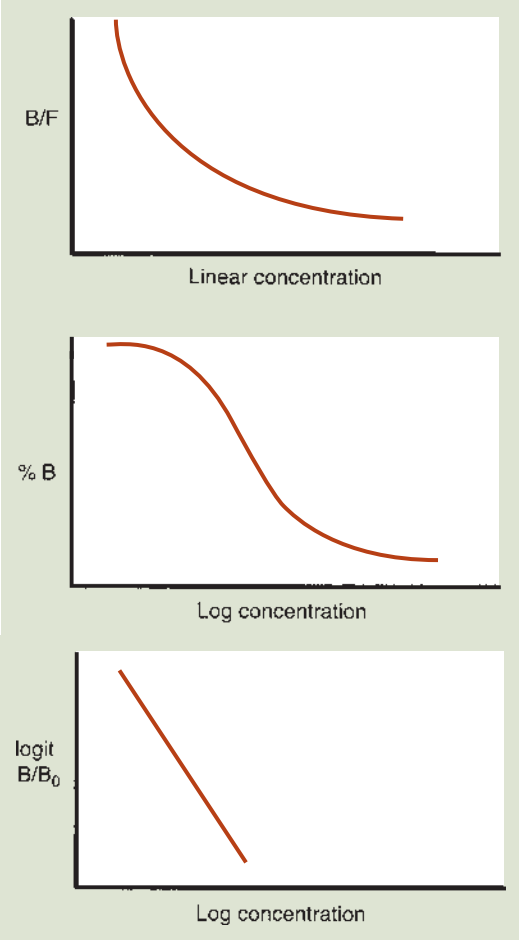


Figure 5.11 Dose–response curves in a competitive assay. B, bound labeled antigen; F, free labeled antigen; B_0, maximum binding; % B, $B/B_0 \times 100$.

values derived from a B/B_0 value greater than 0.9 or less than 0.1 should be interpreted with caution. When the same data are displayed using the logit-log plot, it is easy to overlook the error at either end of thestraight line.

Noncompetitive Immunoassays. **Noncompetitive immunoassays**, also known as sandwich assays, use a labeled reagent Ab to detect the Ag. Excess labeled Ab is required to ensure that the labeled Ab reagent does not limit the reaction. The concentration of the Ag is directly proportional to the bound labeled Ab as shown in **Figure 5.12**. The relationship is linear up to a limit and then may be subject to the high-dose **hook effect**. The hook effect (also sometimes known as prozone) is an immunologic phenomenon that occurs when excess analyte overwhelms the test system, causing a false result.

In the sandwich assay to detect Ag, immobilized unlabeled Ab captures the Ag. After washing to remove unreacted molecules, the labeled detector Ab is added. After another washing to remove free labeled detector Ab, the signal from the bound labeled Ab is proportional to the Ag captured. This format relies on the ability of the Ab reagent to react with a single epitope on the Ag. The specificity and quantity of mAbs have allowed the rapid expansion of diverse assays. A schematic is shown in **Figure 5.13**.

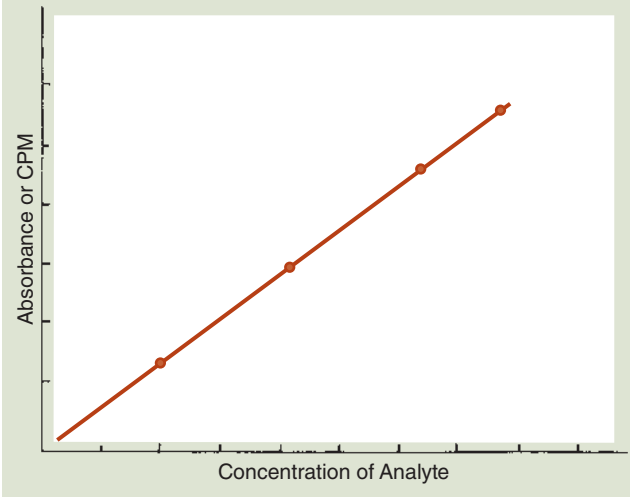


Figure 5.12 Dose–response curve in a noncompetitive immunoassay. CPM, counts per minute.

Separation Techniques

All automated immunoassays require that free labeled reactant be distinguished from bound labeled reactant. The most common separation technique is the use of paramagnetic particles that can quickly be immobilized to a solid phase by application of a magnetic field. Separation is accomplished by wash steps that occur while the magnetic particles are immobilized by a magnet. Coated microwells may also be used, but rarely. In **heterogeneous** assays, physical separation is necessary and is achieved by interaction with a **solid phase**. The better the separation of bound from free reactant, the more reliable the assay will be. The labeled, unbound analyte is separated or washed away, and the remaining labeled, bound analyte is measured. In contrast, **homogeneous** assays, in which the activity or expression of the label depends on whether the labeled reactant is free or bound, do not require a physical separation step. Therefore, no wash step is required.

Adsorption. The binding of the capture Ab to paramagnetic particles is the most common method used by automated immunoassay analyzers. Separation

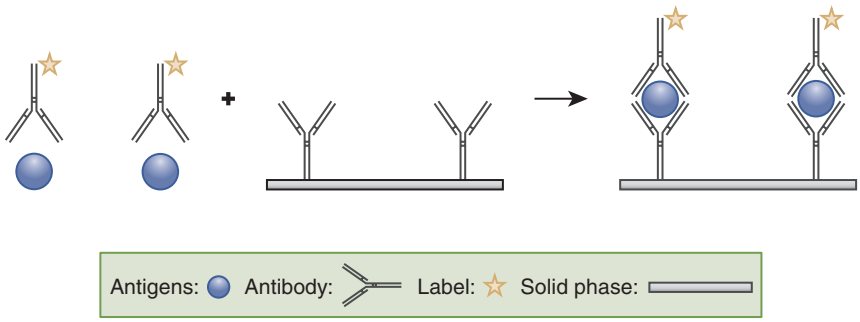


Figure 5.13 Two-site noncompetitive sandwich assay to detect antigen. Immobilized antibody captures the antigen. Then, labeled antibody is added, binds to the captured antigen, and is detected.

takes place by applying a powerful magnet, thereby adhering bound analytes to each reaction chamber (**Figure 5.14**). Unbound constituents and labels are removed by aspiration. Thereafter, the magnet is removed that enables the pellet to reconstitute and additional reagents added in order to generate the analytical signal (usually a chemiluminescent reaction).

Solid Phase. The use of a *solid phase* to immobilize reagent Ab or Ag provides a method to separate free from bound labeled reactant after washing. The solid-phase support is an inert surface to which reagent Ag or Ab is attached. The solid-phase support may be, but is not limited to, polystyrene surfaces, membranes, and magnetic beads. The immobilized Ag or Ab may be adsorbed or covalently bound to the solid-phase support; covalent linkage prevents spontaneous release of the immobilized Ag or Ab. Immunoassays using solid-phase separation are easier to perform and to automate and require less manipulation and time to perform than other immunoassays. However, a relatively large amount of reagent Ab or Ag is required to coat the solid-phase surface, and consistent coverage of the solid phase is difficult to achieve.

Interferences with Sandwich Immunoassays. While an advantage of sandwich-type immunoassays is the production of linear calibration curves, the disadvantage is that these assays are subjected to false-positive and false-negative interferences. Normally, the target analyte is required for the production of a positive analytical signal (**Figure 5.15A**). However, if the sample contains unusual Abs, such as human anti-mouse antibodies (HAMA) or heterophile Abs, they can bind to both the capture and labeled Abs, producing an analytical signal in the absence of the analyte (**Figure 5.15B**). Individuals who are exposed to mouse Ags can develop HAMA Abs that recognize monoclonal Abs derived from murine cell lines as Ags.[22] Heterophile Abs are formed from patients who have autoimmune disease and other disorders.[23] Although the principle

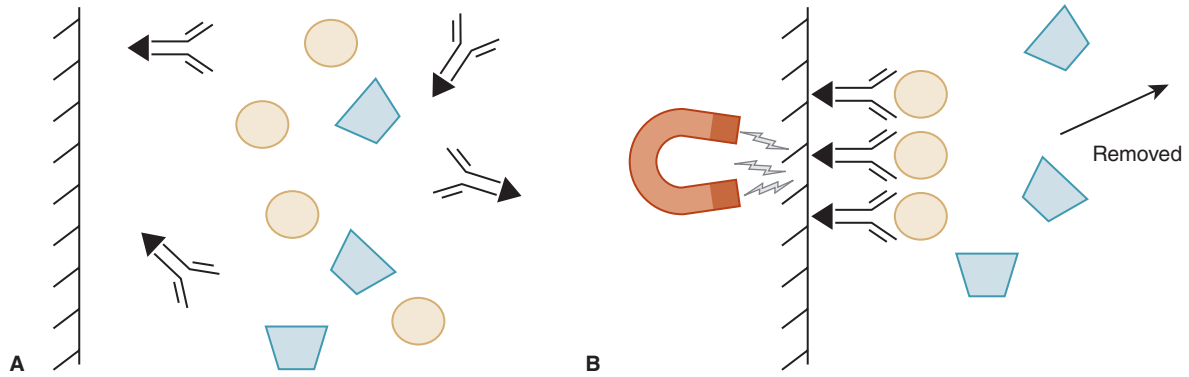


Figure 5.14 Separation by paramagnetic particles. (**A**) Addition of sample containing the analyte (*circle*) and other constituent (*rhomboid*) to a capture antibody containing a paramagnetic particle (*solid triangle*). (**B**) Application of a magnet to adhere capture antibodies and analyte to the side of the reaction chamber. The unbound constituent is removed by aspiration. Removal of the magnet and addition of signal antibody facilitate measurement (not shown).

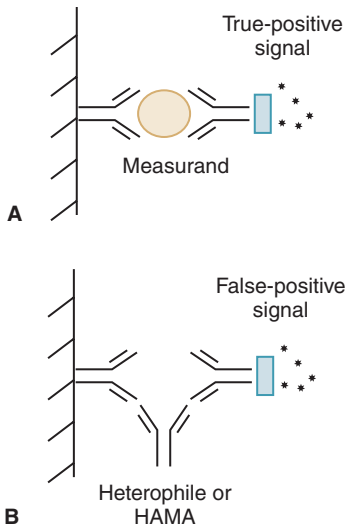


Figure 5.15 Heterophile or human anti-mouse interference. **(A)** Analytical signal in the presence of the analyte. The capture antibody (*left*), attached to a solid support, binds to the analyte (*circle*) at one epitope of the analyte. The labeled antibody (*right*) is added and binds to the analyte at a second analyte epitope. **(B)** False-positive signal in the presence of an interfering antibody. The interfering antibody (*gray middle*) binds to both the capture (*left*) and label (*right*) antibodies, forming a "sandwich" in the absence of the analyte causing a false-positive signal.

is different, the *hook effect* is similar to the prozone effect in that excess concentrations of the Ag from the sample reduce the analytical signal.[24] As shown in **Figure 5.16**, excess Ag binds to free labeled Ab, prohibiting the labeled Ab to bind to the capture Ab (via the analyte), and the resulting signal is reduced after the wash step. Analytes that are present in very low and high concentrations are subject to the hook effect. Commercial assays must be designed to either be immune to the hook effect or produce a warning flag to alert the analyst that the sample must be diluted in order to obtain an accurate result. In anticipation of the need for sample dilution, some laboratories automatically perform dilutions of a sample from a patient known to have high analyte values. The hook effect is less common in contemporary immunoassays due to better design of assays.

Examples of Labeled Immunoassays

Particle-enhanced turbidimetric inhibition immunoassay is a homogeneous competitive immunoassay in which low molecular weight Hps bound to particles compete with unlabeled analyte for the specific Ab. The extent of particle agglutination is inversely proportional to the concentration of

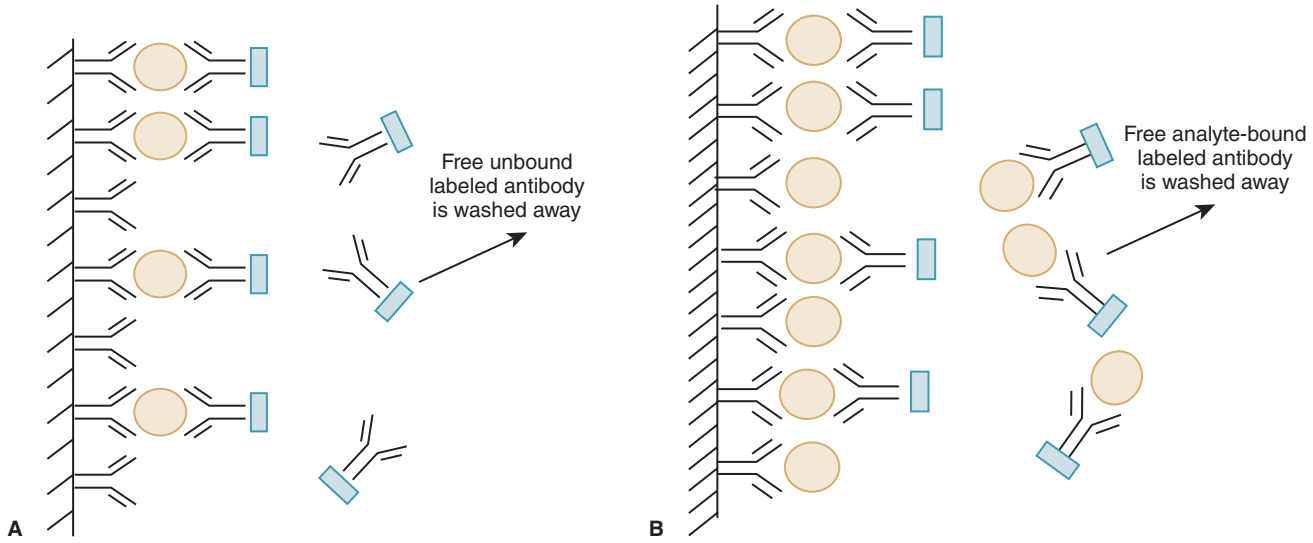


Figure 5.16 The hook effect. (**A**) Analytical signal in the presence of the analyte at a concentration that is within the dynamic range of the assay. There is an excess amount of capture (*left*) and labeled (*right*) antibodies such that all analytes are bound (none are free in solution). Excess free unbound labeled antibody is washed away, and the resulting signal is proportional to the number of captured labeled antibody (4 units in this example). (**B**) Analytical signal in the presence of the analyte at a concentration that is above the dynamic range of the assay. All of the capture antibody is bound with the analyte. The excess antigen is found free in solution and binds to excess labeled antibody found free in solution. These labeled antibodies cannot bind to the analyte-bound capture antibody because the site is already occupied with the analyte. The free analyte-bound labeled antibody is washed away. This leaves 4 units remaining, the same number as in example A, despite the much higher analyte concentration.

CASE STUDY 5.2, PART 2

Miles begins troubleshooting, starting his investigation into the discrepant results by looking at the raw data from the analyzer, which are shown here:

© dotshock/Shutterstock.

Raw Data	Dilution	Result
IgG	1:400	1490 mg/dL
	1:2000	(1450 mg/dL)
G1	1:100	713 mg/dL
	1:400	(<1120 mg/dL)
G2	1:20	101 mg/dL
	1:100	(<325 mg/dL)
G3	1:100	60.7 mg/dL
	1:400	(56.3 mg/dL)
G4	1:100	30.7 mg/dL
	1:400	(>242 mg/dL)
	1:2000	737 mg/dL
	1:8000	(702 mg/dL)

5. Does the total IgG dilutions match within 10%?
6. Does the G4 dilutions match within 10%?
7. What is a possible explanation for these results?

unlabeled analyte and is assessed by measuring the change in transmitted light.[25]

Enzyme-linked immunosorbent assays (ELISAs), a popular group of heterogeneous immunoassays, have an enzyme label and use a solid phase as the separation technique. Four formats are available: a competitive assay using labeled Ag, a competitive assay using labeled Ab, a noncompetitive assay to detect Ag, and a noncompetitive assay to detect Ab. ELISAs are widely used in clinical research as there are commercial assays available to hundreds of analytes. If the analyte has clinical value, an automated version would be made available to the clinical laboratorians. If the assay is only used for research purposes, for example, cytokine analysis, then ELISAs are the technique of choice because they can be easily produced by the manufacturers and most research laboratories have ELISA plate readers. The disadvantage of ELISAs is that they typically use enzyme or fluorescence detection, which is not as sensitive as chemiluminescence or radiodetection. The assays are more labor intensive than modern clinical assays, although some laboratories have automated plate washing and reading stations to improve workflow if a high volume of testing is needed.

One of the earliest homogeneous assays was enzyme multiplied immunoassay technique (EMIT), an enzyme immunoassay (Siemens Healthcare Corp).[26] As shown in **Figure 5.17**, the reactants in most test systems include an enzyme-labeled Ag (commonly, a low molecular weight analyte, such as a drug), an Ab directed against the Ag, the substrate, and test Ag. The enzyme is catalytically active when the labeled Ag is free (not bound to the Ab). It is thought that when the Ab combines with the labeled Ag, the Ab sterically hinders the enzyme. The conformational changes that occur during Ag–Ab interaction inhibit the enzyme activity. In this homogeneous assay, the unlabeled Ag in the sample competes with the labeled Ag for the Ab-binding sites; as the concentration of unlabeled Ag increases, less enzyme labeled Ag can bind to the Ab. Therefore, more labeled Ag is free, and the enzymatic activity is greater.

Cloned enzyme donor immunoassays (CEDIAs) are competitive, homogeneous assays in which the genetically engineered label is β-galactosidase

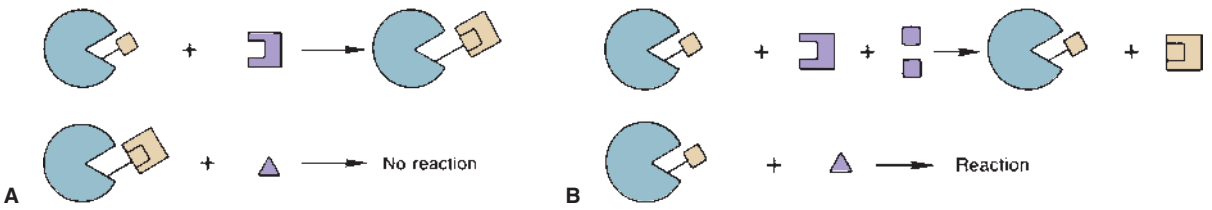


Figure 5.17 Enzyme-multiplied immunoassay technique. **(A)** When enzyme-labeled antigen is bound to the antibody, the enzyme activity is inhibited. **(B)** Free patient antigen binds to the antibody and prevents antibody binding to the labeled antigen. The substrate indicates the amount of free labeled antigen.

(Microgenics Corp).[27] The enzyme is in two inactive pieces: the enzyme acceptor and the enzyme donor. When these two pieces bind together, enzyme activity is restored. In the assay, the Ag labeled with the enzyme donor and the unlabeled Ag in the sample compete for specific Ab-binding sites. When the Ab binds to the labeled Ag, the enzyme acceptor cannot bind to the enzyme donor; therefore, the enzyme is not restored and the enzyme is inactive. More unlabeled Ag in the sample results in more enzyme activity.

Fluorescence excitation transfer immunoassay is a competitive, homogeneous immunoassay using two fluorophores (such as fluorescein and rhodamine).[28] When the two labels are in close proximity, the emitted light from fluorescein will be absorbed by rhodamine. Thus, the emission from fluorescein is quenched. Fluorescein-labeled Ag and unlabeled Ag compete for rhodamine-labeled Ab. More unlabeled Ag lessens the amount of fluorescein-labeled Ag that binds; therefore, more fluorescence is present (less quenching).

Fluorescence polarization immunoassay (FPIA) is another assay that uses a fluorescent label.[29] This homogeneous immunoassay uses polarized light to excite the fluorescent label. Polarized light is created when light passes through special filters and consists of parallel light waves oriented in one plane. When polarized light is used to excite a fluorescent label, the emitted light could be polarized or depolarized. Small molecules, such as free fluorescent labeled Hp, rotate rapidly and randomly, interrupting the polarized light. Larger molecules, such as those created when the fluorescent-labeled Hp binds to an Ab, rotate more slowly and emit polarized light parallel to the excitation polarized light. The polarized light is measured at a 90° angle compared with the path of the excitation light. In a competitive FPIA, fluorescent-labeled Hp and unlabeled Hp in the sample compete for limited Ab sites. When no unlabeled Hp is present, the labeled Hp binds maximally to the Ab, creating large complexes that rotate slowly and emit a high level of polarized light. When Hp is present, it competes with the labeled Hp for the Ab sites; as the Hp concentration increases, more labeled Hp is displaced and is free. The free labeled Hp rotates rapidly and emits less polarized light. The degree of labeled Hp displacement is inversely related to the amount of unlabeled Hp present.

Dissociation-enhanced lanthanide fluoroimmunoassay (DELFIA) is an automated system (Thermo Fisher Scientific) that measures time-delayed fluorescence from the label europium. The assay can be designed as a competitive, heterogeneous assay or a noncompetitive (sandwich), heterogeneous assay.[30]

Total Laboratory Automation

Automated analyzers are now commonplace in clinical laboratories, and the current focus and rapidly progressing areas in automation are nonanalytic automation and automating laboratory workflows from sample input to final result.[31,32] Automation of the pre-analytic, analytic, and post-analytic phase is referred to as **total laboratory automation (TLA)**. While most TLA is still vendor-specific, some automation equipment vendors are developing open architecture components that provide more flexibility in automation implementation.[33] An example of a commercial TLA system is shown in **Figure 5.18**.

Preanalytic Phase

Preparation of the sample for analysis has been and remains a manual process in most laboratories. The clotting time (if using serum), centrifugation, and the transferring of the sample to an analyzer cup (unless using primary tube sampling) can cause delays and expenses in the testing process. One alternative to manual preparation is to automate this process by using **robotics**, or front-end automation, to "handle"

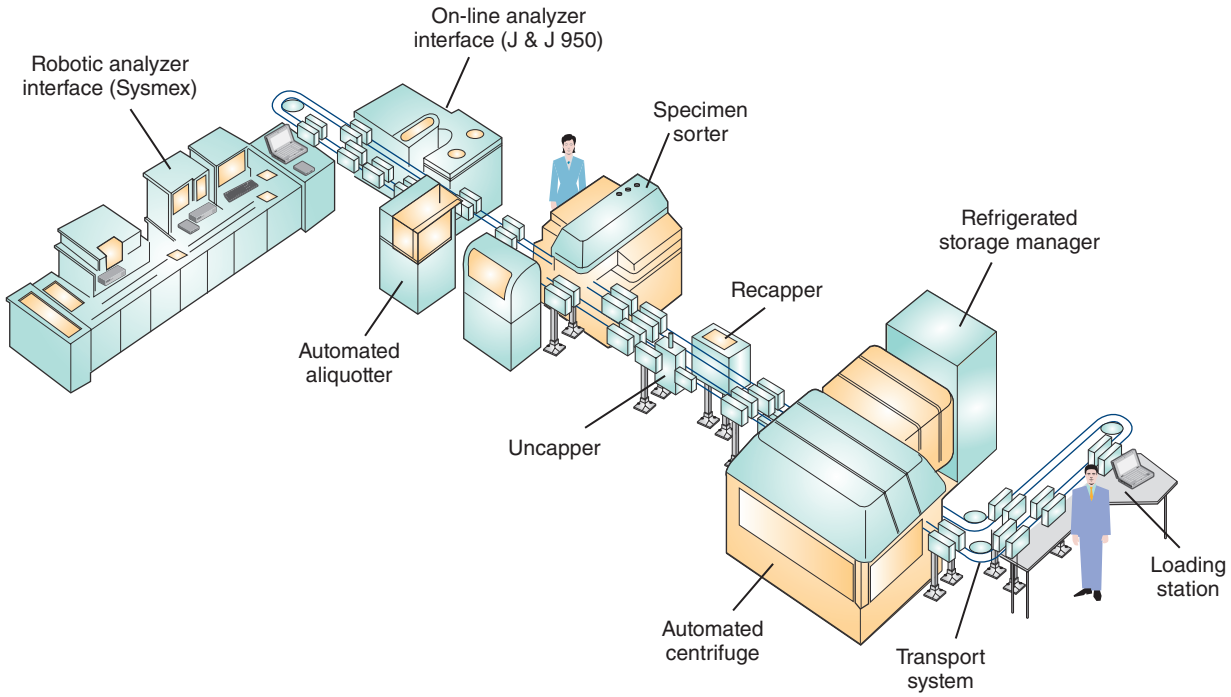


Figure 5.18 Schematic of total laboratory automation system.

Courtesy of Cerner Labotix.

the specimen through these steps and load the specimen onto the analyzer.

Automated processes are gradually replacing manual handling and presentation of the sample to the analyzer. Increasing efficiency while decreasing costs has been a major impetus for laboratories to start integrating some aspects of TLA into their operations. Conceptually, TLA refers to automated devices and robots integrated with existing analyzers to perform all phases of laboratory testing. Most attention to date has been devoted to development of the front-end systems that can identify and label specimens, centrifuge the specimen and prepare aliquots, and sort and deliver samples to the analyzer or to storage.[34] Back-end systems may include removal of specimens from the analyzer and transport to storage, retrieval from storage for retesting, realiquoting, or disposal, as well as comprehensive management of the data from the analyzer and interfacing with the LIS.

Dr. Masahide Sasaki installed the first fully automated clinical laboratory in the world at Kochi Medical School in Japan[35]; since then, the concept has gradually, but steadily, become a reality in the United States. The University of Nebraska and the University of Virginia have been pioneers for TLA system development. In 1992, a prototype of a laboratory automation platform was developed at the University of Nebraska, the key components being a conveyance system, bar-coded specimens, a computer software package to control specimen movement and tracking, and coordination of robots with the instruments as work cells.[36] Some of the first automated laboratories in the United States have reported their experiences with front-end automation with a wealth of information for others interested in the technology.[37,38] The first hospital laboratory to install an automated system was the University of Virginia Hospital in Charlottesville in 1995. Their Medical Automation Research Center cooperated with Johnson & Johnson and Coulter Corporation to use a VITROS 950 attached to a Coulter/IDS "U" lane for direct sampling from a specimen conveyor without using intervening robotics.[39] The first commercially available turnkey system was the Hitachi Clinical Laboratory Automation System (Boehringer-Mannheim Diagnostics; now Roche Diagnostics). It couples the Hitachi line of analyzers to a conveyor belt system to provide a completely operational system with all interfaces.[40]

Robotics and front-end automation are changing the face of the clinical laboratory.[41] Much of the benefit derivable from TLA can be realized merely by automating the front end. The planning, implementation, and performance evaluation of an automated transport and sorting system by a large reference laboratory have been described in detail.[42,43] Several instrument manufacturers are currently working on or are already marketing interfacing front-end

devices together with software for their own chemistry analyzers. Johnson & Johnson introduced the VITROS 950 AT (Automation Technology) system in 1995 with an open architecture design to allow laboratories to select from many front-end automation systems rather than being locked into a proprietary interface. A Lab-Track interface is now available on the Dimension RxL (Siemens) that is compatible with major laboratory automation vendors and allows for direct sampling from a track system. Also, the technology now exists for microcentrifugal separators to be integrated into clinical chemistry analyzers.[44] Several other systems are now on the market, including the Advia LabCell system (Siemens), which uses a modular approach to automation. The Power Processor Core System (Beckman Coulter) performs sorting, centrifugation, and cap removal. The enGen Series Automation System (Ortho-Clinical Diagnostics) provides sorting, centrifugation, uncapping, and sample archiving functions and interface directly with a VITROS 950 AT analyzer. The instruments listed in **Table 5.5** are examples of current TLA solutions offered commercially. Much of the benefit of TLA is derived from automation of the front-end processing steps. Therefore, several manufacturers have developed stand-alone, automated front-end processing systems. The Genesis FE500 (Tecan) is an example of a stand-alone front-end system that can centrifuge, uncap, aliquot into a labeled pour-off tube, and sort into analyzer racks. Systems with similar functionality are available from Labotix, Motoman, and PVT. An example of one such system is shown in **Figure 5.19**. Stand-alone automated sample uncappers and recappers are available from PVT and Sarstedt. These latter devices are less flexible than the complete stand-alone front-end systems and require samples to be presented to them in racks that will

Table 5.5 Summary of Features for Selected Laboratory Automation Systems and Work Cells

	Abbott Diagnostics	Beckman Coulter	Inpeco	Roche Diagnostics Corp.	Siemens Healthineers	Yaskawa America, Motoman Robotics Division
Instrument Name	**Accelerator a3600**	**Power Processor**	**FlexLab Automation**	**Cobas Connection Modules (CCM)**	**Aptio Automation**	**AutoSorter 1200**
First year sold in the United States	2013	1998	2008	2016	2013	2013
Pre-analytic processor/total laboratory automation	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/Yes	Yes/No
Automated centrifugation available	Yes	Yes	Yes	Yes	Yes	No
Automated input/accessioning available	Yes	Yes	Yes	Yes	Yes	Yes
Automated decapping available	Yes	Yes	Yes	Yes	Yes	No
Automated sorting available	Yes	Yes	Yes	Yes	Yes	Yes
Automated aliquoting available	Yes	Yes	Yes	Yes	Yes	No
Automated recapper or sealer available	Yes, recapper and sealer	Yes, recapper	Yes, recapper and sealer	Yes, sealer	Yes, recapper and sealer	No
Automated storage and retrieval available	Yes	Yes	Yes	Yes	Yes	Yes

Information obtained from *CAP Today*. 2017;September:42–64.

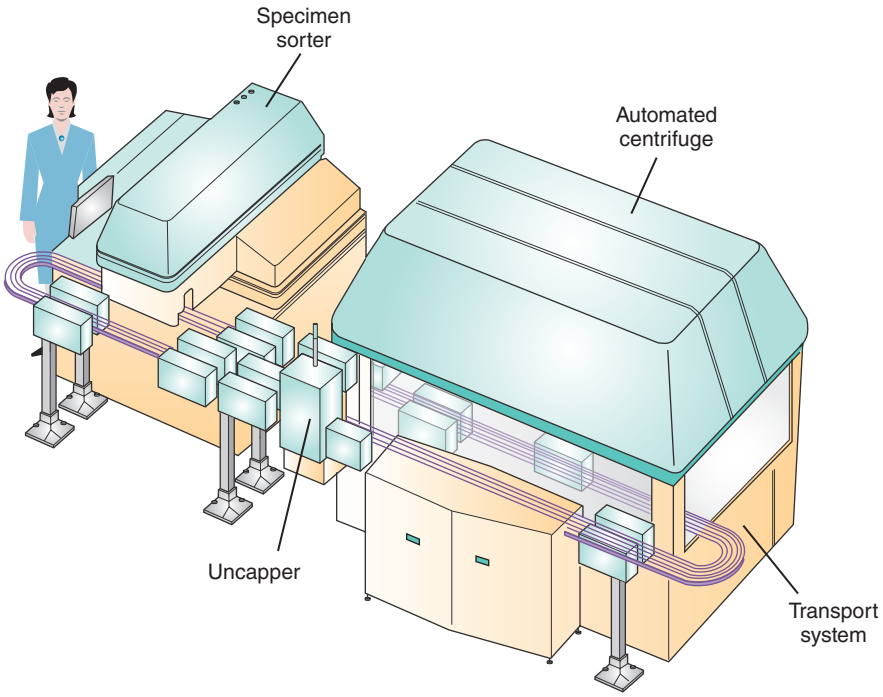


Figure 5.19 Schematic of pre-analytic automation system.

Courtesy of Cerner Labotix.

work with a single analyzer. Some laboratories have taken a modular approach with devices for only certain automated functions. Ciba-Corning Clinical Laboratories installed Coulter/IDS robotic systems in several regional laboratories.[39] Recently, a thawing–mixing work cell that is compatible with a track system in a referral laboratory has been described.[45] The bottom line is that robotics and front-end automation are here to stay. As more and more clinical laboratories reengineer for TLA, they are building core laboratories containing all of their automated analyzers as the necessary first step to link the different instruments more easily into one TLA system.[46]

Analytic Phase

There have been changes and improvements that are now common to many automated chemistry and immunoassay analyzers. They include ever smaller microsampling and reagent dispensing with multiple additions possible from randomly replaced reagents; expanded onboard and total test menus, especially drugs and hormones; accelerated reaction times with chemistries for faster throughput and lower dwell time; higher resolution optics with grating monochromators and diode arrays for polychromatic analysis; improved flow through electrodes; enhanced user-friendly interactive software for quality control, maintenance, and diagnostics; integrated modems for online troubleshooting; LIS-interfacing data management systems; reduced frequencies of calibration and controls; automated modes for calibration, dilution, rerun, and maintenance; as well as ergonomic and physical design improvements for operator ease, serviceability, and maintenance reduction. The features and specifications of five mid-volume and five high-volume systems are summarized in Table 5.1. In addition to the improvements in automated analyses listed above, significant efficiency can be gained through modular analytics consisting of either multiple chemistry analyzers or connected chemistry and immunoassay analyzers. This functionality removes the need to split samples, performs onboard dilutions, and reprograms repeat testing, which reduces the need for operator intervention during testing.

Postanalytic Phase

Specimen Storage and Retrieval

Post-analytic specimen storage may be integrated into total laboratory automation. Refrigerated storage units capable of holding thousands of specimen tubes can function as a post-analytic holding area for specimens prior to being discarded. Bidirectional track systems between the analyzers and storage units make the process of physician "add-ons," where laboratory tests are added to the original order after initial results are obtained, an automated process. Specimens can also be automatically pulled from

the storage unit for repeat or reflex testing using the automation/LIS software.

Data Management

Although most of the attention in recent years in TLA concept has been devoted to front-end systems for sample handling, several manufacturers have been developing and enhancing back-end handling of data. Bidirectional communication between the analyzer(s) and the host computer or LIS has become an absolutely essential link to request tests and enter patient demographics, automatically transfer this customized information to the analyzer(s), as well as post the results in the patient's record. Evaluation and management of data from the time of analysis until posting have become more sophisticated and automated with the integration of work station managers into the entire communication system.[47] Most data management devices are personal computer–based modules with manufacturers' proprietary software that interfaces with one or more of their analyzers and the host LIS. They offer automated management of quality control data with storage and evaluation of quality control results against the laboratory's predefined quality control perimeters with multiple plotting, displaying, and reporting capabilities. Review and editing of patient results before verification and transmission to the host are enhanced by user-defined perimeters for reportable range limits, panic value limits, delta checks, and quality control comparisons for clinical change, repeat testing, and algorithm analysis. Reagent inventory and quality control, along with monitoring of instrument functions, are also managed by the workstation's software. Most LIS vendors have interfacing software available for all the major chemistry analyzers.

Some data handling needs associated with automation cannot be adequately handled by most current LISs and require specialized **middleware**. For example, most current analyzers are capable of assessing the degree of sample hemolysis (H), icterus (I), and lipemia (L) (see Table 5.1). The use of these automated serum indices to assess hemolysis, icterus, and lipemia on automated analyzers and determine specimen acceptability based on logic in the middleware data management systems has revolutionized automated assessment of specimen integrity.

However, making this information available and useful to the laboratorian in an automated fashion requires additional manipulation of the data. Ideally, the tests ordered on the sample, the threshold for interference of each test by each of the three agents, and whether the interference is positive or negative need to be determined. In the case of lipemia, the results for affected tests need to be held until the sample can be clarified and the tests rerun. One company, Data Innovations, has developed a middleware system called Instrument Manager, which links the analyzer to the LIS and provides the ability for the user to define rules for release of information to the LIS. In addition, flags can be displayed to the instrument operator to perform additional operations, such as sample clarification and reanalysis. The ability to fully automate data review using rules-based analysis is a key factor in moving toward TLA. The use of "autoverification" capabilities found with many post-analytic LISs has contributed to a significant reduction in result turnaround time.

Future Trends in Automation

Total laboratory automation continues to evolve at a rapid pace in the 21st century. With most of the same forces driving the automation market as those

CASE STUDY 5.1, PART 3

Remember Mía who is training on the Roche Cobas.

3. What does I.H stand for? What does I.L stand for?

Mía referenced the SOP to look up the AST assay. Under interpretation, the Abs flag was explained and stated if the AST results > 700 U/L and < 7000 U/L, then the analyzer will automatically dilute X10.

AST 1078H > Abs AST 1314H

4. What AST value should be reported?

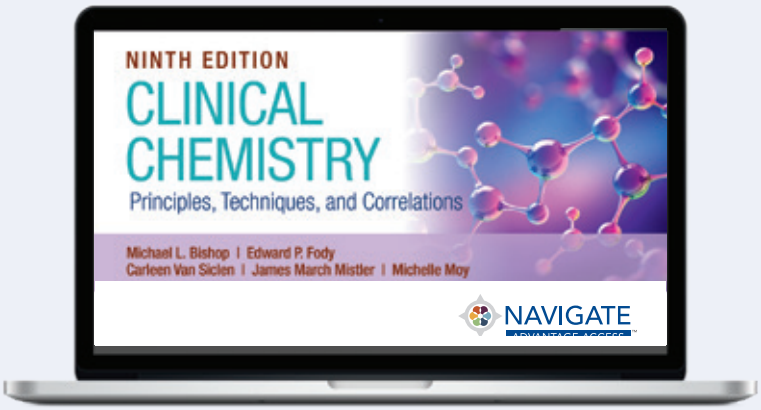
© Ariel Skelley/DigitalVision/Getty Images.

discussed in this chapter, analyzers will continue to perform more cost effectively and efficiently. Effective communications among all automation stakeholders for a given project are key to successful implementation.[48]

In the coming years, more system and workflow integration will occur with robotics and data management for more inclusive TLA.[49] The incorporation of artificial intelligence and machine learning into analytic systems will likely evolve and expand within the clinical laboratory.[50,51] This will greatly advance the technologies of robotics, digital processing of data, computer-assisted diagnosis, and data integration with electronic patient records.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Hodnett J. Automated analyzers have surpassed the test of time. *Adv Med Lab.* 1994;6:8.
2. Eisenwiener H, Keller M. Absorbance measurement in cuvettes lying longitudinal to the light beam. *Clin Chem.* 1979;25:117–121.
3. Schoeff LE, Williams RH. *Principles of Laboratory Instruments*. St. Louis, MO: Mosby–Yearbook; 1993.
4. Jacobs E, Simson E. Point of care testing and laboratory automation: the total picture of diagnostic testing at the beginning of the next century. *Clin Lab News.* 1999;25(12): 12–14.
5. Boyce N. Why hospitals are moving to core labs. *Clin Lab News.* 1996;22:1–2.
6. Boyce N. Why labs should discourage routine testing. *Clin Lab News.* 1996;22:1, 9.
7. Burtis CA. Factors influencing evaporation from sample cups, and assessment of their effect on analytic error. *Clin Chem.* 1975;21:1907–1917.
8. Burtis CA, Watson JS. Design and evaluation of an antievaporative cover for use with liquid containers. *Clin Chem.* 1992;38:768–775.
9. Dudley RF. Chemiluminescence immunoassay: an alternative to RIA. *Lab Med.* 1990;21:216.
10. Berson SA, Yalow RS. Assay of plasma insulin in human subjects by immunological methods. *Nature.* 1959;184: 1648–1649.
11. Sheehan C. An overview of antigen–antibody interaction and its detection. In: Sheehan C, ed. *Clinical Immunology: Principles and Laboratory Diagnosis*. 2nd ed. Philadelphia, PA: Lippincott-Raven Publishers; 1997:109.
12. Kohler G, Milstein C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature.* 1975; 256:495.
13. Kusnetz J, Mansberg HP. Optical considerations: nephelometry. In: Ritchie RF, ed. *Automated Immunoanalysis. Part 1*. New York: Marcel Dekker; 1978.
14. Engvall E, Perlmann P. Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G. *Immunochemistry.* 1971;8:871.
15. Van Weemen BK, Schuurs AHWM. Immunoassay using antigen-enzyme conjugates. *FEBS Lett.* 1971;15:232.
16. Nakamura RM, Bylund DJ. Fluorescence immunoassays. In: Rose NR, deMarcario EC, Folds JD, et al., eds. *Manual of Clinical Laboratory Immunology*. 5th ed. Washington, DC: ASM Press; 1997:39.
17. Diamandis EP, Evangelista A, Pollack A, et al. Time-resolved fluoroimmunoassays with europium chelates as labels. *Am Clin Lab.* 1989;8(8):26.
18. Kricka LJ. Chemiluminescent and bioluminescent techniques. *Clin Chem.* 1991;37:1472.
19. Kricka LJ. Selected strategies for improving sensitivity and reliability of immunoassays. *Clin Chem.* 1994;40:347.
20. Bronstein I, Juo RR, Voyta JC. Novel chemiluminescent adamantyl 1,2-dioxetane enzyme substrates. In: Stanley PE, Kricka LJ, eds. *Bioluminescence and Chemiluminescence: Current Status*. Chichester: Wiley; 1991:73.

21. Edwards B, Sparks A, Voyta JC, et al. New chemiluminescent dioxetane enzyme substrates. In: Campbell AK, Kricka LJ, Stanley PE, eds. *Bioluminescence and Chemiluminescence: Fundamentals and Applied Aspects*. Chichester: Wiley; 1994:56–59.
22. Kricka LJ. Human anti-animal antibody interferences in immunological assays. *Clin Chem.* 1999;45:942.
23. Kaplan IV, Levinson SS. When is a heterophile antibody not a heterophile antibody? When it is an antibody against a specific immunogen. *Clin Chem.* 1999;45:616.
24. Butch AW. Dilution protocols for detection of hook effects/prozone phenomenon. *Clin Chem.* 2000;46:1719.
25. Litchfield WJ. Shell-core particles for the turbidimetric immunoassays. In: Ngo TT, ed. *Nonisotopic Immunoassay*. New York, NY: Plenum Press; 1988.
26. Rubenstein KE, Schneider RS, Ullman EF. "Homogeneous" enzyme immunoassay. A new immunochemical technique. *Biochem Biophys Res Commun.* 1972;47:846.
27. Henderson DR, Freidman SB, Harris JD, et al. CEDIA, a new homogeneous immunoassay system. *Clin Chem.* 1986;32:1637.
28. Ullman EF, Schwartzberg M, Rubinstein KD. Fluorescent excitation transfer assay: a general method for determination of antigen. *J Biol Chem.* 1976;251:4172.
29. Dandliker WB, Kelly RJ, Dandiker BJ, et al. Fluorescence polarization immunoassay: theory and experimental methods. *Immunochemistry.* 1973;10:219.
30. Diamandis EP. Immunoassays with time-resolved fluorescence spectroscopy: principles and applications. *Clin Biochem.* 1988;21:139.
31. Armbruster, D.A., Overcash, D.R., Reyes, J. Clinical chemistry laboratory automation in the 21st century – *Amat Victoria curam* (Victory loves careful preparation). *Clin Biochem Rev.* 2014;35:143–153.
32. Hawker, C.D. Nonanalytic laboratory automation: A quarter century of progress. *Clin Chem.* 2017;63:1074–1082.
33. Douglas L. Redefining automation: perspectives on today's clinical laboratory. *Clin Lab News.* 1997;23:48–49.
34. Felder R. Front-end automation. In: Kost G, ed. *Handbook of Clinical Laboratory Automation and Robotics*. New York, NY: Wiley; 1995.
35. Sasaki M. A fully automated clinical laboratory. *Lab Inform Mgmt.* 1993;21:159–168.
36. Markin R, Sasaki M. A laboratory automation platform: the next robotic step. *MLO Med Lab Obs.* 1992;24:24–29.
37. Bauer S, Teplitz C. Total laboratory automation: a view of the 21st century. *MLO Med Lab Obs.* 1995;27:22–25.
38. Bauer S, Teplitz C. Laboratory automation, part 2. Total lab automation: system design. *MLO Med Lab Obs.* 1995;27:44–50.
39. Felder R. Cost justifying laboratory automation. *Clin Lab News.* 1996;22:10–11, 17.
40. Felder R. Laboratory automation: strategies and possibilities. *Clin Lab News.* 1996;22:10–11.
41. Boyd J, Felder R, Savory J. Robotics and the changing face of the clinical laboratory. *Clin Chem.* 1996;42:1901–1910.
42. Hawker CD, Garr SB, Hamilton LT, et al. Automated transport and sorting system in a large reference laboratory: part 1. Evaluation of needs and alternatives and development of a plan. *Clin Chem.* 2002;48:1751–1760.
43. Hawker CD, Roberts WL, Garr SB, et al. Automated transport and sorting system in a large reference laboratory: part 2. Implementation of the system and performance measures over three years. *Clin Chem.* 2002;48:1761–1767.
44. Richardson P, Molloy J, Ravenhall R, et al. High speed centrifugal separator for rapid online sample clarification in biotechnology. *J Biotechnol.* 1996;49:111–118.
45. Hawker CD, Roberts WL, DaSilva A, et al. Development and validation of an automated thawing and mixing workcell. *Clin Chem.* 2007;53:2209–2211.
46. Zenie F. Re-engineering the laboratory. *J Automat Chem.* 1996;18:135–141.
47. Saboe T. Managing laboratory automation. *J Automat Chem.* 1995;17:83–88.
48. Fisher JA. Laboratory automation: communicating with all stakeholders is the key to success. *Clin Lab News.* 2000;26(7):38–40.
49. Brzezicki L. Workflow integration: does it make sense for your lab? *Adv Lab.* 1996;23:57–62.
50. Place J, Truchaud A, Ozawa K, et al. Use of artificial intelligence in analytical systems for the clinical laboratory. *J Automat Chem.* 1995;17:1–15.
51. Boyce N. Neural networks in the lab: new hope or just hype? *Clin Lab News.* 1997;23:2–3.

PART 2

Analytic Procedures and Clinical Correlation

CHAPTER 6

Proteins and Heme Derivatives

Takara L. Blamires and Kristina Jackson Behan

CHAPTER OUTLINE

KEY TERMS

Aminoacidopathies
Amino acid
Amphoteric
Conjugated or complexed protein
Heme
Hyperproteinemia
Hypoproteinemia
Isoelectric point
Negative acute-phase reactant
Peptide bond
Polypeptide
Positive acute-phase reactant
Primary structure
Protein
Proteinuria
Pyrrole
Quaternary structure
Simple protein
Secondary structure
Tertiary structure
Zwitterion

CHAPTER OBJECTIVES

At the end of this unit of study, the clinical laboratorian should be able to:

- Define key terms associated with amino acids, proteins, and heme derivatives.
- Describe the basic structure and general properties of amino acids.
- Differentiate between essential, conditionally essential, and nonessential amino acids used in protein synthesis.
- Outline the basic structure, functions, and classifications of proteins.
- Explain the processes involved in metabolism, synthesis, and degradation of proteins.
- Diagram the processes involved in heme synthesis.
- Discuss the function and clinical significance of proteins and heme derivatives commonly measured in serum, plasma, urine, cerebrospinal fluid, and other body fluids.
- Specify the specimen requirements and common test methods used in the quantitative analysis of proteins and heme derivatives.
- State reference ranges for serum or plasma total protein and albumin.
- Discuss pathologic and nonpathologic factors that influence total protein and albumin concentrations.
- Identify the fractions of a serum protein electrophoresis and list proteins expected to migrate into each.
- Correlate patient results with associated disease states, disorders, or pre-analytical issues.

CASE STUDY 6.1, PART 1

A young couple, concerned for their 3-week-old infant son, Sean, took him to the emergency department late one night. Upon examination by the physician, the infant had a temperature of 103°F and a runny nose. The physician determined he was suffering from a common cold. In the months that followed, Sean experienced an unusual number of bacterial infections. The prescribed antibiotics cleared up the respiratory infections each time, but Sean continued to have recurring bacterial infections.

© Myroslava Malovana/Shutterstock.

CASE STUDY 6.2, PART 1

Guillermo, a 47-year-old man, had fallen and broken his leg. In the emergency department, he explained his complicated medical history, with type 2 diabetes, peripheral neuropathy, and chronic renal insufficiency. His complete blood count (CBC) showed a normochromic, normocytic anemia.

© Russell Monk/The Image Bank/Getty Images.

CASE STUDY 6.3, PART 1

Fiona, a 36-year-old woman, was seen by her physician with complaints of intermittent blurred vision and numbness and weakness in her left leg that had persisted for longer than 3 weeks. On examination, vertical nystagmus (involuntary back-and-forth or circular movements of eyes) was noted on upward gaze.

© Westend61/Getty Images.

CASE STUDY 6.4, PART 1

Genevieve, concerned for her 2-month-old infant son, Mateo, brought him to his pediatrician due to blisters on his exposed skin.

© Jose Luis Pelaez Inc/DigitalVision/Getty Images.

Proteins are large, complex molecules comprised of 200 to 300 amino acids. In humans, there are 20 different amino acids used in the synthesis of proteins. Some of these amino acids can be synthesized by the body, while others must be obtained through the diet. The sequence of amino acids and the interactions that occur between them direct the structure and function of the protein. Proteins perform a broad range of functions throughout the body; therefore, maintaining an appropriate balance of protein synthesis and degradation is essential. This chapter discusses the general properties and functions of amino acids and proteins, the processes involved in protein synthesis and degradation, the clinical significance of select proteins, and test methods used to analyze proteins in the clinical laboratory.

Amino Acids

Overview

Amino acids are simple organic compounds that serve as the building blocks of proteins, but they are also used in the synthesis of nitrogen-containing nonprotein compounds such as purines, pyrimidines, porphyrins, creatine, histamine, thyroxine, epinephrine, and coenzyme NAD. This chapter will focus on discussion of proteins.

Basic Structure

A single amino acid contains at least one amino group and one carboxyl functional group. The N-terminal end of the amino group ($-NH_2$) and the C-terminal end of the carboxyl group (–COOH) are bonded to an α-carbon forming an amino acid.[1] The basic structure of an amino acid is depicted in **Figure 6.1**. Amino acids are **amphoteric**, meaning they have acidic and basic tendencies. The amino group can accept a proton (basic) forming $-NH_3^+$, while the carboxyl group can donate a proton (acidic) forming $–COO^-$. Since amino acids can have neutral, positive, and negative functional groups, but their net charge is neutral, they are classified as **zwitterions**.

Amino acids differ from one another by the chemical composition of their R group, referred to as side chains.[2] The R groups found on the amino acids used in human protein synthesis are shown in **Table 6.1** as well as the three-letter and one-letter codes used to refer to each. The amino group of one amino acid can be covalently linked with the carboxyl group of another amino acid forming a **peptide bond** (**Figure 6.2**). When a chain of amino acids is linked by peptide bonds, it is known as a **polypeptide**, and a large polypeptide constitutes a protein. Proteins found in human plasma range from 100 to 150 amino acids in the length of their polypeptide chains. The content and arrangement of amino acids in a specific protein are determined by the sequence of nucleotide bases in the gene that encodes for that protein.[2]

Metabolism

About half of the amino acids required by humans for protein synthesis cannot be produced in vivo at a rapid enough rate to support growth.[2] These essential amino acids must be supplied by the diet in the form of proteins. Under normal circumstances, food is mechanically digested in the mouth then transported to the stomach, where chemical digestion occurs.

In the stomach, gastrin stimulates the secretion of hydrochloric acid (HCl) and proteolytic enzymes, such as pepsin, to promote denaturation of the proteins and to catalyze their hydrolysis into

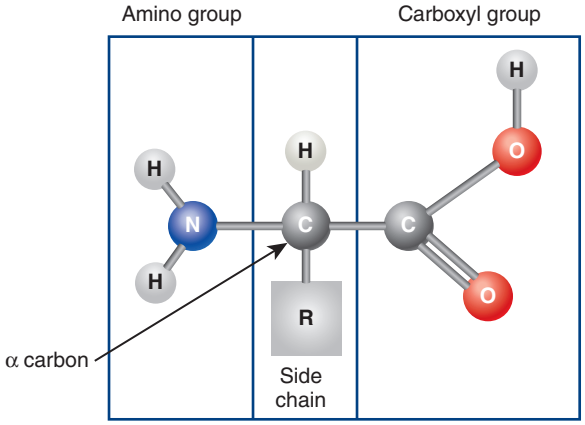


Figure 6.1 General structure of amino acid.

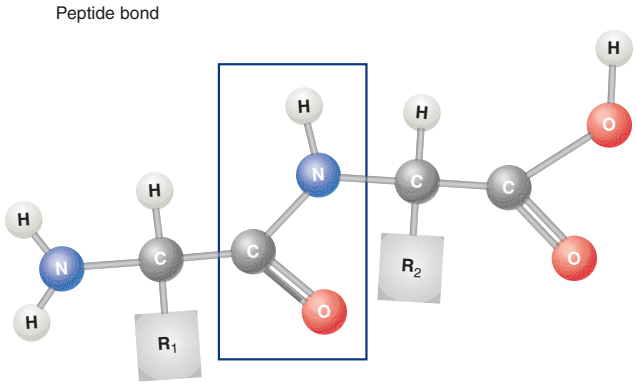
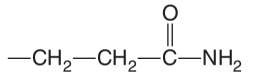
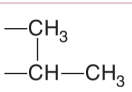
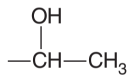
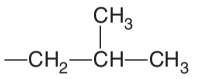
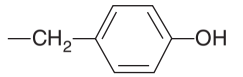
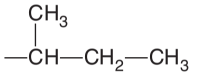
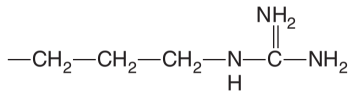
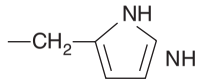
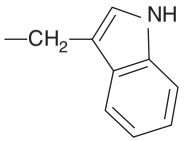
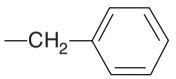
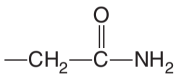
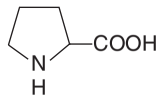


Figure 6.2 Formation of a peptide bond.

Table 6.1 Amino Acids Utilized in Synthesis of Human Proteins

Amino Acid	Three-Letter Code	One-Letter Code	R Group	Amino Acid	Three-Letter Code	One-Letter Code	R Group
Glycine[b]	Gly	G	—H	Glutamine[b]	Gln	Q	
Alanine	Ala	A	$—CH_3$	Serine	Ser	S	
Valine[a]	Val	V		Threonine[a]	Thr	T	
Leucine[a]	Leu	L		Tyrosine[b]	Tyr	Y	
Isoleucine[a]	Ile	I		Lysine[a]	Lys	K	$—CH_2—CH_2—CH_2—CH_2—NH_2$
Cysteine[b]	Cys	C	$—CH_2—SH$	Arginine[b]	Arg	R	
Methionine[a]	Met	M	$—CH_2—CH_2—S—CH_3$	Histidine[a]	His	H	
Tryptophan[a]	Trp	W		Aspartic acid	Asp	D	$—CH_2—COOH$
Phenylalanine[a]	Phe	F		Glutamic acid	Glu	E	$—CH_2—CH_2—COOH$
Asparagine	Asn	N		Proline[b,c]	Pro	P	
Selenocysteine	Sec	U					

The R group is the group attached to the α-carbon.

[a] Essential

[b] Conditionally essential

[c] Exception to attachment of R group as the side chain curves back from the α-carbon to bond to the amine nitrogen

polypeptides. In the small intestine, HCl is neutralized by sodium bicarbonate secreted from the pancreas under the direction of secretin, a digestive hormone, in order to protect the intestinal lining. Secretin and cholecystokinin, another digestive hormone secreted by the small intestine, stimulate the release of proteolytic enzymes, such as trypsin and chymotrypsin, from the pancreas and bile from the gallbladder. These pancreatic enzymes continue the process of digestion by hydrolyzing polypeptides into tripeptides, dipeptides, and the constituent amino acids. **Figure 6.3** illustrates basic processes involved in digestion of dietary proteins.[3]

Dipeptides, tripeptides, and amino acids are absorbed from the intestinal lumen through several cotransporters. Carrier proteins on the luminal surface of the epithelial cell bind with an ion, such as hydrogen or sodium, and a dipeptide, tripeptide, or amino acid, which are then transported into the cell. Peptidases within the cell cleave any remaining peptide bonds resulting in free amino acids. Facilitative transporters on the basolateral surface of the epithelial cell allow transport of the amino acids into the bloodstream, where they will be transported to the liver via the hepatic portal vein. In the liver, the amino acids will be used to synthesize new proteins and nonprotein nitrogen compounds or be used to generate energy through gluconeogenesis or ketogenesis.[3] **Figure 6.4** illustrates the process for absorbing amino acids in the intestines.

Classification

Essential Amino Acids

There are nine essential amino acids: histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine.[4] As mentioned, these amino acids cannot be synthesized in vivo at a fast enough rate, if at all, and must be acquired through dietary intake.

Histidine is needed to help grow and repair body tissues, to maintain the myelin sheaths that protect nerve cells, and to serve as the precursor for several hormones and metabolites essential to renal, gastric, and immune function (e.g., histamine). It also plays an important role in the synthesis of red and white blood cells, in protecting the body from heavy metal toxicity, and serving as a source of carbon atoms in the synthesis of purines for DNA and RNA synthesis.[5]

Isoleucine, *leucine*, and *valine* are branched-chain amino acids, collectively referred to as the branched-chain amino acid group. Isoleucine is concentrated in muscle tissues and is essential for a number of functions such as wound healing, immune function, glucose homeostasis and hemoglobin formation.[6] Leucine is also important in the regulation of blood glucose and wound healing as well as prevention of muscle protein degradation subsequent to trauma.[7] Valine aids in determining the tertiary structure of proteins and promotes mental health and muscle coordination.[8]

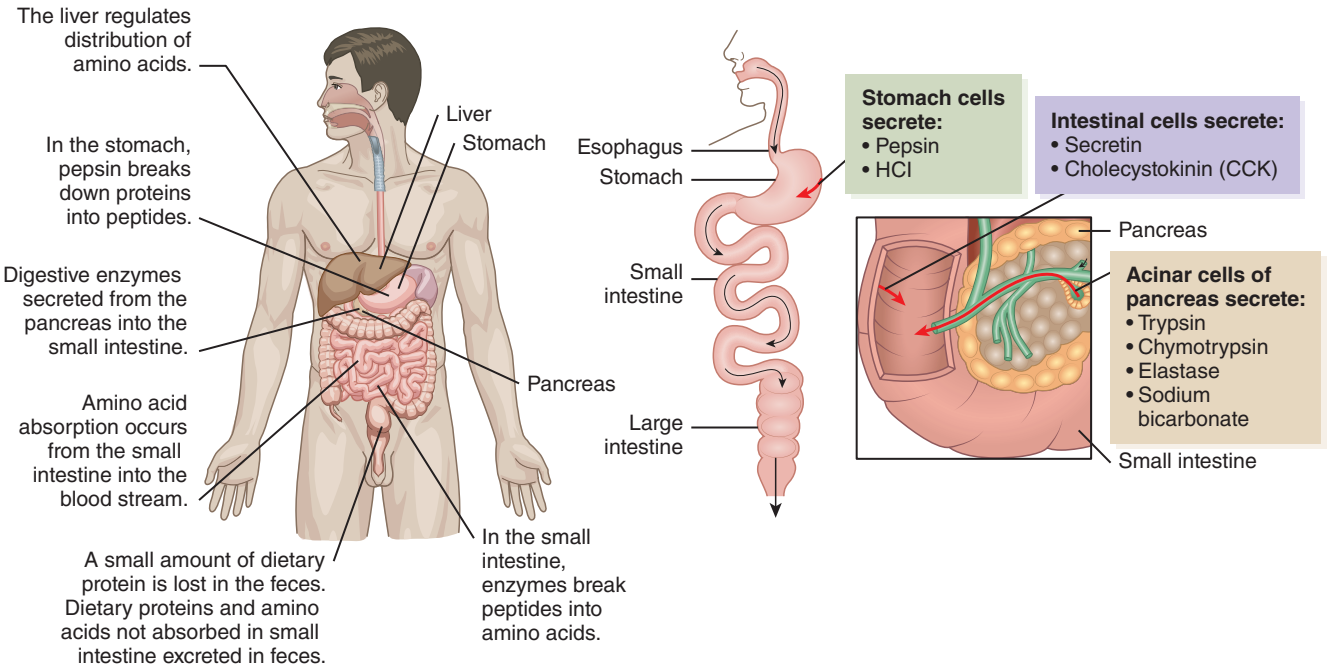


Figure 6.3 Digestion of dietary proteins.

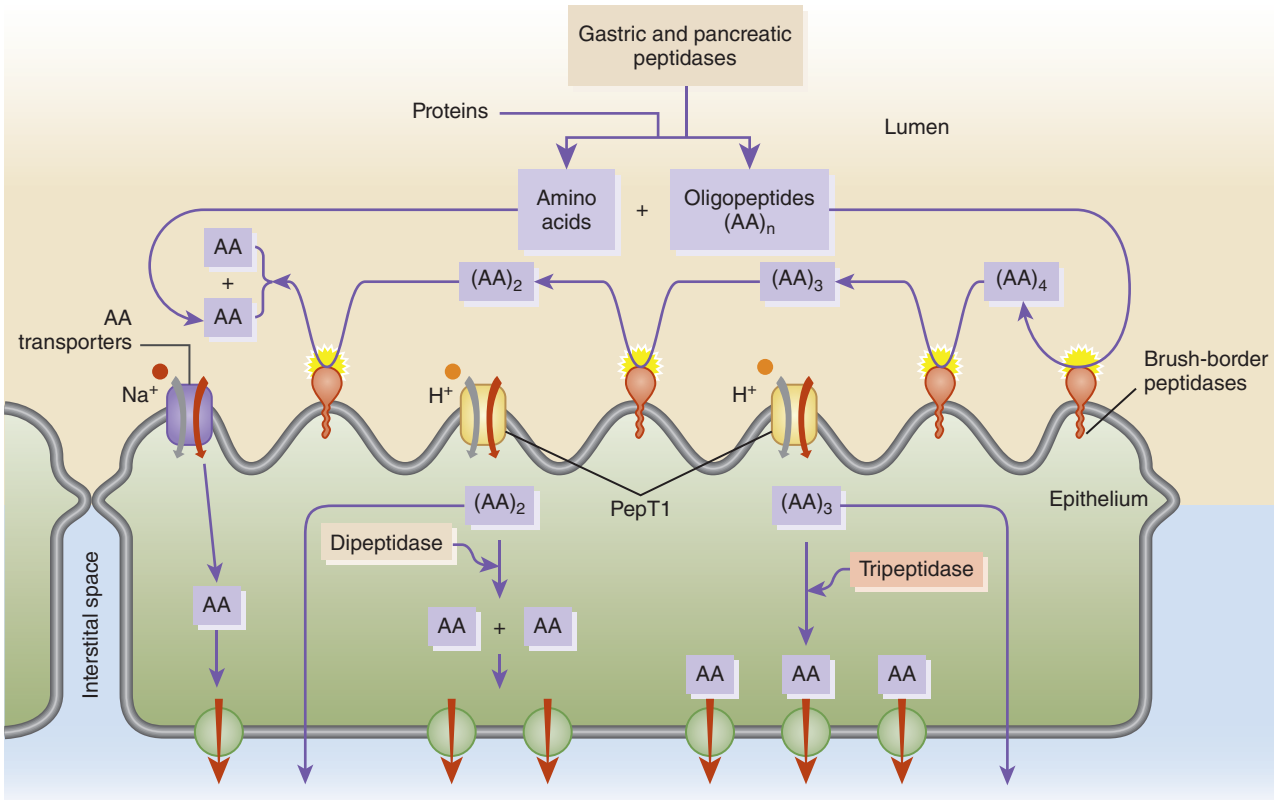


Figure 6.4 Intestinal absorption of amino acids.

Lysine plays a role in the production of antibodies and is required for maintaining healthy tissues. It also helps in the absorption and conservation of calcium and serves an important role in the formation of collagen, a component of cartilage and connective tissue.[9]

Methionine helps to initiate translation of messenger RNA, stabilizes protein structure, and is an important cellular antioxidant. It is an important source of sulfur, which is required for normal metabolism and growth. Methionine also assists in the breakdown of fats, helps to detoxify lead and other heavy metals, helps diminish muscle weakness, and prevents brittle hair. Methionine reacts with adenosine triphosphate (ATP) in the synthesis of many important substances, including epinephrine and choline, and is essential for proper absorption of selenium and zinc.[10]

Phenylalanine is the metabolic precursor for tyrosine, which, in turn, is the precursor for the neurotransmitters dopamine, norepinephrine, and epinephrine, collectively referred to as catecholamines. Catecholamines perform several important roles in maintaining homeostasis through the autonomic nervous system such as promoting alertness and vitality, elevating mood, decreasing pain, and aiding in memory and learning. When deficient, downstream products such as tyrosine will also be decreased.[11]

Threonine is an important component in the formation of collagen, elastin, and tooth enamel. It is also important in the production of neurotransmitters and overall health of the nervous system. Additionally, threonine helps maintain proper protein balance in the body, aids in liver function, and assists in metabolism of porphyrins and fats.[12]

Tryptophan is a metabolic precursor for serotonin and melatonin, which regulate appetite, mood, sleep, and pain. As such, tryptophan is a natural relaxant that helps alleviate insomnia by inducing sleep, soothes anxiety, and reduces depression. Tryptophan is also essential for the production of niacin.[13]

Nonessential Amino Acids

The human body can synthesize adequate amounts of alanine, asparagine, aspartic acid, glutamic acid, selenocysteine, and serine. Since additional dietary intake of these amino acids is unnecessary, they are classified as nonessential.[4]

Alanine is a product of the breakdown of DNA, anserine, and carnosine. It is also formed as a result

of glycolysis in muscle tissue and the conversion of pyruvate, a pivotal compound in carbohydrate metabolism, into α-ketoglutarate. Alanine plays a major role in the transfer of nitrogen from peripheral tissues to the liver for processing and excretion and strengthens the immune system through production of antibodies. As a ketogenic amine, alanine also serves as a source of energy for the central nervous system, brain, and muscle tissues. It also helps in reducing the buildup of toxic substances released when muscle protein is broken down.[14]

Asparagine is derived from aspartic acid and ATP through transamidation, in which an amide group is transferred from one compound to another. It is one of the principal and most abundant amino acids involved in the transport of nitrogen. It is required by the nervous system and plays an important role in the synthesis of ammonia.[15] However, its primary function is in the conversion of one amino acid into another via amination or transamination. Amination is the process by which an amine group is introduced into an organic molecule, and transamination is the transfer of an amino acid to an α-ketoacid.

Aspartic acid, or *aspartate*, is synthesized from oxaloacetate through transamination in the citric acid cycle and the urea cycle. An important role of aspartic acid is to serve as the precursor for several other amino acids, such as asparagine, arginine, lysine, methionine, threonine, and isoleucine, as well as several nucleotides. It also serves as a neurotransmitter and participates in the generation of glucose from non-carbohydrate substrates, a process known as gluconeogenesis.[16]

Glutamic acid, or *glutamate*, is produced from the transamination of amino acids such as alanine and aspartic acid. Glutamic acid serves as a neurotransmitter and has an important role in the metabolism of carbohydrates and fats as well as facilitating amino acid synthesis and degradation.[17]

Selenocysteine, unlike other amino acids present in proteins, is not coded for directly in the genetic code. Rather, it is encoded by a UGA codon, which is normally a stop codon; however, like the other amino acids used by cells, selenocysteine has a specialized transfer RNA (tRNA). As its name implies, selenocysteine is the selenium analogue of cysteine, in which a selenium atom replaces sulfur. Selenocysteine is present in several enzymes, such as formate dehydrogenases, glycine reductases, and some hydrogenases.[18]

Serine is synthesized from 3-phosphoglycerate, which is an intermediate in glycolysis. Serine is needed for the proper metabolism of lipids and fatty acids and plays an important role in the synthetic pathways for pyrimidines, purines, creatine, and porphyrins. It is highly concentrated in all cell membranes, serves as a component of the protective myelin sheaths surrounding nerve fibers, and aids in the production of antibodies.[19]

Conditionally Essential Amino Acids

Arginine, cysteine, glutamine, glycine, proline, and tyrosine are classified as semi-essential, or conditionally essential, amino acids because adults can synthesize adequate amounts to meet the demands of the body.[4] Infants that are born prematurely, as well as individuals that are in severe catabolic distress, may need supplementation of these amino acids.[2] If a conditionally essential amino acid is produced from an essential amino acid, a deficiency in the essential amino acid will likely require supplementation of both. For example, tyrosine, a conditionally essential amino acid, is produced from phenylalanine, an essential amino acid. If dietary intake of phenylalanine is insufficient, tyrosine will also be deficient and may require dietary supplementation.

Arginine is a complex amino acid often found at the catalytic site in proteins and enzymes due to its amine-containing side chain. Arginine plays an important role in cell division, wound healing, stimulation of protein synthesis, immune function, and the release of hormones. Another important role of arginine is in the conversion of ammonia into urea.[20]

Cysteine should not be confused with *cystine*, as they are two different amino acids. Cysteine is potentially toxic, so it is absorbed during digestion as cystine, which is more stable in the gastrointestinal tract and less toxic. Cystine is transported to cells, where it is reduced to two cysteine molecules upon cell entry. Cysteine can also be synthesized in vivo from methionine through a series of enzymatic reactions. Cysteine is an important structural and functional component of many proteins. It is found in beta-keratin and is important in collagen formation. It also has antioxidant properties and involvement in metabolism of other molecules.[21]

Glutamine is synthesized from glutamic acid through the addition of an ammonia group and can donate the ammonia group to form urea, which is excreted by the kidneys. As such, glutamine plays an integral role in regulation of ammonia, which is considered a toxic substance. Additionally, glutamine has many other important functions, including renal maintenance of the acid–base balance, providing fuel for a healthy digestive tract, and acting as the basis of the building blocks for synthesis of RNA and DNA.

Glutamine is also a source of cellular energy and aids in immune function.[22]

Glycine is synthesized from the amino acid serine. It is essential for synthesis of nucleic acids, bile acids, proteins, peptides, purines, ATP, porphyrins, hemoglobin, glutathione, creatine, bile salts, glucose, glycogen, and other amino acids. Glycine is also an inhibitory neurotransmitter in the central nervous system and a metal-complexing agent. Additionally, glycine limits muscle degeneration, improves glycogen storage, promotes healing, and is utilized by the liver in the detoxification of compounds.[1]

Proline is produced from glutamic acid and other amino acids. It serves as the precursor for hydroxyproline, which is manufactured into collagen, tendons, ligaments, and cardiac tissue. Proline is also involved in wound healing, especially that of cartilage, and in the strengthening of joints, tendons, and cardiac tissue.[23]

Tyrosine is synthesized from phenylalanine and serves as a precursor for adrenal and thyroid hormones. Tyrosine stimulates metabolism and the nervous system, acts as a mood elevator, and aids in function of the adrenal, thyroid, and pituitary glands.[24]

Aminoacidopathies

Aminoacidopathies are a class of inborn errors of metabolism in which an enzyme defect inhibits the body's ability to metabolize certain amino acids. The abnormalities exist either in the activity of a specific enzyme in the metabolic pathway or in the membrane transport system for the amino acid.[25] Aminoacidopathies can cause severe medical complications, such as brain damage, due to the accumulation of toxic amino acids or their by-products in the blood and tissues. Due to the severity of these complications, newborn screening tests are routinely performed to aid in the early diagnosis of numerous inborn errors of metabolism. Refer to Chapter 24, *Pregnancy and Prenatal Testing*, for additional information on aminoacidopathies as well as maternal and newborn screening.

Methods of Analysis

Analysis for amino acids may be performed on urine samples to screen for disorders affecting amino acid transport and on plasma samples to diagnose and monitor aminoacidopathies. Amino acid concentrations may also be evaluated in the investigation of conditions involving the liver, endocrine glands, gastrointestinal tract, and kidneys and in patients with severe burns, muscular, neurologic, or neoplastic diseases.[1] Cerebrospinal fluid (CSF) may also be analyzed to aid in the diagnosis of select neurotransmitter disorders.

Blood samples for amino acid analysis should be drawn after 6 to 8 hours of fasting to avoid the effect of absorbed amino acids originating from dietary proteins. The sample should be collected in a heparinized tube and the plasma removed from the cells within 2 hours of collection. Care should be taken to avoid aspirating the platelet and white cell layer to prevent contamination with amino acids from these cells. For example, the levels of aspartic acid and glutamic acid in white blood cells are about 100 times higher than plasma levels. Hemolyzed samples are unacceptable for the same reason.[26] For quantitation of urinary amino acids, a well-mixed aliquot from a first morning collection is preferred.[27] For CSF, samples should be immediately centrifuged and the supernatant transferred to a new tube in order to separate it from the cellular material.[28] For all sample types, testing should be performed immediately. If testing is to be delayed, samples can be refrigerated up to 24 hours or frozen for up to 1 month.

Techniques, such as high-performance liquid chromatography coupled with tandem mass spectrometry (HPLC-MS/MS) may be used to quantitate amino acids and their metabolites in the patient sample.[29] MS/MS methods are considered to have higher specificity and greater sensitivity allowing for detection of lower concentrations of the amino acid(s) and an earlier diagnosis. Genetic assays using DNA analysis may also be performed to aid in the diagnosis and detection of carrier status in families with an inborn error of metabolism. Refer to Chapter 4, *Analytic Techniques*, for more information on HPLC-MS/MS.

Proteins

Overview

Proteins catalyze almost all reactions in living cells, thereby controlling virtually all cellular processes. Some of their major functions include catalyzing biochemical reactions as enzymes, transporting metals such as iron and copper, acting as receptors for hormones, providing structure and support to cells, and participating in the immune response as antibodies.[4]

Basic Structure

Proteins consist of the elements carbon, oxygen, hydrogen, nitrogen, and sulfur. The fact that proteins contain nitrogen sets them apart from pure carbohydrates and lipids, which do not contain

nitrogen atoms. Proteins are polymers built from one or more unbranched chains of amino acids. A typical protein contains 200 to 300 amino acids, but some are much smaller (peptides) or larger (e.g., titin has 27,000 to 35,000 amino acids). Proteins are considered macromolecules and range from approximately 6000 daltons for insulin to several million daltons for some structural proteins.[3]

The four distinct levels of a protein's structure: primary, secondary, tertiary, and quaternary are depicted in **Figure 6.5**. A protein's **primary structure** refers to the number, type, and sequence of amino acids in the

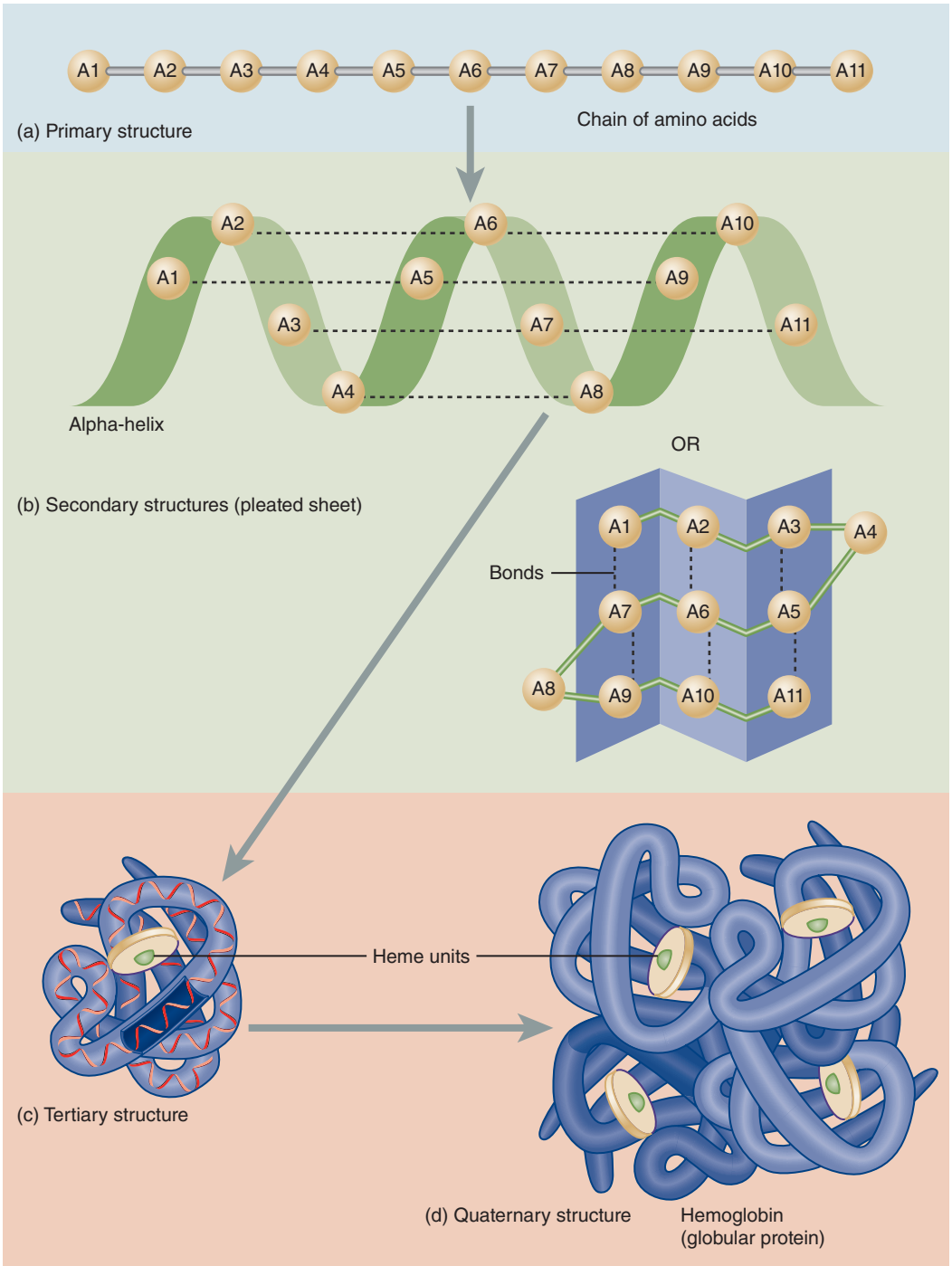


Figure 6.5 Four Levels of Protein Structure. (**A**) Primary structure: The sequence of amino acids linked by peptide bonds. (**B**) Secondary structure: Amino acids near to each other interact through hydrogen bonds. Secondary structure of α helix and β pleated sheet are shown. (**C**) Tertiary structure: Three-dimensional structure of the protein due to hydrophobic effect, ionic attraction, hydrogen bonds, and disulfide bonds. (**D**) Quaternary structure: Molecular association of more than one peptide/protein to form a dimer, trimer, tetramer, etc. by noncovalent bonds.

polypeptide chain.[4] In order to function properly, proteins must have the correct sequence of amino acids. For example, when the amino acid valine is substituted for glutamic acid in the β-chain of hemoglobin A, hemoglobin S is formed, resulting in sickle cell anemia. The location of certain amino acids in the primary structure will determine the secondary, tertiary, and quaternary structure of the protein.

Secondary structure refers to commonly formed arrangements stabilized by hydrogen bonds between nearby amino acids within the protein. Two of the main types of secondary structure are the α-helix and the β-pleated sheet, with most serum proteins forming a helix. Secondary structures add new properties to a protein such as strength and flexibility.[4]

Tertiary structure refers to the overall shape, or conformation, of the protein molecule. The conformation is known as the *fold*, or the spatial relationship of the secondary structures to one another. Tertiary structures are three dimensional and result from the interaction of side chains, which are stabilized through the hydrophobic effect, ionic attraction, hydrogen bonds, and disulfide bonds. The function and physical and chemical properties of a protein are related to its tertiary structure.[4]

Quaternary structure is the shape or structure that results from the interaction of more than one protein molecule, or protein subunits, referred to as a multimer, that functions as a single unit. Multimers are held together by noncovalent forces such as hydrogen bonds and electrostatic interactions and occur as dimers (two subunits), trimers (three subunits), tetramers (four subunits), etc.[4] Not all proteins have a quaternary structure. One example of a protein with a quaternary structure is hemoglobin, which has a tetramer globin composed of two α and two β subunits.

When the secondary, tertiary, or quaternary structure of a protein is disturbed, the protein may lose its functional and chemical characteristics. This loss of its native, or naturally occurring, folded structure is called *denaturation*. Denaturation can be caused by heat, hydrolysis by strong acid or alkali, enzymatic action, exposure to urea or other substances, or exposure to ultraviolet light.

General Chemical Properties

The structure of a protein directly affects its chemical properties as each amino acid side chain has differing properties.[30] Proteins contain many ionizable groups on the side chains of their amino acids as well as on their N- and C-terminal ends. As a result, proteins can be positively and negatively charged. For example, the side chains of lysine, arginine, and histidine include basic groups (proton acceptors), whereas acidic groups (proton donors) are found on the side chains of glutamic acid and aspartic acid. The acid or base groups that are not involved in the peptide bond can exist in different charged forms depending on the pH of the surrounding environment (**Figure 6.6**). The pH of the solution,

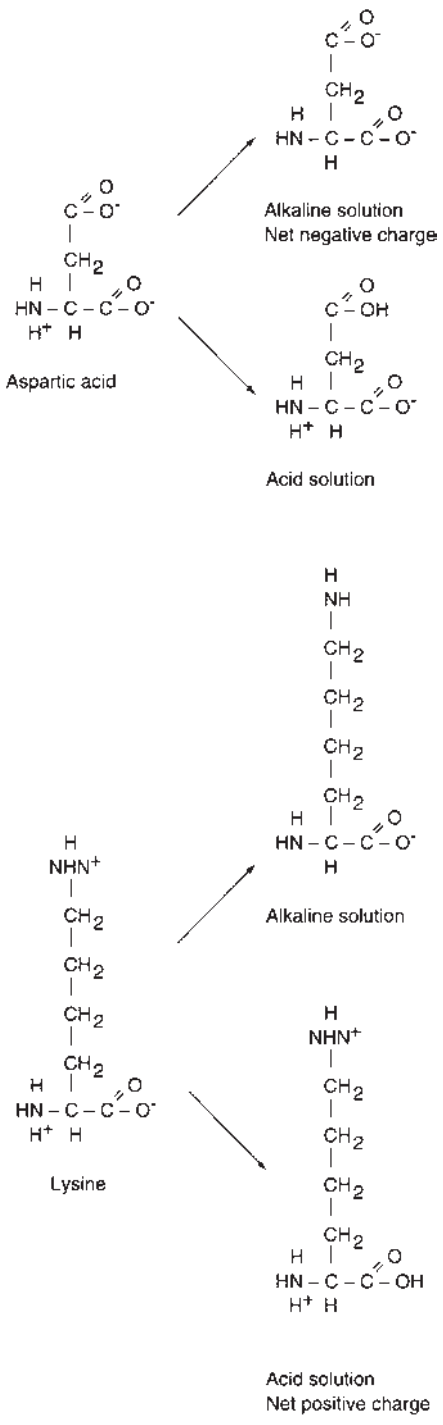


Figure 6.6 Charged states of amino acids. The charge of the protein depends upon the pH of its environment.

the pK_a of the side chain, and the side chain's environment influence the charge on each side chain. The relationship between pH, pK_a, and charge for individual amino acids can be described by the Henderson-Hasselbalch equation:

$$\text{pH} = \text{p}K_a + \log \frac{[\text{conjugate base}]}{[\text{conjugate acid}]} \qquad \textbf{(Eq. 6-1)}$$

In general terms, as the pH of a solution increases, becoming more alkaline, deprotonation of the acidic and basic groups occurs; carboxyl groups are converted to carboxylate anions (R–COOH to R–COO^-) and ammonium groups are converted to amino groups (R–NH_3+ to R–NH_2). The pH at which an amino acid or protein has no net charge is known as its **isoelectric point (pI)**. When the pH is greater than the p*I*, the protein has a net negative charge, and when the pH is less than the p*I*, the protein has a net positive charge. As such, the p*I* is the point at which the number of positively charged groups equals the number of negatively charged groups in a protein. If a protein is placed in a solution that has a pH greater than the p*I*, the protein will become negatively charged. If a protein is placed in a solution that has a pH less than the p*I*, the protein will become positively charged. Proteins differ in their p*I* values, but most occur in the pH range of 5.5 to 8.0. Because proteins carry different net charges depending on the pH of their environment, the difference in net charge at a given pH is the basis for several procedures used to separate and quantify proteins.

A protein's solubility is also dependent on the charge on its surface, with its lowest solubility at its p*I*, where the positive and negative charges balance resulting in a net charge of zero. If there is a charge at the protein surface, the protein is referred to as hydrophilic and prefers to interact with water rather than with other protein molecules, making it more soluble. Without a net surface charge, the protein is less soluble and protein–protein interactions and precipitation are more likely. Proteins in the blood require a pH in the range of 7.35 to 7.45 to remain soluble. In the laboratory, the relative solubilities of proteins may be used to separate and quantify them.

Synthesis

Plasma proteins are predominantly synthesized in the liver and secreted into circulation. Immunoglobulins are an exception, as they are synthesized by B lymphocytes. It is the information encoded in genes that provides each protein with its own unique amino acid sequence. The amino acid sequence of a polypeptide chain is determined by a corresponding sequence of bases (guanine, cytosine, adenine, and thymine) in the DNA contained in the specific gene. This genetic code is a set of three nucleotides known as codons, with each three-nucleotide combination standing for a specific amino acid. Because DNA contains four nucleotides, the total number of possible codons is 64; therefore, some redundancy in the genetic code allows for some amino acids to be specified by more than one codon. Double-stranded DNA unfolds in the nucleus, and one strand is used as a template for the formation of a complementary strand of mRNA in a process known as *transcription*. The mRNA is manufactured in the cell nucleus and then translocated across the nuclear membrane into the cytoplasm for *translation*.[30]

In the process of translation, the mRNA strand is used as a template for protein synthesis by the ribosome. The mRNA is surrounded by the ribosome and is read three nucleotides at a time by matching each codon to its base pairing anticodon located on a tRNA molecule. tRNA is a short chain of RNA that occurs freely in the cytoplasm and brings amino acids to the ribosome. Each amino acid has a specific tRNA that contains three bases (anticodon) corresponding to the three bases in the mRNA (codon). The tRNA carries its particular amino acid to the ribosome and attaches to the mRNA in accordance with the matching codon. As each new amino acid is added, the preceding amino acid is transferred onto the amino group of the new amino acid and enzymes located in the ribosome form a peptide bond. The tRNA is released into the cytoplasm, where it can pick up another amino acid, and the cycle repeats. In this manner, the amino acids are aligned in sequence and linked by peptide bonds to form a polypeptide. This process continues until a stop codon is reached signaling that all amino acids have been joined in the specific sequence to form the polypeptide chain. When the terminal codon is reached, the peptide chain is released and the ribosome and mRNA dissociate.[30] **Figure 6.7** illustrates basic protein synthesis through transcription and translation.

Protein targeting is the process by which proteins are directed to the location they are needed (translocated) during or after translation. This occurs under the direction of signal peptides, known as signal recognition particles (SRPs), that occur at their N-terminus. In the absence of an SRP, the protein will be synthesized and remain in the cytosol. Proteins with an SRP that signals *post-translational translocation* are synthesized in the cytosol then moved to a

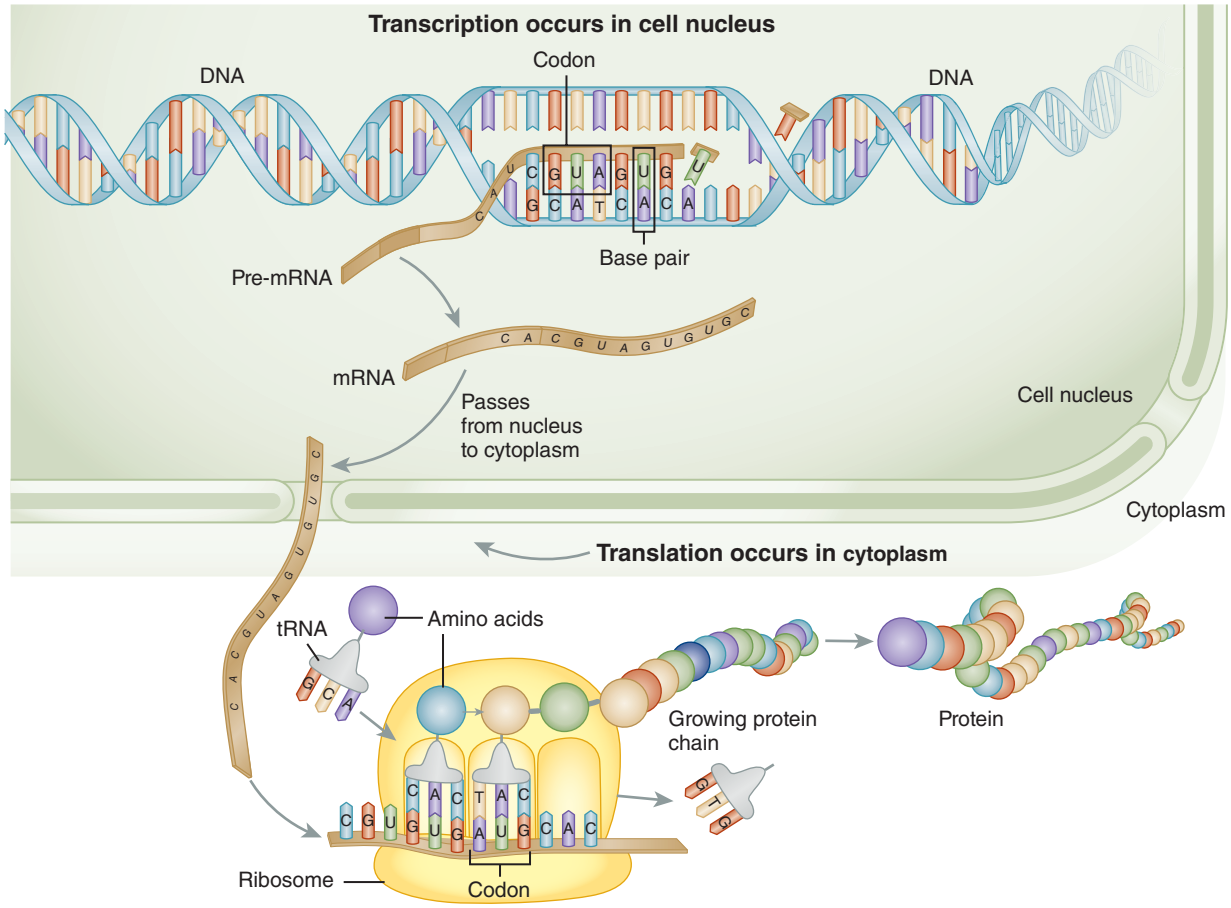


Figure 6.7 Protein synthesis through transcription and translation.

nonendomembane location such as the nucleus or peroxisomes. Intracellular proteins are generally synthesized in this fashion, whereas proteins destined for the endomembrane system or secretion from the liver are made on ribosomes attached to the rough endoplasmic reticulum. Translation for these proteins also begins in the cytosol, but is paused by SRPs signaling *co-translational translocation*. After attaching to the endoplasmic reticulum, the SRP is cleaved, allowing translation to resume. The polypeptide is delivered through a protein channel into the lumen of the endoplasmic reticulum. Additional SRPs can direct the polypeptide to organelles in the endomembrane system or to the cell membrane for secretion into circulation.[31]

Catabolism and Nitrogen Balance

Unlike fats and carbohydrates, nitrogen-containing compounds, such as proteins, have no designated storage. Most proteins in the body are repetitively synthesized (anabolism) and then degraded (catabolism) allowing for efficient recycling of amino acids. These processes result in a turnover of about 125 to 220 g of protein each day, with the rate of individual proteins varying widely. For example, plasma proteins and most intracellular proteins are rapidly degraded, having half-lives of hours or days, whereas some of the structural proteins, such as collagen, are metabolically stable and have half-lives of years.

In health, nitrogen balance is maintained by equal intake and excretion of amino acids. Pregnant women, growing children, and adults recovering from major illness are often in positive nitrogen balance because their nitrogen intake exceeds their loss. When more nitrogen is excreted than incorporated, an individual is in negative nitrogen balance, which may occur in conditions associated with excessive tissue destruction, such as burns, wasting diseases, continual high fevers, or starvation.

When in excess, amino acids can be converted to urea for excretion by the kidneys and into glucose or ketones to be used as an alternative source of energy. These catabolic processes typically involve *transamination* and *oxidative deamination* of the amino acid, which primarily occurs in the liver. Transamination involves

transferring an amino group from the α-amino acid to the keto-carbon of an α-ketoglutarate resulting in the production of glutamate and an α-ketoacid. These reversible reactions are catalyzed by a group of intracellular enzymes known as *aminotransferases* or *transaminases*. Glutamate is then deaminated, catalyzed by glutamate dehydrogenase, forming ammonia and α-ketoglutarate. For some amino acids, such as serine and threonine, only deamination is required to produce ammonia and the α-ketoglutarate. In the hepatocytes, ammonia is then converted to urea by the urea cycle, which is less toxic and can be excreted in urine.[32]

The resultant ketoacids enter into the Krebs cycle (citric acid cycle) to release stored energy derived from carbohydrates, fats, and proteins. Glucogenic amino acids generate precursors of glucose, such as pyruvate or citric acid cycle intermediates. Examples include alanine, which can be deaminated to pyruvate; arginine, which is converted to α-ketoglutarate; and aspartic acid, which is converted to oxaloacetate. Ketogenic amino acids, such as leucine and lysine, are degraded to acetyl-CoA or acetoacetyl-CoA and form ketone bodies. Isoleucine, phenylalanine, tryptophan, tyrosine, and threonine are both ketogenic and glucogenic.[32]

Hormonal Regulation of Protein Metabolism

In some tissues, protein metabolism is regulated through interaction of several hormones including insulin, growth hormone (GH), insulin-like growth factor I (IGF-1), sex hormones (testosterone/estrogen), glucocorticoids (e.g. cortisol), glucagon, catecholamines, and thyroxine.[33] The effects of each of these hormones on protein synthesis and/or degradation are listed in **Table 6.2**.

Table 6.2 Hormones Regulating Protein Metabolism

Hormone	Primary Influence on Protein Metabolism
Insulin	Suppresses proteolysis
Growth Hormone	Stimulates protein synthesis
Insulin-like Growth Factor-1	Inhibits proteolysis in fasting state, but stimulates protein synthesis in fed state
Glucocorticoids	Stimulates protein degradation in muscles
Glucagon	Promotes uptake of amino acids in the liver
Sex Hormones	Increases protein deposition in tissues
Catecholamines	Increases production of gluconeogenic amino acids
Thyroxine	Increases the basal metabolic rate

Classification

Proteins are often discussed or classified based on their structure, composition, and function. **Simple proteins** contain peptide chains composed of only amino acids and may be globular or fibrous in shape. Globular proteins are globe-like, have symmetrical proteins that are soluble in water, and commonly function as transporters, enzymes, and messengers. Examples of globular proteins are albumin, hemoglobin, and the immunoglobulins, IgG, IgA, and IgM. Fibrous proteins form long protein filaments or subunits, are asymmetrical and usually inert, and are generally water-insoluble due to their hydrophobic R groups. They provide structure to cells, such as connective tissues, tendons, bone, and muscle. Examples of fibrous proteins include troponin and collagen.

Conjugated or **complex proteins** consist of a protein and a nonprotein. The non-amino part of a conjugated protein is generally referred to as the prosthetic group, which may be a lipid, carbohydrate, porphyrin, metal, etc. In a conjugated protein, it is the prosthetic group that defines the characteristics of the protein. Examples of conjugated proteins include metalloproteins, glycoproteins, lipoproteins, and nucleoproteins. Metalloproteins have a metal ion attached to the protein, either directly, as in ferritin, which contains iron, and ceruloplasmin, which contains copper, or as a complex metal such as hemoglobin. Lipoproteins, including high-density lipoproteins (HDLs) and very low-density lipoproteins (VLDLs), are composed of a protein and a lipid, such as cholesterol or triglyceride. There are several terms used to describe conjugated proteins that are bound to carbohydrates. In general, molecules, such as haptoglobin and α_1-antitrypsin, that have a composition of 10% to 40% carbohydrate are referred to as glycoproteins. When the percentage of carbohydrate is greater than 40%, the protein conjugate is referred to as a mucoprotein or

proteoglycan. An example of a mucoprotein is mucin, which is a lubricant that protects body surfaces from friction or erosion. Nucleoproteins are proteins that are combined with nucleic acids. Chromatin is an example of a nucleoprotein as it is a complex of DNA and protein.

Enzymes are proteins that catalyze biochemical reactions. They are normally found intracellularly and are released into the bloodstream when tissue damage occurs, making enzyme measurements an important diagnostic tool. Aminotransferases, dehydrogenases, and phosphatases are just a few examples of enzyme groups routinely analyzed in the clinical laboratory to evaluate possible tissue damage.

Hormones are chemical messenger proteins that control the action(s) of specific cells or organs. Hormones directly affect growth and development, metabolism, sexual function, reproduction, and behavior. Examples of hormones commonly measured in the clinical laboratory include insulin, testosterone, growth hormone, thyroid stimulating hormone, and cortisol.

Many proteins serve as transporters for molecules such as hormones, vitamins, minerals, and lipids, across a biologic membrane. Hemoglobin, albumin, ceruloplasmin, haptoglobin, and transferrin are examples of transport proteins.

Immunoglobulins, or antibodies, are proteins that are produced by B cells (lymphocytes) in the bone marrow. Immunoglobulins mediate the humoral immune response to identify and neutralize foreign antigens. Examples of immunoglobulins of clinical importance are IgG, IgM, IgE, IgD, and IgA.

Structural proteins are fibrous proteins that provide structure to many cells and tissues throughout the body, such as muscle, tendons, and bone matrix. Examples include collagen, elastin, and keratin.

Storage proteins serve as reservoirs for metal ions and amino acids so they can be stored without causing harm to the cell and released when needed. An example is ferritin, which stores iron in the hepatocyte until it is needed in the synthesis of hemoglobin.

Some proteins serve as an *energy source* for tissues and muscle. Creatine is one example of an energy source protein, as it helps to supply energy to cells throughout the body but is primarily found in muscle tissue. Proteins provide up to 20% of the total energy required daily by the body.

Maintaining *water distribution* throughout the compartments of the body is another important function of proteins. Due to their size, plasma proteins cannot cross the capillary membrane. As a result of colloid osmotic pressure, water is absorbed from the tissue into the capillary. When the concentration of plasma proteins is significantly decreased, the decrease in osmotic pressure results in increased levels of interstitial fluid and edema in the tissues. This often occurs in renal disease when proteins are inappropriately excreted in urine and plasma protein concentrations are decreased.

Proteins also play an important role in maintaining the *acid–base balance* by serving as buffers to maintain pH. They are also involved in *hemostasis* by participating in the formation and dissolution of blood clots.

Plasma Proteins

Plasma proteins are commonly analyzed in the clinical laboratory and can be divided into two major groups: albumin and globulins. Routine analysis of blood specimens will typically include measurement of total protein and albumin and calculation of the albumin-to-globulin (A/G) ratio; however, there are many other clinically important proteins that may be measured when indicated by clinical presentation. Some of the more significant plasma proteins and their function, structure, and relation to disease states are discussed below. The characteristics of select plasma proteins are listed in **Table 6.3**.

Prealbumin

Prealbumin, or transthyretin, is so named because it migrates before albumin in classic serum protein electrophoresis (SPE). It can also be separated using high-resolution electrophoresis (HRE) or immunoelectrophoresis techniques. Prealbumin is a transport protein for the thyroid hormones, thyroxine (T_4) and triiodothyronine (T_3). It also forms a complex with retinol-binding protein to transport retinol (vitamin A) and is rich in the amino acid tryptophan. Serum and plasma concentrations may be decreased in hepatic damage due to decreased protein synthesis, during an acute-phase inflammatory response, or as a result of tissue necrosis. A low concentration may also indicate poor nutritional status. Diets deficient in protein may not provide sufficient amino acids for protein synthesis by the liver, resulting in decreased plasma concentrations of prealbumin, albumin, and globulins. Prealbumin has a half-life of approximately 2 days, so blood concentrations decrease rapidly when protein synthesis is inhibited. Blood concentrations may be

Table 6.3 Characteristics of Select Plasma Proteins

	Reference Value (Adult, g/L)	Molecular Mass (Da)	Isoelectric Point, p/	Electrophoretic Mobility pH 8.6	Clinical Usage
Prealbumin	0.1–0.4	55,000	4.7	7.6	Indicator of nutrition; binds thyroid hormones and retinol-binding protein
Albumin	35–55	66,300	4.9	5.9	Binds bilirubin, steroids, fatty acids; major contributor to oncotic pressure
α_1-Globulins					
α_1-Antitrypsin	2–4	53,000	4.0	5.4	Acute-phase reactant; protease inhibitor
α_1-Fetoprotein	1×10^5	76,000	2.7	6.1	Principal fetal protein; elevated levels indicate risk for spina bifida
α_1-Acid glycoprotein (orosomucoid)	0.55–1.4	44,000	2.8–3.8	5.2	Acute-phase reactant; transport of drugs and hormones
α_1-Lipoprotein	2.5–3.9	200,000		4.4–5.4	Lipid transport (HDL)
α_1-Antichymotrypsin	0.3–0.6	68,000		Inter α	Inhibits serine proteinases
Inter-α–trypsin inhibitor	0.2–0.7	160,000		Inter α	Inhibits serine proteinases
Gc-globulin	0.2–0.55	59,000		Inter α	Transports vitamin D and binds actin
α_2-Globulins					
Haptoglobin Type 1-1	1.0–2.2	100,000		4.5	Acute-phase reactant
Haptoglobin Type 2-1	1.6–3.0	200,000	4.1	3.5–4.0	Binds hemoglobin
Haptoglobin Type 2-2	1.2–1.6	400,000		3.5–4.0	Binds hemoglobin
Ceruloplasmin	0.15–0.60	134,000	4.4	4.6	Acute-phase reactant, oxidase activity; contains copper
α_2-Macroglobulin	1.5–4.2	725,000	5.4	4.2	Inhibits proteases
β-*G*lobulins					
Pre-β lipoprotein	1.5–2.3	250,000		3.4–4.3	Transports lipids (primarily VLDL triglyceride)
Transferrin	2.04–3.60	76,000	5.9	3.1	Transports iron
Hemopexin	0.5–1.0	57,000–80,000		3.1	Binds and transports heme
β-Lipoprotein	2.5–4.4	3,000,000		3.1	Transports lipids (primarily LDL cholesterol)
β_2-Microglobulin (B2M)	0.001–0.002	11,800		β_2	Component of human leukocyte antigens (HLA)

	Reference Value (Adult, g/L)	Molecular Mass (Da)	Isoelectric Point, pI	Electrophoretic Mobility pH 8.6	Clinical Usage
C4 complement	0.20–0.65	206,000		0.8–1.4	Immune response
C3 complement	0.55–1.80	180,000		0.8–1.4	Immune response
C1q complement	0.15	400,000			Immune response
Fibrinogen	2.0–4.5	341,000	5.8	2.1	Precursor of fibrin clot
C-Reactive Protein	0.01	118,000	6.2		Acute-phase reactant; promotes phagocytosis
γ-Globulins					
IgG	8.0–12.0	150,000	5.8–7.3	0.5–2.6	Antibodies
IgA	0.7–3.12	180,000		2.1	Secretory antibodies
IgM	0.5–2.80	900,000		2.1	Antibodies (early response)
IgD	0.005–0.2	170,000		1.9	Antibodies
IgE	6×10^4	190,000		2.3	Antibodies; promotes release of histamine (allergies)

increased in patients receiving steroid therapy, who have issues with alcohol abuse, or who are in chronic renal failure.[34]

Albumin

Albumin is synthesized in the liver at a rate of 9 to 12 g/day. Because it is decreased during an acute-phase response, it is classified as a **negative acute-phase reactant**, meaning it decreases in the presence of acute disease states, such as inflammation, infection or injury. It is the most abundant plasma protein and is also found in significant amounts in the extravascular (interstitial) space. Interestingly, the amount of extravascular albumin exceeds the intravascular amount by about 30%; however, the concentration of albumin in plasma (albumin mass/plasma volume) is much greater. Albumin leaves the bloodstream at a rate of 4% to 5% of the total intravascular concentration per hour. This rate of movement, known as the transcapillary escape rate, measures the systemic capillary efflux of albumin. Due to its high presence in the plasma, albumin is responsible for nearly 80% of the colloid osmotic pressure, meaning it is the primary protein involved in maintaining the fluid balance between the intravascular and extravascular spaces. Albumin also plays an important role in maintaining a homeostatic pH by serving as a buffer in the circulation.[35]

Another key function of albumin is its ability to bind and transport a large variety of substances throughout the body. There are four binding sites on albumin, each with varying specificities for the different substances requiring transport in circulation. For example, albumin is involved in the transport of thyroid hormones, unconjugated bilirubin, fat-soluble hormones, iron, fatty acids, calcium (Ca^{2+}), magnesium (Mg^{2+}), and many drugs such as salicylic acid (aspirin). Albumin also binds well with certain dyes, which is the basis for several methods used to quantitate it in fluids such as serum, plasma, and urine.[35]

Decreased blood concentrations of albumin are most commonly associated with an acute inflammatory response as albumin is a negative acute-phase reactant. Liver and kidney disease may also result in low blood albumin concentrations. The liver is a primary site for protein synthesis; damage to hepatocytes, such as in liver cirrhosis, may result in decreased protein synthesis. Note that an increase in globulins occurs in early liver cirrhosis, which balances the loss in albumin to give a total protein concentration within acceptable limits.[36] Albumin is normally excreted in very small amounts by the kidneys; however, increased renal loss of albumin, termed albuminuria, commonly occurs in renal disease. This increased excretion occurs when the glomerulus no longer restricts the passage of proteins from the blood into the ultrafiltrate, as occurs

in nephrotic syndrome and glomerular damage associated with diabetes mellitus.[37]

Low albumin concentrations may also be the result of malnutrition and malabsorption in which an inadequate ingestion of proteins or amino acid-rich foods results in decreased protein synthesis by the liver. Less commonly, low blood albumin concentrations occur as a result of hypothyroidism, burns or exfoliative dermatitis, dilution by excessive intake (polydipsia), or infusion of intravenous liquids.[36] Albumin may be redistributed by hemodilution, increased capillary permeability, or decreased lymphatic clearance. In sepsis, there is a profound reduction in plasma albumin associated with marked fluid shifts. Mutations resulting from an autosomal recessive trait can cause an absence of albumin, known as *analbuminemia*. *Bisalbuminemia*, which is inherited or acquired, is characterized by the presence of two albumin bands during serum protein electrophoresis instead of the single band usually observed. Elevated albumin concentrations are generally considered clinically insignificant as they most likely indicate dehydration or excessive albumin infusion.

Another important clinical application for measuring albumin is to determine the percent glycated albumin as a means to monitor short-term efficacy of therapies for diabetes mellitus.[38] Typically, glycated hemoglobin is measured to assess glucose concentrations over the past 2 to 3 months for patients with diabetes mellitus, as the lifespan of a red blood cell is approximately 120 days. While this is the preferred test for monitoring long-term therapy, it may not be appropriate for patients with certain red blood cell disorders. For this patient population, glycated albumin or fructosamine determinations may be monitored instead. Fructosamine is a measure of the non-enzymatic glycation of circulating plasma proteins, including albumin, globulins, and lipoproteins. The half-life of serum albumin is approximately 20 days, making it a better indicator of short-term glycemic control.[39] Methods used to determine the concentration of glycated albumin in serum often use affinity chromatography based on specific interaction of boronic acids with the glycated proteins.

Globulins

There are four major types, or fractions, of globulins designated as α_1, α_2, β, and γ based on their electrophoretic mobilities. Each fraction consists of a number of different proteins with individual functions. Select globulins from each fraction are discussed below.

α_1-Globulins

α_1-*Antitrypsin* (AAT) is a glycoprotein synthesized predominantly in the liver. Its main function is to inhibit neutrophil elastase, a protease released from neutrophils and macrophages during an infection. Mutations in the *SERPINA1* gene can lead to a deficiency of α_1-antitrypsin or an abnormal form of the protein that does not properly control neutrophil elastase. As a result, damage to the alveoli can lead to emphysema. The abnormal form of α_1-antitrypsin can also accumulate in the liver, causing cirrhosis. α_1-Antitrypsin is a **positive acute-phase reactant**, meaning it increases in the presence of an acute onset of disease states, such as inflammation, infections, or injuries; therefore, increased levels are seen in inflammatory reactions, as well as in pregnancy and contraceptive use. Decreased α_1-antitrypsin levels are most often identified by the absence of an α_1-globulin band on serum protein electrophoresis because α_1-antitrypsin accounts for approximately 90% of the α_1-globulin fraction.[40]

α_1-*Fetoprotein* (AFP) is synthesized in utero by the developing embryo and then by the fetal liver and gastrointestinal tract. AFP concentrations decrease gradually after birth, reaching adult concentrations by 8 to 12 months of age. As a positive acute-phase reactant, increases of AFP are seen in inflammatory conditions.[41] Maternal serum with an elevated AFP concentration is associated with increased likelihood of spina bifida, neural tube defects, abdominal wall defects, anencephaly, general fetal distress, and the presence of multiple gestation. Low levels of maternal AFP indicate an increased risk for trisomy 21 (Down syndrome) and trisomy 18 (Edwards syndrome).[42] Refer to Chapter 24, *Pregnancy and Prenatal Testing*, for additional information on maternal screening of AFP. AFP may also be used as a tumor marker, as elevated concentrations are associated with hepatocellular carcinoma, metastatic liver disease, and nonseminomatous germ cell tumors. AFP's utility as a tumor marker will be discussed further in Chapter 28, *Tumor Markers*.

α_1-*Acid glycoprotein* (AGP), or orosomucoid, is a positive acute-phase reactant produced primarily by the liver. Its physiological functions include maintaining the barrier function of capillaries, regulating immunity, mediating sphingomyelin metabolism, and acting as a transport protein. AGP elevates as a result of stress, inflammation, tissue damage, acute myocardial infarction (AMI), trauma, pregnancy, cancer, pneumonia, rheumatoid arthritis, and surgery.[43]

α_2-Globulins

Haptoglobin (Hp) is an α_2-glycoprotein synthesized in the liver as a tetramer consisting of two α and two β chains. It is a positive acute-phase reactant that increases in many inflammatory diseases, such as ulcerative colitis, acute rheumatic disease, acute myocardial infarction, and severe infection. The primary function of haptoglobin is to bind free hemoglobin to prevent the loss of its constituent, iron, into the urine. When haptoglobin and hemoglobin attach, mononuclear phagocytic cells, predominantly in the spleen, remove the haptoglobin–hemoglobin complex from circulation. The hemoglobin constituents, iron and amino acids, can be recycled, but haptoglobin is destroyed in the process. For this reason, haptoglobin concentrations are primarily used to evaluate possible hemolytic anemias and to aid in distinguishing intravascular hemolysis from extravascular hemolysis. Patients with a hemolytic anemia have a decreased haptoglobin concentration due to intravascular hemolysis. In contrast, the haptoglobin concentration would be normal if the anemia is related to extravascular destruction of red blood cells in organs such as the spleen and liver. In extravascular hemolysis, the hemoglobin is not released into the bloodstream, and the haptoglobin therefore stays intact. In the case of in vitro hemolysis, the haptoglobin concentration would appear normal, indicating hemolysis occurred during or after the collection. If haptoglobin concentrations are decreased without any sign of hemolytic anemia, it is possible the liver is not producing adequate amounts of haptoglobin.[44]

Ceruloplasmin is a copper-containing, α_2-glycoprotein synthesized in the liver. As a positive acute-phase reactant, it is frequently elevated in inflammation, severe infection, and tissue damage and may be increased with some cancers. Ceruloplasmin may also be increased during pregnancy and in patients who are taking estrogen, oral contraceptives, and medications such as carbamazepine, phenobarbital, and valproic acid. The primary function of ceruloplasmin is to serve as a transport protein for copper. Approximately 90% of copper in circulation is bound to ceruloplasmin, with the remaining 10% bound to albumin. Ceruloplasmin is primarily measured along with blood and urine copper concentrations to aid in the diagnosis of Wilson's disease, an inherited autosomal recessive disorder. Patients with Wilson's disease generally have decreased concentrations of ceruloplasmin, increased serum concentrations of free copper, and increased urinary excretion of copper.[45] Due to the limited availability of ceruloplasmin, free copper is deposited in the liver, brain, and other organs, which results in hepatic cirrhosis and neurologic damage. Copper is also deposited in the cornea producing the characteristic Kayser-Fleischer rings. Low ceruloplasmin is also seen in malnutrition, malabsorption, severe liver disease, nephrotic syndrome, and Menkes' syndrome, in which a decreased absorption of copper results in a decrease in ceruloplasmin.[46]

α_2-Macroglobulin, a major component of the α_2-globulin fraction, is a tetramer of four identical subunits synthesized by the liver. Alpha$_2$-macroglobulin is a protease inhibitor. It inhibits proteases such as pepsin, trypsin, thrombin, and plasmin, which reduces the accessibility of the protease functional sites but does not completely inactivate them. After binding with proteases, α_2-macroglobulin is removed by the mononuclear phagocytic system. Additional functions of α_2-macroglobulin include regulation of growth factors and cytokines as well as chaperoning of misfolded proteins. In nephrotic syndrome, serum concentrations of α_2-macroglobulin may appear increased relative to other proteins. Its large size inhibits filtration at the renal glomeruli, making measurement of α_2-macroglobulin useful in evaluation of renal disease and damage to the glomeruli. Increased concentrations may be seen during pregnancy, while using hormone-based contraceptives, and in patients with renal disease secondary to diabetes mellitus or hepatorenal syndromes.[47]

β-Globulins

Transferrin, a negative acute-phase glycoprotein synthesized by the liver, is the major component of the β-globulin fraction on protein electrophoresis. Transferrin's primary function is to bind and transport iron to and from storage sites, such as the liver and bone marrow, and to prevent iron from being inappropriately deposited in other tissues. Because transferrin is made in the liver, it may be decreased in patients with liver disease or when dietary intake of proteins is insufficient. A low transferrin concentration may also be due to excessive loss through the kidneys, as occurs in protein-losing disorders such as nephrotic syndrome. Many conditions, including infection, inflammation, and malignancy, lead to decreased transferrin levels, as it is a negative acute-phase reactant. Low transferrin concentrations can lead to anemia due to decreased delivery of iron to the bone marrow and subsequent impairment of hemoglobin synthesis. A deficiency of plasma transferrin may also result in the

inappropriate accumulation and precipitation of iron in tissues as hemosiderin. An increase in transferrin is generally associated with iron-deficiency anemia as a compensatory mechanism to promote iron absorption and mobilization. Similarly, transferrin concentrations decrease in patients with iron overload. As such, transferrin concentrations are routinely measured to aid in determining the cause of an anemia, to evaluate processes associated with iron metabolism, and to determine the total iron-binding capacity of the blood.[48] Additional information on iron metabolism and iron status studies can be found in Chapter 27, *Trace Elements, Toxic Elements, and Vitamins*.

Hemopexin is an acute-phase β-globulin, synthesized in the parenchymal cells of the liver, whose main function is to bind with free **heme**. When heme is released during the breakdown of hemoglobin, myoglobin, or catalase, it binds to hemopexin. The heme–hemopexin complex is carried to the liver, where it is destroyed, releasing the heme components to be recycled. In this manner, hemopexin preserves the body's iron and amino acid stores and prevents oxidative damage by free heme. Because hemopexin is destroyed during this process, low concentrations are associated with intravascular hemolytic conditions such as hemolytic anemias. Increased concentrations of hemopexin are found in inflammation, diabetes mellitus, Duchenne-type muscular dystrophy, and some malignancies, especially melanomas.[49]

Lipoproteins are complexes of proteins and lipids whose function is to transport cholesterol, triglycerides, and phospholipids in the bloodstream. Lipoproteins are subclassified according to their apolipoprotein and lipid content into chylomicrons, very-low-density lipoproteins (VLDL), intermediate-density lipoproteins (IDL), low-density lipoproteins (LDL), and high-density lipoproteins (HDL). In high-resolution electrophoresis, HDL migrates between the albumin and α_1-globulin bands, VLDL migrates at the beginning of the β-globulin band (pre-β), and LDL appears as a separate band in the β-globulin region.[50] A detailed discussion of the structure, function, and laboratory methods used to measure lipoproteins can be found in Chapter 10, *Lipids and Lipoproteins*.

β_2-Microglobulin (B2M) is the light chain component of the major histocompatibility complex or human leukocyte antigen (HLA). This protein is found on the surface of most nucleated cells and is present in high concentrations on lymphocytes. Because of its small size (molecular weight: 11,800 kDa), B2M is filtered by the renal glomeruli, but is almost completely reabsorbed and catabolized in the proximal tubules. Elevated plasma concentrations reflect impaired renal clearance or overproduction, which occurs in a number of inflammatory diseases, such as rheumatoid arthritis and systemic lupus erythematosus (SLE). In patients with human immunodeficiency virus (HIV), a high B2M level in the absence of renal failure indicates a large lymphocyte turnover rate, suggesting the virus is killing lymphocytes. Elevated urine concentrations of B2M are associated with renal tubular damage or disease due to decreased reabsorption.[51]

The *complement system* is a natural defense mechanism against infections. These proteins are synthesized in the liver as single polypeptide chains and circulate in the blood as non-functional precursors. *Complement C3* is the most abundant complement protein in human plasma, with *complement C4* being the second most abundant. In the classic pathway, activation of these proteins begins when the first complement factor, C1q, binds to an antigen–antibody complex. Each complement protein (C2–C9) is then activated sequentially and can bind to the membrane of the cell to which the antigen–antibody complex is bound, eventually leading to cell lysis. An alternate pathway for complement activation, known as the properdin pathway, bypasses early components of the cascade allowing the process to begin with C3. This pathway is triggered by different substances and does not require the presence of an antibody; however, the lytic attack on the cell membrane follows the same sequence of C5 to C9. Increased levels of both C3 and C4 are linked to acute inflammatory disease and tissue inflammation. In contrast, decreased levels of C3 are typically associated with autoimmune disease, neonatal respiratory distress syndrome, bacteremia, tissue injury, and chronic hepatitis. Complement C3 has been found to be important in the pathogenesis of age-related macular degeneration, and this finding further underscores the influence of the complement pathway in the pathogenesis of this disease. Decreased levels of C4 are seen in disseminated intravascular coagulation (DIC), acute glomerulonephritis, chronic hepatitis, and SLE. Inherited deficiencies of individual complement proteins have also been described.[52]

Fibrinogen is one of the largest proteins in blood plasma. It is synthesized in the liver and classified as a glycoprotein due to its considerable carbohydrate content. The function of fibrinogen is to form a fibrin clot when activated by thrombin; therefore, all fibrinogen is virtually removed in the clotting process and should not be present in serum specimens. Fibrinogen is a positive acute-phase reactant and increases significantly during an inflammatory

process. Fibrinogen levels also rise with pregnancy and the use of oral contraceptives. Decreased values generally reflect extensive coagulation during which fibrinogen is consumed. On electrophoresis, fibrinogen can be seen as a small, distinct band between the β- and γ-globulin regions and indicates use of plasma instead of serum.[53]

C-reactive protein (CRP) is synthesized in the liver and is one of the first acute-phase reactants to rise in response to inflammatory disease. CRP rises sharply whenever there is tissue inflammation and has been demonstrated through many studies to be important in the development of atherosclerosis. Atherosclerosis, in addition to being a disease of lipid accumulation, also represents a chronic inflammatory process, which leads to an elevated CRP concentration. Elevated levels of CRP stimulate the production of tissue factor, which initiates coagulation, activates complement, and binds to LDL in the atherosclerotic plaque. As a result, high-sensitivity CRP (hsCRP) concentrations are measured during the evaluation of arteriosclerosis and as a risk indicator for cardiovascular disease. CRP is not tissue-specific; however, it does have value as a general indicator of inflammation. Normally, there are minimal amounts of CRP in blood. A high or increasing amount suggests an acute infection or inflammation. Results above 1 mg/dL are considered high for CRP, and most infections and inflammatory processes result in levels above 10 mg/dL. In cases of inflammatory rheumatic diseases, such as rheumatoid arthritis and SLE, CRP is used to assess the effectiveness of treatment and monitor periods of disease eruption. However, even in known cases of inflammatory disease, a low CRP concentration is possible and is not indicative of the absence of inflammation. CRP is also significantly elevated in acute rheumatic fever, bacterial infections, myocardial infarctions, carcinomatosis, gout, and viral infections.[54]

High-sensitivity CRP (hsCRP) refers to a monoclonal antibody–based test method that can detect CRP at levels below 1 mg/L. This test is most commonly used to determine risk of cardiovascular disease (CVD) as high levels of CRP consistently predict recurrent coronary events in patients with unstable angina and AMI.[54] A detailed discussion of cardiovascular risk factors and markers of cardiac damage can be found in Chapter 20, *Cardiac Function*.

γ-Globulins

Immunoglobulins (Igs), or antibodies, are glycoproteins composed predominately of protein with a small percentage of carbohydrate (14–18%). Immunoglobulins are produced by white blood cells, known as B lymphocytes, that confer humoral immunity. These proteins consist of two identical heavy (H) chains and two identical light (L) chains linked by two disulfide bonds. They can be in the form of a monomer (one unit), dimer (two units), or a pentamer (five units). There are five classes, or isotypes, of immunoglobulins: IgG, IgA, IgM, IgD, and IgE. This classification is based on the type of heavy chain the immunoglobin possesses, which include γ, α, μ, δ, and ε, respectively. There are two types of light chains, kappa (κ) and lambda (λ). For example, IgG has two γ-type heavy chains and two identical light chains, either κ or λ. Each heavy chain has two regions; the constant region and the variable region. The constant region is identical in all antibodies of the same isotype but differs in antibodies of different isotypes. The variable region of the heavy chain differs in antibodies produced by different B lymphocytes but is the same for all antibodies produced by a single B lymphocyte or B-lymphocyte clone. The ratio of κ to λ chains is 2:1, which is sometimes used as a marker of immune abnormalities.[55]

The N-terminal regions of the heavy and light chains exhibit highly variable amino acid composition referred to as VH and VL, respectively. This variable region is involved in antigen binding. Similar to the variable region, the constant domains of light and heavy chains are referred to as CL and CH, respectively. The constant regions are involved in complement binding, placental passage, and binding to cell membranes.[55] **Figure 6.8** illustrates the structure and forms of an immunoglobulin.

Genes for the variable regions contain three distinct types of segments encoded in the human genome. For example, the immunoglobulin heavy chain region contains 65 *Variable* (V) genes plus 27 *Diversity* (D) genes and six functional *Joining* (J) genes. The light chains also possess numerous V and J genes, but do not have D genes. By the mechanism of DNA rearrangement of these regional genes, it is possible to generate an antibody repertoire of more than 10^7 possible combinations. V(D)J recombination is a genetic mechanism that randomly selects and assembles segments of genes encoding specific proteins, which generates a diverse repertoire of T-cell receptor and immunoglobulin molecules. These molecules are necessary for the recognition of diverse antigens from bacterial, viral, and parasitic pathogens and from dysfunctional cells such as tumor cells.[55]

Immunoglobulin class switching, or isotype switching, is a biologic mechanism that changes an

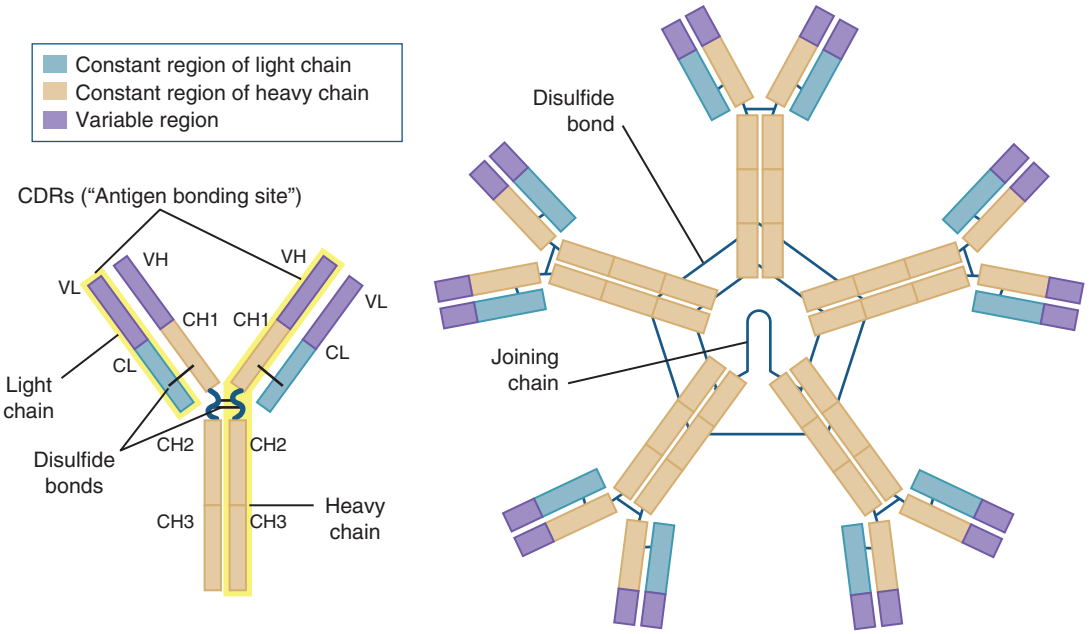


Figure 6.8 Immunoglobulin structure of IgG and IgM.

antibody from one class to another.[56] For example, an isotype IgM could be changed to an isotype IgG. This process occurs after activation of the B lymphocyte, which allows the cell to produce different classes of antibody. Only the constant region of the antibody heavy chain changes during class switching. Because the variable region does not change, class switching does not affect the antigens that are bound by the antibody. Instead, the antibody retains affinity for the same antigens but can interact with different effector molecules, which are regulatory molecules that bind to a protein and alter their activity.

The antibody molecule has a "Y" shape, with the top being the site that binds antigen, and, therefore, recognizes specific foreign objects. This region of the antibody is called the Fab (fragment, antigen binding) region. It is composed of one constant and one variable domain from each heavy and light chain of the antibody. The base of the Y is called the fragment crystallizable (Fc) region and is composed of two heavy chains that contribute two or three constant domains depending on the class of the antibody. By binding to specific proteins, the Fc region ensures that each antibody generates an appropriate immune response for a given antigen. The Fc region also binds to various cell receptors, such as Fc receptors, and other immune molecules, such as complement proteins. By doing this, it mediates different physiologic effects including opsonization, cell lysis, and degranulation of mast cells, basophils, and eosinophils.[55]

Immunoglobulin G (IgG) is the most abundant class of antibodies found in blood plasma and lymph. IgG antibodies act on bacteria, fungi, viruses, and foreign particles by agglutination, opsonization, and complement activation and by neutralizing toxins. IgG is increased in liver disease, infections, IgG myeloma, parasitic disease, and many rheumatic diseases. Decreased IgG levels are associated with acquired immunodeficiency, an increased susceptibility to infections, hereditary deficiency, protein-losing states, and non-IgG myeloma. Immunoglobulins are not synthesized to any extent by the developing fetus; however, IgG can cross the placenta, so any IgG present in a newborn's serum was synthesized by the mother.[55]

Immunoglobulin A (IgA) is the main immunoglobulin found in mucous secretions, including tears, saliva, colostrum, vaginal fluid, and secretions from the respiratory and gastrointestinal mucosa, and is also found in small amounts in blood. It is made by B lymphocytes. IgA exists in two isotypes, IgA1 and IgA2. While IgA1 and IgA2 are both found in serum and mucosa, IgA1 predominates in serum, while IgA1 and IgA2 tend to be evenly distributed in the mucosa. IgA is also classified, based upon location, as serum IgA or secretory IgA. Secretory IgAs

are polymers of two to four IgA monomers linked by a joining protein (J chain), which is a polypeptide-containing cysteine and a secretory component. It is different structurally from other immunoglobulins. Secretory IgA is resistant to enzyme degradation and remains active in the digestive and respiratory tracts to provide antibody protection in body secretions. Increases in serum IgA are found in liver disease, infections, and autoimmune diseases, while decreases are found in impaired protein synthesis and immunodeficiency.[57]

Immunglobulin M (IgM) is the first antibody to appear in response to antigenic stimulation and is present in B lymphocytes. IgM is a pentamer and contains a J chain. Both anti-A and anti-B are naturally occurring IgM antibodies as part of the ABO blood group system. An increased IgM concentration is found in bacterial infections, toxoplasmosis, primary biliary cirrhosis, cytomegalovirus, rubella, herpes, and various fungal diseases. A monoclonal increase in IgM antibodies is seen in Waldenström's macroglobulinemia as a spike in the vicinity of the late β zone on protein electrophoresis. Decreases are seen in protein-losing conditions and hereditary immunodeficiencies. IgM cannot cross the placenta, and it is the only immunoglobulin synthesized by the neonate.[58]

Immunoglobulin D (IgD) molecules are present on the surface of most, but not all, B lymphocytes early in their development, though little IgD is ever released into the circulation. IgD may help regulate B lymphocyte function; however, the specific function of circulating IgD is largely unknown. Its concentration is typically increased in infections, liver disease, and connective tissue disorders.[55]

Immunoglobulin E (IgE) is produced by B lymphocytes; increased concentrations are associated with allergic and anaphylactic reactions, autoimmune processes, and parasitic infections. In contrast to other immunoglobulins, the concentration of IgE in the circulation is very low. An elevated concentration is not diagnostic of any single condition but is observed in many inflammatory and infectious diseases, including asthma and hay fever. Monoclonal increases are seen in IgE myeloma, but these are rare.[59]

CASE STUDY 6.1, PART 2

Remember Sean, the 3-week-old infant who was seen in the emergency department. Sean is now admitted to the hospital with pneumonia. The physician orders a number of laboratory tests to evaluate the child's immune system. Hematology results show that the child does not have anemia and that his white blood cell count is slightly elevated compared to the reference range. Additional testing reveals normal levels of B lymphocytes and T lymphocytes. Chemistry results for total protein, albumin, and immunoglobulin levels are listed in **Case Study Table 6.1**.

Case Study Table 6.1 Laboratory Results

Test	Result	Reference Range
Total protein	8.7 g/dL	1 y: 5.4–7.5 g/dL
Albumin	3.8 g/dL	1–3 y: 3.4–4.2 g/dL
IgG	153 mg/dL	1–3 y: 507–1407 mg/dL
IgM	576 mg/dL	1–3 y: 18–171 mg/dL
IgA	11 mg/dL	1–3 y: 63–298 mg/dL
IgD	0 mg/dL	Newborn to adult: 0–8 mg/dL
IgE	1 kU/L	0–5 mos. <13 kU/L

© Myroslava Malovana/Shutterstock.

1. Which serum immunoglobulin results are abnormal?
2. Which immunoglobulin is the first type to increase following antigenic stimulation?
3. How do the laboratory results correlate with recurring bacterial infections?

Other Proteins of Clinical Significance

Troponin

Cardiac troponin (cTn) represents a complex of regulatory proteins: troponin C (TnC), troponin I (cTnI), and troponin T (cTnT), that are specific to cardiac muscle. As such, they are considered the gold standard for diagnosis of acute coronary syndrome (ACS), in which the blood supply to the heart muscle is suddenly impeded. It is recommended that cTn be measured in all patients presenting with symptoms suggestive of ACS in conjunction with physical examination and electrocardiogram. Because of the specificity of cTn for myocardial damage, a single blood cTn measurement above the decision limit, along with clinical evidence, is indicative of myocardial injury.[60] A detailed discussion of cardiac troponins and other markers used to evaluate cardiac damage can be found in Chapter 20, *Cardiac Function*.

Natriuretic Peptides

Natriuretic peptides are a family of structurally related hormones that include atrial natriuretic peptide (ANP), B-type (or brain) natriuretic peptide (BNP), C-type natriuretic peptide (CNP), and dendroaspis natriuretic peptide (DNP). B-type natriuretic peptides are produced initially as the prohormone, proBNP, which is cleaved by the enzymes corin and furin into N-terminal (NT)-proBNP, a biologically inert portion, and BNP, the biologically active portion. NT-proBNP and BNP are found in greatest concentration in the left ventricular myocardium. They are also detectable in atrial tissue and the myocardium of the right ventricle. Natriuretic peptides are neurohormones that affect body fluid homeostasis, through natriuresis and diuresis, and blood pressure, through decreased angiotensin II and norepinephrine synthesis (see Chapter 11, *Electrolytes*). Elevated concentrations of BNP and NT-proBNP are associated with cardiac diseases such as congestive heart failure and acute coronary syndrome.[61] Additional discussion of cardiac markers can be found in Chapter 20, *Cardiac Function*.

Fibronectin

Fibronectin is a glycoprotein composed of two nearly identical subunits. Although fibronectin is the product of a single gene, the resulting protein can exist in multiple forms due to alternate splicing of a single pre-mRNA. The variants demonstrate a wide variety of cellular interactions including roles in cell adhesion, tissue differentiation, growth, and wound healing. These proteins are found in plasma and on cell surfaces and can be synthesized by hepatocytes, endothelial cells, peritoneal macrophages, and fibroblasts. Fetal fibronectin (fFN) is used to help predict the short-term risk of premature delivery. fFN is produced at the boundary between the amniotic sac and the lining of the uterus. It is thought that fFN helps maintain the integrity of this boundary and adherence of the placenta to the uterus. fFN is normally detectable in amniotic fluid and placental tissue during early pregnancy, but in a normal pregnancy, it is not detectable after 24 weeks of gestation. Elevated fFN concentrations between 22 and 36 weeks of gestation reflect a disturbance at the uteroplacental junction and are associated with an increased risk of preterm labor and delivery.[62,63] See Chapter 24, *Pregnancy and Prenatal Testing*, for additional information.

Cross-Linked C-Telopeptide

Cross-linked C-telopeptide (CTX) is a proteolytic fragment of type I collagen crosslinked at the N- and C-terminal ends of the molecule and formed during bone resorption. CTX measurements cannot replace bone mineral density testing in the diagnosis of osteoporosis, but serum determinations may be performed to measure bone resorption and monitor response to antiresorptive therapies[64] in postmenopausal women treated for osteoporosis.

Cystatin C

Cystatin C is a low-molecular-weight protease inhibitor produced by all nucleated cells. Due to its small size, it is freely filtered by the glomerulus but is then almost completely reabsorbed and catabolized by the proximal tubular cells. Unlike creatinine, which is routinely measured to evaluate the glomerular filtration rate, cystatin C concentrations are not affected by muscle mass, gender, or age. As such, cystatin C may be used to evaluate glomerular function in individuals for whom creatinine measurements may be misleading, such as patients with cirrhosis, obesity, malnutrition, or who have a reduced muscle mass.[65] See Chapter 7, *Nonprotein Nitrogen Compounds*, for more information on markers used to evaluate renal function.

Plasma Total Protein Abnormalities

Total protein measurements include all plasma proteins and are used in the evaluation of nutritional status, kidney disease, liver disease, and many other conditions (**Table 6.4**). If total protein results are abnormal, further testing must be performed to identify which protein fraction is abnormal so that a specific diagnosis can be made.

Hypoproteinemia

Hypoproteinemia refers to conditions in which the serum or plasma total protein concentration is below the reference range. General causes of hypoproteinemia include excessive loss, decreased synthesis, and increased protein catabolism. Excessive loss of plasma proteins may occur as a result of renal disease, such as nephrotic syndrome, in which increased glomerular permeability allows inappropriate passage of proteins into the glomerular ultrafiltrate for excretion. Plasma proteins may also be lost due to leakage into the gastrointestinal tract during inflammation of the digestive system, through blood loss (e.g., wounds, internal bleeding), or from extensive burns. Hypoproteinemia may also be the result of decreased intake of proteins and amino acids associated with malnutrition, maldigestion, or malabsorption syndromes. Liver disease can also cause low serum protein concentrations as the liver is a primary site for protein synthesis. Inherited immunodeficiency disorders, such as hypogammaglobulinemia and x-linked agammaglobulinemia, in which antibody production is diminished, also results in decreased blood total protein concentrations. Accelerated catabolism of proteins, as occurs in burns, trauma, or other injuries, is another cause of hypoproteinemia.[66]

Table 6.4 Reference Ranges for Serum/Plasma Protein Analyses

TOTAL PROTEIN—SERUM, PLASMA	
Adult	6.5–8.3 g/dL (65–83 g/L)
ALBUMIN—SERUM, PLASMA	
Adult	3.5–5.5 g/dL (35–55 g/L)
ALBUMIN-GLOBULIN RATIO (A/G)—SERUM, PLASMA	
Adult	1.1–1.8

Hyperproteinemia

Hyperproteinemia, or an increase in total plasma proteins, is not generally associated with a specific disease state, but rather, it is more likely the result of dehydration. In dehydration, excessive water is lost from the vascular system, but proteins, due to their size, remain within the blood vessels. Although the absolute quantity of proteins in the blood remains unchanged, the concentration appears elevated due to a decreased volume of water, which is the solvent. In addition to decreased fluid intake, dehydration may result from a variety of conditions including vomiting, diarrhea, excessive sweating, diabetic acidosis, and hypoaldosteronism.[66]

Hyperproteinemia may also be the result of excessive production of proteins, primarily of the γ-globulins. These disorders may be characterized by the appearance of a monoclonal protein, or paraprotein, in the plasma, which often appears in the urine as well. Monoclonal proteins are intact immunoglobulin molecules or, occasionally, κ or λ light chains only. Many disorders, including chronic inflammatory states, collagen vascular disorders, and other neoplasms, may be associated with paraproteins. Waldenström's macroglobulinemia is a rare type of slow-growing, non-Hodgkin lymphoma associated with IgM paraprotein; however, the most well-known monoclonal disorder is multiple myeloma, in which neoplastic B lymphocytes proliferate in the bone marrow.[67,68] Paraproteins in multiple myeloma may reach plasma concentrations of several grams per deciliter and are usually IgG, IgA, or κ or λ light chains.

Polyclonal increases in immunoglobulins, which would be represented by increases in both κ and λ light chains, are seen in the patient's serum and urine in many chronic diseases. During inflammation and infection, positive acute-phase reactants, such as C-reactive protein, increase and negative acute reactants, such as albumin, decrease.[69]

Methods of Analysis

Protein concentrations may be useful in evaluating patients for nutritional status, gastrointestinal disease, kidney disease, liver disease, and many other conditions. A number of different test methods are used in the clinical laboratory to qualitatively and quantitatively determine protein concentrations in body fluids. These methods rely on the various physical and chemical properties of proteins. **Table 6.5** summarizes the properties used for separation and analysis of proteins. Additional information

Table 6.5 Physical Properties of Proteins Used in Separation and Analysis

Property	Applications
Molecular Size	Gel chromatography Ultracentrifugation Gradient pore electrophoresis Mass spectrometry
Solubility	Protein precipitation Turbidimetry Nephelometry
Charge	Serum protein electrophoresis Isoelectric focusing Immunofixation electrophoresis Ion exchange chromatography
Molecular Interactions	Affinity chromatography Immunoassays Immunoelectrophoresis Immunofixation Dye-binding

on specific analytical techniques can be found in Chapter 4, *Analytic Techniques*.

Sample Requirements

Both serum and plasma are acceptable specimens. Serum samples collected in a serum separator tube are preferred for protein analyses due to the presence of fibrinogen in plasma. Plasma samples should be collected using a plasma separator tube. Serum samples should be allowed to clot completely at room temperature. Samples should be centrifuged within two hours to separate the serum/plasma from the cells.[70]

Total Protein

Total protein concentrations are lower at birth but reach adult levels (6.5–8.3 g/dL) by 3 years of age. Total protein concentrations decrease in pregnancy and with age. In terms of pre-analytical variables, hemolyzed specimens are generally considered unacceptable because the release of red blood cell proteins can falsely elevate the total protein result. When patients are drawn in a recumbent position, the reference range is slightly lower due to reduction in plasma albumin and fluid shift toward the extracellular compartments. Total protein concentrations above the reference range indicate hyperproteinemia, and those below the reference range indicate hypoproteinemia. Methods for the determination of total protein are described below and summarized in **Table 6.6**.

Biuret Method

The biuret reaction is widely used in the determination of total protein concentrations in serum or plasma. The reaction is based on the principle that in an alkaline medium and the presence of at least two peptide bonds, cupric ions (Cu^{2+}) will complex with groups involved in the peptide bond to form a violet-colored

Table 6.6 Analytical Methods to Measure Proteins

Method	Principle	Comment
Biuret	Formation of violet-colored chelate between Cu^{2+} ions and peptide bonds	Routine method; requires at least two peptide bonds and an alkaline medium
Turbidimetry & Nephelometry	Formation of aggregates by protein precipitation or antibody-binding that affect light scatter	Immunoturbidimetric and immunonephelometric methods used to quantitate specific proteins
Dye binding	Dye binds to protein causing a spectral shift in the absorbance maximum of the dye	Used in automated protein measurements and to quantitate proteins separated by electrophoresis
Protein electrophoresis	Migration of proteins based on their density and charge under the influence of an electric field	Used to separate and quantify protein fractions in serum, urine, and cerebrospinal fluid
Ultraviolet (UV) absorption	Proteins in solution absorb UV light with maximum absorbances of 200 and 280 nm	Limited utility for protein mixtures due to variable absorption characteristics; potential use in measurement of specific fractions after separation by another method[71]

chelate. The reagent also contains sodium potassium tartrate, which forms a complex with cupric ions to prevent their precipitation in the alkaline solution, and potassium iodide, which acts as an antioxidant. The absorbance of the colored chelate is measured at 540 nm using a spectrophotometer. The color and resultant absorbance are proportional to the number of peptide bonds present and reflect the total protein concentration of the specimen. The color of the reaction varies from a pink to a reddish-violet color because when small peptides react, the color of the chelate has a different shade than that seen with larger peptides.[66]

Turbidimetry and Nephelometry

Turbidimetry and nephelometry are used to measure the concentration of a solution with particles. In the case of proteins, the particles may be protein precipitated by the addition of a reagent or large protein-antibody complexes. Turbidimetry uses a spectrophotometer to measure the amount of light transmitted through the particulate solution. Nephelometers are similar to spectrophotometers, except they measure the amount of light scattered by the particles. The detectors are placed at various forward angles as well as 90° to the incident light. These techniques are described further in Chapter 4, *Analytic Techniques*.

Dye Binding

Dye binding methods are based on the ability of most proteins to bind dyes, causing a spectral shift in the absorbance maximum, although the affinity with which they bind may vary. Several dyes, including bromophenol blue, Ponceau S, amido black, lissamine green, and Coomassie brilliant blue, may be used to stain protein bands after electrophoresis.[66]

Ultraviolet Absorption

The concentration of a protein can be estimated by its ultraviolet (UV) absorption at 280 nm. This is due to the property of the aromatic amino acids tryptophan and tyrosine to absorb UV light with a maximum absorbance at 280 nm. Phenylalanine absorbs UV light with a maximum of 255 nm, but not at 280 nm. UV absorption measurement is used to estimate protein concentration in a solution, protein contamination in DNA extraction preparations, and to investigate protein folding, reduction, and interactions with ligands. The utilization of UV absorption to measure peptide bonds is limited to peptide fragments.[71–73]

Albumin

The most widely used methods for determining albumin are dye-binding procedures utilizing bromocresol green (BCG) or bromocresol purple (BCP). The pH of the solution is adjusted so that albumin is positively charged and will bind to an anionic dye by electrostatic forces. When bound to albumin, the dye has a different absorption maximum than the free dye. The concentration of albumin is then calculated by measuring the absorbance of the albumin–dye complex, which is proportional to the specimen's albumin concentration.[66]

BCG methods are affected by interfering substances such as hemolysis, lipemia, and icterus. Hemoglobin binds to BCG and can lead to falsely elevated albumin values in hemolyzed samples, therefore, grossly hemolyzed specimens should be rejected and recollected. Lipemic specimens should be ultracentrifuged, while icteric specimens may be diluted with saline. A limitation of the BCG method is that it overestimates albumin values due to interference from acute-phase reactants. Interference also occurs due to binding of BCG to α_1 and α_2 globulins in the patient sample. The specificity of the reaction for albumin can be improved by taking absorbance readings within a short, standardized range (<5 minutes) as longer incubations increase binding of BCG with other plasma proteins.[74]

BCP binds specifically to albumin and has fewer interfering substances than BCG, but does appear to be affected by the presence of bilirubin. In patients with renal insufficiency or those undergoing hemodialysis, the BCP method underestimates albumin concentrations. Both methods are easily automated, but BCP may be preferred due to its increased specificity for albumin.[74]

When evaluating total serum protein concentrations, it may be clinically useful to determine the albumin and globulin fractions and compare them as a ratio (A/G ratio). The globulin fraction is not measured but is calculated from measurements of total protein and albumin as seen in **Equation 6-2**.

$$A/G\ ratio = \frac{(Albumin)}{(Globulin)} = \frac{(Albumin)}{(Total\ Protein - Albumin)} \quad \text{(Eq. 6-2)}$$

A decreased A/G ratio may be the result of decreased albumin or increased globulins. Low serum albumin concentrations may be the result of decreased protein synthesis by the liver or renal loss of proteins as might occur in nephrotic syndrome. Globulins may be elevated due to conditions such

CASE STUDY 6.1, PART 3

Remember Sean, the 3-week-old infant that was seen in the emergency department and was admitted to the hospital.

© Myroslava Malovana/Shutterstock.

4. Calculate the A/G ratio based on his laboratory results. Use the adult reference range found in Table 6.4 to determine if his result is normal or not.

as multiple myeloma, chronic inflammatory disorders, or rheumatoid arthritis. An elevated A/G ratio is generally associated with decreased globulin synthesis as an increased albumin is likely the result of dehydration, which would affect all analyte measurements.

Immunochemical Assays

Specific proteins may be identified and quantified by immunochemical assays (immunoassays) in which the reaction of the antigen and its complementary antibody is measured. There are several different types of immunoassays used in the clinical laboratory including radioimmunoassay, enzyme immunoassay, chemiluminescent immunoassay, and enzyme-linked immunosorbent assays. Even lateral flow immunochromatographic methods are now commonly used for point-of-care testing for proteins.[66] A detailed discussion of the various types of immunoassays can be found in Chapter 4, *Analytic Techniques*.

Electrophoresis

When an abnormality in the total protein concentration is found, protein electrophoresis may be performed to identify which fraction(s) is involved. Electrophoresis separates proteins on the basis of their electric charge and density across a support media such as an agarose gel or cellulose acetate. Proteins, when placed in an electric current, will move according to their charge, which is determined by the pH of the surrounding buffer. At a pH greater than its p*I*, the protein is negatively charged (anion) and will migrate toward the anode, which is the positive terminal. At a pH less than its p*I*, the protein will be positively charged (cation) and will migrate toward the negatively charged terminal, the cathode. The speed of migration can be estimated from the difference between the p*I* of the protein and the pH of the buffer. The greater the difference between the pH of the buffer and the p*I*, the greater the magnitude of the net charge of that protein, and the faster it will move in the electric field. In addition to the size, shape, and charge of the protein, the velocity of the migration also depends on the electric field strength, reaction temperature, composition, pH, and ionic strength of the buffer. The electrophoretic mobility (μ) of a specific protein may be calculated using the following equation, where s is the distance traveled in centimeters, t is time of migration in seconds, and F is the field strength in V/cm.[66]

$$\mu = \frac{(s/t)}{F} \qquad \textbf{(Eq. 6.3)}$$

Serum Protein Electrophoresis

In the standard performance of serum protein electrophoresis (SPE), serum specimens are applied to a support medium such as an agarose gel or cellulose acetate plate. The support medium is then placed into an electrophoretic chamber containing an alkaline buffer (pH 8.4–8.8) with the application end close to the cathode. An electrical current is applied to separate the proteins. All major serum proteins carry a net negative charge at pH 8.6 and will migrate toward the anode. Using standard SPE methods, serum proteins separate into five fractions (bands or zones): albumin, α_1-globulins, α_2-globulins, β-globulins, and γ-globulins. Some laboratories may prefer to use a split-β protein gel that allows separation of β-globulins into two fractions: β_1 (hemopexin and transferrin) and β_2 (β lipoproteins and complement).[66]

After separation, the gel/plate is stained allowing visualization of the bands. A variety of dyes may be used including Ponceau S, amido black, Coomassie blue, and acid blue stain. The amount of dye bound in a particular fraction depends on the concentration of proteins present. After destaining (clearing) and drying the gel/plate, it is placed in a densitometer. A densitometer is a type of spectrophotometer used to measure light transmission through a solid, such as the gel/plate, at a specific wavelength. The measurements are graphed as an electropherogram, and the area under each peak can be used to quantitate the relative percent and absolute concentration of each band relative to the total protein concentration.[66] **Figure 6.9** shows a normal electropherogram and electrophoresis pattern on cellulose acetate. Reference ranges are listed in **Table 6.7**.

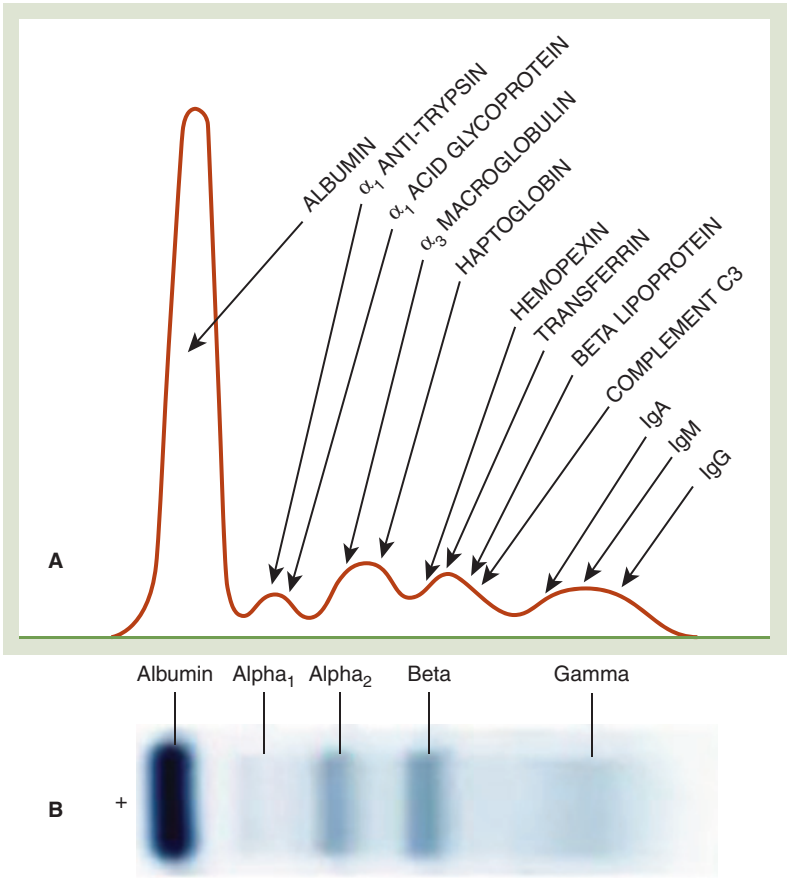


Figure 6.9 Classic 5-banded separations of serum proteins of reference sample. (**A**) Electropherogram pattern from densitometer, showing the major peaks: albumin, α_1, α_2, β, and γ. Arrows indicate the major serum proteins in each peak. (**B**) Matching electrophoresis pattern on cellulose acetate, fixed and stained.

Used with permission from Helena Laboratories.

The specimen of choice for SPE is serum collected in a serum separator tube, but a red top, no additive tube may also be acceptable. The sample should be centrifuged, and the serum separated from the cells if testing is delayed. Analytical errors will occur if plasma is substituted for serum or if the serum is hemolyzed. Inadvertent use of plasma specimens will result in a narrow band between the β and γ regions because of the presence of fibrinogen. The presence of free hemoglobin can cause a small band in the late α_2 or early β zone, and the presence of hemoglobin–haptoglobin complexes will cause a small band in the α_2 zone.

Table 6.7 Reference Ranges for Serum Protein Electrophoresis

ALBUMIN	53%–65% of total protein (3.5–5.5 g/dL)
α_1-Globulins	2.5%–5.0% of total protein (0.1–0.3 g/dL)
α_2-Globulins	7.0%–13.0% of total protein (0.6–1.0 g/dL)
β-Globulins	8.0%–14.0% of total protein (0.7–1.1 g/dL)
γ-Globulins	12.0%–22.0% of total protein (0.8–1.6 g/dL)

When reviewing serum protein electrophoresis (SPE) patterns, the relative proportions of the fractions may be useful in determining specific disorders, as well as information about the homogeneity of a fraction (**Figure 6.10**).

Nephrotic syndrome is characterized by inappropriate passage of proteins through the renal glomeruli due to increased permeability of the basement membrane. As a result, significant amounts of albumin, low-molecular-weight proteins, and IgG are excreted in the urine leading to decreased blood concentrations. Due to its large size, α_2-macroglobulin is not filtered and appears elevated compared to the other plasma proteins. As such, SPE patterns for patients

CASE STUDY 6.2, PART 2

Remember Guillermo, the 47-year-old man who had fallen and broken his leg. The radiograph of his ankle showed bone loss. Based on admitting chemistry test results, the provider ordered a serum protein electrophoresis.

1. Compare the image of the electrophoresis gel (Figure A) to the reference pattern in Figure 6.9. What protein fraction shows an increase?
2. What additional test should be ordered to identify the increased protein?

A

with nephrotic syndrome show a dramatic decrease in the relative amounts of albumin, α_1-globulins, β-globulins, and γ-globulins and an apparent increase in the α_2-globulin fraction (**Figure 6.10A**).[75]

Inflammatory conditions will also result in an abnormal SPEP due to increases in positive acute-phase reactants and decreases in negative acute-phase reactants. **Figure 6.10B** shows an acute-phase reactant pattern, which is associated with trauma, burns, myocardial infarction, toxicosis, and acute infections. Positive acute-phase reactants, such as α_1-antitrypsin, haptoglobin, ceruloplasmin, α_2-macroglobulin and CRP, increase within days of the event resulting in an increase in the relative amounts of the α_1- and α_2-globulin fractions. Chronic inflammatory conditions, such as chronic infections, malignancy, autoimmune diseases, and liver disease, will also demonstrate increases in the α_1- and α_2-globulin fractions. However, SPE will also show a decrease in albumin and an increase in γ-globulins.[69]

As illustrated in **Figure 6.10C**, a decreased or absent α_1-globulin fraction is associated with α_1-antitrypsin deficiency, as α_1-antitrypsin accounts for approximately 90% of the α_1-globulin fraction.[76]

Immunodeficiency disorders, such as hypogammaglobulinemia (**Figure 6.10D**) and agammaglobulinemia, are associated with low or absent γ-globulin fractions, respectively, due to impaired production of immunoglobulins.[77,78] In contrast, gammopathies (hypergammaglobinemias) are the result of increased production of immunoglobulins. Monoclonal gammopathies are disorders involving proliferation of a single B lymphocyte clone, such as occurs in multiple myeloma. This increase results in a single homogenous spike (M protein) in the γ-globulin region (**Figure 6.10E**).[68] A broad increase in the γ-globulin region, as seen in **Figure 6.10F**, is suggestive of a polyclonal gammopathy which may be associated with conditions such as infections, hematologic disorders, and autoimmune disorders.[79]

The electrophoretic pattern associated with severe liver disease shows a relative decrease in albumin, α_1-globulins, α_2-globulins, and β-globulins, with a relative increase in γ-globulins (**Figure 6.10G**). This pattern is anticipated as the majority of plasma proteins are synthesized in the liver, whereas γ-globulins are produced by B lymphocytes. The γ-globulin fraction also rises in infectious hepatitis due to increases in IgG and IgM production. In chronic diseases of the liver, such as cirrhosis, increases in fast-moving γ-globulins, such a IgA, prevent resolution of the β- and γ-globulin bands, resulting in a β–γ bridge (**Figure 6.10H**).[80]

High-Resolution Protein Electrophoresis

High-resolution protein electrophoresis (HRE) uses a higher voltage coupled with a cooling system in the electrophoretic apparatus and a more concentrated buffer. Standard SPE separates proteins into five distinct bands. By modifying the electrophoretic parameters using HRE, proteins can be separated into as many as 15 distinct bands. After staining, each band is compared with the same band on a reference pattern to evaluate its color density, appearance, and migration rate. Careful evaluation for the appearance of abnormal bands or regions of density is also important. As with routine protein electrophoresis, the patterns may be scanned with a densitometer to obtain semiquantitative estimates of the protein found in each band. HRE is particularly useful in detecting small monoclonal bands and differentiating unusual bands or prominent increases of normal bands that can be confused with a monoclonal gammopathy.[81]

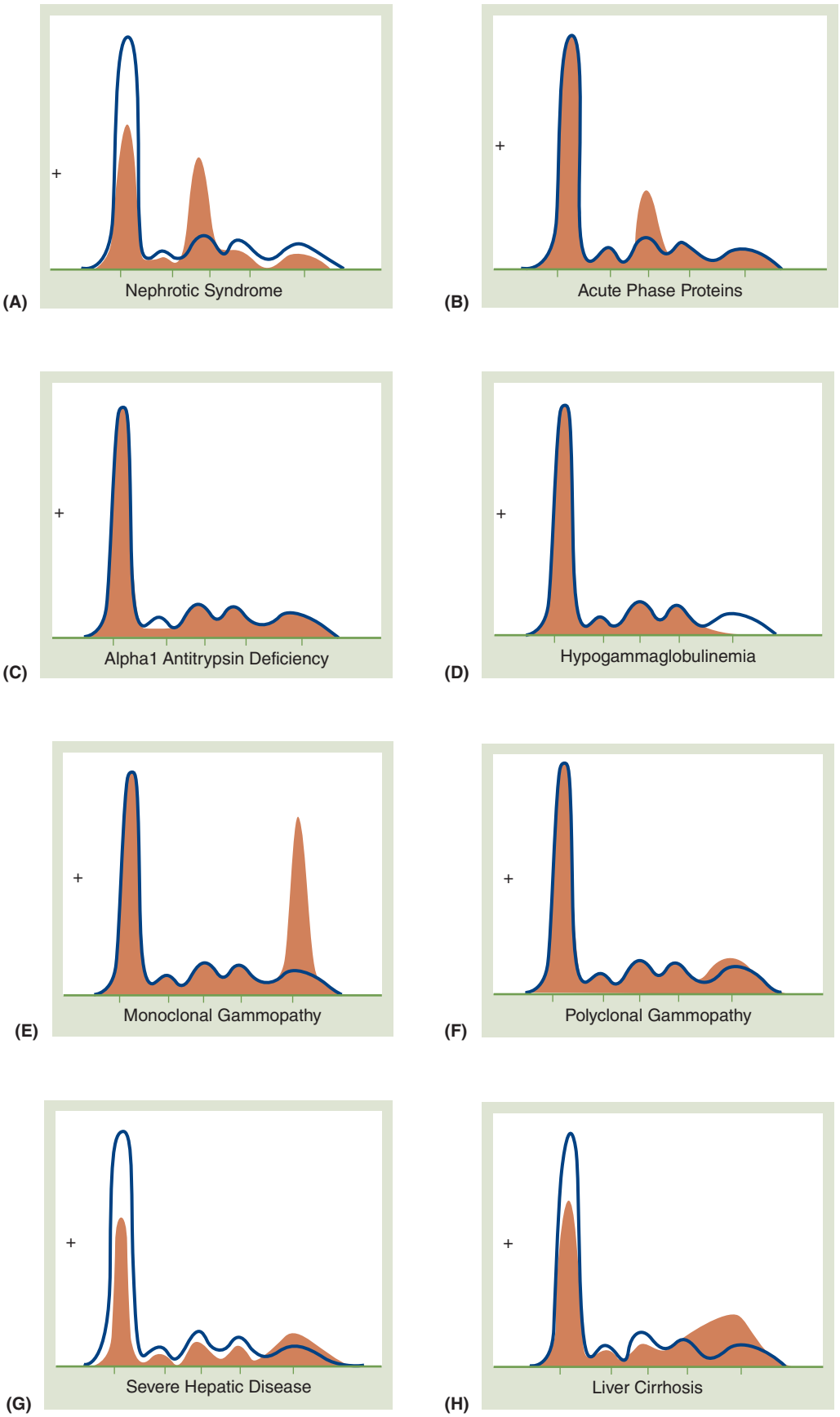


Figure 6.10 Selected patterns of protein electrophoresis. (**A**) Nephrotic syndrome; (**B**) acute-phase proteins; (**C**) α-1 antitrypsin deficiency; (**D**) hypogammaglobulinemia; (**E**) monoclonal gammopathy; (**F**) polyclonal gammopathy; (**G**) severe hepatic disease; and (**H**) liver cirrhosis.

Used with permission from Helena Laboratories.

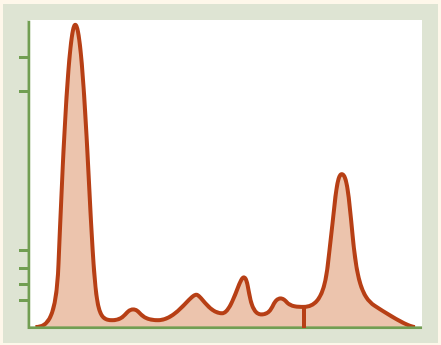
CASE STUDY 6.2, PART 3

Remember Guillermo, the 47-year-old man who had fallen and broken his leg.

3. Compare the image of Guillermo's electropherogram from the densitometer (Figure B) to the reference patterns in Figure 6.10. Which pattern looks the most similar?

© Russell Monk/The Image Bank/Getty Images.

B

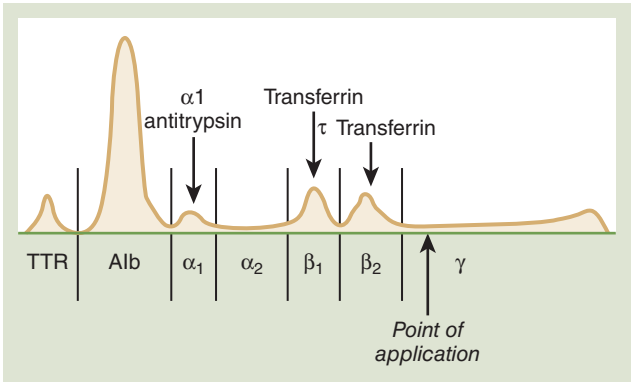


Used with permission from Helena Laboratories.

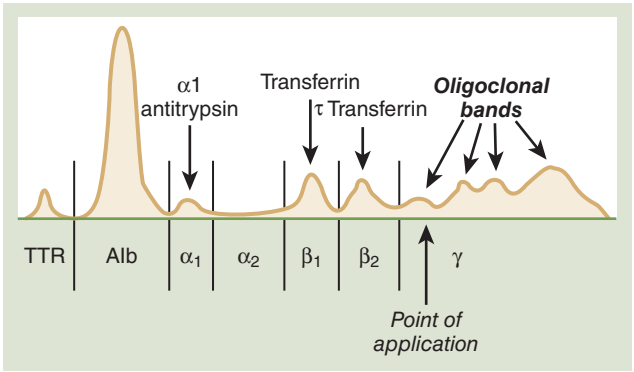
HRE may also be used to separate proteins in other body fluids, such as urine and cerebrospinal fluid (CSF), after samples have been concentrated. Normal and abnormal CSF HRE patterns are shown in **Figure 6.11** demonstrating the presence of three oligoclonal bands in the γ-region, which is characteristically seen in patients with multiple sclerosis.

Capillary Electrophoresis

Capillary electrophoresis is a collection of techniques in which separation takes place inside a silica capillary. The capillaries are typically 30 to 50 cm long with an internal diameter between 25 and 100 μm. The capillaries are filled with a conducting solution, usually an aqueous buffer; the detection end of the capillary is grounded; and the sample injection end is connected to a high-voltage power supply. When a positive voltage is applied, the positively charged buffer molecules flow to the detection end, which is negatively charged relative to the injection end. The net flow of the buffer toward the detector is called electro-osmotic flow (EOF). When a specimen is injected, all molecules move toward the negatively charged detector end of the capillary due to the EOF. However, the negatively charged molecules in the specimen migrate back toward the positively charged injector end, which is referred to as *electrophoretic mobility*. EOF is usually stronger than electrophoretic mobility and, eventually, all ions will migrate to the detector end, but with different net mobilities based on their differences in size and charge. The separated molecules are detected by their absorbance as they



(A)



(B)

Figure 6.11 High-resolution electrophoresis pattern of CSF. (**A**) Normal pattern. (**B**) Abnormal pattern demonstrating the presence of oligoclonal bands in the γ-region. TTR: transthyretin.

Adapted from Brunzel, N. A. (2013). Fundamentals of Urine and Body Fluid Analysis - E-Book. United Kingdom: Elsevier Health Sciences.

pass through a small window near the detection end of the capillary. Use of the capillaries allows heat to be effectively dissipated, which means that higher operating voltages can be used thereby improving separation and analysis times.

Isoelectric Focusing

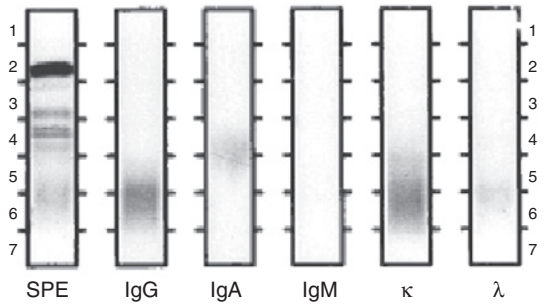
In isoelectric focusing (IEF), proteins are separated on the basis of their isoelectric point. IEF uses polyacrylamide or agarose gel mediums, which contain a pH gradient. The pH gradient is established by the incorporation of small polyanions and polycations in the gel. In the presence of an electric field, the varying p*I*s of the polyions cause them to seek their place in the gradient and to remain there. When a protein is electrophoresed in the gel, it migrates to a place where the pH is the same as its p*I*, resulting in a neutral charge. The protein becomes focused there because, if it should diffuse in either direction, it will gain a net charge. When this occurs, the electric current once again carries it back to where its pH is the same as its p*I*. Clinical applications of IEF include phenotyping of α_1-antitrypsin deficiencies, determination of genetic variants of enzymes and hemoglobins, detection of paraproteins in serum, oligoclonal bands in CSF, and isoenzyme determinations.

Immunoelectrophoresis

Immunoelectrophoresis (IEP) is a two-step process in which the sample (serum or urine) first undergoes electrophoresis to separate the proteins by their density and charge. After electrophoresis, antibodies are applied and allowed to diffuse into the plate. A precipitin arc will form where the antibody and antigen interact. The antibodies used in the test procedure are selected based on what protein(s) are being investigated. As an example, IEP has been used in the evaluation of monoclonal gammopathies, such as multiple myeloma, Waldenström's macroglobulinemia, and monoclonal gammopathy of undetermined significance (MGUS). It is essential to the diagnosis and treatment to identify which immunoglobulin class (heavy chain) and type (light chain) is affected.

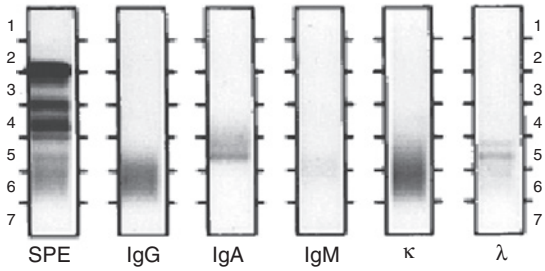
Immunofixation Electrophoresis

Immunofixation electrophoresis (IFE) has replaced the use of IEP in the clinical laboratory. However, the principle is similar. Proteins in serum, urine, or CSF samples are first separated using HRE. Antiseras (anti-IgG, anti-IgA, anti-IgM, anti-κ, and anti-λ) are then added and react with their complementary heavy or light chains forming insoluble antigen–antibody complexes that precipitate. A protein stain, such as acid blue or acid violet, is applied to enhance visualization of the bands. **Figure 6.12A** shows a normal serum IFE pattern. IFE is often performed as a reflex test when a monoclonal spike presents during serum protein electrophoresis, as it allows for identification of the specific M protein, such as the IgA λ monoclonal band seen in **Figure 6.12B** and the IgM κ monoclonal band in **Figure 6.12C**.



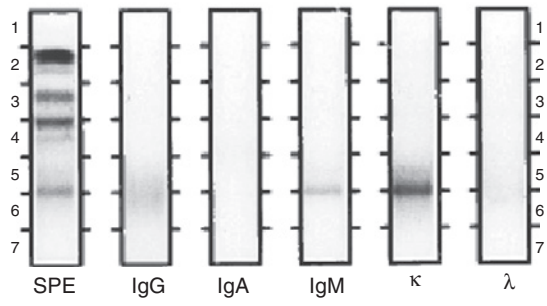
Normal IFE electrophoresis

(A)



IFE electrophoresis showing IgA λmonoclonal band

(B)



IFE electrophoresis showing IgM κmonoclonal band

(C)

Figure 6.12 Patterns of immunofixation electrophoresis (IFE). (**A**) Normal serum IFE; (**B**) IFE showing the IgA λ monoclonal band; and (**C**) IFE showing the IgM κ monoclonal band.

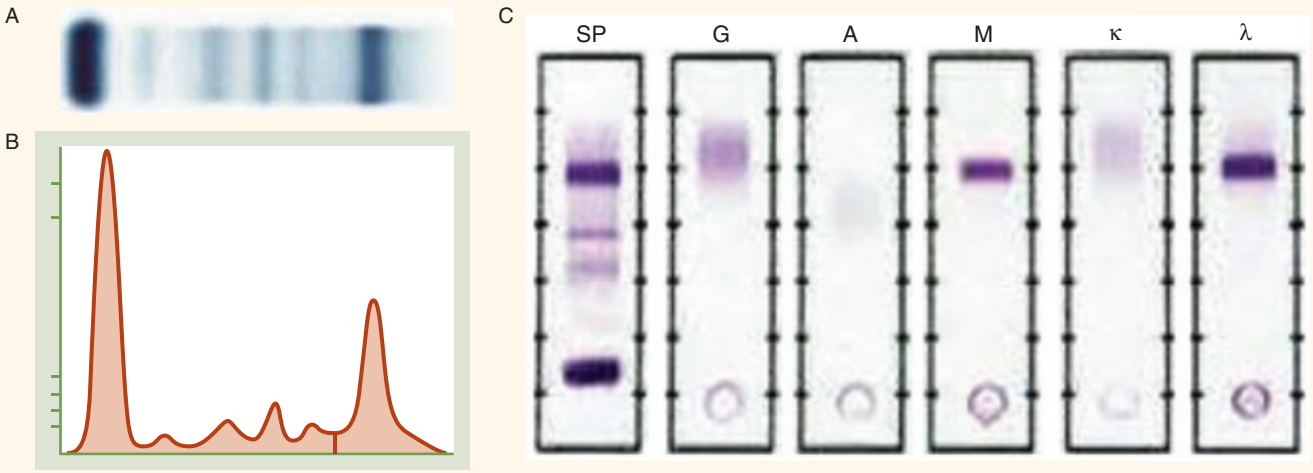
CASE STUDY 6.2, PART 4

Remember Guillermo, the 47-year-old man who had fallen and broken his leg. His provider ordered an IFE and the results are now available.

4. Evaluate the image of Guillermo's serum immunofixation electrophoresis in Figure C. Figure A is the serum protein electrophoresis (SPE). If you turn Figure C 90° to the right, it will look like the SPE pattern in Figure A. What immunoglobulin heavy chain is prominent? What light chain is in the same location and has similar staining intensity?

5. How would this gammopathy be classified?

© Russell Monk/The Image Bank/Getty Images.



(A–C) Used with permission from Helena Laboratories.

Proteins in Other Body Fluids

In addition to blood samples, many other body fluids are received by the clinical laboratory for the determination of their protein content including urine, cerebrospinal fluid, tears, and serous fluids. This chapter focuses on discussion of the analysis of urine and CSF.

Urine

In health, high-molecular-weight plasma proteins are not found in urine as they cannot easily penetrate the glomerular filtration barrier. In contrast, low-molecular-weight proteins, such as globulins, are readily filtered but are generally not found in urine as they are mostly reabsorbed in the renal tubules. Due to its moderate weight and high plasma concentration, small amounts of albumin may be filtered at the glomerulus and excreted in the urine. Some urinary proteins are directly produced in the kidney, such as uromodulin, urokinase, and secretory IgA, and contribute to total urine measurements.

Increased amounts of protein in urine, or **proteinuria**, may be the result of several mechanisms including protein overflow from the plasma, increased permeability of the glomeruli due to disease or damage, decreased tubular reabsorption, increased synthesis of renal proteins, and post-renal inflammation (**Table 6.8**). Contamination from other sources such as the vagina and prostate may also contribute to urine total protein measurements.

Qualitative Screening

The reference values or ranges for urinary proteins are highly method dependent, ranging from 100 to 250 mg every 24 hours. Qualitative screening for proteinuria is commonly performed on a random or spot urine sample using a reagent test strip. These methods are based on the change of an indicator dye in the presence of proteins, known as the *protein error of indicators*. These methods are more specific to albumin and may yield a negative result despite the presence of other proteins such as globulins, myoglobin, hemoglobin, and immunoglobulins if they are in low concentrations.

Table 6.8 Reference Ranges for Urine Protein Analyses

TOTAL PROTEIN—URINE, 24-HOUR	
Adult	<150 mg/24 hours
PROTEIN/CREATININE RATIO—URINE, 24-HOUR	
Adult	≤0.114 mg/mg creatinine

Immunochemical-based reagent test strips are also available for microalbumin, which are designed to detect even lower concentrations of albumin than the routine test strips. Testing for microalbumin is useful for monitoring patients at risk for renal disease as the presence of albumin serves as an early indicator of increased permeability of the glomeruli.[82] Refer to Chapter 23, *Body Fluid Analysis*, for additional information on chemical examination of urine using reagent test strips for total protein and microalbumin.

Precipitation Methods

Precipitation methods using sulfosalicylic acid, trichloroacetic acid, or benzethonium chloride have been used to determine total protein concentrations in urine by assessing resultant turbidity. These methods are sensitive, but the reagent does not react equally with each protein fraction. This is particularly true of sulfosalicylic acid, which produces four times more turbidity with albumin than with α-globulins.

Quantitative Testing

Most quantitative assays require timed urine collections of 12 to 24 hours to account for diurnal variation. The total volume of the timed specimen is measured and recorded as results are generally reported in terms of the mass of protein excreted during a 24-hour period. For random or incomplete collections, the creatinine concentration may also be quantitated and the value reported as a ratio of urine total protein to creatinine.

Urine and cerebrospinal fluid have much lower concentrations of total protein compared to serum and plasma and therefore require test methods with greater sensitivity. Several modifications to the biuret reaction have been used to increase test sensitivity including precipitation of the proteins, dissolution of the protein precipitate, and then measurement of the resultant color formation with biuret reagent. Another procedure uses the Folin-Ciocalteu reagent, which is a mixture of phosphomolydate and phosphotungstate, frequently called phenol reagent because it oxidizes phenolic compounds. The reagent changes color from yellow to blue during reaction with tyrosine, tryptophan, and histidine residues in protein. This method is about ten times more sensitive than the biuret method. Lowry et al. increased the sensitivity of the Folin-Ciocalteu reaction by incorporating a biuret reaction as the initial step (Folin-Lowry method). After the cupric ions bind to the peptide bonds, the Folin-Ciocalteu reagent is added. As the Cu^{2+}–protein complex is oxidized, the reagent is reduced, forming the chromogens tungsten blue and molybdenum blue. This increased the sensitivity a hundredfold greater than that of the biuret method alone.

Dye-binding methods using pyrogallol red may be used to determine the quantitative total protein content of urine. Pyrogallol red is combined with molybdate to form a red complex. In the presence of proteins, a blue-purple complex is formed with a maximum absorbance at 600 nm. The resulting increase in absorbance is measured and used to determine the total protein concentration of the specimen. While this method is used to measure urine total proteins, dye-binding assays are not sensitive enough for urine microalbumin testing, and immunochemical assays are still preferred. **Table 6.9** summarizes the methods for measurement of urinary total protein.

Electrophoresis

When an elevated urine total protein concentration is identified, additional testing, such as high-resolution protein electrophoresis or capillary electrophoresis, is needed to characterize the specific protein(s) involved. A first morning urine collection is preferred as it yields a concentrated sample. Random and 24-hour collections may also be acceptable specimens, but might require concentration prior to analysis depending on the test method. Evaluation of the electropherogram may show a predominance of albumin suggestive of glomerular disease, an elevation of low molecular weight proteins in the α_1-, α_2-, or β-fractions associated with overflow proteinuria or renal tubular impairment, or an increase in the γ-region indicating an elevation in intact immunoglobulins and/or free light chains.[83] The primary reason urine protein electrophoresis may be performed is to determine if the patient has

Table 6.9 Urine Protein Methods

Method	Principle	Comment
Turbidimetric methods (sulfosalicylic acid, trichloroacetic acid, or benzethonium chloride)	Urine protein is precipitated as fine particles; turbidity is measured spectrophotometrically	Rapid, easy to use; unequal sensitivity for individual proteins
Biuret	Proteins are concentrated by precipitation, redissolved in alkali, then reacted with Cu^{2+}; Cu^{2+} forms colored complex with peptide bonds	Accurate
Folin-Lowry	Initial biuret reaction; oxidation of tyrosine, tryptophan, and histidine residues by phenol reagent (mixture of phosphotungstic and phosphomolybdic acids); measurement of resultant blue color	Very sensitive
Dye binding (pyrogallol red)	Protein binds to dye, causes shift in absorption maximum	Used in automated methods

a light chain myeloma producing excessive free light chains, otherwise referred to as Bence Jones proteins. As mentioned, most low-molecular-weight proteins filtered at the glomerulus will be reabsorbed in the renal tubules. However, when free light chains are produced in excess, the rate of tubular reabsorption is exceeded, and they are excreted in the urine. Following electrophoresis, immunofixation may be performed to further distinguish if the band in the γ-region is the result of intact immunoglobulins (IgG, IgM, IgA) or from the presence of free light chains (κ, λ). Immunoassays are also now available to measure free light chains in serum, which offers an alternative method for diagnosing and monitoring patients with monoclonal gammopathies.[83,84]

Cerebrospinal Fluid

CSF is formed in the choroid plexus of the ventricles of the brain by ultrafiltration of blood plasma. Total protein and glucose measurements are routinely performed on CSF in addition to a differential cell count, culture, and sensitivity studies. The reference range for CSF total protein for patients greater than 1 month of age is 14 to 45 mg/dL (**Table 6.10**).

An abnormally increased CSF total protein is generally associated with increased permeability of the capillary endothelial barrier through which ultrafiltration occurs. Examples of such conditions include meningitis, traumatic lumbar puncture, multiple sclerosis, obstruction, neoplasm, disk herniation, and cerebral infarction. The degree of permeability of the blood–brain barrier can be evaluated by measuring the CSF albumin and comparing it with simultaneous measurement of serum albumin. Albumin is used as the reference protein for permeability because it is not synthesized to any degree in the central nervous system. The reference range for the CSF-to-serum albumin ratio is 2.7 to 7.3; a value above the reference range may indicate damage to the blood–brain barrier. Low CSF protein values are found in hyperthyroidism and when fluid is leaking from the central nervous system.[85]

The concentration of total protein in CSF may be determined by the same methods referred to earlier in the discussion on urinary proteins. Although total protein concentrations in the CSF are informative,

Table 6.10 Reference Ranges for CSF Protein Analyses

TOTAL PROTEIN—CSF	
0–7 days	40–120 mg/dL
8 days–1 month	20–40 mg/dL
>1 month	14–45 mg/dL
ALBUMIN—CSF	
	0–35 mg/dL
CSF—SERUM ALBUMIN RATIO	
	2.7–7.3
IgG INDEX	
	0.26–0.70

electrophoresis may be performed to characterize individual protein fractions. The pattern of proteins present can be visualized best by using a concentrated CSF specimen. A normal CSF pattern shows prealbumin (transthyretin), a prominent albumin band, an α_1-globulin band composed predominantly of α_1-antitrypsin, an α_2-globulin band consisting primarily of haptoglobin and ceruloplasmin, a β_1 band composed principally of transferrin, and a CSF-specific transferrin that is deficient in carbohydrate, referred to as τ protein, in the β_2 zone. The globulin present in the γ band is typically IgG with a small amount of IgA.[85]

Electrophoretic patterns of CSF from patients with multiple sclerosis demonstrate multiple, distinct oligoclonal bands in the γ region (**Figure 6.11B**). The presence of discrete bands in the γ region of CSF that are not also present in the serum is consistent with production of IgG in the CSF. These bands cannot be seen on routine cellulose acetate electrophoresis, but require a high-resolution technique using agarose. More than 90% of patients with multiple sclerosis demonstrate the presence of oligoclonal bands, although oligoclonal bands have also been found in inflammatory conditions and infectious neurologic diseases, such as Guillain-Barré syndrome, bacterial meningitis, viral encephalitis, subacute sclerosing panencephalitis, and neurosyphilis.[86]

To differentiate elevated CSF IgG concentrations due to local CNS production from leakage of plasma into the CSF, the laboratory may compare CSF and serum IgG concentrations with reference to albumin in a value known as the IgG index. The CSF albumin concentration corrects for increased permeability of the blood–brain barrier. To identify the source of an elevated CSF IgG level, the IgG index can be calculated as follows:

$$\text{IgG Index} = \frac{\left(\text{CSF IgG}\left(\frac{\text{mg}}{\text{dL}}\right) \div \text{Serum IgG}\left(\frac{\text{g}}{\text{dL}}\right)\right)}{\left(\text{CSF albumin}\left(\frac{\text{mg}}{\text{dL}}\right) \div \text{Serum albumin}\left(\frac{\text{g}}{\text{dL}}\right)\right)} \quad \textbf{(Eq. 6.4)}$$

The reference range for the IgG index is 0.26 to 0.70. A high IgG index is indicative of local CNS production of IgG, whereas a low IgG index is suggestive of hypergammaglobulinemia or low serum albumin.[85]

In the investigation of multiple sclerosis, myelin basic proteins present in the CSF may also be assayed because these proteins can provide an index of active demyelination. Myelin basic proteins are constituents of myelin, the sheath that surrounds many of the CNS axons. In very active demyelination, concentrations of myelin basic proteins of 17 to 100 ng/mL are found using an enzyme-linked immunosorbent assay (ELISA). In slow demyelination, values of 6 to 16 ng/mL occur, and in remission, the values are less than 4 ng/mL. In addition to multiple sclerosis, other conditions that induce CNS demyelination and elevated concentrations of myelin basic protein include meningoencephalitis, SLE, diabetes mellitus, and chronic renal failure.[85]

Amyloids are insoluble fibrous protein aggregates formed due to an alteration in their secondary structure known as β-pleated sheets. Amyloidosis refers to conditions in which amyloids are abnormally deposited in organs and tissues, including the heart, blood vessels, brain, peripheral nerves, kidneys, liver, spleen, and intestines causing localized or widespread organ failure. Deposition of amyloid in the brain is

CASE STUDY 6.3, PART 2

Remember Fiona, the 36-year-old woman, was seen by her physician with complaints of intermittent blurred vision. CSF was collected via lumbar puncture and sent to the laboratory along with a paired serum sample. The CSF specimen was clear and colorless with a normal cell count. The CSF total protein concentration was 49 mg/dL with an IgG of 8.1 mg/dL. Electrophoresis of Fiona's serum and CSF revealed more than two oligoclonal bands in the CSF electropherogram (seen in Figure 6.11) and a polyclonal pattern on SPE.

1. What is the significance of the CSF protein bands indicated by the arrows?
2. What conditions would produce this type of CSF protein electrophoresis pattern?
3. What other tests would be helpful in the investigation of this patient's diagnosis?
4. What laboratory test can be useful for monitoring the course of this patient's condition?

associated with the development of Alzheimer's. As such, measurements of cerebrospinal fluid for amyloid β42 (Aβ42) may aid in differentiating Alzheimer's disease from other forms of dementia. CSF concentrations are decreased in patients with Alzheimer's due to deposition in the brain.

Hemoglobin

Hemoglobin is classified as a transport protein. Its role is to transport oxygen from the lungs to the tissues, and transport carbon dioxide back to the lungs. It is one of the major buffering systems in the body. Hemoglobin is produced in red blood cells during their development in the bone marrow. Hemoglobin has a mass of 64,500 daltons, making it similar in mass to albumin. Hemoglobin is a tetramer, having two α-like and two β-like globin chains. Each chain holds one heme molecule, which reversibly binds oxygen, dependent on the PO_2 of the surrounding tissue. The major secondary structure of the globin chains is α helix. The helices surround the heme to form the tertiary structure. The globin chains first form homodimers, then come together as a tetramer to form the quaternary structure (**Figure 6.13A**).

Different globin chains are expressed over the course of development from separate genes on chromosomes 11 and 16. The relative proportion of the different chains varies in embryos, fetuses, newborns, and adults. The major hemoglobin in fetuses is HbF, which consists of two α and two γ chains. In adults the major hemoglobin is HbA, which consists of two α and two β chains. HbA accounts for up to 97% of the total hemoglobin. The remainder is comprised of HbA_2 and a small amount of HbF. HbA_2 consists of two α and δ chains.

Hemoglobin concentration is measured in peripheral blood as part of a complete blood count, along with a red blood cell (RBC) count. A decreased hemoglobin concentration is indicative of anemia. *Hemoglobinopathies* are qualitative defects caused by mutations in the DNA of the globin chain that produce a structurally different form leading to a decrease in red blood cell survival. For example, HbS is a mutated form of HbA, causing sickle cell anemia. *Thalassemias* are quantitative defects due to mutations that reduce the synthesis of normal hemoglobin. This chapter will focus on discussion of disorders in the production of heme classified as *porphyrias*.

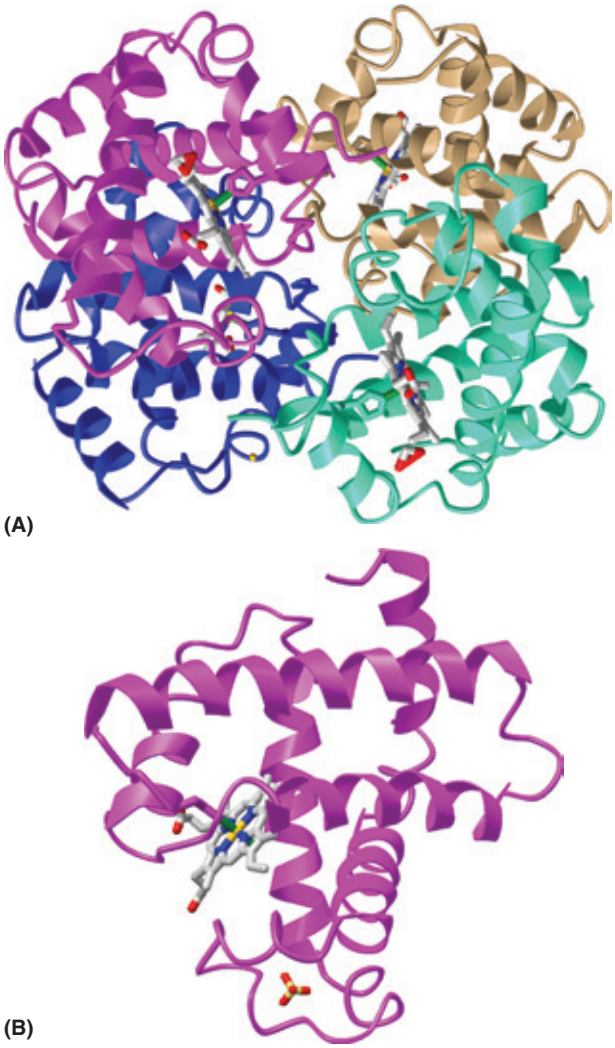


Figure 6.13 Molecular structure of hemoglobin compared to myoglobin. (**A**) Hemoglobin. The four different colors are four proteins that have formed a tetramer. (**B**) Myoglobin. Myoglobin is a monomer. The secondary structure is predominantly α helix.

Wang J, Youkharibache P, Zhang D, Lanczycki CJ, Geer RC, Madej T, Phan L, Ward M, Lu S, Marchler GH, Wang Y, Bryant SH, Geer LY, Marchler-Bauer A. iCn3D, a Web-based 3D Viewer for Sharing 1D/2D/3D Representations of Biomolecular Structures. *Bioinformatics*. 2020 Jan 1;36(1):131-135. (Epub 2019 June 20.) doi: 10.1093/bioinformatics/btz502

Synthesis and Degradation of Hemoglobin

Hemoglobin synthesis occurs in the immature RBCs located in the bone marrow. Newly synthesized heme exits the mitochondria and complexes with four globin molecules in the cell cytoplasm. Iron is transported from storage sites to the developing RBCs, where it is inserted into heme and is available to bind oxygen.

Red blood cells may lyse in the blood vessels (intravascular hemolysis) or be degraded outside of the circulatory system within the phagocytic cells of the spleen, liver, and bone marrow (extravascular

hemolysis) releasing hemoglobin. In intravascular hemolysis, circulating haptoglobin and hemopexin will bind with free hemoglobin and free heme, respectively, and transport them to the liver so their constituent parts can be recycled. Iron is reclaimed and stored until it is needed (see Chapter 27, *Trace Elements, Toxic Elements, and Vitamins*), and the globin chains are broken down so the amino acids can be recycled. Heme is converted to bilirubin and urobilinogen through several steps (see Chapter 19, *Liver Function*), and laboratory measurement of these degradation products can help determine the underlying cause for increased RBC destruction.

Myoglobin

Myoglobin is the primary oxygen-carrying protein found in striated skeletal and cardiac muscle, accounting for approximately 2% of total muscle protein. It can reversibly bind oxygen and requires a very low oxygen tension to release the bound oxygen. Myoglobin transports oxygen from the muscle cell membrane to the mitochondria, and it serves as an extra reserve of oxygen to help exercising muscle maintain activity longer. Myoglobin is a monomer, a single-chain globular protein with a mass of 16,700 daltons. Its secondary and tertiary structure is similar to hemoglobin. Much of the secondary structure is α helix, surrounding a heme prosthetic group, shown in **Figure 6.13B**.

When striated muscle is damaged, myoglobin is released into the bloodstream resulting in elevated blood and urine concentrations. Elevations are found shortly after an acute myocardial infarction (see Chapter 20, *Cardiac Function*). Elevations are also seen in conditions in which skeletal muscle is damaged, such as progressive muscular dystrophy,[87] crush injuries, and rhabdomyolysis.[88] Myoglobin is a nephrotoxin, and in severe muscle injury, concentrations of myoglobin can rise very quickly causing damage to the kidneys. Renal failure can also elevate blood concentrations of serum myoglobin due to impaired filtration at the glomerulus.[89] **Table 6.11** lists some causes of serum myoglobin elevation. Myoglobin levels in urine are normally very low or not detected. Increased levels of myoglobin in urine will make the urine appear reddish-brown. Myoglobin will cross-react with the hemoglobin pad on the urine dipstick and cause a positive reaction. Myoglobin is measured in urine and blood by immunometric methods.

Table 6.11 Causes of Myoglobin Elevations

Acute myocardial infarction	Angina without infarction
Rhabdomyolysis	Multiple fractures; muscle trauma
Renal failure	Myopathies
Vigorous exercise	Intramuscular injections
Open heart surgery	Tonic–clonic seizures
Electric shock	Arterial thrombosis
Certain toxins	Malignant hyperthermia
Muscular dystrophy	Systemic lupus erythematosus

Heme

Heme is not a protein; it is a prosthetic group found in hemoglobin, myoglobin, chlorophyll, cytochromes, and several other enzymes. Heme is a flat cyclic tetrapyrrole, with a central ferrous iron atom (Fe^{2+}) that reversibly binds oxygen (**Figure 6.14**).

Heme is produced by a series of eight enzymatic reactions within the mitochondria and cytosol of all cells. It is most concentrated in RBCs and liver cells. Intermediates in the enzymatic pathway are collectively referred to as porphyrins. A mutation in any of

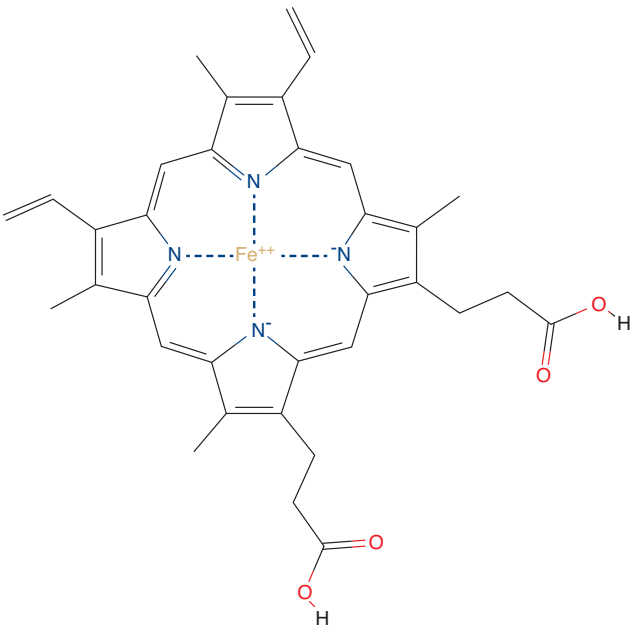


Figure 6.14 Structure of heme.

the enzymes in this pathway leads to an accumulation of specific porphyrins. The symptoms of porphyrias that manifest are related to which porphyrins are elevated. Therefore, porphyrias are classified as either erythropoietic or hepatic.[90]

Porphyrins are photoactive due to extensive conjugation of the tetrapyrrole ring, which is highlighted in the diagram of heme shown in Figure 6.14. Their dark red color is due to a strong absorbance in the visible region of the spectrum. The compounds absorb light of 400 nm wavelength and emit a characteristic orange-red fluorescence between 600 and 650 nm.[91] This property produces some of the manifestations of disease, such as photosensitivity, and provides a means for detection of the compounds in body fluids tested in the laboratory. Aqueous solubility of porphyrins varies with the number of carboxylic acid substituents present in the porphyrin compound. Because of the variable solubilities of the porphyrins, selection of an appropriate sample type for analysis is essential. Uroporphyrin (Uro) has eight carboxylic acid groups. It is most soluble in water and is excreted by the kidneys. Coproporphyrin (Copro) has four carboxylic acid substituents and intermediate solubility and is found in blood, urine, and feces. Protoorphyrins (Proto) with only two carboxyl groups are the least water soluble and are not found in urine. They are present in the blood and are excreted in the feces.

Synthesis of Heme

Heme biosynthesis is diagrammed in **Figure 6.15**. In the first step in the pathway, glycine and succinyl coenzyme A are condensed to form δ-aminolevulinic acid (ALA) in the mitochondrion. In the next reaction, the porphobilinogen synthase catalyzes condensation of two molecules of ALA to form porphobilinogen (PBG). Polymerization of four molecules of PBG creates the linear tetrapyrrole, hydroxymethylbilane (HMB). This process is catalyzed by the enzyme hydroxymethylbilane synthase, or PBG deaminase. Uroporphyrinogen III synthase catalyzes intramolecular rearrangement and ring closure resulting in the cyclic isomer, uroporphyrinogen III. Sequential removal of one carboxyl group from each **pyrrole** ring in the cyclic tetrapyrrole is catalyzed by uroporphyrinogen decarboxylase. The hepta-, hexa-, penta-, and tetra-carboxyl intermediates are formed in succession. In subsequent reactions, coproporphyrinogen oxidase catalyzes the oxidation of two additional carboxylic acid substituents to form protoporphyrinogen IX, which is oxidized to protoporphyrin in a reaction catalyzed by protoporphyrinogen oxidase. In the final step of biosynthesis, ferrochelatase catalyzes the insertion of ferrous iron into the porphyrin ring to form heme. The process of heme biosynthesis is regulated by negative feedback of heme itself.

Disorders of Heme Biosynthesis

The porphyrias are a group of rare inherited or acquired metabolic disorders caused by loss or gain of function mutations in the enzymes responsible for heme biosynthesis. The defective enzyme results in overproduction, accumulation, and excretion of toxic precursor compounds and porphyrins. Conditions corresponding to enzyme abnormalities have been identified in every step of heme synthesis. **Table 6.12** lists the porphyrias, their associated enzyme deficiencies,

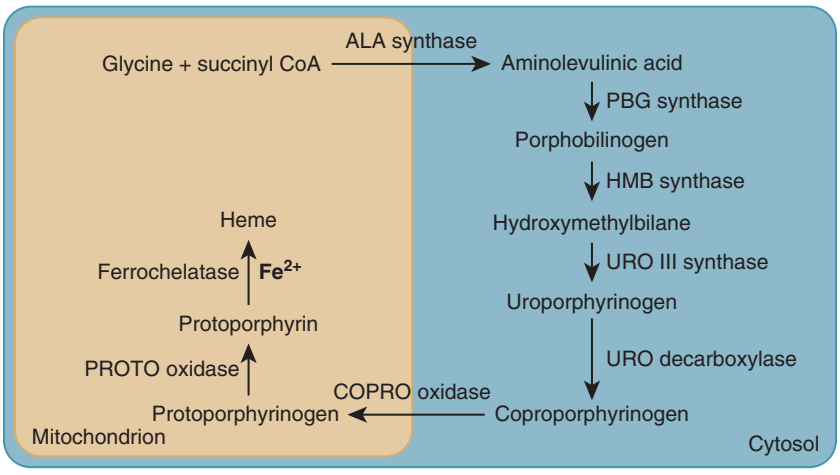


Figure 6.15 Heme biosynthesis.

the intermediates that accumulate, and the body fluids that might be collected for analysis.

The porphyrias can be categorized as either *erythropoietic* (CP and EPP) or *hepatic* (AIP, ADP, VP, HCP, PCT, and HEP). Porphyrias may also be classified based on their associated symptoms as *acute* or *cutaneous* (non-acute). Acute porphyrias, which include ADP, AIP, HCP, and VP, are associated with neurological symptoms and acute attacks of abdominal pain. The cutaneous porphyrias, CEP, PCT, EPP, and XLPP, present as chronic conditions. Exposure to sunlight results in photosensitization of the skin due to accumulation of excess porphyrins in tissues. The porphyrins absorb light and react, producing blistering skin lesions on sun-exposed areas that are characteristic of cutaneous porphyrias. Chronic photosensitivity is also a symptom in VP and HCP. Many of the porphyrias have an autosomal dominant inheritance pattern, meaning that a mutation of one gene causes disease, which leads to a 50% reduction in activity of the affected enzyme. Exceptions are X-linked protoporphyria (XLPP) and CEP, both are autosomal recessive. Some individuals demonstrate an enzyme deficiency but do not experience clinical or biochemical manifestations of porphyria indicating that other factors, such as an increased demand for heme biosynthesis, must be present to cause disease expression.[90]

Other Conditions Associated with Increased Porphyrins

Secondary porphyrinurias are acquired conditions associated with increased excretion of urinary porphyrins.[92] These disorders are not the result of an inherited biochemical defect in heme synthesis, but are secondary to another disorder or toxin that interferes with heme formation or metabolism. Impaired heme synthesis may result from functional changes due to liver diseases such as hepatitis, cirrhosis, hereditary tyrosinemia, and obstructive jaundice.[93,94] Urinary porphyrins are also increased in inherited disorders of bilirubin metabolism such as Dubin Johnson, Rotor, and Gilbert syndromes.[95] Heavy metal poisoning, such as lead toxicity, can also lead to excess urinary porphyrin excretion in the absence of porphyria.

Zinc Protoporphyrin

When iron stores are insufficient to meet the demands of heme synthesis, erythrocyte protoporphyrin can assimilate zinc instead of iron to

Table 6.12 Heme Biosynthesis: Metabolites and Diseases

		Accumulated Substrates		
Porphyria	**Associated Enzyme**	**Precursors**	**Porphyrins**	**Affected Body Fluids**
X-linked protoporphyria (XLPP)	ALA synthase	—	Proto	Blood, feces
ALA dehydratase deficiency porphyria (ADP or ALDP)	PBG synthase	ALA	Copro III	Urine
Acute intermittent porphyria (AIP)	Hydroxymethylbilane synthase	PBG	Uro	Urine
Porphyria cutanea tarda (PCT)	Uroporphyrinogen decarboxylase	—	Uro, Hepta, Hexa, Penta	Urine, plasma, feces
Hereditary coproporphyria (HCP)	Coproporphyrinogen oxidase	PBG	Copro III	Urine, plasma, feces
Variegate porphyria (VP)	Protoporphyrinogen oxidase	PBG	Proto, Copro III	Blood, plasma, feces
Congenital erythropoietic porphyria (CEP)	Uroporphyrinogen III synthase	—	Uro I, Copro I	Urine, blood, plasma, feces
Erythropoietic protoporphyria (EPP)	Ferrochelatase	—	Proto	Blood, plasma, feces

ADP, ALA dehydratase deficiency porphyria; AIP, acute intermittent porphyria; ALA, δ-aminolevulinic acid; CEP, congenital erythropoietic porphyria; Copro, coproporphyrin(ogen); EPP, erythropoietic protoporphyria; HCP, hereditary coproporphyria; Hepta, heptacarboxyl porphyrin(ogen); Hexa, hexacarboxyl porphyrin(ogen); PBG, porphobilinogen; PCT, porphyria cutanea tarda; Penta, pentacarboxyl porphyrin(ogen); Proto, protoporphyrin(ogen); Uro, uroporphyrin(ogen); VP, variegate porphyria; XLPP, X-linked protoporphyria.

become zinc protoporphyrin (ZPP). As such, ZPP has been suggested as a biomarker to screen for iron deficiency in the absence of lead toxicity.[96,97] Lead inhibits enzymatic formation of PBG, which increases concentrations of ALA and coproporphyrinogen oxidase function leading to an accumulation of Copro III. Lead also interferes with insertion of ferrous iron into heme, causing ZPP to be formed by incorporation of zinc ions into the protoporphyrin ring instead. Some symptoms of lead intoxication are similar to those found in acute porphyria, so identification of the underlying cause is essential. If lead poisoning is suspected, measurement of whole blood lead concentrations is reommended.[98]

Clinical Application

Testing for porphyrins and porphyrin precursors is used to diagnose and monitor porphyrias and other disorders affecting heme metabolism. Methods available for identification of porphyrias include genetic assays to detect DNA mutations, functional and quantitative assays for the enzymes responsible for heme biosynthesis, and biochemical assays to identify and quantify porphyrins and precursor compounds. Clinical diagnostic tests for porphyrin disorders are based most often on identification of characteristic patterns of excess porphyrin precursors and intermediate compounds in body fluids (Table 6.11). The specimen to be tested (urine, blood, or feces) depends on the solubility of the metabolic intermediates associated with particular disease symptoms.

Testing for Porphyrin Disorders

When an acute porphyria is suspected, initial testing will likely include measurement of a random or timed urine collection for PBG. If the symptoms are due to an acute porphyria, increased PBG concentrations will be detected in the urine specimen. The urine may actually have a deep red color, referred to as port wine.[99] Follow-up testing for a positive PBG test may include measurement of the deficient enzyme, hydroxymethylbilane synthase (PBG deaminase) to identify AIP, and analysis of fecal porphyrins to differentiate among the acute porphyrias (AIP, VP, and HCP). PBG may not be increased in asymptomatic individuals; therefore negative results do not exclude porphyria completely, and testing should be repeated on a 24-hour urine specimen or when acute symptoms recur.

Initial testing for cutaneous disease requires measurement of porphyrins and PBG in a random or timed urine collection. The suspicion for CEP in infants may begin when the caregiver notices red urine in the diaper.[100] Porphyrins can also be detected on a wet diaper by characteristic fluorescence of the urine when exposed to long wave UV light.[90] Increased excretion of porphyrins in a characteristic pattern particular for each disease is diagnostic for PCT and CEP. Testing for PBG is negative in these conditions. Suspected EPP or XLPP is evaluated by assessing erythrocyte porphyrins in whole blood.[101]

Porphyrin analysis is typically limited to large reference laboratories or "porphyrin centers." Important considerations for porphyrin analyses include appropriate selection of which test to perform and careful attention to specimen collection, storage, and transport requirements.

Analytical Methods

Testing for the precursor compounds, PBG and ALA, utilizes ion exchange chromatography to remove interferences and isolate each compound. Detection is accomplished by the addition of 4-dimethylaminobenzaldehyde (Ehrlich's reagent) to produce a characteristic rose-red color that is detected spectrophotometrically. The principle is that of the Watson-Schwartz assay.[102] The absorbance spectrum has a maximum at approximately 555 nm and a shoulder at 525 nm. Current assays allow quantitative measurement of analyte concentrations.

Porphyrin intermediates may be evaluated using HPLC with fluorescence detection.[92] Urine is acidified to enhance the natural fluorescence of the porphyrins, and an aliquot is injected into the HPLC system. Measured porphyrin concentrations can be normalized for random urine collections by expressing them as a ratio to creatinine. Results for 24-hour collections are expressed as excretion per day.

Fecal porphyrin analysis using HPLC with fluorescence detection allows separation and quantification of coproporphyrin isomers and protoporphyrin to differentiate the acute porphyrias: AIP, HCP, and VP. An aliquot of the fecal specimen is lyophilized and reconstituted in acidic solution to enhance the natural porphyrin fluorescence before it is subjected to chromatographic analysis. Testing for protoporphyrin to determine presence of EPP or XLPP is performed by scanning fluorescence of a plasma specimen and by analysis of whole blood.

CASE STUDY 6.4, PART 2

Remember Mateo, the 2-month-old infant with blisters. His mother explains that their current visit is because Mateo developed multiple blisters on his exposed skin when they had only been outside for a short period of time, and his urine-soaked diaper appears reddish in color. The pediatrician checks his current diaper with long-wave UV light and notes fluorescence.

© Jose Luis Pelaez Inc/DigitalVision/Getty Images.

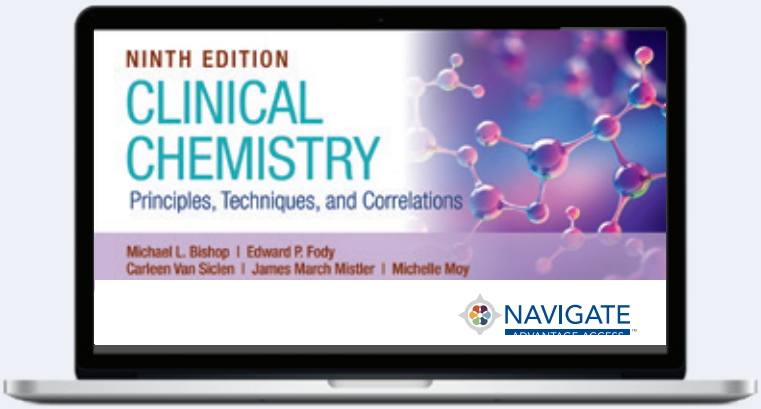
1. What porphyria should be suspected?
2. How is this porphyria classified?
3. What is the first analyte that should be measured, and in which body fluid?
4. Mateo's mother was instructed to collect a urine and a stool sample and bring them to the laboratory for analysis for Uro I and Copro I. A blood sample was also collected and sent at the same time. Only the blood sample was positive. What preanalytical error could lead to this false result?

Specimen Requirements

Specimens must be protected from light to prevent degradation. Porphyrins and precursor compounds are stable in unpreserved urine at 4°C for up to 48 hours and can be frozen (–20°C) for several weeks. Dilute urine (creatinine <25 mg/dL) is not adequate for analysis. Porphyrins in fecal specimens must be stored frozen (–20°C). Whole blood collected for protoporphyrin measurement should be anticoagulated with EDTA and stored in the dark at 4°C.[101]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Shen F, Sergi C. Biochemistry, amino acid synthesis and degradation. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
2. Lopez MJ, Mohiuddin SS. Biochemistry, essential amino acids. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
3. Morris AL, Mohiuddin SS. Biochemistry, nutrients. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
4. LaPelusa A, Kaushik R. Physiology, proteins. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
5. National Center for Biotechnology Information. PubChem Compound Summary for CID 6274, Histidine. https://pubchem.ncbi.nlm.nih.gov/compound/Histidine. Accessed March 21, 2021.
6. National Center for Biotechnology Information. PubChem Compound Summary for CID 6306, l-Isoleucine. https://pubchem.ncbi.nlm.nih.gov/compound/l-Isoleucine. Accessed March 21, 2021.
7. National Center for Biotechnology Information. PubChem Compound Summary for CID 6106, Leucine. https://pubchem.ncbi.nlm.nih.gov/compound/Leucine. Accessed March 21, 2021.
8. National Center for Biotechnology Information. PubChem Compound Summary for CID 6287, Valine. https://pubchem

.ncbi.nlm.nih.gov/compound/Valine. Accessed March 21, 2021.
9. National Center for Biotechnology Information. PubChem Compound Summary for CID 5962, Lysine. https://pubchem.ncbi.nlm.nih.gov/compound/Lysine. Accessed March 21, 2021.
10. National Center for Biotechnology Information. PubChem Compound Summary for CID 876, DL-Methionine. https://pubchem.ncbi.nlm.nih.gov/compound/DL-Methionine. Accessed March 21, 2021.
11. National Center for Biotechnology Information. PubChem Compound Summary for CID 6140, Phenylalanine. https://pubchem.ncbi.nlm.nih.gov/compound/Phenylalanine. Accessed March 21, 2021.
12. National Center for Biotechnology Information. PubChem Compound Summary for CID 6288, L-Threonine. https://pubchem.ncbi.nlm.nih.gov/compound/L-Threonine. Accessed March 21, 2021.
13. National Center for Biotechnology Information. PubChem Compound Summary for CID 6305, Tryptophan. https://pubchem.ncbi.nlm.nih.gov/compound/Tryptophan. Accessed March 21, 2021.
14. National Center for Biotechnology Information. PubChem Compound Summary for CID 5950, Alanine. https://pubchem.ncbi.nlm.nih.gov/compound/Alanine. Accessed March 21, 2021.
15. National Center for Biotechnology Information. PubChem Compound Summary for CID 6267, Asparagine. https://pubchem.ncbi.nlm.nih.gov/compound/Asparagine. Accessed March 21, 2021.
16. National Center for Biotechnology Information. PubChem Compound Summary for CID 5960, Aspartic acid. https://pubchem.ncbi.nlm.nih.gov/compound/Aspartic-acid. Accessed March 21, 2021.
17. National Center for Biotechnology Information. PubChem Compound Summary for CID 33032, Glutamic acid. https://pubchem.ncbi.nlm.nih.gov/compound/Glutamic-acid. Accessed March 21, 2021.
18. National Center for Biotechnology Information. PubChem Compound Summary for CID 6326983, Selenocysteine. https://pubchem.ncbi.nlm.nih.gov/compound/Selenocysteine. Accessed March 21, 2021.
19. National Center for Biotechnology Information. PubChem Compound Summary for CID 5951, Serine. https://pubchem.ncbi.nlm.nih.gov/compound/Serine. Accessed March 21, 2021.
20. National Center for Biotechnology Information. PubChem Compound Summary for CID 6322, Arginine. https://pubchem.ncbi.nlm.nih.gov/compound/Arginine. Accessed March 21, 2021.
21. National Center for Biotechnology Information. PubChem Compound Summary for CID 594, Cysteine. https://pubchem.ncbi.nlm.nih.gov/compound/Cysteine. Accessed March 21, 2021.
22. National Center for Biotechnology Information. PubChem Compound Summary for CID 5961, Glutamine. https://pubchem.ncbi.nlm.nih.gov/compound/Glutamine. Accessed March 21, 2021.
23. National Center for Biotechnology Information. PubChem Compound Summary for CID 145742, Proline. https://pubchem.ncbi.nlm.nih.gov/compound/Proline. Accessed March 21, 2021.
24. National Center for Biotechnology Information. PubChem Compound Summary for CID 6057, Tyrosine. https://pubchem.ncbi.nlm.nih.gov/compound/Tyrosine. Accessed March 21, 2021.
25. Jeanmonod R, Asuka E, Jeanmonod D. Inborn errors of metabolism. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
26. Amino Acids Quantitative by LC-MS/MS, Plasma. ARUP Lab Test Directory. https://ltd.aruplab.com/Tests/Pub/2009389. Published November 18, 2019. Accessed March 21, 2021.
27. Amino Acids Quantitative by LC-MS/MS, Urine. ARUP Lab Test Directory. https://ltd.aruplab.com/Tests/Pub/2009419. Published November 18, 2019. Accessed March 21, 2021.
28. Amino Acids Quantitative by LC-MS/MS, CSF. ARUP Lab Test Directory. https://ltd.aruplab.com/Tests/Pub/0080137. Published November 18, 2019. Accessed March 21, 2021.
29. Bijarnia-Mahay S, Kapoor S. Testing modalities for inborn errors of metabolism — what a clinician needs to know? *Indian Pediatr.* 2019;56(9):757–766. doi:10.1007/s13312-019-1637-9
30. Sanvictores T, Farci F. Biochemistry, primary protein structure. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
31. Karamyshev AL, Tikhonova EB, Karamysheva ZN. Translational control of secretory proteins in health and disease. *Int. J. Mol. Sci.* 2020; 21(7):2538. https://doi.org/10.3390/ijms21072538
32. Gurina TS, Mohiuddin SS. Biochemistry, protein catabolism. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
33. Umpleby AM, Russell-Jones DL. The hormonal control of protein metabolism. *Best Pract Res Clin Endocrinol Metab.* 1996;10(4):551–570. doi:10.1016/s0950-351x(96)80711-7
34. Dag Z, Koseoglu H, Kekilli M. The use of prealbumin as a predictor of malnutrition in cirrhotic patients and the effect of nutritional support in patients with low prealbumin levels. *Turk J Med Sci.* 2020;44(2):398–404. doi:10.3906/sag-1910-27
35. Moman RN, Gupta N, Varacallo M. Physiology, albumin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; September 22, 2020.
36. Gounden V, Vashisht R, Jialal I. Hypoalbuminemia. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
37. Prasad RM, Tikaria R. Microalbuminuria. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
38. Goyal R, Jialal I. Diabetes mellitus type 2. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
39. Gounden V, Ngu M, Anastasopoulou C, Jialal I. Fructosamine. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
40. Meseeha M, Attia M. Alpha 1 antitrypsin deficiency. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
41. Mizejewski GJ. Alpha-fetoprotein (AFP) and inflammation: Is AFP an acute and/or chronic phase reactant? *J Hematol Thrombo Dis*. 2015;03(01). doi:10.4172/2329-8790.1000191
42. Mizejewski GJ. Biological roles of alpha-fetoprotein during pregnancy and perinatal development. *Exp Biol Med.* 2004;229(6):439–463. doi:10.1177/153537020422900602
43. Luo Z, Lei H, Sun Y, Liu X, Su D-F. Orosomucoid, an acute response protein with multiple modulating activities. *J Physiol Biochem*. 2015;71(2):329–340. doi:10.1007/s13105-015-0389-9

44. Shih AWY, McFarlane A, Verhovsek M. Haptoglobin testing in hemolysis: measurement and interpretation. *Am J Hematol*. 2013;89(4):443–447. doi:10.1002/ajh.23623
45. McMillin GA, Travis JJ, Hunt JW. Direct measurement of free copper in serum or plasma ultrafiltrate. *Am J Clin Pathol*. 2009;131(2):160–165. doi:10.1309/ajcp7z9kbfinvgyf
46. Lopez MJ, Royer A, Shah NJ. Biochemistry, ceruloplasmin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
47. Cater JH, Wilson MR, Wyatt AR. Alpha-2-macroglobulin, a hypochlorite-regulated chaperone and immune system modulator. *Oxid Med Cell Longev*. 2019;2019:1–9. doi:10.1155/2019/5410657
48. Ogun AS, Adeyinka A. Biochemistry, transferrin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
49. Vanacore R, Eskew JD, Sung L, Davis T, Smith A. Safe coordinated trafficking of heme and iron with copper maintain cell homeostasis: modules from the hemopexin system. *BioMetals*. 2019;32(3):355–367. doi:10.1007/s10534-019-00194-4
50. Lent-Schochet D, Jialal I. Biochemistry, lipoprotein metabolism. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
51. Li L, Dong M, Wang X-G. The implication and significance of beta 2 microglobulin. *Chin Med J (Engl)*. 2016;129(4):448–455. doi:10.4103/0366-6999.176084
52. Bardhan M, Kaushik R. Physiology, Complement Cascade. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
53. Kaur J, Jain A. Fibrinogen. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
54. Nehring SM, Goyal A, Bansal P, Patel BC. C reactive protein. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
55. Justiz Vaillant AA, Jamal Z, Ramphul K. Immunoglobulin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
56. Aziz M, Iheanacho F, Hashmi MF. Physiology, antibody. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
57. Rawla P, Joseph N. IgA deficiency. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
58. Sathe A, Cusick JK. Biochemistry, Immunoglobulin M. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
59. Godwin L, Sinawe H, Crane JS. Biochemistry, immunoglobulin E. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
60. Stark M, Kerndt CC, Sharma S. Troponin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
61. Novack ML, Zevitz ME. Natriuretic peptide B type test. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
62. Maurer LM, Ma W, Mosher DF. Dynamic structure of plasma fibronectin. *Crit Rev Biochem Mol*. 2016;51(4):213–227. doi:10.1080/10409238.2016.1184224
63. Ramakrishnan N, Auger K, Jialal I. Biochemistry, adiponectin. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
64. Williams C, Sapra A. Osteoporosis markers. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
65. Gounden V, Bhatt H, Jialal I. Renal function tests. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
66. Bertholf RL. Proteins and albumin. *Lab Med*. 2014;45(1):25–41. doi:10.1309/lmkrnrgw5j03apzq
67. Kaseb H, Gonzalez-Mosquera LF, Parsi M, Mewawalla P. Lymphoplasmacytic lymphoma. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
68. Albagoush SA, Azevedo AM. Multiple myeloma. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
69. Gulhar R, Ashraf MA, Jialal I. Physiology, acute phase reactants. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
70. Protein, Total, Serum or Plasma. ARUP Lab Test Directory. https://ltd.aruplab.com/Tests/Pub/0020029. Accessed March 21, 2021.
71. Becklin RR, Desiderio DM. The amount of ultraviolet absorbance in a synthetic peptide is directly proportional to its number of peptide bonds. *Anal Lett*. 1995;28(12):2175–2190. doi:10.1080/00032719508000037
72. Buckingham L. *Molecular Diagnostics*. 2nd ed. Philadelphia, PA: F.A. Davis, 2012;822.
73. Biter AB, Pollet J, Chen WH, Strych U, Hotez PJ, Bottazzi ME. A method to probe protein structure from UV absorbance spectra. *Anal Biochem*. 2019;587:113450. doi:10.1016/j.ab.2019.113450
74. Garcia Moreira V, Beridze Vaktangova N, Martinez Gago MD, Laborda Gonzalez B, Garcia Alonso S, Fernandez Rodriguez E. Overestimation of albumin measured by bromocresol green vs bromocresol purple method: influence of acute-phase globulins. *Lab Med*. 2018;49(4):355–361. doi:10.1093/labmed/lmy020
75. Tapia C, Bashir K. Nephrotic syndrome. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
76. Meseeha M, Attia M. Alpha 1 antitrypsin deficiency. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
77. Huq M, Bhatnagar NK, Hostoffer RW. Hypogammaglobulinemia. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
78. Mazhar M, Waseem M. Agammaglobulinemia. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2021.
79. Zhao EJ, Carruthers MN, Li CH, Mattman A, Chen LY. Conditions associated with polyclonal hypergammaglobulinemia in the IgG4-related disease era: a retrospective study from a hematology tertiary care center. *Haematologica*. 2020;105(3). doi:10.3324/haematol.2019.219725
80. Sharma B, John S. Hepatic cirrhosis. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
81. Keren D. High resolution electrophoresis aids detection of gammopathies. *Clin Chem News*. 1989;14:26–29.
82. Brunzel NA. Chemical examination of urine. In: *Fundamentals of Urine and Body Fluid Analysis*. 4th Edition. St. Louis, MI: Elsevier; 2018:91–94.
83. Ramakrishnan N, Jialal I. Bence-Jones Protein. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; 2020.
84. Jenkins MA. Serum and urine electrophoresis for detection and identification of monoclonal proteins. *Clin Biochem Rev*. 2009;30(3):119–122.
85. Strasinger SK, Schaub DLM. Cerebrospinal fluid. In: *Urinalysis and Body Fluids*. 7th Edition. Philadelphia, PA: F.A. Davis Company; 2021:266–268.
86. Link H, Huang YM. Oligoclonal bands in multiple sclerosis cerebrospinal fluid: an update on methodology and clinical usefulness. *J Neuroimmunol*. 2006;180(1-2):17–28. doi:10.1016/j.jneuroim.2006.07.006
87. Rouillon J, Lefebvre T, Denard J, et al. High urinary ferritin reflects myoglobin iron evacuation in DMD patients. *Neuromuscul Disord*. 2018 Jul;28(7):564–571. doi: 10.1016/j.nmd.2018.03.008

88. Chen CY, Lin YR, Zhao LL, Yang WC, Chang YJ, Wu HP. Clinical factors in predicting acute renal failure caused by rhabdomyolysis in the ED. *Am J Emerg Med*. 2013 Jul;31(7):1062–1066. doi: 10.1016/j.ajem.2013.03.047
89. Wu AH, Laios I, Green S, et al. Immunoassays for serum and urine myoglobin; myoglobin clearance assessed as a risk factor for acute renal failure. *Clin Chem*. 1994;40(5):796–802.
90. Phillips JD. Heme biosynthesis and the porphyrias. *Mol Genet Metab*. 2019;128(3):164–177. doi:10.1016/j.ymgme.1092.04.008
91. Goodwin RG, Kell WJ, Laidler P, et al. Photosensitivity and acute liver injury in myeloproliferative disorder secondary to late-onset protoporphyria caused by deletion of a ferrochelatase gene in hematopoietic cells. *Blood*. 2006;107:60–62.
92. Badminton MN, Whatley SD, Aarsand AK. Porphyrias and porphyrias. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 8th Edition. St. Louis, MI: Elsevier; 2019.
93. Kaczynski J, Hansson G, Wallerstedt S. Increased porphyrins in primary liver cancer mainly reflect a parallel liver disease. *Gastroenterol Res Pract*. 2009;2009:1–6.
94. Bissell DM, Lai JC, Meister RK, Blanc PD. Role of delta-aminolevulinic acid in the symptoms of acute porphyria. *Am J Med*. 2015;128(3):313–317. doi: 10.1016/j.amjmed.2014.10.026
95. Chowdhury JR, Wolkoff AW, Chowdhury NR, Arias IM. Hereditary jaundice and disorders of bilirubin metabolism. In: Valle D, Beaudet AL, Vogelstein B, et al., eds. *The Online Metabolic and Molecular Bases of Inherited Disease*. New York: McGraw-Hill Education. 2019. http://ommbid.mhmedical.com/. Accessed February 2021.
96. Magge H, Sprinz P, Adams WG, Drainoni ML, Meyers A. Zinc protoporphyrin and iron deficiency screening: trends and therapeutic response in an urban pediatric center. *JAMA Pediatr.* 2013;167(4):361–367. doi: 10.1001/jamapediatrics.2013.751
97. Füllenbach C, Stein P, Glaser P, et al. Screening for iron deficiency in surgical patients based on noninvasive zinc protoporphyrin measurements. *Transfusion*. 2020 Jan;60(1):62–72. doi: 10.1111/trf.15577
98. Childhood Lead Poisoning Prevention. CDC Centers for Disease Control and Prevention. https://www.cdc.gov/nceh/lead/lab/default.htm. Accessed February 2021.
99. Waldman SD. Acute intermittent porphyria. In: *Atlas of Uncommon Pain Syndromes*. 3rd Edition. Elsevier; 2014:205–206.
100. Woolf J, Marsden JT, Degg T, et al. Best practice guidelines on first-line laboratory testing for porphyria. *Ann Clin Biochem*. 2017;54(2):188–198. doi: 10.1177/0004563216667965
101. Porphyrias. ARUP Laboratories Consult. https://arupconsult.com/content/porphyrias. Accessed February 2021.
102. Watson CJ, Schwartz S. A simple test for urinary porphobilinogen. *Exp Biol Med (Maywood)*. 1941;47:393–394.

Nonprotein Nitrogen Compounds

Shawn H.W. Luby

CHAPTER OUTLINE

KEY TERMS

Ammonia
Azotemia
Coupled enzymatic method
Creatine
Creatinine
Creatinine clearance
Estimated glomerular filtration rate
Glomerular filtration rate
Gout
Hyperammonemia
Hyperuricemia
Hypouricemia
Postrenal
Prerenal
Protein-free filtrate
Reabsorption
Secretion
Urea
Urea nitrogen/creatinine ratio
Uremia or uremic syndrome
Uric acid

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- List the nonprotein nitrogen components of the blood, list their relative physiologic concentrations, and recognize their chemical structures.
- Describe the biosynthesis and excretion of urea, uric acid, creatinine, creatine, and ammonia.
- State the specimen collection, transport, and storage requirements necessary for determinations of urea, uric acid, creatinine, creatine, and ammonia.
- Discuss the methodology for the determination of urea, uric acid, creatinine, creatine, and ammonia in plasma and urine. Identify sources of error and variability in these methods and describe the effects on the clinical utility of the laboratory measurements.

- State the reference ranges for urea, uric acid, creatinine, and ammonia in plasma and urine. State the effects of age and gender on these values.
- Perform calculations to convert laboratory results between systems of measurement discussed in this chapter.
- Describe the major pathological conditions associated with increased and decreased plasma concentrations of urea, uric acid, creatinine, creatine, and ammonia.
- Describe the use of the urea nitrogen/creatinine ratio to distinguish prerenal, renal, and postrenal causes of uremia.
- Relate the solubility of uric acid to the pathologic consequences of increased plasma uric acid.
- Explain the use and limitations of serum creatinine for calculations of estimated glomerular filtration rate.
- Describe the toxic effects related to an increased plasma ammonia concentration.
- Suggest possible clinical conditions associated with test results, given patient values for urea, uric acid, creatinine, and ammonia and supporting clinical history.

The term nonprotein nitrogen (NPN) originated in the early days of clinical chemistry when analytic methodologies required the removal of protein from a specimen before analysis. The concentration of nitrogen-containing compounds in a **protein-free filtrate** was then quantified spectrophotometrically by converting nitrogen to **ammonia**. This was followed by a subsequent reaction with Nessler's reagent ($K_2[HgI_4]$) to produce a yellow color.[1] Though it was technically difficult to perform, the method provided an accurate determination of total NPN concentration. Although determination of total urinary nitrogen is of value in the assessment of nitrogen balance for nutritional management,[2] more useful clinical information is obtained by analyzing a patient's specimen for individual nitrogen-containing compounds.

Numerous compounds of clinical interest are included in the NPN fraction of plasma and urine. The most abundant of these are listed in **Table 7.1**.[3] The majority of these compounds arise from the catabolism of proteins and nucleic acids. The biochemistry, clinical utility, and analytical methods for measurement of the NPN compounds urea, uric acid, creatinine, creatine, and ammonia are presented in this chapter.

Table 7.1 Clinically Significant Nonprotein Nitrogen Compounds

Compound	Approximate Plasma Concentration (% of Total NPN)	Approximate Urine Concentration (% of Excreted Nitrogen)
Urea	45–50	86.0
Amino acids	25	—
Uric acid	10	1.7
Creatinine	5	4.5
Creatine	1–2	—
Ammonia	0.2	2.8

Urea

The NPN compound present in highest concentration in the blood is **urea**. Urea is the major excretory product of protein metabolism; it is formed in the liver from amino groups ($-NH_2$) and free ammonia generated

CASE STUDY 7.1, PART 1

Leroy is a 65-year-old recent retiree living alone and managing his COPD and heart disease. Recently, Leroy has had a hard time keeping up with his daily tasks, including shopping and meal preparation. Today, his daughter stopped by with groceries and noticed he seemed to be moving more slowly than normal. He was also notably confused about her visit, though the visits had been a weekly occurrence for the past 2 months. After conferring with her family members, Leroy's daughter, Sheila, brings Leroy to the emergency department. The examining physician confirms that Leroy is dehydrated and experiencing confusion and exhaustion upon small amounts of exertion. The physician also notes that Leroy is having difficulty breathing and is wincing when he stands to walk. Upon inquiry, the daughter mentions that the pain when Leroy walks "seems to come and go."

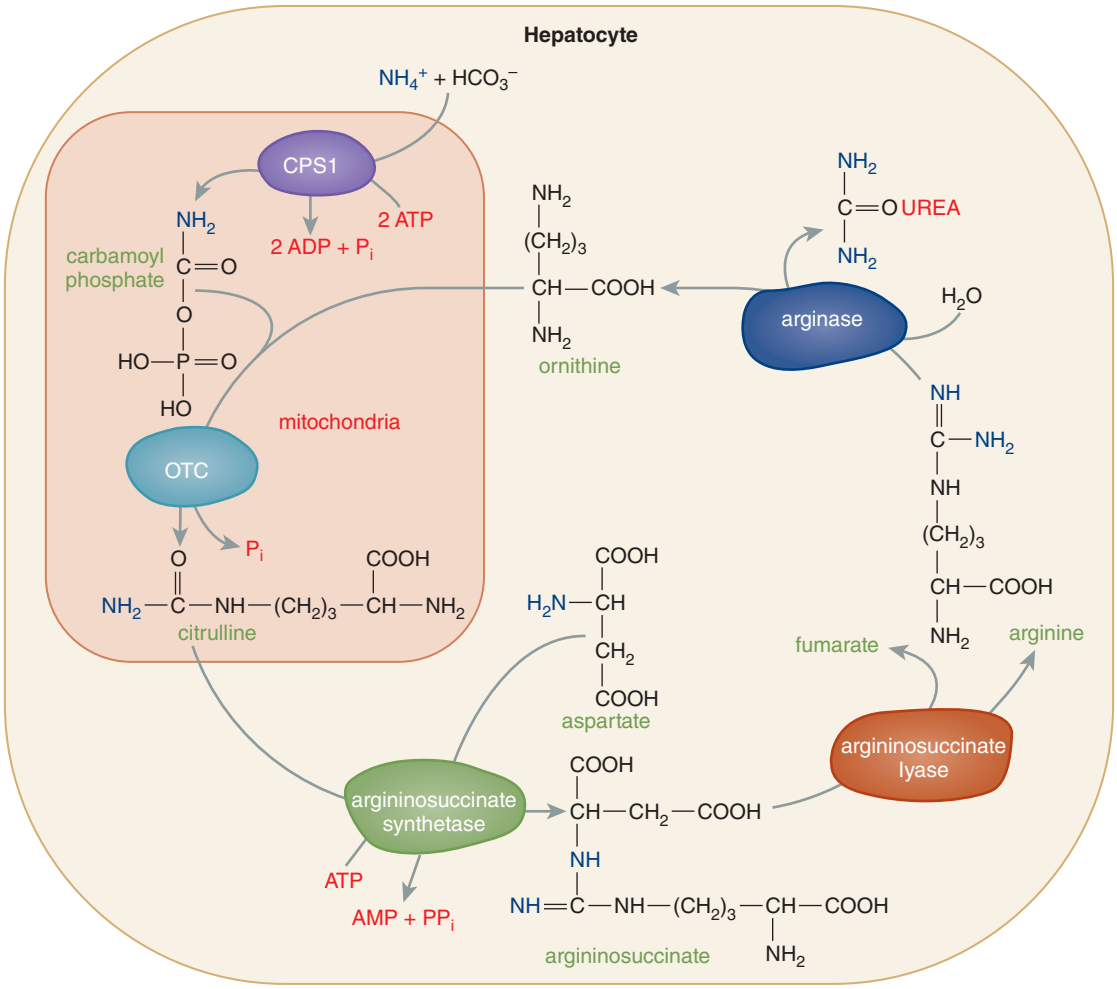


Figure 7.1 Urea cycle. The urea cycle takes place in hepatocytes and is a compartmentalized process between the mitochondria and cytosol. Toxic ammonia is bound to CO_2 and completes a series of steps, ultimately becoming urea, which is sent to the kidneys for excretion.

during protein catabolism.[4] This enzymatically catalyzed process is termed the urea cycle and is an important pathway to reduce the levels of ammonia in the blood, as ammonia is toxic to cells. In the liver, ammonia is bound with CO_2 to form carbamoyl phosphate, which enters the urea cycle and ultimately becomes urea (**Figure 7.1**). Since historic assays for urea were based on the measurement of nitrogen via blood samples, the term blood urea nitrogen (BUN) has commonly been used to refer to urea determination, but urea nitrogen (urea N) is a more appropriate term.[5]

Biochemistry

Protein metabolism produces amino acids that can be oxidized to produce energy or stored as fat and glycogen. During protein metabolism, nitrogen is released, converted to urea, and excreted as a waste product. Following synthesis in the liver, urea is carried in the blood to the kidney and readily filtered from the plasma by the glomerulus. Most of the urea in the glomerular filtrate is excreted in the urine, although some urea is reabsorbed by passive diffusion during passage of the filtrate through the renal tubules.[6] The amount reabsorbed depends on the urine flow rate and extent of hydration. Small quantities of urea (<10% of the total) are excreted through the gastrointestinal (GI) tract and skin. The concentration of urea in the plasma is determined by the protein content of the diet, the rate of protein catabolism, and renal function and perfusion.[4]

Clinical Application

Measurement of urea is used to evaluate renal function, to assess hydration status, to determine nitrogen balance, to aid in the diagnosis of renal disease, and to verify adequacy of dialysis.[4]

Measurements of urea were originally performed on a protein-free filtrate of whole blood and based on

measuring the amount of nitrogen. Current analytic methods have retained this custom, and urea is often reported in terms of nitrogen concentration rather than urea concentration. Urea concentration can be converted to urea concentration by multiplying by 2.14, as follows:

$$\frac{1 \text{ mg urea N}}{\text{dL}} \times \frac{1 \text{ mmol N}}{14 \text{ mg N}} \times \frac{1 \text{ mmol urea}}{2 \text{ mmol N}} \times \frac{60 \text{ mg urea}}{1 \text{ mmol urea}} = \frac{2.14 \text{ mg urea}}{\text{dL}} \quad \text{(Eq. 7.1)}$$

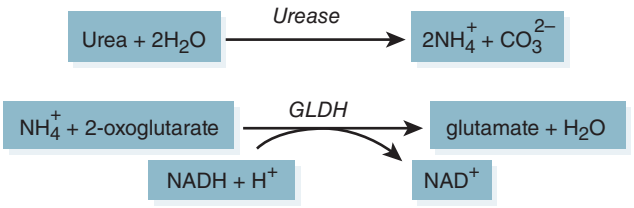
In the International System of Units (SI), urea is reported in units of millimoles per liter. Urea concentration in milligrams per deciliter may be converted to urea concentration in millimoles per liter by multiplying by 0.36.[7]

Analytical Methods

Several analytic approaches have been used to assay urea. Enzymatic methods are used most frequently in clinical laboratories.[8] The enzyme urease (urea amidohydrolase, EC 3.5.1.5) catalyzes hydrolysis of urea in the sample, and the ammonium ion (NH_4^+). produced in the reaction is quantified.[9] The most common method couples the urease reaction with glutamate dehydrogenase (GLDH, EC 1.4.1.3), where the conversion of nicotinamide adenine dinucleotide (reduced, NADH) at 340 nm is measured (**Figure 7.2**).[10]

Ammonium from the urease reaction can also be measured by the color change associated with a pH indicator. This approach has been incorporated into instruments using liquid reagents, a multilayer film format, and reagent strips.[8] A method that uses an electrode to measure the rate of increase in conductivity as ammonium ions are produced from urea is in use in approximately 15% of laboratories in the United States.[8] Because the rate of change in conductivity is measured, ammonia contamination is not a problem, as it is in other methods.

A reference method using isotope dilution mass spectrometry (IDMS) has been developed.[11] Analytic methods are summarized in **Table 7.2**.



GLDH = glutamate dehydrogenase (EC 1.4.1.3)

Figure 7.2 Enzymatic assay for urea.

Specimen Requirements

Urea concentration may be measured in plasma, serum, or urine. If plasma is collected, ammonium ions and high concentrations of sodium citrate and sodium fluoride must be avoided, as citrate and fluoride inhibit urease.[7] Although the protein content

Table 7.2 Summary of Analytic Methods—Urea

Enzymatic Methods		
Methods use a similar first step—catalyzed by urease	Enzymatic production of ammonium ion (NH_4^+) from urea	See Figure 7.2
GLDH-coupled enzymatic	Enzymatic reaction of (NH_4^+) 2-oxoglutarate, and NADH to form glutamate and NAD^+	Used on many automated instruments; best as a kinetic measurement
Indicator dye	(NH_4^+) + pH indicator → color change	Used in automated systems, multilayer film reagents, and dry reagent strips
Conductometric	Conversion of unionized urea to (NH_4^+) and CO_3^{2-} results in increased conductivity	Specific and rapid
Other Methods		
Isotope dilution mass spectrometry	Detection of characteristic fragments following ionization; quantification using isotopically labeled compound	Proposed reference method

of the diet influences urea production, the effect of a single protein-containing meal on urea concentration is minimal and a fasting sample is not generally required. Hemolyzed samples should not be used. Urea is susceptible to bacterial decomposition, so specimens (particularly urine) that cannot be analyzed within a few hours should be refrigerated. Timed urine specimens should be refrigerated during the collection period. Methods for plasma or serum may require modification for use with urine specimens because of high urea concentration and the presence of endogenous ammonia.[7]

Reference Ranges

Blood Urea Nitrogen (BUN)[7]		
Adult		
Plasma or serum	6–20 mg/dL	2.1–7.1 mmol/L
Urine, 24 h	12–20 g/d	0.43–0.71 mol urea/d

Pathophysiology

An elevated concentration of urea in the blood is called **azotemia**. A very high plasma urea concentration accompanied by renal failure is termed **uremia** or **uremic syndrome**. Uremia can present with a wide variety of symptoms including fatigue, nausea or vomiting, and generalized confusion and can eventually become fatal if not treated by dialysis or renal transplantation. Conditions resulting in increased plasma urea concentration are classified into three main categories according to cause: **prerenal**, renal, and **postrenal**.[6]

Prerenal azotemia occurs as a result of reduced renal blood flow. Less blood is delivered to the kidney and so consequently, less urea is filtered. Causative factors include congestive heart failure, shock, hemorrhage, dehydration, and other factors resulting in a significant decrease in blood volume. The amount of protein metabolism also induces prerenal changes in blood urea concentration. A regular high-protein diet or increased protein catabolism, such as occurs in stress, fever, major illness, corticosteroid therapy, and GI hemorrhage, may increase the urea concentration.

Renal causes of elevated urea include acute and chronic renal failure, glomerular nephritis, tubular necrosis, and other intrinsic renal disease (see Chapter 21, *Renal Function*). Decreased renal function causes an increase in plasma urea concentration as a result of compromised urea excretion.

Postrenal azotemia can be due to obstruction of urine flow anywhere in the urinary tract by renal calculi, tumors of the bladder or prostate, or severe infection.

A decrease in urea can also be seen in certain circumstances. The major causes of decreased plasma urea concentration include low protein intake and severe liver disease. Additionally, plasma urea concentration can decrease during late pregnancy and in infancy as a result of increased protein synthesis. The conditions affecting plasma urea concentration are summarized in **Table 7.3**.

Differentiation of the cause of abnormal urea concentration is aided by calculation of the **urea nitrogen/creatinine** (**BUN/creatinine**) **ratio**, which is normally 10:1 to 20:1. Prerenal conditions tend to elevate the BUN/creatinine ratio. This is due to an elevation in plasma urea, while plasma creatinine remains normal. Prerenal conditions that result in a lowered BUN/creatinine ratio are associated with decreased urea production as seen in low protein intake and severe liver disease. An increase in both plasma BUN and creatinine, which in turn shows a "normal" BUN/creatinine ratio, is usually seen in renal conditions as both BUN and creatinine are affected. A high BUN/creatinine ratio with an elevated creatinine is usually seen in postrenal conditions involving urinary flow obstruction.[6]

CASE STUDY 7.1, PART 2

Let's turn back to Leroy. The examining physician wants to determine the reason for Leroy's confusion and lethargy. The doctor orders a chest x-ray and CT scan of the Leroy's lungs and orders the following laboratory tests: BUN, Creatinine, Uric Acid, pH, O_2 sat (%), pCO_2 (mm Hg), and pO_2 (mm Hg).

1. Consider the physician's laboratory orders and imaging studies. This information, along with Leroy's clinical history and current presentation, suggests that the physician is considering what acute diagnoses?

The physician provides an additional assessment of Leroy's ankles and collects a joint aspirate for cell count and analysis. Leroy is then admitted to the 24-hour observation unit.

Table 7.3 Causes of Abnormal Plasma Urea Concentration

Increased Concentration	
Prerenal	Congestive heart failure Shock, hemorrhage Dehydration Increased protein catabolism High-protein diet
Renal	Acute and chronic renal failure Renal disease, including glomerular nephritis and tubular necrosis
Postrenal	Urinary tract obstruction
Decreased Concentration	
	Low protein intake Severe vomiting and diarrhea Liver disease Pregnancy

Uric Acid

In humans and other higher mammals, **uric acid** is the final product of catabolism of purine nucleic acids.[12] Although it is filtered by the glomerulus and secreted by the distal tubules into the urine, most uric acid is reabsorbed in the proximal tubules, maintaining a steady physiological concentration.[13] Uric acid is relatively insoluble in plasma and, at high concentrations, can be deposited in the joints and tissue, causing painful inflammation. The inability of humans and other higher primates to further metabolize uric acid to the more soluble allantoin is attributed to a series of progressive evolutionary events ultimately resulting in the prevention of uricase production, the enzyme responsible for further uric acid degradation (**Figure 7.3**).[12]

Biochemistry

Uric acid is primarily formed in the liver as the end product of purine metabolism, either through the breakdown of dietary nucleic acids or from cellular turnover. Uric acid is transported in the plasma from the liver to the kidney, where it is filtered by the glomerulus. **Reabsorption** of 98% to 100% of the uric acid from the glomerular filtrate occurs in the proximal tubules. Small amounts of uric acid are secreted by the distal tubules into the urine. Renal excretion accounts for about 70% of uric acid elimination; the remainder passes into the GI tract and is degraded by bacterial enzymes.[13] Interestingly, an increase in affinity for uric acid in renal transport proteins has been identified as occurring concurrently with the cessation of enzymatic uricase production in humans, adding to the theory that increased uric acid levels provide evolutionary advantage to humans and other higher primates.[14]

Nearly all of the uric acid in plasma is present as monosodium urate. At the pH of plasma (pH ~ 7), urate is relatively insoluble; at concentrations greater than 6.8 mg/dL, the plasma is saturated. As a result, urate crystals may form and precipitate in the tissues. In acidic urine (pH < 5.75), uric acid is the predominant species, and uric acid crystals may form.[13]

Clinical Application

Uric acid is usually measured to confirm diagnosis and monitor treatment of **gout**, an arthritic condition characterized by the precipitation of uric acid crystal deposition in joints and tissues. It is also used to assess and prevent uric acid nephropathy during chemotherapeutic treatment, to assess inherited disorders of purine metabolism, to detect kidney dysfunction, and to assist in the diagnosis of renal calculi.[15]

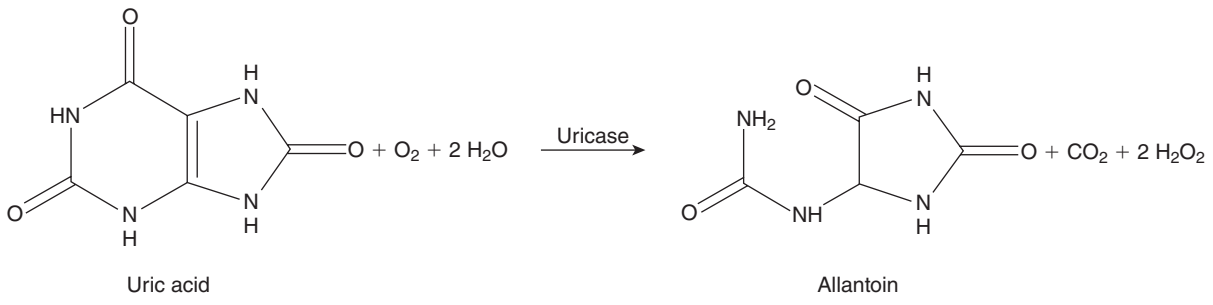


Figure 7.3 Conversion of uric acid to allantoin.

Analytical Methods

Uric acid is readily oxidized to allantoin and, therefore, can function as a reducing agent in chemical reactions. This property was exploited in early analytic procedures for the determination of uric acid. The most common method of this type is the Caraway method, which is based on the oxidation of uric acid in a protein-free filtrate, with subsequent reduction of phosphotungstic acid in alkaline solution to tungsten blue.[16] The method lacks specificity.

Methods using uricase (urate oxidase, EC 1.7.3.3), the enzyme that catalyzes the oxidation of uric acid to allantoin, are more specific and are used almost exclusively in clinical laboratories. The simplest of these methods measures the differential absorption of uric acid and allantoin at 293 nm.[17] The difference in absorbance before and after incubation with uricase is proportional to the uric acid concentration. Proteins can cause high background absorbance, reducing sensitivity; hemoglobin and xanthine can cause negative interference.[18]

Coupled enzymatic methods measure uric acids levels by measuring the hydrogen peroxide produced as uric acid is converted to allantoin.[19] Peroxidase or catalase (EC 1.11.1.6) is then used to catalyze a chemical indicator reaction. The color produced is proportional to the quantity of uric acid in the specimen. Enzymatic methods of this kind have been adapted for use on traditional wet chemistry analyzers and for dry chemistry slide analyzers. Bilirubin and ascorbic acid, which destroy peroxide if present in sufficient quantity, can interfere. Commercial reagent preparations often include potassium ferricyanide and ascorbate oxidase to minimize these interferences.[20]

HPLC (high-performance liquid chromatography) methods, typically using UV detection, have been developed.[21] IDMS has been proposed as a candidate reference method.[22] Analytic methods are summarized in **Table 7.4**.

Specimen Requirements

Uric acid may be measured in heparinized plasma, serum, or urine. Serum should be removed from cells as quickly as possible to prevent dilution by intracellular contents. Diet can affect uric acid concentration overall, but a recent meal has no significant effect; therefore, a fasting specimen is unnecessary. Gross lipemia (seen in individuals with elevated triglyceride levels) should be avoided. High bilirubin concentration may falsely decrease results obtained by peroxidase methods. Significant hemolysis, with concomitant glutathione release, may result in low values. Drugs such as salicylates and thiazides have been shown to increase values for uric acid.[23]

Uric acid is stable in plasma or serum after red blood cells have been removed. Serum samples may be stored refrigerated for 3 to 5 days. Ethylenediaminetetraacetic acid (EDTA) or fluoride additives should not be used for specimens that will be tested by a uricase method. Urine collections must be alkaline (pH 8).[7]

Table 7.4 Summary of Analytic Methods—Uric Acid

Chemical Methods		
Phosphotungstic acid	In carbonate solution (Na_2CO_3/OH^-), uric acid + $H_3PW_{12}O_{40}$ + O_2 → allantoin + tungsten blue + CO_2	Nonspecific; requires protein removal
Enzymatic Methods		
Similar first step—catalyzed by uricase	Enzymatic production of allantoin from uric acid	Uric acid + O_2 + $2H_2O$ → Allantoin + CO_2 + $2H_2O_2$
Coupled enzymatic—peroxidase	H_2O_2 + indicator dye → colored compound	Readily automated; reducing agents interfere
Spectrophotometric	Decrease in absorbance at 293 nm measured	Hemoglobin and xanthine interfere
Other Methods		
Isotope dilution mass spectrometry	Detection of characteristic fragments following ionization; quantification using isotopically labeled compound	Proposed reference method

Reference Ranges

Uric Acid (Uricase Method)[7]		
Patient	**Sample**	**Reference Range**
Adult Male	Plasma or Serum	3.5–7.2 mg/dL (0.21–0.43 mmol/L))
Adult Female	Plasma or Serum	2.6–6.0 mg/dL (0.16–0.36 mmol/L)
Child	Plasma or Serum	2.0–5.5 mg/dL (0.12–0.33 mmol/L)
Adult Male or Female	Urine, 24 h	24 h 250–750 mg/dL (1.5–4.4 mmol/dL)

Results expressed in conventional units of milligrams per deciliter can be converted to SI units using the molecular mass of uric acid (168 g/mol).

Pathophysiology

While there is growing evidence highlighting the physiological benefits of elevated uric acid levels in humans in comparison with other mammals, abnormally increased plasma uric acid concentration are seen in a variety of clinical conditions including gout, increased catabolism of nucleic acids, and renal disease.[13]

Gout is a disease found primarily in men and is usually first diagnosed between 30 and 50 years of age. Affected individuals have pain and inflammation of the joints caused by precipitation of sodium urates. The most common joint affected is the big toe; however, other commonly involved joints include ankles, knees, and fingers. In 25% to 30% of these patients, **hyperuricemia** is a result of overproduction of uric acid and may be exacerbated by a purine-rich diet, drugs, and alcohol. Genetic variations in urate transport proteins have also been shown to influence uric acid retention, independent of diet and/or drug and alcohol use.[13] Plasma uric acid concentration in affected individuals is usually greater than 6.0 mg/dL. Patients with gout are susceptible to the formation of renal calculi, although not all persons with abnormally high serum urate concentrations develop this complication. In women, urate concentration rises after menopause. Postmenopausal women may develop hyperuricemia and gout. In severe cases, deposits of crystalline uric acid and urates called tophi form in tissue, causing deformities.[13]

Another common cause of elevated plasma uric acid concentration is increased metabolism of cell nuclei, as occurs in patients on chemotherapy for such proliferative diseases as leukemia, lymphoma, multiple myeloma, and polycythemia. Monitoring uric acid concentration in these patients is important to avoid nephrotoxicity. Allopurinol (Zyloprim), which inhibits xanthine oxidase (EC 1.1.3.22), an enzyme in the uric acid synthesis pathway, is used for treatment.

Patients with hemolytic or megaloblastic anemia may exhibit elevated uric acid concentration. Increased urate concentrations can occur as a result of increased tissue catabolism due to inadequate dietary intake (starvation).[13]

Inherited disorders of purine metabolism are associated with significant increases in physiological uric acid concentrations.[13] Lesch-Nyhan syndrome is an X-linked genetic disorder (seen only in males) caused by the complete deficiency of hypoxanthine–guanine phosphoribosyltransferase (EC 2.4.2.8), an important enzyme in the biosynthesis of purines. Lack of this enzyme prevents the reutilization of purine bases in the nucleotide salvage pathway and results in increased de novo synthesis of purine nucleotides and high plasma and urine concentrations of uric acid. Neurologic symptoms, severe cognitive impairment, and self-mutilation characterize this extremely rare disease. Mutations in the first enzyme in the purine synthesis pathway, phosphoribosylpyrophosphate

CASE STUDY 7.1, PART 3

Recall Leroy. He was immediately placed on supportive oxygen and given diuretics and intravenous (IV) fluids to promote diuresis. Later that evening, Leroy's nurse noted that the laboratory identified crystals in Leroy's joint fluid.

2. Based on Leroy's clinical presentation, and on the findings of the joint fluid analysis, what would you expect his uric acid level to be?

© digitalskillet/Shutterstock.

synthetase (EC 2.7.6.1), also cause elevated uric acid concentration. Increased uric acid is found secondary to glycogen storage disease (deficiency of glucose-6-phosphatase, EC 3.1.3.9) and fructose intolerance (deficiency of fructose-1-phosphate aldolase, EC 2.1.2.13). Metabolites such as lactate and triglycerides are produced in excess and compete with urate for renal excretion in these diseases.

Hyperuricemia as a result of decreased uric acid excretion is a common feature of toxemia of pregnancy (preeclampsia) and lactic acidosis, presumably as a result of competition for binding sites in the renal tubules. Chronic renal disease causes elevated uric acid concentration because filtration and **secretion** are impaired.

Uric acid nephrolithiasis, the formation of kidney stones (renal calculi), may occur due to a variety of predisposing factors and conditions.[24] In acidic urine, the relatively insoluble uric acid precipitates to form calculi, which can cause intense flank pain. The stones may be dissolved by alkalinization of the urine or treated by increased fluid intake and administration of xanthine oxidase inhibitors to reduce uric acid production.

Hypouricemia is less common than hyperuricemia and is usually secondary to severe liver disease or defective tubular reabsorption, as in Fanconi syndrome (a disorder of reabsorption in the proximal convoluted tubules of the kidney).[25] Decreased plasma uric acid can be caused by overtreatment with allopurinol during chemotherapy, or treatment with 6-mercaptopurine or azathioprine, inhibitors of de novo purine synthesis. Some studies have shown an association between low uric acid concentrations and neurodegenerative conditions such as Alzheimer's and Parkinson's diseases.[26]

Conditions known to affect plasma urate concentrations are shown in **Table 7.5**.

Table 7.5 Causes of Abnormal Plasma Uric Acid Concentration

Increased Concentration
Gout
Treatment of myeloproliferative disease with cytotoxic drugs
Hemolytic and proliferative processes
Purine-rich diet
Increased tissue catabolism or starvation
Enzyme deficiencies
Lesch-Nyhan syndrome (hypoxanthine guanine phosphoribosyltransferase deficiency)
Phosphoribosylpyrophosphate synthetase deficiency
Glycogen storage disease type I (glucose-6-phosphatase deficiency)
Fructose intolerance (fructose-1-phosphate aldolase deficiency)
Toxemia of pregnancy
Lactic acidosis
Chronic renal disease
Drugs and poisons
Chemotherapy with azathioprine or 6-mercaptopurine
Decreased Concentration
Liver disease
Defective tubular reabsorption (Fanconi syndrome)
Overtreatment with allopurinol

Creatinine/Creatine

Creatinine is formed from **creatine** and creatine phosphate in muscle and is excreted into the plasma at a constant rate related to muscle mass. Plasma creatinine is inversely related to **glomerular filtration rate (GFR)** and, although an imperfect measure, it is commonly used to assess renal filtration function.[5]

Biochemistry

Creatine is synthesized primarily in the liver from arginine, glycine, and methionine.[27] It is then transported to other tissues, such as muscle, where it is converted to creatine phosphate, which serves as a high energy source. Creatine phosphate undergoes the spontaneous loss of phosphoric acid while creatine loses water, each independently forming the cyclic compound creatinine, which diffuses into the plasma and is excreted in the urine. The structures and relationship of these compounds are shown in **Figure 7.4**. Creatinine is released into the circulation at a relatively constant rate that has been shown to be proportional to an individual's muscle mass. It is removed from the circulation by glomerular filtration and excreted in the urine. Small amounts of creatinine are secreted by the proximal tubule and reabsorbed by the renal tubules.[5] Daily creatinine excretion is reasonably stable.

Clinical Application

Measurement of creatinine concentration is used to determine the sufficiency of kidney function, to determine the severity of kidney damage, and to monitor the progression of kidney disease.[27]

Plasma creatinine concentration is a function of relative muscle mass, the rate of creatine turnover, and renal function. The amount of creatinine in the bloodstream is reasonably stable, although increased protein content of the diet may influence the plasma

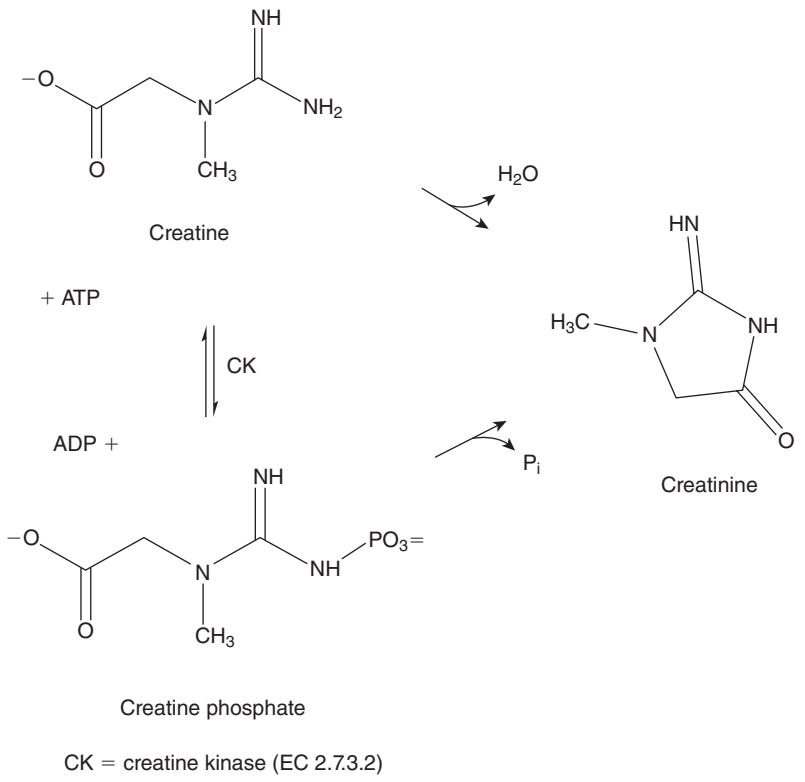


Figure 7.4 Interconversion of creatine, creatine phosphate, and creatinine.

concentration. Because of the constancy of endogenous production, urinary creatinine excretion has been used as a measure with which to compare additional urinary constituents, expressed as a ratio of analyte to creatinine quantity rather than as an independent concentration.[27]

Creatinine clearance (CrCl), a measure of the amount of creatinine eliminated from the blood by the kidneys, and GFR are used to gauge renal function.[28] The GFR is the volume of plasma filtered (V) by the glomerulus per unit of time (t):

$$\text{GFR} = \frac{V}{t} \qquad \textbf{(Eq. 7.2)}$$

Assuming a substance, S, can be measured and is freely filtered at the glomerulus and neither secreted nor reabsorbed by the tubules, the volume of plasma filtered would be equal to the mass of S filtered (M_S) divided by its plasma concentration (P_S):

$$V = \frac{M_S}{P_S} \qquad \textbf{(Eq. 7.3)}$$

The mass of S filtered is equal to the product of its urine concentration (U_S) and the urine volume (V_U):

$$M_S = U_S V_S \qquad \textbf{(Eq. 7.4)}$$

CASE STUDY 7.1, PART 4

Recall Leroy. His laboratory results showed an elevated urea N/creatinine ratio with normal creatinine. His pH was slightly acidic, with increased pCO_2 and decreased pO_2 levels.

3. What is the most likely cause of Leroy's elevated urea N/creatinine ratio?

If the urine and plasma concentrations of *S*, the volume of urine collected, and the time over which the sample was collected are known, the GFR can be calculated:

$$\text{GFR} = \frac{U_S V_U}{P_S t} \qquad \textbf{(Eq. 7.5)}$$

The clearance of a substance is the volume of plasma from which that substance is removed per unit time. The formula for CrCl is given as follows, where U_{Cr} is urine creatinine concentration and P_{Cr} is plasma creatinine concentration:

$$\text{CrCl} = \frac{U_{Cr} V_U}{P_{Cr} t} \qquad \textbf{(Eq. 7.6)}$$

CrCl is usually reported in units of mL/min and can be corrected for body surface area. CrCl requires both a venous blood draw and a 24-hour urine collection specimen. In addition to the error potential during urine collection, CrCl is limited in that it overestimates GFR due to both renal tubular reabsorption of creatinine in small amounts, and the active renal tubular generation and secretion of creatinine (see Chapter 21, *Renal Function*). Despite these limitations, CrCl remains a reasonable approximation of GFR.[28]

While the observed relationship between plasma creatinine and GFR make the analyte a good endogenous filtration marker, the measurement of plasma creatinine does not provide sufficient sensitivity for the detection of mild renal dysfunction. Measured creatinine concentration used in combination with other variables in one of several empirically determined equations provides a better assessment of renal disease, in part because the equations estimate GFR, not just CrCl.

Clinical laboratories have been strongly encouraged to report an estimated GFR when serum creatinine is ordered as a means to increase identification of kidney disease and improve patient care.[29] In order to address method variation in creatinine determination, and thus improve the accuracy of **estimated glomerular filtration rate (eGFR)** calculations, the National Kidney and Disease Education Program recommends that creatinine assays used in the determination of eGFR be calibrated and traceable to an IDMS method.[30] Initially, the abbreviated Modification of Diet in Renal Disease (MDRD) equation was advocated. The current iteration of the MDRD equation includes four variables in the calculation; serum creatinine concentration, age, gender (sex), and race (African American/Black vs. non-African American/Black race). The equation has been validated for individuals between the ages of 17 and 70 and makes the assumption that all filtered creatinine is excreted.[31]

Using an IDMS-traceable method, the MDRD equation for eGFR is

$$\text{eGFR (mL/min/1.73m}^2) = 175 \times (S_{Cr})^{-1.154} \times (\text{age})^{-0.203} \times (0.742 \text{ if female}) \times (1.210 \text{ if African American}) \qquad \textbf{(Eq. 7.7)}$$

where S_{Cr} is serum (plasma) creatinine concentration in mg/dL, age is in years, and results are normalized to a standard body surface area (1.73 m^2).[31]

In an effort to further improve the accuracy of eGFR equations, the Chronic Kidney Disease Epidemiology (CKDEPI) Collaboration published the CKD-EPI equation in 2009, introducing cystatin-C alongside creatinine as a measurable variable in eGFR calculations.[32]

Recent criticisms questioning the use of a race coefficient or variable in the MDRD and CKD-EPI equation highlight the approximate 16% overestimation of eGFR in patients categorized as Black versus non-Black individuals.[33] Previous equations failed to provide adequate support for the assertion that creatinine production was specific and distinguishable based on an individual's race. Additionally, the use of race-based variables has been implicated in the disparity of care for Black individuals, in part due to higher calculated eGFR values, potentially delaying treatment and underestimating the degree of kidney damage.[34]

In response, and in examination of current practices, the National Kidney Foundation, alongside the American Society for Nephrology, developed the NKF-ASN Task Force on Reassessing the Inclusion of Race in Diagnosing Kidney Disease. In March of 2021, the NKF-ASN task force issued a statement calling for the cessation of race identification and inclusion as a coefficient in eGFR calculations.[35] Additional methods utilizing new biomarkers, without the inclusion of race as a calculated variable, have been proposed, and it is anticipated the NKF-ASN task force will issue additional guidance on eGFR calculations in the near future.[34]

Pediatric assessment of eGFR utilizes the modified Schwartz equation, or "Bedside Schwartz" to calculate eGFR in children younger than 18 years old.[36] The equation is

$$\text{eGFR (mL/min/1.73 m}^2) = (0.41 \times \text{height})/S_{Cr} \qquad \textbf{(Eq. 7.8)}$$

Height is measured in cm and S_{cr} is serum (plasma) creatinine concentration in mg/dL. This equation improves upon the original Schwartz equation, which overestimates the GFR by 20% to 40% for creatinine concentrations measured by a traceable method.[37]

Analytical Methods

Creatinine

The methods most frequently used to measure creatinine are based on the Jaffe reaction first described in 1886.[38] In this reaction, creatinine reacts with picric acid in alkaline solution to form a red-orange chromogen. The reaction was adopted for the measurement of blood creatinine by Folin and Wu in 1919.[39] The reaction is nonspecific and subject to positive interference by a large number of compounds, including acetoacetate, acetone, ascorbate, glucose, and pyruvate. More accurate results are obtained when creatinine in a protein-free filtrate is adsorbed onto Fuller's earth (aluminum magnesium silicate) or Lloyd's reagent (sodium aluminum silicate), then eluted and reacted with alkaline picrate.[40] This method is time consuming and not readily automated; it is not routinely used.

Two approaches have been used to increase the specificity of assay methods for creatinine: a kinetic Jaffe method and reaction with various enzymes. In the kinetic Jaffe method, serum is mixed with alkaline picrate and the rate of change in absorbance is measured.[41] Although this method eliminates some of the nonspecific reactants, it is subject to interference by α-keto acids and cephalosporins.[42] Bilirubin and hemoglobin may cause a negative bias, probably a result of their destruction in the strong base used. The kinetic Jaffe method is used routinely despite these problems because it is inexpensive, rapid, and easy to perform.

Additional efforts to enhance the specificity of the Jaffe reaction utilize **coupled enzymatic methods**. A method using creatininase (creatinine amidohydrolase, EC 3.5.2.10), creatinase (creatine amidinohydrolase, EC 3.5.3.3), sarcosine oxidase (EC 1.5.3.1), and peroxidase (EC 1.11.1.7) was adapted for use on a dry slide analyzer.[43]

IDMS is used as a reference method.[30] Assays used on automated analyzers are designated as "traceable" (calibrated) to an IDMS method. Analytic methods for creatinine are summarized in **Table 7.6**.

Specimen Requirements

Creatinine may be measured in plasma, serum, or urine. Hemolyzed and icteric samples should be avoided, particularly if a Jaffe method is used. Lipemic samples may produce erroneous results in some methods. A fasting sample is not required, although high protein ingestion may transiently elevate serum concentrations. Urine should be refrigerated after collection or frozen if longer storage than 4 days is required.[7]

Sources of Error

Ascorbate, glucose, α-keto acids, and uric acid may increase creatinine concentration measured by the Jaffe reaction, especially at temperatures above 30°C. This interference is significantly decreased when kinetic measurement is applied. Depending on the concentration of reactants and measuring time, interference from α-keto acids may persist in kinetic Jaffe methods. Some of these substances interfere in enzymatic methods for creatinine measurement. Bilirubin causes a negative bias in both Jaffe and enzymatic methods. Ascorbate will interfere in enzymatic methods that use peroxidase as a reagent.[7]

Patients taking cephalosporin antibiotics may have falsely elevated results when the Jaffe reaction is used. Other drugs have been shown to increase

CASE STUDY 7.1, PART 5

Recall Leroy. He remains in inpatient treatment for 2 more days, receiving IV fluids, diuretics, and supplemental oxygen. During this time his cognition and energy levels continue to improve; however, he still reports sharp pain in his ankles when walking.

4. What laboratory levels should be monitored to confirm the efficacy of his inpatient treatment?
5. What additional clinical concerns should be addressed?
6. What additional treatment options should be considered?

© digitalskillet/Shutterstock.

Table 7.6 Summary of Analytic Methods—Creatinine

Chemical Methods Based on Jaffe Reaction		
Jaffe reaction	In alkaline solution, creatinine + picrate → red-orange complex	
Jaffe kinetic	Jaffe reaction performed directly on sample; detection of color formation timed to avoid interference of noncreatinine chromogens	Positive bias from α-keto acids and cephalosporins; requires automated equipment for precision
Jaffe with adsorbent	Creatinine in protein-free filtrate adsorbed onto Fuller's earth (aluminum magnesium silicate), then reacted with alkaline picrate to form colored complex	Adsorbent improves specificity; previously considered reference method
Jaffe without adsorbent	Creatinine in protein-free filtrate reacts with alkaline picrate to form colored complex	Positive bias from ascorbic acid, glucose, glutathione, α-keto acids, uric acid, and cephalosporins
Enzymatic Methods		
Creatininase–H_2O_2	In a series of enzymatically catalyzed reactions, creatinine is hydrolyzed to creatine, which is converted to sarcosine and urea. Sarcosine is oxidized to glycine, CH_2O, and H_2O_2. Peroxidase-catalyzed oxidation of a colorless substrate produces a colored product + H_2O	Adapted for use as dry slide method; potential to replace Jaffe; no interference from acetoacetate or cephalosporins; some positive bias due to lidocaine
Creatininase–CK	In a series of reactions catalyzed by the enzymes creatininase, creatine kinase, pyruvate kinase, and lactate dehydrogenase, NAD^+ is produced and measured as a decrease in absorbance	Lacks sensitivity; not used widely
Other Methods		
Isotope dilution mass spectrometry	Detection of characteristic fragments following ionization; quantification using isotopically labeled compound	Highly specific; accepted reference method

creatinine results. Dopamine, in particular, is known to affect both enzymatic and Jaffe methods. Lidocaine causes a positive bias in some enzymatic methods.[18]

Creatine

The traditional method for creatine measurement relies on the analysis of the sample using an end-point Jaffe method for creatinine before and after it is heated in acid solution. Heating converts creatine to creatinine, and the difference between the two sample measurements is the creatine concentration. High temperatures may result in the formation of additional chromogens, and the precision of this method is poor. Several enzymatic methods have been developed; one is the creatininase assay. The initial enzyme is omitted and creatine kinase (EC 2.7.3.2), pyruvate kinase (EC 2.1.7.40), and lactate dehydrogenase (EC 1.1.1.27) are coupled to produce a measurable colored product.[44] Creatine can be measured by HPLC.[45]

Reference Ranges

Creatinine

Adult	Plasma or Serum	Jaffe Method	Enzymatic Method
Male		0.9–1.3 mg/dL	0.6–1.1 mg/dL
		(80–115 μmol/L)	(53–97 μmol/L)
Female		0.6–1.1 mg/dL	0.5–0.8 mg/dL
		(53–97 μmol/L)	(44–71 μmol/L)
Child		0.3–0.7 mg/dL	0.0–0.6 mg/dL
		(27–62 μmol/L)	(0–53 μmol/L)
Adult	Urine, 24 h		
Male		800–2000 mg/dL (7.1–17.7 mmol/d)	
Female		600–1800 mg/dL (5.3–15.9 mmol/d)	

CASE STUDY 7.1, PART 6

Recall Leroy. After 2 days of treatment, he is feeling much better, and his laboratory results are essentially normal. Leroy has agreed to move in with his daughter upon discharge so that she can better support him with his daily needs. He has been referred to a respiratory therapist for additional support with his COPD symptoms and has been prescribed allopurinol for gout. Additionally, the physician requests that Leroy follow up with a primary physician within 3 weeks of his discharge and to report to the emergency department if he has any further episodes of confusion or difficulties breathing.

Reference ranges vary with assay type, age, and gender.[7] Creatinine concentration decreases with age beginning in the 5th decade of life.

Pathophysiology

Creatinine

Elevated creatinine concentration is associated with abnormal renal function, especially as it relates to glomerular function. Plasma concentration of creatinine is inversely proportional to the clearance of creatinine. Therefore, when plasma creatinine concentration is elevated, GFR is decreased, indicating renal damage. Plasma creatinine is a relatively insensitive marker and may not be measurably increased until renal function has deteriorated more than 50%.[5]

Creatine

In muscle disease such as muscular dystrophy, poliomyelitis, hyperthyroidism, and trauma, both plasma creatine and urinary creatinine are often elevated. Plasma creatinine concentrations are usually normal in these patients. Measurement of creatine kinase is typically used for the diagnosis of muscle disease in place of creatine because analytic methods for creatine are not readily available in most clinical laboratories. Plasma creatine concentration is not elevated in renal disease.[5]

Ammonia

Ammonia is produced in the deamination of amino acids during protein metabolism.[2] It is removed from the circulation and converted to urea in the liver. Free ammonia is extremely toxic to human cells; however, ammonia is present in the plasma in low concentrations.

Biochemistry

Ammonia (NH_3) is produced in the catabolism of amino acids and by bacterial metabolism in the lumen of the intestine.[46] Some endogenous ammonia results from anaerobic metabolic reactions that occur in skeletal muscle during exercise. Ammonia is consumed by the parenchymal cells of the liver in the Krebs-Henseleit or urea cycle to produce urea, a nontoxic compound that is excreted in the urine. At normal physiologic pH, most ammonia in the blood exists as ammonium ion (NH_4^+). **Figure 7.5** shows the pH-dependent equilibrium between NH_3 and (NH_4^+). Ammonia is excreted as ammonium ion by the kidney and acts to buffer urine.[47]

Clinical Application

Clinical conditions in which blood ammonia concentration provides useful information are hepatic failure, Reye's syndrome, and inherited deficiencies of urea cycle enzymes. Severe liver disease is the most common cause of disturbed ammonia metabolism. The monitoring of blood ammonia may be used to determine prognosis, although correlation between the extent of hepatic encephalopathy and plasma ammonia concentration is not always consistent. Arterial ammonia concentration is a better indicator of the severity of disease.[48]

Reye's syndrome, occurring most commonly in children, is a serious disease that can be fatal. Frequently, the disease is preceded by a viral infection and the administration of aspirin. Reye's

$$NH_4^+ + H_2O \rightleftharpoons NH_3 + H_3O^+$$

Figure 7.5 Interconversion of ammonium ion and ammonia.

syndrome is an acute metabolic disorder of the liver, and autopsy findings show severe fatty infiltration of that organ.[49] Blood ammonia concentration can be correlated with both the severity of the disease and prognosis; as ammonia levels rise, the prognosis decreases. Survival reaches 100% if plasma NH_3 concentration remains below five times normal.[50]

Ammonia is of use in the diagnosis of inherited deficiency of urea cycle enzymes. Testing should be considered for any neonate with unexplained nausea, vomiting, or neurological deterioration associated with feeding.[51]

Blood ammonia can also be used to monitor hyperalimentation therapy, and urine ammonia determination can be used to confirm the ability of the kidneys to produce ammonia.[7]

Analytical Methods

The accurate laboratory measurement of ammonia in plasma is complicated by its low concentration, instability, and pervasive contamination. Two approaches have been used for the measurement of plasma ammonia. One is a two-step approach in which ammonia is isolated from the sample and then assayed. The second involves direct measurement of ammonia by an enzymatic method or ion selective electrode. Current assays usually detect either NH_3 or (NH_4^+).

One of the first analytic methods for ammonia, developed by Conway in 1935, exploited the volatility of ammonia to separate the compound in a microdiffusion chamber.[52] Ammonia gas from the sample diffuses into a separate compartment and is absorbed in a solution containing a pH indicator. The amount of ammonia is determined by titration.

Ammonia can be measured by an enzymatic method using GLDH. This method is convenient and the most common technique used currently.[53] The decrease in absorbance at 340 nm as nicotinamide adenine dinucleotide phosphate (reduced, NADPH) is consumed in the reaction proportionally to the ammonia concentration in the specimen. NADPH is the preferred coenzyme because it is used specifically by GLDH; NADH will participate in reactions of other endogenous substrates, such as pyruvate. Adenosine diphosphate is added to the reaction mixture to increase the rate of the reaction and to stabilize GLDH.[54] This method is used on many automated systems and is available as a prepared kit from numerous manufacturers.

A dry slide automated system uses a thin-film colorimetric assay.[55] In this method, ammonia reacts with an indicator to produce a colored compound that is detected spectrophotometrically.

Direct measurement using an ion-selective electrode has been developed.[56] The electrode measures the change in pH of an ammonium chloride solution as ammonia diffuses across a semipermeable membrane. Analytic methods for ammonia are summarized in **Table 7.7**.

Specimen Requirements

Careful specimen handling is extremely important for plasma ammonia assays. Whole blood ammonia concentration increases rapidly following specimen collection because of in vitro amino acid deamination. Venous blood should be obtained without trauma and placed on wet ice immediately. Heparin and EDTA are suitable anticoagulants. Commercial collection containers should be evaluated for ammonia

Table 7.7 Summary of Analytic Methods—Ammonia

Chemical Methods		
Ion-selective electrode	Diffusion of NH_3 through selective membrane into NH_4Cl causes a pH change, which is measured potentiometrically	Good accuracy and precision; membrane stability may be a problem
Spectrophotometric	NH_3 + bromophenol blue → blue color	
Enzymatic Methods		
Catalyzed by GLDH	Enzymatic reaction of NH^+, 2-oxoglutarate, and NADPH to form glutamate and $NADP^+$, which is detected spectrophotometrically	Most common on automated instruments; accurate and precise

interference before a new lot is put into use. Samples should be centrifuged at 0 to 4°C within 20 minutes of collection and the plasma removed. Specimens should be assayed as soon as possible or frozen. Frozen plasma is stable for several days at –20°C. Erythrocytes contain two to three times as much ammonia as plasma; hemolysis should be avoided.

Cigarette smoking by the patient is a significant source of ammonia contamination. It is recommended that patients do not smoke for several hours before a specimen is collected.[7]

Many substances influence the in vivo ammonia concentration.[18] Ammonium salts, asparaginase, barbiturates, diuretics, ethanol, hyperalimentation, narcotic analgesics, and some other drugs may increase ammonia in plasma. Diphenhydramine, *Lactobacillus acidophilus*, lactulose, levodopa, and several antibiotics decrease concentrations. Glucose at concentrations greater than 600 mg/dL (33 mmol/L) interferes in dry slide methods.

Sources of Error

Ammonia contamination is a potential problem in the laboratory measurement of ammonia. Precautions must be taken to minimize contamination in the laboratory in which the assay is performed. Eliminating sources of ammonia contamination can significantly improve the accuracy of ammonia assay results. Sources of contamination include tobacco smoke, urine, and ammonia in detergents, glassware, reagents, and water.

The ammonia content of serum-based control material is unstable. Frozen aliquots of human serum albumin containing known amounts of ammonium chloride or ammonium sulfate may be used. Solutions containing known amounts of ammonium sulfate are commercially available.

Reference Ranges

Values obtained vary somewhat with the method used.[8] Higher concentrations are seen in newborns.

Ammonia

Patient	Sample	Conventional Units	SI Units
Adult	Plasma	19–60 µg/dL	11–35 µmol/L
	Urine, 24 h	140–1500 mg N/d	10–107 mmol N/d
Child (10 d to 2 y)	Plasma	68–136 µg/dL	40–80 µmol/L

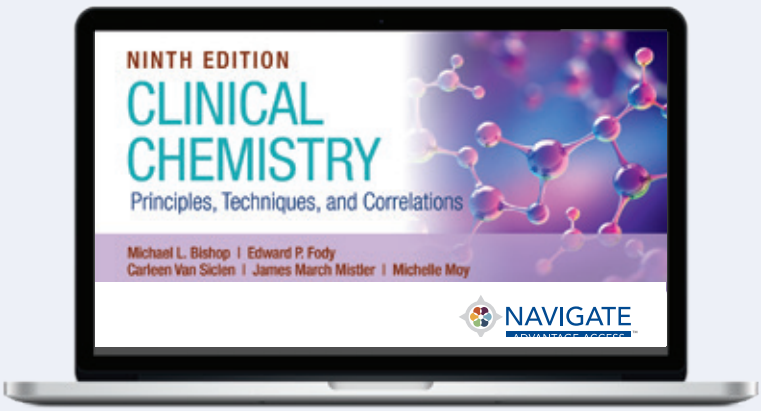
Pathophysiology

In severe liver disease in which there is significant collateral circulation (as in cirrhosis) or if parenchymal liver cell function is severely impaired, ammonia is not removed from the circulation, and blood concentration increases. High concentrations of NH_3 are neurotoxic and often associated with encephalopathy. Ammonia alters metabolic processes in the brain, which results in accumulation of toxic species, and impairs astrocyte function.[57]

Hyperammonemia is associated with inherited deficiency of urea cycle enzymes.[51] Measurement of plasma ammonia is important in the diagnosis and monitoring of these inherited metabolic disorders (see Chapter 30, *Newborn and Pediatric Clinical Chemistry*).

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Gentzkow CJ. An accurate method for determination of blood urea nitrogen by direct nesslerization. *J Biol Chem.* 1942;143:531–544.
2. Matthews DE. Proteins and amino acids. In: Shils ME, ed. *Modern Nutrition in Health and Disease*. 10th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2006:23–61.
3. Montgomery R, Conway TW, Spector AA, et al. *Biochemistry: A Case-Oriented Approach*. 6th ed. St. Louis, MO: Mosby; 1996.
4. Weiner ID, Mitch WE, Sands JM. Urea and ammonia metabolism and the control of renal nitrogen excretion. *Clin J Am Soc Nephrol*. 2015;10(8):1444–1458. doi:10.2215/CJN.10311013
5. Lamb EJ, Price CP. Kidney function tests. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 5th ed. St. Louis, MO: Elsevier Saunders; 2012;669–707.
6. Winter WE. The kidney. In: Laposata M, ed. *Laboratory Medicine: the Diagnosis of Disease in the Clinical Laboratory*. 2nd ed. New York: McGraw-Hill Education; 2014. http://accessmedicine.mhmedical.com. Accessed February 2016.
7. Wu AHB. *Tietz Clinical Guide to Laboratory Tests*. 4th ed. St. Louis, MO: WB Saunders; 2006.
8. SURVEYS 2015: C-B Chemistry/Therapeutic Drug Monitoring and U-B Urine Chemistry. Northfield, IL: College of American Pathologists; 2015.
9. Berthelot MPE. Violet d'aniline. *Repert Chim Appl.* 1859;1:284.
10. Talke H, Schubert GE. Enzymatic urea determination in the blood and serum in the warburg optical test. *Klin Wochenschr.* 1965;43:174–175.
11. Kessler A, Siekmann L. Measurement of urea in human serum by isotope dilution mass spectrometry: a reference procedure. *Clin Chem.* 1999;45:1523–1529.
12. Kratzer JT, Lanaspa MA, Murphy MN, et al. Evolutionary history and metabolic insights of ancient mammalian uricases. *Proc Natl Acad Sci USA*. 2014;111(10):3763–3768. doi:10.1073/pnas.1320393111
13. Burns CM, Wortmann RL. Disorders of purine and pyrimidine metabolism. In: Kasper DL, Fauci AS, Hauser SL, Longo DL, Jameson JL, Loscalzo J, eds. *Harrison's Principles of Internal Medicine*. 19th ed. New York: McGraw-Hill Medical; 2015. http://accessmedicine.mhmedical.com. Accessed February 2016.
14. Tan PK, Farrar JE, Gaucher EA, Miner JN. Coevolution of URAT1 and uricase during primate evolution: Implications for serum urate homeostasis and gout. *Mol Biol Evol*. 2016;33(9):2193–2200. doi:10.1093/molbev/msw116
15. Uric acid. Lab Tests Online. https://labtestsonline.org/understanding/analytes/uric-acid. Reviewed June 25, 2013. Accessed March 2021.
16. Caraway WT, Hald PM. Uric acid. In: Seligson D, ed. *Standard Methods of Clinical Chemistry*. Vol. 4. New York, NY: Academic Press; 1963:239–247.
17. Feichtmeier TV, Wrenn HT. Direct determination of uric acid using uricase. *Am J Clin Pathol.* 1955;25:833–839.
18. Young DS. *Effects of Preanalytical Variables on Clinical Laboratory Tests*. 3rd ed. Washington, DC: American Association for Clinical Chemistry Press; 2007.
19. Gochman N, Schmitz JM. Automated determination of uric acid with use of a uricase-peroxidase system. *Clin Chem.* 1971;17:1154–1159.
20. Hullin DA, McGrane MT. Effect of bilirubin on uricase-peroxidase coupled reactions. Implications for urate measurement in clinical samples and external quality assessment schemes. *Ann Clin Biochem*. 1991;28 (Pt 1): 98–100. doi:10.1177/000456329102800117
21. Cooper N, Khosravan R, Erdmann C, et al. Quantification of uric acid, xanthine and hypoxanthine in human serum by HPLC for pharmacodynamic studies. *J Chromatogr B.* 2006;837:1–10.
22. Ellerbe P, Cohen A, Welch MJ, et al. Determination of serum uric acid by isotope dilution mass spectrometry as a new candidate reference method. *Anal Chem.* 1990;62: 2173–2177.
23. Young DS. *AACC Effects on Clinical Laboratory Tests: Drugs, Disease, Herbs and Natural Products*. https://www.clinfx.com. Published 2013. Accessed February 2016.
24. Uric acid nephrolithiasis. *UpToDate*. http://www.uptodate.com. Updated January 7, 2016. Accessed February 2016.
25. Medical Dictionary: MedlinePlus. U.S. National Library of Medicine. https://www.nlm.nih.gov/medlineplus/mplusdictionary.html. Page updated April 18, 2012. Accessed February 2016.
26. Kutzing MK, Firestein BL. Altered uric acid levels and disease states. *J Pharmacol Exp Ther.* 2008;324:1–7.
27. Creatinine. Lab Tests Online. https://labtestsonline.org/understanding/analytes/creatinine. Reviewed August 31, 2015. Accessed March 2021.
28. Creatinine Clearance. Lab Tests Online. https://labtestsonline.org/tests/creatinine-clearance. Accessed March 25, 2021.
29. National Institutes of Health | National Institute of Diabetes and Digestive and Kidney Diseases. National Kidney Disease Education Program. https://www.niddk.nih.gov/health-information/community-health-outreach/information-clearinghouses/nkdep. Accessed March 25, 2021.
30. Myers GL, Miller WG, Coresh J, et al. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. *Clin Chem.* 2006;52:5–18.
31. Levey AS, Coresh J, Greene T, et al. Expressing the modification of diet in renal disease study equation for estimating glomerular filtration rate with standardized serum creatinine vales. *Clin Chem.* 2007;53:766–772.
32. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150:604–612.
33. Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. *JAMA*. 2019;322(2):113–114. doi:10.1001/jama.2019.5774
34. Eneanya ND, Kostelanetz S, Mendu ML. Race-free biomarkers to quantify kidney function: Health equity lessons learned from population-based research. *Am J Kidney Dis*. 2021;77(5):667–669. doi:10.1053/j.ajkd.2020.12.001
35. National Kidney Foundation. Removing race from estimates of kidney function. https://www.kidney.org/news/removing-race-estimates-kidney-function. Accessed March 25, 2021.
36. Schwartz GJ, Munoz A, Schneider MF, et al. New equations to estimate GFR in children with CKD. *J Am Soc Nephrol.* 2009;20:629–637.
37. National Kidney Foundation. Creatinine-based "Bedside Schwartz" equation (2009). https://www.kidney.org. Accessed March 25, 2021.

38. Jaffe M. Uber den niederschlag welchen pikrinsaure in normalen harn erzeugt und uber eine neue reaktion des kreatinins. *Z Physiol Chem.* 1886;10:391.
39. Folin O, Wu H. System of blood analysis. *J Biol Chem.* 1919;31:81.
40. Haeckel R. Assay of creatinine in serum with use of Fuller's earth to remove interferents. *Clin Chem.* 1981;27:179–183.
41. Larsen K. Creatinine assay by a reaction kinetic principle. *Clin Chim Acta.* 1972;41:209–217.
42. Bowers LD, Wong ET. Kinetic serum creatinine assays. II. A critical evaluation and review. *Clin Chem.* 1980;26:555–561.
43. Creatinine Test Methodology. *Kodak Ektachem Clinical Chemistry Products*. Rochester, NY: Eastman Kodak Company; 1992: Pub No MP2-49.
44. Beyer C. Creatine measurement in serum and urine with an automated enzymatic method. *Clin Chem.* 1993;39: 1613–1619.
45. Smith-Palmer T. Separation methods applicable to urinary creatine and creatinine. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2002;781:93–106.
46. Dufour DR. Liver disease. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 5th ed. St. Louis, MO: Elsevier Saunders; 2012:1637–1693.
47. Hood JL, Scott MG. Physiology and disorders of water, electrolyte, and acid-base metabolism. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 5th ed. St. Louis, MO: Elsevier Saunders; 2012:1609–1635.
48. Bhatia V, Singh R, Acharya SK. Predictive value of arterial ammonia for complications and outcome in acute liver failure. *Gut.* 2006;55:98–104.
49. Chalela JA. Acute toxic-metabolic encephalopathy in children. *UpToDate*. http://www.uptodate.com. Updated June 22, 2015. Accessed February 2016.
50. Fitzgerald JF, Clark JH, Angelides AG, et al. The prognostic significance of peak ammonia levels in Reye's syndrome. *Pediatrics.* 1982;70:997–1000.
51. Pasquali M, Longo N. Newborn screening and inborn errors of metabolism. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 5th ed. St. Louis, MO: Elsevier Saunders; 2012:2045–2082.
52. Conway EJ, O'Malley E. Microdiffusion methods: ammonia and urea using buffered absorbents (revised methods for ranges greater than 10 µg N). *Biochem J.* 1942;36:655–661.
53. Mondzac A, Ehrlich GE, Seegmiller JE. An enzymatic determination of ammonia in biological fluids. *J Lab Clin Med.* 1965;66:526–531.
54. Ammonia Test Methodology. Roche/Hitachi Products. Indianapolis, IN: Roche Diagnostics Corporation; 2001.
55. VITROS AMON Slides. Instructions for Use. VITROS Chemistry Products. Rochester, NY: Ortho-Clinical Diagnostics; 2002:Pub No MP2–90.
56. Willems D, Steenssens W. Ammonia determined in plasma with a selective electrode. *Clin Chem.* 1988;34:2372.
57. Albrecht J, Zielinska M, Norenberg MD. Glutamine as a mediator of ammonia neurotoxicity: A critical appraisal. *Biochem Pharmacol.* 2010;80:1303–1308.

Enzymes

Donna J. Spannaus-Martin

CHAPTER OUTLINE

KEY TERMS

Activation energy
Activators
Apoenzyme
Coenzyme
Cofactor
Enzyme–substrate (ES) complex
First-order kinetics
Holoenzyme
Hydrolase
International unit (IU)
Isoenzyme
Isoform
Isomerases
Kinetic assays
Ligases
Lyases
Michaelis-Menten constant
Oxidoreductase
Transferases
Zero-order kinetics
Zymogen

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define the term enzyme, including physical composition and structure.
- Classify enzymes according to the International Union of Biochemistry.
- List the different factors affecting the rate of an enzymatic reaction.
- Diagram enzyme kinetics, including zero-order and first-order kinetics.
- Explain why the measurement of plasma enzyme concentrations is clinically useful.
- Differentiate which enzymes are useful in the diagnosis of various disorders, including cardiac, hepatic, bone, and muscle malignancies, and acute pancreatitis.
- Discuss the tissue sources, diagnostic significance, clinical assays, and the sources of error for the following enzymes: creatine kinase, lactate

dehydrogenase, aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, acid phosphatase, γ-glutamyltransferase, 5'-nucleotidase, amylase, lipase, and glucose-6-phosphate dehydrogenase.

- Evaluate patient plasma enzyme concentrations in relation to disease states.
- State the clinical importance for detecting macroenzymes.
- Specify the role of enzymes in drug metabolism.

Enzymes are specific proteins that catalyze biochemical reactions without altering the equilibrium point of the reaction or being consumed or changed in composition. The other substances in the reaction, *substrates*, are converted to products. The catalyzed reactions are frequently specific and essential to physiologic functions, such as the hydration of carbon dioxide, muscle contraction, nutrient degradation, and energy use. Found in all body tissues, enzymes frequently appear in the serum following cellular injury, or sometimes from degraded cells, in smaller amounts. Certain enzymes, such as those that facilitate coagulation, are specific to plasma and, therefore, are present in significant concentrations in plasma. Plasma or serum enzyme levels are often useful in the diagnosis of particular diseases or physiologic abnormalities. This chapter discusses the general properties and principles of enzymes, aspects relating to the clinical diagnostic significance of specific physiologic enzymes, and assay methods for those enzymes.

General Properties and Definitions

Enzymes catalyze many specific reactions. These reactions are facilitated by the enzyme structure and several other factors. As a protein, each enzyme contains a specific amino acid sequence (*primary structure*), with the resultant polypeptide chains adopting bends and turns (*secondary structure*), and folding into a three-dimensional structure (*tertiary structure*) that results in structural features such as binding cavities. If an enzyme contains more than one polypeptide unit, each polypeptide is called a subunit and the *quaternary structure* refers to the binding and interactions between the subunits. Each enzyme contains an *active site*, often a water-free cavity, where the substrate interacts with particular amino acid residues. An *allosteric site,* a cavity other than the active site, binds regulator molecules and, thereby, may be influentially significant to the basic enzyme structure.

Even though a particular enzyme maintains the same catalytic function throughout the body, different forms of that enzyme may exist in various types of tissue within the same individual. The different forms may differ in select physical properties, such as electrophoretic mobility, solubility, or resistance to inactivation. The term **isoenzyme** is generally used when discussing such forms of the enzymes, although the International Union of Biochemistry (IUB) suggests restricting this term to multiple forms of similar genetic origin. An **isoform** results when an enzyme is subject to posttranslational modifications with a functional group added to an amino acid. Isoenzymes and isoforms contribute to heterogeneity in properties and function of enzymes because these measured properties are influenced by changes in amino acid chemistry and the resulting changes in structural features.

CASE STUDY 8.1, PART 1

Let's meet Carl, a 57-year-old moderately overweight White male. Carl visits his family physician with a symptom of "indigestion" of 5 days' duration. He has also had bouts of sweating, malaise, and headache. His blood pressure is 140/105 mm Hg; his family history includes a father who died at age 62 of acute myocardial infarction (AMI) secondary to diabetes mellitus.

In addition to the basic enzyme structure, a nonprotein molecule, called a **cofactor**, may be necessary for enzyme activity. Inorganic cofactors, such as metal ions, are called **activators**. A **coenzyme** is an organic cofactor, such as nicotinamide adenine dinucleotide (NADH). When bound tightly to the enzyme, the coenzyme is called a *prosthetic group*. The protein portion (**apoenzyme**), binding its respective coenzyme, forms a complete and active system, a **holoenzyme**.

Some enzymes, mostly digestive enzymes, are originally secreted from the organ of production in a structurally inactive form, called a *proenzyme* or **zymogen**. The zymogen form is thus prevented from acting until needed. The zymogen can be activated by other enzymes that alter the structure of the proenzyme, typically with the removal of some sequence that masks the active form. For example, this mechanism prevents digestive enzymes from digesting the tissue where they are synthesized.

Enzyme Classification and Nomenclature

To standardize enzyme nomenclature, the Enzyme Commission (EC) of the IUB adopted a classification system in 1961, with the standards revised in 1972 and 1978. The IUB system assigns a *systematic name* to each enzyme, defining the substrate acted on, the reaction catalyzed, and, possibly, the name of any coenzyme involved in the reaction. Because many systematic names are lengthy, a more usable, trivial, *recommended name* is also assigned by the IUB system.[1]

In addition to naming enzymes, the IUB system identifies each enzyme by an EC numerical code containing four digits separated by decimal points. The first digit places the enzyme in one of the following six classes:

1. **Oxidoreductases** catalyze an oxidation–reduction reaction between two substrates.
2. **Transferases** catalyze the transfer of a group other than hydrogen from one substrate to another.
3. **Hydrolases** catalyze hydrolysis of various bonds.
4. **Lyases** catalyze removal of groups from substrates without hydrolysis; the product contains double bonds.
5. **Isomerases** catalyze the interconversion of geometric, optical, or positional isomers.
6. **Ligases** catalyze the joining of two substrate molecules, coupled with breaking of the pyrophosphate bond in adenosine triphosphate (ATP) or a similar compound.

The second and third digits of the EC code number represent the subclass and sub-subclass of the enzyme, respectively, divisions that are made according to criteria specific to the enzymes in the class. The final number is the serial number specific to each enzyme in a sub-subclass. **Table 8.1** provides the EC code numbers, as well as the systematic and recommended names, for enzymes frequently measured in the clinical laboratory.

Table 8.1 Classification of Frequently Quantified Enzymes

Class	Recommended Name	Common Abbreviation	Standard Abbreviation	EC Code No.	Systematic Name
Oxidoreductases	Lactate dehydrogenase	LD	LD	1.1.1.27	L-Lactate:NAD^+ oxidoreductase
	Glucose-6-phosphate dehydrogenase	G-6-PDH	G-6-PD	1.1.1.49	D-Glucose-6-phosphate:$NADP^+$ 1-oxidoreductase
	Glutamate dehydrogenase	GLD	GLD	1.4.1.3	L-Glutamate:NAD(P) oxidoreductase, deaminase
Transferases	Aspartate amino transferase	GOT (glutamate oxaloacetate transaminase)	AST	2.6.1.1	L-Aspartate:2-oxoglutarate aminotransferase
	Alanine amino-transferase	GPT (glutamate transaminase)	ALT	2.6.1.2	L-Alanine:2-oxoglutarate aminotransferase

(continues)

Table 8.1 Classification of Frequently Quantified Enzymes *(continued)*

Class	Recommended Name	Common Abbreviation	Standard Abbreviation	EC Code No.	Systematic Name
	Creatine kinase	CPK (creatine phosphokinase)	CK	2.7.3.2	ATP:creatine *N*-phosphotransferase
	γ-Glutamyl-transferase	GGTP	GGT	2.3.2.2	(5-Glutamyl)peptide: amino acid-5-glutamyltransferase
	Glutathione-*S*-transferase	α-GST	GST	2.5.1.18	Glutathione transferase
	Glycogen phosphorylase	GP	GP	2.4.1.1	1,4-D-Glucan: orthophosphate α-D-glucosyltransferase
	Pyruvate kinase	PK	PK	2.7.1.40	Pyruvate kinase
Hydrolases	Alkaline phosphatase	ALP	ALP	3.1.3.1	Orthophosphoric monoester phosphohydrolase (alkaline optimum)
	Acid phosphatase	ACP	ACP	3.1.3.2	Orthophosphoric monoester phosphohydrolase (acid optimum)
	α-Amylase	AMY	AMS	3.2.1.1	1,4-D-Glucan glucanohydrolase
	Cholinesterase	PCHE	CHE	3.1.1.8	Acetylcholine acyl-hydrolase
	Chymotrypsin	CHY	CHY	3.4.21.1	Chymotrypsin
	Elastase-1	E1	E1	3.4.21.36	Elastase
	5′-Nucleotidase	NTP	5NT	3.1.3.5	5′-Ribonucleotide phosphohydrolase
	Triacylglycerol lipase		LPS	3.1.1.3	Triacylglycerol acylhydrolase
	Trypsin	TRY	TRY	3.4.21.4	Trypsin
Lyases	Aldolase	ALD	ALD	4.1.2.13	D-D-Fructose-1,6-bisdiphosphate D-glyceraldehyde3-phosphate-lyase
Isomerases	Triosephosphate isomerase	TPI	TPI	5.3.1.1	Triosephosphate isomerase
Ligase	Glutathione synthetase	GSH-S	GSH-S	6.3.2.3	Glutathione synthase

Data from Competence Assurance, ASMT. *Enzymology, An Educational Program.* Bethesda, MD: RMI Corporation; 1980.

Table 8.1 also lists common and standard abbreviations for commonly analyzed enzymes. Without IUB recommendation, capital letters have been used as a convenience to identify enzymes. The common abbreviations, sometimes developed from previously accepted names for the enzymes, were used until the standard abbreviations listed in the table were developed.[2,3] These standard abbreviations are used in the United States and are used later in this chapter to indicate specific enzymes.

Enzyme Kinetics

Enzyme Catalysis

Some chemical reactions will occur at a slow rate if there is not enough kinetic energy to drive the reaction to the formation of products (uncatalyzed reaction). Other chemical reactions can occur spontaneously if the free energy or available kinetic energy is higher for the reactants than for the products. An obstacle that slows the reaction is often the formation of the transition state of the reaction. The *transition state* is the intermediate state formed after the reactants exist, but before the product is formed and energy is required for the formation of the transition state and conversion to product. The reaction proceeds toward the lower energy if a sufficient number of the reactant molecules possess enough excess energy to break their chemical bonds and collide to form new bonds. The excess energy needed to induce the transition state is called **activation energy** of the reaction and is the energy required to raise all molecules in one (1) mole of a compound at a certain temperature to the transition state at the peak of the energy barrier. Reactants possessing enough energy to overcome the energy barrier result in product formation.

One way to provide more energy for a reaction is to increase the temperature, which will increase intermolecular collisions; however, this does not normally occur physiologically. Enzymes catalyze physiologic reactions by lowering the activation energy level that the reactants (*substrates*) require to form the transition state, allowing the reaction to occur (**Figure 8.1**). The reaction may then more readily achieve a state of equilibrium; the extent to which the reaction proceeds to product depends on the number of substrate molecules that pass the energy barrier.

The general relationship among enzyme, substrate, and product may be represented as follows:

$$E + S \rightarrow ES \rightarrow E + P \qquad \text{(Eq. 8.1)}$$

where E is enzyme, S is substrate, ES is enzyme–substrate complex, and P is product.

The ES complex is a physical association of a substrate to the active site of an enzyme. The structural arrangement of amino acid residues within the enzyme directs the formation of secondary structures and guides the formation of its three-dimensional active site. Enzymes are induced to form a tighter active structure after the binding of substrate drives a rearrangement, resulting in enhanced substrate binding; this is the *Induced Fit* model of enzyme action.

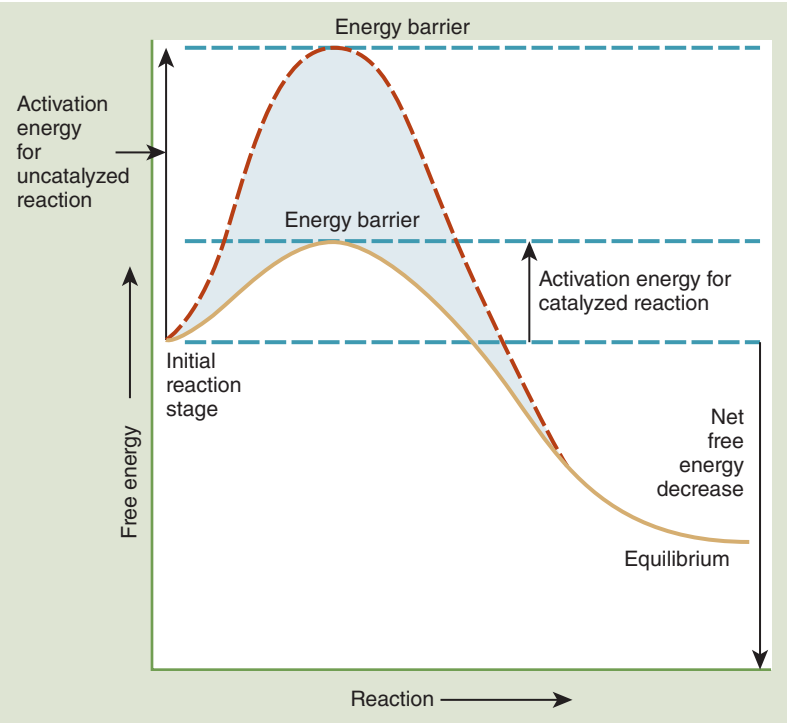


Figure 8.1 Energy versus progression of reaction, indicating the energy barrier that the substrate must surpass to react with and without enzyme catalysis. The enzyme considerably reduces the free energy needed to activate the reaction.

The transition state for the ES complex has a lower energy of activation to form the transition state than S alone, so that the reaction proceeds more effectively after the ES complex is formed. An actual reaction may involve several substrates and products.

Different enzymes show specificity to substrates in different extents or respects. Certain enzymes exhibit *absolute specificity*, meaning that the enzyme combines with only one substrate and catalyzes only the one corresponding reaction. Other enzymes are *group specific* because they combine with all substrates containing a particular functional group, such as a phosphate ester. Still other enzymes are specific to chemical bonds and thereby exhibit *bond specificity.*

Stereoisomeric specificity refers to enzymes that predominantly combine with only one optical isomer of a certain compound. For example, an enzyme may bind more than one molecule of substrate, and this may occur in a cooperative fashion. Binding of one substrate molecule, therefore, may facilitate binding of additional substrate molecules. Likewise, some enzymes may act on more than one substrate and have preferred binding order for the substrates. In some examples, binding may occur via an ordered binding sequence in which one substrate must bind first before the second substrate can bind for the reaction to proceed. Other enzymes show a random order of substrate binding in which either of two substrates can bind first, followed by the other substrate driving the reaction to form the product.

Factors That Influence Enzymatic Reactions

Substrate Concentration

The rate at which an enzymatic reaction proceeds and whether the forward or reverse reaction occurs can be influenced by several reaction conditions. One major influence on enzymatic reactions is substrate concentration. In 1913, Leonor Michaelis and Maud Menten hypothesized the role of substrate concentration in formation of the **enzyme–substrate (ES) complex**. According to their hypothesis, represented in **Figure 8.2**, the substrate readily binds to free enzyme at a low substrate concentration. The reaction rate steadily increases when more substrate is added as substrate increasingly associates with enzyme. During this process, the reaction is following **first-order kinetics** because the rate of the reaction is directly proportional to substrate concentration. If the substrate concentration is high enough to saturate all available enzyme, the reaction

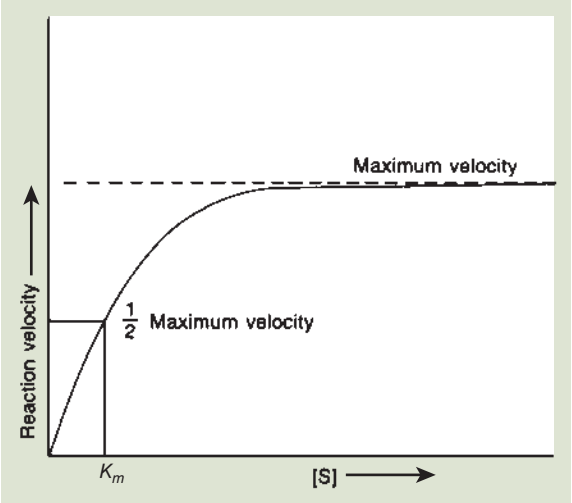


Figure 8.2 Michaelis-Menten curve of velocity versus substrate concentration for enzymatic reaction. K_m is the substrate concentration at which the reaction velocity is half of the maximum level.

velocity has reached its maximum. When product is formed, the resultant free enzyme immediately becomes available to combine with available free substrate. At this point the reaction is in **zero-order kinetics**, and the reaction rate depends only on enzyme concentration.

The Michaelis-Menten hypothesis describes the relationship between reaction velocity and substrate concentration, which is represented mathematically as follows:

$$V = \frac{V_{max}[S]}{K_m + [S]} \qquad \text{(Eq. 8.2)}$$

where V is measured velocity of reaction, V_{max} is maximum velocity, [S] is substrate concentration, and K_m is Michaelis-Menten constant of enzyme for a specific substrate.

The **Michaelis-Menten constant** (K_m), derived from the theory of Michaelis and Menten, is a constant for a specific enzyme and substrate under defined reaction conditions and can provide an expression of the relationship between the velocity of an enzymatic reaction and substrate concentration. The assumptions are made that an equilibrium among E, S, ES, and P is established rapidly and that the reverse reaction, E + P → ES, r is negligible. The rate-limiting step is the formation of product and enzyme from the ES complex (ES → E + P). When reached, maximum velocity is fixed, and the reaction rate is a function of only the enzyme concentration. A specific variable, the Michaelis Constant, or K_m, is characteristic of each enzyme-substrate pair under the reaction conditions

being used and is determined by a series of rate constants for the reaction. As designated in Figure 8.2, K_m is specifically the substrate concentration at which the enzyme yields half the possible maximum velocity. Therefore, K_m reflects the amount of substrate needed for a particular enzymatic reaction.

$$V = \frac{V_{max}[S]}{K_m + [S]}$$

Theoretically, V_{max} and then K_m could be determined from the plot in Figure 8.2. However, V_{max} is difficult to determine from the hyperbolic plot and often not actually achieved in enzymatic reactions. A more accurate and convenient determination of V_{max} and K_m is made through mathematical manipulations of the Michaelis-Menten equation to yield a linear equation. The Lineweaver-Burk plot, a double reciprocal plot of the Michaelis-Menten equation, yields a straight line (**Figure 8.3**) after the reciprocal is taken of both the substrate concentration term and the velocity term of the Michaelis-Menton reaction. The equation becomes:

$$\frac{1}{V} = \frac{K_m}{V_{max}} \times \frac{1}{[S]} + \frac{1}{V_{max}} \qquad \text{(Eq. 8.3)}$$

Plotting the reciprocal of the velocity $\left(\frac{1}{V}\right)$ on the y-axis vs. the reciprocal of the substrate concentration $\left(\frac{1}{[S]}\right)$ on the x-axis provides a straight line with a slope of $\frac{K_m}{V_{max}}$. The intercept on the y-axis is the value of $\frac{1}{V_{max}}$ for the reaction and the intercept on the x-axis provides a value of $\frac{-1}{K_m}$ for the reaction.

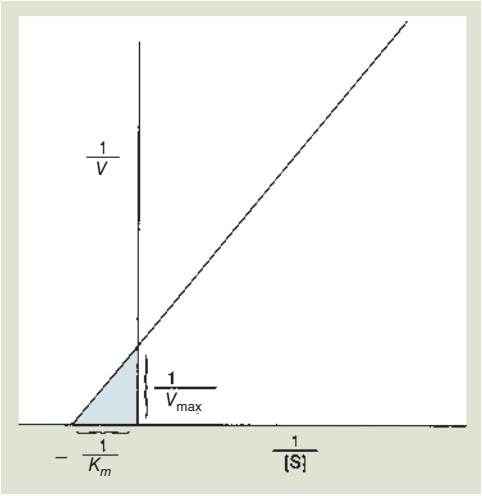


Figure 8.3 Lineweaver-Burk transformation of Michaelis-Menten curve. V_{max} is the reciprocal of the *x*-intercept of the straight line. K_m is the negative reciprocal of the *x*-intercept of the same line.

Enzyme Concentration

Because enzymes catalyze physiologic reactions, the enzyme concentration affects the rate of the catalyzed reaction. As long as the substrate concentration exceeds the enzyme concentration, the velocity of the reaction is proportional to the enzyme concentration. The higher the enzyme concentration, the faster the reaction will proceed because more enzyme is available to bind substrate.

pH

Enzymes are proteins that carry net molecular charges because many of the amino acid side chains carry a functional group capable of carrying an ionization charge. As such, changes in the pH of the solution can influence the ionization state of the enzyme. Changes in pH may also denature an enzyme or influence its ionic state, resulting in structural changes or a change in the charge on amino acid residues in the active site, affecting enzyme function. Hence, each enzyme operates within a specific pH range and maximally at a specific pH. Most physiologic enzymatic reactions occur in the pH range of 7.0 to 8.0, but some enzymes are active in wider pH ranges than others. In the laboratory, the pH for a reaction is carefully controlled at the optimal pH by means of appropriate buffer solutions.

Temperature

Increasing temperature usually increases the rate of a chemical reaction by increasing the movement of molecules, the rate at which intermolecular collisions occur, and the energy available for the reaction enabling achieving the transition state and providing sufficient activation energy. Typically, for each 10°C increase in temperature, the rate of the reaction will approximately double. This is likewise the case with enzymatic reactions, unless the temperature is high enough to cause denaturation of the protein structure of the enzyme.

Each enzyme functions optimally at a particular temperature, which is influenced by other reaction variables, especially the total time for the reaction. The optimal temperature is usually close to that of the physiologic environment of the enzyme; however,

some denaturation may occur at the human physiologic temperature of 37°C. The rate of denaturation increases as the temperature increases and usually is significant at 40°C to 50°C. Enzymes show various dependencies on temperature related to the thermodynamic properties involved in their secondary and tertiary structures.

Because low temperatures render enzymes reversibly inactive, many serum or plasma specimens for enzyme measurement are refrigerated or frozen to prevent activity loss until analysis. Storage procedures may vary from enzyme to enzyme because of individual stability characteristics. Repeated freezing and thawing, however, tends to denature protein and should be avoided.

Because of their temperature sensitivity, enzymes should be analyzed under strictly controlled temperature conditions. Incubation temperatures should be accurate within ±0.1°C. Laboratories usually attempt to establish an analysis temperature for routine enzyme measurement of 25°C, 30°C, or 37°C. Attempts to establish a universal temperature for enzyme analysis have been unsuccessful, and therefore, reference ranges for enzyme levels may vary significantly among laboratories. In the United States, however, 37°C is most commonly used.

Cofactors

Simple cofactors are nonprotein entities that often bind to particular enzymes and provide a necessary function before a reaction occurs. Common *activators* (inorganic cofactors) are metallic (Ca^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Zn^{2+}, and K^{+}). The activator may be essential for the reaction or may only enhance the reaction rate in proportion with concentration. Activators function via various mechanisms, including alternating the spatial configuration of the enzyme for proper substrate binding, linking substrate to the enzyme or coenzyme, or undergoing oxidation or reduction.

Organic cofactors, called coenzymes, include compounds such as nucleotide phosphates and vitamins, and often serve a more specialize function for enzymes. When bound tightly to the enzyme, coenzymes are called *prosthetic groups*. Coenzymes also often serve as second substrates for enzymatic reactions. For example, as a cofactor, nicotinamide adenine dinucleotide as NAD^{+} may be reduced to NADH in a reaction in which the primary substrate is oxidized. Increasing coenzyme concentration will increase the velocity of an enzymatic reaction in a manner synonymous with increasing substrate concentration. When quantifying an enzyme that requires a particular cofactor, that cofactor should always be provided in excess so that the extent of the reaction does not depend on the concentration of the cofactor.

Inhibitors

The primary basis for an enzymatic reaction is the initial association with its substrate. Enzymatic reactions may not progress normally if a particular substance, an *inhibitor*, interferes with the reaction. Enzyme–inhibitor interactions are reversible, reaching an equilibrium depending on the concentration of the inhibitor. A compound that shares some structural feature found in the substrate will typically physically bind to the same form of the enzyme that the substrate binds, often in the active site of an enzyme, and compete with the substrate for a place in the active site. Such inhibitors are called *competitive inhibitors*, usually binding to the free form of the enzyme, E. With a substrate concentration significantly higher than the concentration of the competitive inhibitor, the inhibition is reversible because the substrate is more likely than the inhibitor to bind in the active site and the enzyme structure has not been destroyed by the inhibitor.

Enzymes may also be sensitive to the presence of other compounds that are called *noncompetitive inhibitors*. Most commonly, these inhibitors associate with enzymes at a place other than the active site showing allosteric inhibition. Noncompetitive inhibitors may allow substrate binding but will inhibit the formation of product. In this situation, the inhibitor may bind to both the enzyme (E) and enzyme–substrate complex (ES) with the same affinity. A related type of inhibitor may be identified as a *mixed inhibitor* with the capacity to bind both the E and ES forms of the enzyme, but with different affinities for the two forms. A third pattern of inhibition, *uncompetitive inhibition*, may be observed in which the inhibitor binds only to the ES complex; increasing substrate concentration results in more ES complexes to which the inhibitor binds and, thereby, increases the inhibition. The enzyme–substrate–inhibitor complex (ESI) does not yield product. In these situations, binding of an inhibitor to the free E or ES complex is defined by an equilibrium constant for the formation of an EI or ESI complex.

The various types of inhibitors may be distinguished by the changes identified in the double-reciprocal plots obtained for the enzyme reaction data. Each of the reversible types of inhibition are unique with respect to effects on the V_{max} and K_m parameters of enzymatic reactions (**Figure 8.4**).

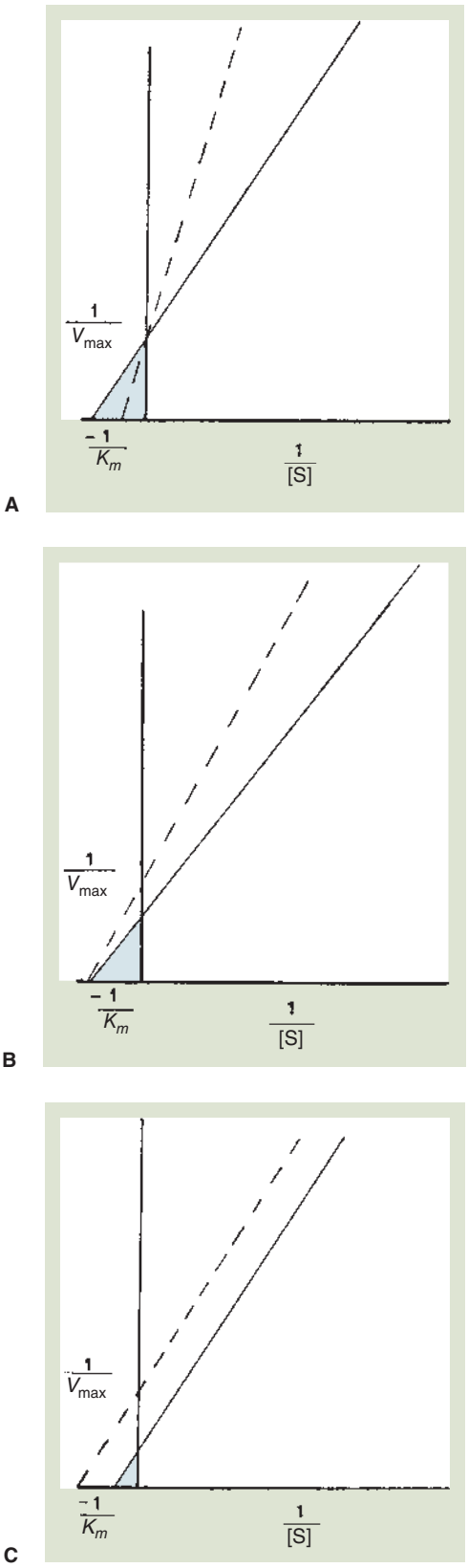


Figure 8.4 Normal Lineweaver-Burk plot (*solid line*) compared with each type of enzyme inhibition (*dotted line*). **(A)** Competitive inhibition V_{max} unaltered; K_m appears increased. **(B)** Noncompetitive inhibition V_{max} decreased, K_m unchanged. **(C)** Uncompetitive inhibition V_{max} decreased; K_m appears decreased.

For competitive inhibition, the primary effect is on the interaction of the enzyme with the substrate. Although the K_m is a constant for each enzyme–substrate pair and cannot be altered, and because the amount of substrate needed to achieve a particular velocity is higher in the presence of a competing inhibitor, the K_m appears to increase with the presence of the inhibitor. The effect of the inhibitor can be overcome by adding excess substrate to bind the enzyme. The amount of the inhibitor will become negligible by comparison, and the reaction will reach the same maximum velocity as an uninhibited reaction.

The substrate and inhibitor may both be bound to an enzyme in a noncompetitive inhibition. The inhibitor may bind the E form or associate with the ES form and inhibit the formation of product with the affinity of the inhibitor for E and ES being the same. Thus, for noncompetitive inhibition, the maximum reaction velocity will not be achieved, and the value measured will be decreased. Inhibitor binding does not influence substrate binding, so the value of K_m is unchanged. Mixed type noncompetitive inhibitors have the potential to bind both free E and the ES complex with different affinities so changes in both the apparent V_{max} and the apparent K_m values may be measurable.

Because uncompetitive inhibition requires the formation of an ES complex, increasing substrate concentration can also effectively increase inhibition. Therefore, maximum velocity equal to that of an uninhibited reaction cannot be achieved, and the K_m appears to be decreased from the diminished functional significance of the ES complex.

Irreversible inhibitors are also known and are better called inactivators. These compounds may bind to proteins and in a time-dependent manner react with an amino acid side chain in the protein. Loss of enzyme activity may result from a reaction at the active site or at another site that causes a structural change in the protein. Such inactivators typically bind the enzyme independent of substrate, so increasing the concentration of substrate will not prevent or reverse inactivation.

Measurement of Enzyme Activity

Enzymes are usually present in very small quantities in biologic fluids and are often difficult to isolate from similar compounds, so a convenient method of enzyme quantification is needed for the measurement of catalytic activity, which is then related to concentration. Common methods might photometrically measure an increase in product concentration, a decrease

in substrate concentration, a decrease in coenzyme concentration, or an increase in the concentration of an altered coenzyme.

When all substrates and any coenzyme are present in excess in an enzymatic reaction, the amount of substrate or coenzyme used or the amount of product or altered coenzyme formed will depend only on the amount of enzyme present to catalyze the reaction. Using enzyme-catalyzed reactions to measure enzyme concentrations, therefore, is always performed in zero-order kinetics, with the substrate in sufficient excess to ensure that no more than 20% of the available substrate is converted to product. Any coenzymes also must be in excess. When these conditions are met, the only factor affecting the rate of the reaction will be the concentration of the enzyme in the specimen being analyzed.

NADH is a coenzyme frequently measured in the laboratory. NADH absorbs light at 340 nm, whereas NAD^+ does not, and a change in absorbance at 340 nm is easily measured. Such reactions follow the change of available NADH at 340 nm, with the rate calculated using the molar absorptivity (6.22×10^3 liter $\cdot$ mol^{-1} cm^{-1}) of NADH, and lastly conversion to units of enzyme activity. NAD^+ or NADH is often convenient as a reagent for a *coupled enzyme* assay when neither NAD^+ nor NADH is a coenzyme for the initial reaction. In coupled enzyme assays, one or more enzymes are added in excess as reagents and multiple reactions are catalyzed. After the enzyme under analysis catalyzes its specific reaction, a product of that reaction becomes the substrate for the reaction catalyzed by an intermediate auxiliary enzyme. The product of the intermediate reaction becomes the substrate for the final reaction, which is catalyzed by an indicator enzyme and commonly involves the conversion of NAD^+ to NADH or vice versa. For many assays, the second enzymatic reaction provides an indicator, in which case the auxiliary enzyme is also the indicator enzyme.

When performing an enzyme quantification in zero-order kinetics, inhibitors must not be present, and other variables that may influence the rate of the reaction must be carefully controlled. All substrates and cofactors must be present in excess, so substrate concentration does not change the kinetics to first or second order. (Second-order kinetic reactions occur when concentrations of two different substrates are needed to produce a product, so both substrates affect the rate of reaction. Second-order reactions are not used in the clinical laboratory and therefore not discussed in this chapter.) A constant pH should be maintained by means of an appropriate buffer solution. The temperature should be constant within ±0.1°C throughout the assay at a temperature at which the enzyme is active (usually 25°C, 30°C, or 37°C).

During the progress of the reaction, the analysis time must be carefully designed and selected. When the enzyme is initially introduced to the substrate to start the reaction, the high substrate concentration steadily combines with available enzyme, driving the forward reaction to product formation. As the enzyme becomes effectively saturated, the rates of product formation, release of enzyme, and recombination with more substrate reach an equilibrium such that the net reaction proceeds linearly. After a time, usually 6 to 8 minutes after reaction initiation, the reaction rate decreases as the substrate is depleted. For some enzyme reactions, the amount of product present may be sufficient to bind to the enzyme in its active site and inhibit the reaction with the substrate. Hence, meaningful measurement of enzyme activity is performed during the linear phase of the reaction.

One of two general methods may be used to measure the extent of an enzymatic reaction: (1) fixed-time and (2) continuous-monitoring or kinetic assay. In the *fixed-time method*, the reactants are combined, the reaction proceeds for a designated time, the reaction is stopped (usually by inactivating the enzyme with a weak acid), and a measurement is made of the amount of reaction that has occurred. The reaction is designed to be linear over the reaction time, achieved by first evaluating the reaction for various times aiming to find a suitable time for adequate measurement of the reaction, typically by substrate loss or product appearance. The greater the reaction measurement, the more enzyme is present.

In *continuous-monitoring* or **kinetic assays**, multiple measurements, usually of absorbance change, are made during the reaction, either at specific time intervals (usually every 30 or 60 seconds) or continuously by a continuous-recording spectrophotometer. Collection of data at specific time intervals is essentially a series of stopped-time assays to construct a linear plot of the reaction. These assays are advantageous over fixed-time methods because the linearity of the reaction may be more adequately verified. If absorbance is measured at intervals, several data points are necessary to increase the accuracy of assessed linearity. Continuous measurements are preferred because any deviation from linearity is readily observable.

The most common cause of deviation from linearity occurs when the amount of enzyme is so elevated that all substrate is used rapidly in the reaction time,

effectively making the linear phase for the reaction quite short. With continuous monitoring, the laboratorian may observe a sudden decrease in the reaction rate (deviation from zero-order kinetics) of a particular determination and may repeat the determination using less patient sample. When possible, the analysis should be repeated using less patient sample rather than diluting the sample in order to eliminate the possibility of the diluent interfering with the reaction. The result obtained will still be multiplied by any dilutional factor. (Sample dilution with saline may be necessary to minimize negative effects in analysis caused by hemolysis or lipemia). Less sample should catalyze less reaction over the same time period to provide for a longer reaction time being observed as linear. Enzyme activity measurements may not be accurate if storage conditions compromise integrity of the protein, if enzyme inhibitors are present, or if required cofactors are not present.

Calculation of Enzyme Activity

When enzymes are quantified relative to their activity rather than a direct measurement of concentration, the units used to report enzyme levels are activity units. The definition for the activity unit must consider conditions that may alter results (e.g., pH, temperature, substrate). Historically, specific method developers frequently established their own units for reporting results and often named the units after themselves (i.e., Bodansky and King units). To standardize the system of reporting quantitative results, the EC defined the **international unit (IU)** as the amount of enzyme that will catalyze the reaction of 1 µmol of substrate per minute while also including descriptions of the specified conditions of temperature, pH, substrates, and activators used in the reaction. Such details are provided to enable other laboratorians to replicate the results. Because specified conditions may vary among laboratories, reference ranges are still often laboratory specific. Enzyme concentration is usually expressed in units per liter (IU/L). The unit of enzyme activity recognized by the International System of Units (Système International d'Unités [SI]) is the katal (mol/s). The mole is the unit for substrate concentration, and the unit of time is the second. Enzyme concentration is then expressed as katals per liter (kat/L) (1.0 IU = 17 nkat).

Measurement of Enzyme Mass

Understanding these units of enzyme activity also requires knowledge of the amount of enzyme available in the reaction, more specifically, the amount of enzyme present as mass or concentration. The relationship between enzyme activity and enzyme quantity is generally linear but should be verified with the conditions used for each enzyme under study. The mass of enzymes may also be determined by electrophoretic techniques, which provide resolution of isoenzymes and isoforms. Immunoassay methodologies that quantify enzyme concentration by mass are also available and are routinely used for quantification of some enzymes, such as creatine kinase isoenzyme CK-MB. Immunoassays may overestimate active enzyme due to possible cross-reactivity with inactive enzymes, such as zymogens, inactive isoenzymes, macroenzymes, or partially digested enzyme.

Ensuring the accuracy of enzyme measurements has long been a concern of laboratorians. The Clinical Laboratory Improvement Amendment of 1988 (CLIA 88) has established guidelines for quality control and proficiency testing for all laboratories. Problems with quality control materials for enzyme testing have been a significant issue. Differences between clinical specimens and control samples include species of origin of the enzyme, integrity of the molecular species, isoenzyme forms, matrix of the solution, addition of preservatives, and lyophilization processes. Many studies have been conducted to ensure accurate enzyme measurements and good quality control materials.[4]

Enzymes as Reagents

Enzymes may be used as reagents to measure many nonenzymatic constituents in serum. For example, glucose, cholesterol, and uric acid are examples of biomolecules frequently quantified by means of enzymatic reactions that measure the concentration of the analyte by following the reaction catalyzed by an enzyme. Enzymes are also used as reagents in methods to quantify analytes that are substrates for the corresponding enzyme. One example, lactate dehydrogenase (LD), may be a reagent when lactate or pyruvate concentrations are evaluated. For such methods, the enzyme is added in excess in a quantity sufficient to provide a complete reaction in a short period.

The nature of the enzyme being used may allow the development of novel approaches for the analysis. Immobilized enzymes are chemically linked to adsorbents, such as agarose or certain types of cellulose, via covalent bonds often through azide groups, diazo, and triazine. When substrate is passed through the preparation, the product is retrieved and analyzed, and the enzyme is present

and free to react with more substrate. Immobilized enzymes are convenient for batch analyses and are more stable than enzymes in a solution. Enzymes are also commonly used as reagents in competitive and noncompetitive immunoassays, such as those used to measure human immunodeficiency virus (HIV) antibodies, therapeutic drugs, and cancer antigens. Enzymes often are covalently linked (conjugated) to a secondary antibody to reflect the amount of the primary antigen. Commonly used enzymes include horseradish peroxidase, alkaline phosphatase (ALP), glucose-6-phosphate dehydrogenase (G6PD), and β-galactosidase. The enzyme in these assays functions as an indicator that reflects either the presence or absence of the analyte.

Enzymes of Clinical Significance

Table 8.2 lists the major enzymes of clinical significance, including their systematic names and clinical significance. The most frequently analyzed clinical enzymes are discussed in this chapter with respect to tissue source, diagnostic significance, assay method, source of error, and reference range.

Creatine Kinase

CK is an enzyme with a molecular weight of approximately 82,000 that is generally associated with ATP regeneration in contractile or transport systems. Its predominant physiologic function occurs in muscle cells, where it is involved in the storage of high-energy creatine phosphate. Every contraction cycle of muscle results in creatine phosphate use with the production of ATP. This results in relatively constant levels of muscle ATP. The reversible reaction catalyzed by CK is shown in **Equation 8.4**.

$$\text{Creatine} + ATP \overset{CK}{\rightleftharpoons} \text{Creatine phosphate} + \text{ADP}$$

(**Eq. 8.4**)

Tissue Source

CK is widely distributed in tissue, with highest activities found in skeletal muscle, heart muscle, and brain tissue. CK is present in much smaller quantities in other tissues, including the bladder, placenta, gastrointestinal tract, thyroid, uterus, kidney, lung, prostate, spleen, liver, and pancreas.

Diagnostic Significance

Due to the high concentrations of CK in muscle tissue, plasma CK levels are frequently elevated in disorders of cardiac and skeletal muscle (myocardial infarction [MI], rhabdomyolysis, and muscular dystrophy). The CK level is considered a sensitive indicator of acute myocardial infarction (AMI) and muscular dystrophy, particularly the Duchenne type. Extreme elevations of CK occur in Duchenne-type muscular dystrophy, with values reaching 50 to 100 times the upper limit of normal (ULN). Although total CK levels are sensitive

CASE STUDY 8.1, PART 2

Remember Carl. At the doctor's office, an electrocardiogram revealed changes from one performed 6 months earlier. The results of the patient's blood work are as follows:

Test	Result	Reference Range
CK	129 U/L	(30–60)
CK-MB	4%	(<6%)
LD	280 U/L	(100–225)
LD	Isoenzymes	LD-1 > LD-2
AST	35 U/L	(5–30)
TnT	12 ng/L	(<= 15 ng/L male; <= 10 mg/L female)

© Tony Wear/Shutterstock.

1. Can a diagnosis of AMI be ruled out in this patient?
2. What further cardiac markers should be run on this patient?
3. Should this patient be admitted to the hospital?

Table 8.2 Major Enzymes of Clinical Significance

Enzyme	Clinical Significance
Acid phosphatase (ACP)	Prostatic carcinoma
Alanine aminotransferase (ALT)	Hepatic disorder
Aldolase (ALD)	Skeletal muscle disorder
Alkaline phosphatase (ALP)	Hepatic disorder
	Bone disorder
Amylase (AMY)	Acute pancreatitis
Angiotensin-converting enzyme (ACE)	Blood pressure regulation
Aspartate aminotransferase (AST)	Myocardial infarction
	Hepatic disorder
	Skeletal muscle disorder
Chymotrypsin (CHY)	Chronic pancreatitis insufficiency
Creatine kinase (CK)	Myocardial infarction
	Skeletal muscle disorder
Elastase-1 (E1)	Chronic pancreatitis insufficiency
Glucose-6-phosphate dehydrogenase (G-6-PD)	Drug-induced hemolytic anemia
Glutamate dehydrogenase (GLD)	Hepatic disorder
γ-Glutamyltransferase (GGT)	Hepatic disorder
Glutathione-*S*-transferase (GST)	Hepatic disorder
Glycogen phosphorylase (GP)	Acute myocardial infarction
Lactate dehydrogenase (LD)	Myocardial infarction
	Hepatic disorder
	Hemolysis
	Carcinoma
Lipase (LPS)	Acute pancreatitis
5′-Nucleotidase	Hepatic disorder
Pseudocholinesterase (PChE)	Organophosphate poisoning
	Genetic variants
	Hepatic disorder
	Suxamethonium sensitivity
Pyruvate kinase (PK)	Hemolytic anemia
Trypsin (TRY)	Acute pancreatitis

indicators of these disorders, they are not entirely specific indicators in as much as CK elevation is found in various other abnormal cardiac and skeletal muscle conditions. Levels of CK also vary with muscle mass and, therefore, may depend on gender, degree of physical conditioning, and age.

Other pathophysiologic conditions in which elevated CK levels occur are hypothyroidism, malignant

hyperpyrexia, and Reye's syndrome. **Table 8.3** lists the major disorders associated with abnormal CK levels. Plasma CK levels and CK/progesterone ratio have been useful in the diagnosis of ectopic pregnancies.[5] Total plasma CK levels have also been used as an early diagnostic tool to identify patients with *Vibrio vulnificus* infections.[6] Elevated CK levels are also occasionally seen in central nervous system disorders such as cerebrovascular accident, seizures, nerve degeneration, and central nervous system shock. Damage to the blood–brain barrier must occur to allow enzyme release to the peripheral circulation.

CK occurs as a dimer consisting of two subunits (B and M type subunits) with the subunits combining to form three distinct molecular forms that can be separated readily. The three isoenzymes have been designated as CK-BB (brain type), CK-MB (hybrid type), and CK-MM (muscle type). On electrophoretic separation, CK-BB will migrate fastest toward the anode and is therefore called *CK-1*. CK-BB is followed by CK-MB (*CK-2*) and, finally, by CK-MM (*CK-3*), exhibiting the slowest mobility (**Figure 8.5**). Table 8.3 indicates the tissue localization of the isoenzymes and the major conditions associated with elevated levels. Because enzyme elevation is found in numerous disorders, the separation of total CK into its various isoenzyme fractions is considered a more specific indicator of various disorders than total levels. Typically, the clinical relevance of CK activity depends more on isoenzyme fractionation than on total levels. Other isoforms occur following cleavage of the carboxyl-terminal amino acid from the M subunit by plasma carboxypeptidase N. Three isoforms have been described for CK-MM and two isoforms for CK-MB. CK-MB2 is the tissue isoform and CK-MB1 is the form that has modified by the plasma carboxypeptidase. Normally, CK-MB 1 is the predominant isoform in the plasma and the CK-MB2/CK-MB1 ratio is ≤1. However, following an AMI, the CK-MB2 isoform becomes elevated in the plasma and the ratio increases, and ratios of >1.5 are considered to be diagnostic for myocardial damage.

In the plasma of healthy people, 94% to 100% of the CK isoenzyme is the MM form. Values for the MB isoenzyme range from undetectable to trace (<6% of total CK). CK-BB is present in very small quantities in the plasma of healthy people, but the level measured can depend on the method of detection. Most techniques cannot detect CK-BB in normal plasma.

CK-MM is the major isoenzyme fraction found in striated muscle with a small amount of CK-MB. The majority of CK activity in heart muscle is also attributed to CK-MM, with approximately 20% a

Table 8.3 Creatine Kinase Isoenzymes—Tissue Localization and Sources of Elevation

Isoenzyme	Tissue	Condition
CK-MM	Heart	Myocardial infarction
	Skeletal muscle	Skeletal muscle disorder
		Muscular dystrophy
		Rhabdomyolysis
		Polymyositis
		Hypothyroidism
		Malignant hyperthermia
		Physical activity
		Intramuscular injection
CK-MB	Heart	Myocardial infarction
	Skeletal muscle	Myocardial injury
		Ischemia
		Angina
		Inflammatory heart disease
		Cardiac surgery
		Duchenne-type muscular dystrophy
		Polymyositis
		Malignant hyperthermia
		Reye's syndrome
		Rocky Mountain spotted fever
		Carbon monoxide poisoning
CK-BB	Brain	Central nervous system shock
	Bladder	Anoxic encephalopathy
	Lung	Cerebrovascular accident
	Prostate	Seizure
	Uterus	Placental or uterine trauma
	Colon	Carcinoma
	Stomach	Reye's syndrome
	Thyroid	Carbon monoxide poisoning
		Malignant hyperthermia
		Acute and chronic renal failure

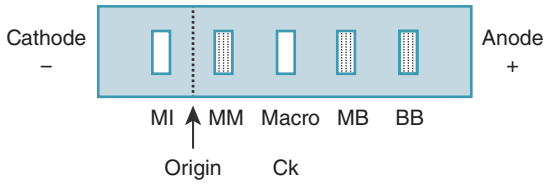


Figure 8.5 Electrophoretic migration pattern of normal and atypical CK isoenzymes.

result of CK-MB.[7] Normal plasma consists of approximately 94% to 100% CK-MM. Injury to both cardiac and skeletal muscle accounts for the majority of cases of CK-MM elevations (Table 8.3). Hypothyroidism results in CK-MM elevations because of the involvement of muscle tissue (increased membrane permeability), the effect of thyroid hormone on enzyme activity, and, possibly, the slower clearance of CK as a result of slower metabolism.

Mild to strenuous activity may contribute to elevated plasma CK levels, as may intramuscular injections. In physical activity, the extent of elevation is variable. However, the degree of exercise in relation to the exercise capacity of the individual is the most important factor in determining the degree of elevation.[8] Patients who are physically well conditioned show lesser degrees of elevation than do patients who are less conditioned. Levels may be elevated for as long as 48 hours following exercise. Plasma CK elevations are generally less than five times ULN following intramuscular injections and usually are not apparent after 48 hours, although elevations may persist for 1 week. The predominant isoenzyme is CK-MM.

The quantity of CK-BB in the tissue (Table 8.3) is usually small and the enzyme has a relatively short half-life (1 to 5 hours), resulting in CK-BB activities that are generally low and transient and not usually measurable when tissue damage occurs. Highest concentrations are found in the central nervous system, the gastrointestinal tract, and the uterus during pregnancy. Although brain tissue has high concentrations of CK, plasma rarely contains CK-BB of brain origin. Because of its molecular size (~82,000 kDa), its passage across the blood–brain barrier is hindered. However, when extensive damage to the brain has occurred, significant amounts of CK-BB can sometimes be detected in the plasma. It has been observed, however, that CK-BB may be significantly elevated in patients with carcinoma of various organs including untreated prostatic carcinoma and other adenocarcinomas. These findings indicate that CK-BB may be a useful tumor-associated marker.[9] The most common causes of CK-BB elevations are central nervous system damage, tumors, childbirth, and the presence of macro-CK, an enzyme–immunoglobulin complex. In most of these cases, the CK-BB level is greater than 5 U/L, usually in the range of 10 to 50 U/L. Other conditions listed in Table 8.3 usually show CK-BB activity below 10 U/L.[10]

The importance of CK isoenzyme separation can be found principally in the detection of myocardial damage. Although cardiac tissue predominantly contains CK-MM as described, the tissue does contain significant quantities of CK-MB, approximately 20% of all CK. Whereas CK-MB is found in small quantities in other tissue, myocardium is essentially the only tissue from which CK-MB enters the plasma in significant quantities. Demonstration of elevated levels of CK-MB greater than or equal to 6% of the total CK is considered a good indicator of myocardial damage, particularly AMI. Following MI, the CK-MB levels begin to rise within 4 to 8 hours, peak at 12 to 24 hours, and return to normal levels within 48 to 72 hours. This time frame must be considered when interpreting CK-MB levels.

Nonenzyme proteins (troponin I and troponin T) have become the preferred markers to detect MI because they are more sensitive and specific markers of myocardial damage. Troponins are released into the bloodstream earlier and persist longer than CK and its isoenzyme CK-MB. The troponin assays appear to be more cost-effective than CK-MB assays and are now usually performed in place of or in addition to CK-MB. The specificity of CK-MB levels in the diagnosis of AMI can be increased if interpreted in conjunction with LD isoenzymes and troponins, and if measured sequentially over a 48-hour period to detect the typical rise and fall of enzyme activity seen in AMI (**Figure 8.6**).

In Duchenne-type muscular dystrophy, there may be some cardiac involvement as well. CK-MB levels in Reye's syndrome also may reflect myocardial damage. Despite the findings of CK-MB levels in disorders other than MI, its presence remains an indicator of AMI.[11] The typical time course of CK-MB elevation following AMI is not found in other conditions.

Numerous reports have been made describing the appearance of unusual CK isoenzyme bands displaying electrophoretic properties that differ from the three major isoenzyme fractions (Figure 8.5).[12–16] These atypical forms are generally of two types and are referred to as macro-CK and mitochondrial CK.

Macro-CK appears to migrate to a position midway between CK-MM and CK-MB. This type of macro-CK largely comprises CK-BB complexed with immunoglobulin. In most instances, the associated

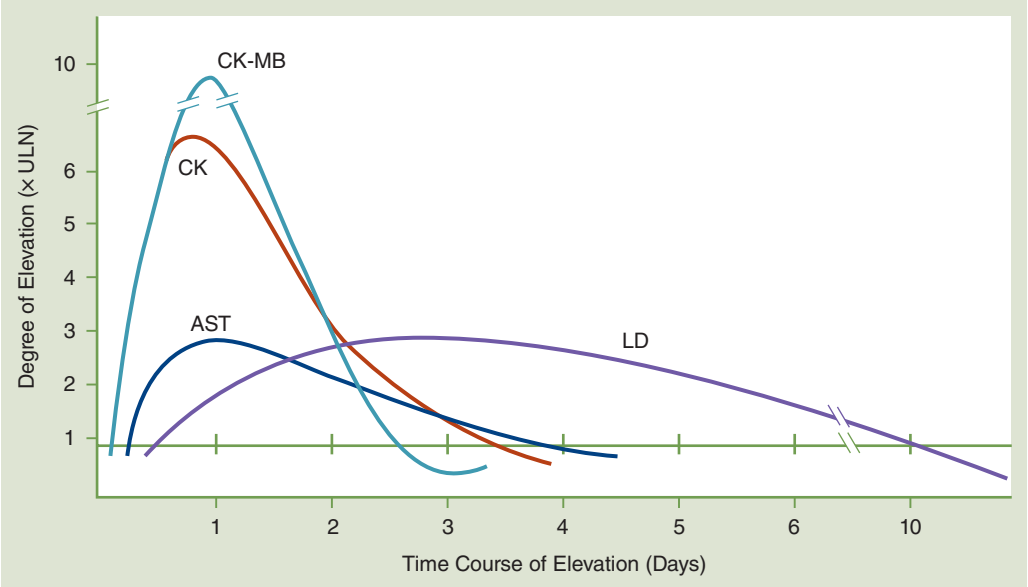


Figure 8.6 Time–activity curves of enzymes in myocardial infarction for AST, CK, CK-MB, and LD. CK, specifically the MB fraction, increases initially, followed by AST and LD. LD is elevated the longest. All enzymes usually return to normal within 10 days. ULN = upper limit of normal.

immunoglobulin is IgG, although a complex with IgA has also been described. The term *macro-CK* has also been used to describe complexes of lipoproteins with CK-MM. The incidence of macro-CK in plasma ranges from 0.8% to 1.6%. Currently, no specific disorder is associated with its presence, although it appears to be age and sex related, occurring most frequently in women older than age 50.

Mitochondrial CK (CK-Mi) is bound to the exterior surface of the inner mitochondrial membranes of muscle, brain, and liver. It migrates to a point cathodal to CK-MM and exists as a dimeric molecule of two identical subunits. It occurs in serum in both the dimeric state and in the form of oligomeric aggregates of high molecular weight (350,000). CK-Mi is not present in normal plasma and is typically not present following MI. The incidence of CK-Mi ranges from 0.8% to 1.7%. For it to be detected in plasma, extensive tissue damage must occur, causing breakdown of the mitochondrion and cell wall. Its presence does not correlate with any specific disease state but appears to be an indicator of severe illness. CK-Mi has been detected in cases of malignant tumor and cardiac abnormalities.

In the absence of any definite correlation between these atypical CK forms and a specific disease state, it seems that their significance relates primarily to the methods used for detecting CK-MB. In certain analytic procedures, these atypical forms may be measured as CK-MB, resulting in erroneously high CK-MB levels.

The identification of these different CK isoenzymes is accomplished using the methods of electrophoresis, ion-exchange chromatography, and several immunoassays. Although mass methods are more sensitive and preferred for quantification of CK-MB, electrophoresis has been commonly used as the reference method for the distinction of the different forms. The electrophoretic properties of the CK isoenzymes are shown in Figure 8.5. Generally, the technique consists of performing electrophoresis on the sample, measuring the reaction using an overlay technique and then visualizing the bands under ultraviolet light. With electrophoresis, the atypical bands can be separated, allowing their detection apart from the three major bands. A complicating feature is often a strongly fluorescent band that appears migrating in close proximity to the CK-BB form. The exact nature of this fluorescence is unknown, but it has been attributed to the binding of fluorescent drugs or bilirubin by albumin.

In addition to visualizing atypical CK bands, other advantages of electrophoresis methods include detecting an unsatisfactory separation and allowing visualization of adenylate kinase (AK). AK is an enzyme released from erythrocytes in hemolyzed samples and appearing as a band cathodal to CK-MM. AK may interfere with chemical or immunoinhibition methods, causing a falsely elevated CK or CK-MB value. Ion-exchange chromatography has

the potential for being more sensitive and precise than electrophoretic procedures performed with good technique. On an unsatisfactory column, however, CK-MM may merge into CK-MB and CK-BB may be eluted with CK-MB. Also, macro-CK may elute with CK-MB.

Immunological methods are also used with specific antibodies against the various subunit types of CK and provide distinction between the isoenzymes. Antibodies against both the M and B subunits have been used jointly to determine CK-MB activity. The sample is first treated with the anti-M antibody, which causes inhibition of activity the M subunit, but not B subunit activity. The remaining CK activity correlates to only the B subunit of both the MB and/or BB isoenzymes present. The major disadvantage of this method is that it detects BB activity, which, although not normally detectable, will cause falsely elevated MB results when BB is present. In addition, the atypical forms of CK-Mi and macro-CK are not inhibited by anti-M antibodies and also may cause erroneous results for MB activity. Immunoassays detect CK-MB reliably with minimal cross-reactivity. Immunoassays measure the concentration of enzyme protein rather than enzymatic activity and can, therefore, detect enzymatically inactive CK-MB. This leads to the possibility of permitting detection of infarction earlier than other methods. A double-antibody immunoinhibition assay is also available. This technique allows differentiation of MB activity due to AK as described and the atypical isoenzymes, resulting in a more specific analytic procedure for CK-MB.[17] Point-of-care assay systems for CM-MB are available but not as widely used as those for troponins.

Assay Enzyme Activity

As indicated by Equation 8.4, CK catalyzes both forward and reverse reactions involving phosphorylation of creatine or ADP. Typically, for analysis of CK activity, this reaction is coupled with other enzyme systems and measured following the change in absorbance at 340 nm with the change in $NADH^+$ coenzyme involved with the coupling enzyme. The forward reaction is coupled with the pyruvate kinase–LD–NADH system and proceeds according to Equation 8.5:

$$\text{Creatine} + \text{ATP} \xrightleftharpoons{\text{CK}} \text{Creatine phosphate} + \text{ADP}$$
$$\text{ADP} + \text{phosphoenolpyruvate} \xrightleftharpoons{\text{PK}} \text{Pyruvate} + \text{ATP}$$
$$\text{Pyruvate} + \text{NADH} + \text{H}^+ \xrightleftharpoons{\text{LD}} \text{Lactate} + \text{NAD}^+$$

(**Eq. 8.5**)

The reverse reaction is coupled with the hexokinase–glucose-6-phosphate dehydrogenase–NADPH system, as indicated in Equation 8.6:

$$\text{Creatine phosphate} + \text{ADP} \xrightleftharpoons{\text{CK}} \text{Creatine ADP}$$
$$\text{ATP} + \text{glucose} \xrightleftharpoons{\text{KH}} \text{ADP} + \text{glucose-6-phosphate}$$
$$\text{Glucose-6-phosphate} + \text{NADP}^+ \xrightleftharpoons{\text{G-6-PD}} \text{6-phosphogluconate} + \text{NADPH}$$

(**Eq. 8.6**)

The reverse reaction proposed by Oliver and modified by Rosalki is the most commonly performed method for total CK in the clinical laboratory.[13] The reaction proceeds two to six times faster than the forward reaction, depending on the assay conditions, and there is less interference from side reactions. The optimal pH for the reverse reaction is 6.8; for the forward reaction, it is 9.0.

CK activity in serum is unstable, being rapidly inactivated because of oxidation of sulfhydryl groups. The oxidation is preventable or reversible with the of sulfhydryl compounds in the reaction mixture; compounds include N-acetylcysteine, mercaptoethanol, thioglycerol, and dithiothreitol.

Source of Error

Hemolysis of plasma samples may be a source of elevated CK activity. Erythrocytes are virtually devoid of CK; however, they are rich in adenylate kinase (AK) activity. AK reacts with ADP to produce ATP, which is then available to participate in the assay reaction, causing falsely elevated CK levels. This interference can occur when hemolysis achieves a release of greater than 320 mg/L hemoglobin in the plasma; this is accompanied by a release of sufficient AK to exhaust the AK inhibitors in the reagent. The release of AK may be detected by electrophoresis allowing visualization of AK appearing as a band cathodal to CK-MM. Trace hemolysis causes little, if any, CK elevation. Plasma should be stored in a dark place because CK is inactivated by light. Activity can be restored after storage in the dark at 4°C for 7 days or at –20°C for 1 month when the assay is conducted using a sulfhydryl activator.[18] Because of the effect of muscular activity and muscle mass on CK levels, it should be noted that people who are physically well trained tend to have elevated baseline levels, whereas patients who are bedridden for prolonged periods may have decreased CK activity.

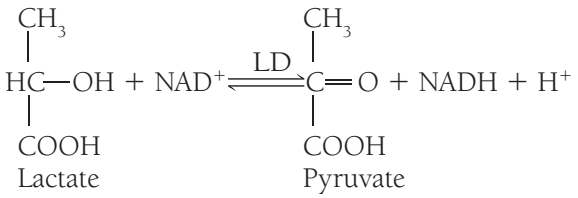
Reference Range

Total CK:
Male, 46–171 U/L (37°C)
Female, 34–145 U/L (37°C)
CK-MB: <5% total CK

The higher values in males are attributed to increased muscle mass. Note that enzyme reference ranges are subject to variation, depending on the method used and the assay conditions.

Lactate Dehydrogenase

LD is an enzyme that catalyzes the interconversion of lactic and pyruvic acids. It is a hydrogen transfer enzyme that uses the coenzyme NAD^+ according to **Equation 8.7**:



(**Eq. 8.7**)

Tissue Source

LD is widely distributed in the body. Higher activities are found in the heart, liver, skeletal muscle, kidney, and erythrocytes, with lesser amounts found in the lung, smooth muscle, and brain.

Diagnostic Significance

Because of its widespread activity in numerous body tissue, plasma LD is elevated in a variety of disorders. Increased plasma levels are found in cardiac, hepatic, skeletal muscle, and renal diseases, as well as in several hematologic and neoplastic disorders. The highest plasma levels of total LD are seen in pernicious anemia and hemolytic disorders. Intramedullary destruction of erythroblasts causes elevation because of the high concentration of LD in erythrocytes. Liver disorders, such as viral hepatitis and cirrhosis, show slight elevations of two to three times ULN. However, other enzymes such as ALP and the transaminases are better plasma makers for liver damage. AMI and pulmonary infarct also show slight elevations of approximately the same degree (2 to 3× ULN). In AMI, plasma LD levels begin to rise within 12 to 24 hours, reach peak levels within 48 to 72 hours, and may remain elevated for 10 days. Skeletal muscle disorders and some leukemias contribute to increased LD levels. One important clinical use for LD is in the prognosis and management of certain tumors. When coupled with α-fetoprotein and human chorionic gonadotropin, LD has been found to be an important serological marker for diagnosis, staging, recurrence, and monitoring of germ cell tumors.[19] In particular, marked elevations can also be observed in most patients with acute lymphoblastic leukemia. LD has been shown to form complexes with immunoglobulins. This macromolecular complex is not associated with any specific clinical abnormality.

LD is found as one of five isoenzyme forms each containing four subunits, a tetramer from of LD. Each tetramer has a molecular weight of 128,000 Da comprised from four subunit polypeptide chains with a molecular weight of 32,000 Da each. There are two different types of subunits designated H (heart) and M (muscle) forms with the polypeptide chains combining to yield five major isoenzymes of the tetramer form. **Table 8.4** indicates the tissue localization of the different LD isoenzymes and the major disorders associated with elevated levels.

Historically, the electrophoretic procedure for the separation of LD isoenzymes has been widely used,

Table 8.4 Lactate Dehydrogenase (LD) Isoenzymes—Tissue Localization and Sources of Elevation

Isoenzyme	Tissue	Disorder
LD-1 (HHHH)	Heart	Myocardial infarction
	Red blood cells	Hemolytic anemia
LD-2 (HHHM)	Heart	Megaloblastic anemia
	Red blood cells	Acute renal infarction
		Hemolyzed specimen
LD-3 (HHMM)	Lung	Pulmonary embolism
	Lymphocytes	Extensive pulmonary pneumonia
	Spleen	Pulmonary pneumonia
	Pancreas	Lymphocytosis
		Acute pancreatitis
LD-4 (HMMM)		Carcinoma
LD-5 (MMMM)	Liver	Hepatic injury or inflammation

particularly in the late diagnosis of MI. However, this procedure was labor intensive and time consuming. Analysis of LD isoenzymes could also be accomplished by immuno-inhibition or chemical inhibition methods, or by differences in substrate affinity. The cardiac troponin I assays have greater sensitivity than lactate dehydrogenase isoenzymes and are a more cost-effective test. Troponin assays have replaced LD isoenzymes in the diagnosis of MI.[20]

In the plasma of healthy individuals, the major isoenzyme fraction is LD-2, followed by LD-1, LD-3, LD-4, and LD-5 (for the isoenzyme ranges, see Table 8.4). LD-1 and LD-2 are present to approximately the same extent in the tissues listed in Table 8.4, with LD-1 higher in cardiac tissue and red blood cells. Therefore, in conditions involving cardiac necrosis (AMI) and intravascular hemolysis, the plasma levels of LD-1 will increase to a point at which they are present in greater concentration than LD-2, unlike the normal pattern in which LD-2 is highest. Elevations of LD-1 have also been shown. Elevation of LD-3 occurs most frequently with pulmonary involvement and is also observed in patients having various carcinomas. The LD-4 and LD-5 isoenzymes are found primarily in liver and skeletal muscle tissue, with LD-5 being the predominant fraction in these tissues. LD-5 levels have greatest clinical significance in the detection of hepatic disorders, particularly intrahepatic disorders. Disorders of skeletal muscle will also reveal elevated LD-5 levels, as depicted in the muscular dystrophies.

Assay for Enzyme Activity

LD catalyzes the interconversion of lactic and pyruvic acids using the coenzyme NAD^+. The reaction sequence is outlined in **Equation 8.8**:

$$\text{Lactate} + NAD^+ \rightleftharpoons \text{Pyruvate} + NADH + H^+ \quad \textbf{(Eq. 8.8)}$$

The reaction can proceed in either a forward (lactate [L]) or reverse (pyruvate [P]) direction. Both reactions have been used in clinical assays. The rate of the reverse reaction is approximately three times faster, allowing smaller sample volumes and shorter reaction times. However, the reverse reaction is more susceptible to substrate exhaustion and loss of linearity. The optimal pH for the forward reaction is 8.3 to 8.9; for the reverse reaction, it is 7.1 to 7.4. With these options, LD is also commonly used to measure lactic and pyruvic acids or as a coupled reaction.

Source of Error

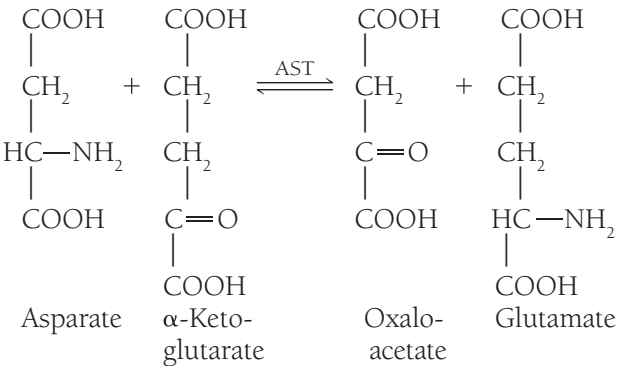
Erythrocytes contain an LD concentration approximately 100 to 150 times that found in plasma. Therefore, any degree of hemolysis should render a sample unacceptable for analysis. LD activity is unstable in plasma regardless of the temperature at which it is stored. If the sample cannot be analyzed immediately, it should be stored at 25°C and analyzed within 48 hours. LD-5 is the most labile isoenzyme. Loss of activity occurs more quickly at 4°C than at 25°C.

Reference Range

LD, 125–220 U/L (37°C)

Aspartate Aminotransferase

Aspartate aminotransferase (AST) is an enzyme belonging to the class of transferases. It is commonly referred to as a *transaminase* and is involved in the transfer of an amino group between aspartate and α-keto acids. The older terminology, *serum glutamic-oxaloacetic transaminase* (*SGOT*), may also be used. Pyridoxal phosphate functions as a coenzyme. The reaction proceeds according to **Equation 8.9**:



(Eq. 8.9)

The transamination reaction is important in intermediary metabolism because of its function in the synthesis and degradation of amino acids. The keto-acids formed by the reaction are ultimately oxidized by the tricarboxylic acid cycle to provide a source of energy.

AST exists in two isoenzyme forms, one located in the cell cytoplasm and the other in the mitochondria. The cytoplasmic isoenzyme is the predominant form occurring in plasma. In hepatocytes, which have the highest concentration of cytoplasmic AST, the intracellular concentration of AST is approximately 7000 times higher than the extracellular concentration. In disorders producing cellular necrosis, the mitochondrial form may be significantly

increased. Isoenzyme analysis of AST is not routinely performed in the clinical laboratory.

Tissue Source

AST is widely distributed in human tissue. The highest concentrations are found in cardiac tissue, liver, and skeletal muscle, with smaller amounts found in the kidney, pancreas, and erythrocytes.

Diagnostic Significance

Because of its wide distribution throughout the body, AST elevations can occur in a number of conditions. However, the clinical use of AST is limited mainly to the evaluation of hepatocellular disorders and skeletal muscle involvement. AST levels are highest in acute hepatocellular disorders. In viral hepatitis, levels may reach 100 times ULN. In cirrhosis, only moderate levels—approximately four times ULN—are detected. Skeletal muscle disorders, such as the muscular dystrophies, and inflammatory conditions can cause increases in AST levels of four to eight times ULN.

Following an AMI, AST levels begin to rise within 6 to 8 hours, peak at 24 hours, and generally return to normal within 5 days. AST elevations are frequently seen in pulmonary embolism. Following congestive heart failure, AST levels also may be increased, probably reflecting liver involvement as a result of inadequate blood supply to that organ.

Assay for Enzyme Activity

Assay methods for AST are generally based on the principle of the Karmen method, which incorporates a coupled enzymatic reaction using malate dehydrogenase (MD) as the indicator reaction to monitor the change in absorbance at 340 nm continuously as NADH is oxidized to NAD^+ (**Eq. 8.10**). The optimal pH is 7.3 to 7.8.

$$\text{Aspartate} + \alpha\text{- ketoglutarate} \xrightleftharpoons{\text{AST}} \text{Oxaloacetate} + \text{glutamate}$$

$$\text{Oxaloacetate} + \text{NADH} + H^+ \rightleftharpoons \text{Malate} + NAD^+$$

(**Eq. 8.10**)

Source of Error

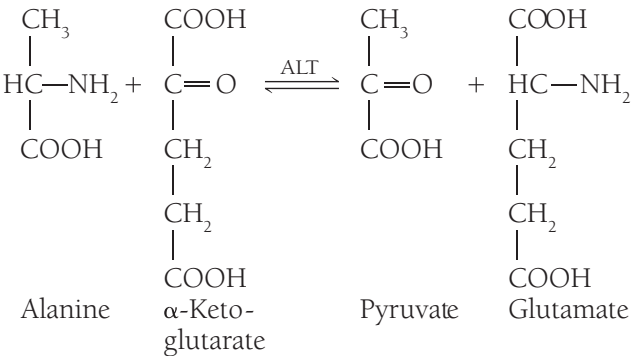
Hemolysis should be avoided because it can dramatically increase plasma AST concentration. AST activity is stable in plasma for 3 to 4 days at refrigerated temperatures.

Reference Range

AST, 5–35 U/L (37°C)

Alanine Aminotransferase

Alanine aminotransferase (ALT) is a transferase with an enzymatic reaction similar to that of AST. Specifically, it catalyzes the transfer of an amino group from alanine to α-ketoglutarate with the formation of glutamate and pyruvate. The older terminology was serum *glutamic-pyruvic transaminase* (*SGPT*). **Equation 8.11** indicates the transferase reaction. Pyridoxal phosphate acts as the coenzyme.



(**Eq. 8.11**)

Tissue Source

ALT is distributed in many tissues, with comparatively high concentrations in the liver. It is considered the more liver-specific enzyme of the two transferases discussed.

Diagnostic Significance

Clinical applications of ALT assays are primarily in the evaluation of hepatic disorders. Higher elevations are found in hepatocellular disorders than in extrahepatic or intrahepatic obstructive disorders. In acute inflammatory conditions of the liver, ALT elevations are frequently higher than those of AST and tend to remain elevated longer because of the longer half-life of ALT in plasma (17 and 47 hours, respectfully).[21]

Cardiac tissue contains a small amount of ALT activity, but the plasma level usually remains normal in AMI unless subsequent liver damage has occurred. ALT levels have historically been compared with levels of AST to help determine the source of an elevated AST level and to detect liver involvement concurrent with myocardial injury.

De Ritis Ratio = AST/ALT

Assay for Enzyme Activity

The typical assay for ALT consists of a coupled enzymatic reaction using LD as the indicator enzyme, which catalyzes the reduction of pyruvate to lactate with the simultaneous oxidation of NADH. The change in absorbance at 340 nm measured continuously is directly proportional to ALT activity. The reaction proceeds according to **Equation 8.12**. The optimal pH is 7.3 to 7.8.

$$\text{Alanine} + \alpha\text{-ketoglutarate} \xrightleftharpoons{\text{ALT}} \text{Pyruvate} + \text{glutamate}$$

$$\text{Pyruvate} + \text{NADH} + \text{H}^+ \xrightleftharpoons{\text{LD}} \text{Lactate} + \text{NAD}^+$$

(**Eq. 8.12**)

Source of Error

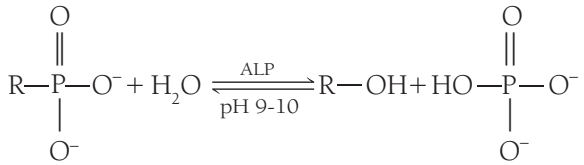
ALT is stable for 3 to 4 days at 4°C. It is relatively unaffected by hemolysis.

Reference Range

ALT, 7–45 U/L (37°C)

Alkaline Phosphatase

ALP belongs to a group of enzymes that catalyze the hydrolysis of various phosphomonoesters at an alkaline pH. Consequently, ALP is a nonspecific enzyme capable of reacting with many different substrates. Specifically, ALP functions to liberate inorganic phosphate from an organic phosphate ester with the concomitant production of an alcohol. The reaction proceeds according to **Equation 8.13**:



(**Eq. 8.13**)

The optimal pH for the reaction is 9.0 to 10.0, but optimal pH varies with the substrate used. ALP requires Mg^{2+} as an activator.

Tissue Source

ALP activity is present on cell surfaces in most human tissue. The highest concentrations are found in the intestine, liver, bone, spleen, placenta, and kidney. In the liver, the enzyme is located on both sinusoidal and bile canalicular membranes; activity in bone is confined to the osteoblasts, those cells involved in the production of bone matrix. The specific location of the enzyme within this tissue accounts for the more predominant elevations in certain disorders.

Diagnostic Significance

Elevations of ALP are of most diagnostic significance in the evaluation of hepatobiliary and bone disorders. In hepatobiliary disorders, elevations are more predominant in obstructive conditions than in hepatocellular disorders; in bone disorders, elevations are observed when there is involvement of osteoblasts.

In biliary tract obstruction, ALP levels range from 3 to 10 times ULN. Increases are primarily a result of increased synthesis of the enzyme induced by the decreased or blocked flow of bile or cholestasis. In contrast, hepatocellular disorders, such as hepatitis and cirrhosis, show only slight increases, usually less than three times ULN. Because of the degree of overlap of ALP elevations that occurs in the various liver disorders, a single elevated ALP level is difficult to interpret. It assumes more diagnostic significance when evaluated along with other tests of hepatic function.

Elevated ALP levels are also observed in various bone disorders. Perhaps the highest elevations of ALP activity occur in Paget's disease (osteitis deformans). Other bone disorders include osteomalacia, rickets, hyperparathyroidism, and osteogenic sarcoma. In addition, increased levels are observed in healing bone fractures and during periods of physiologic bone growth.

In normal pregnancy, increased ALP activity, averaging approximately 1½ times ULN, can be detected between weeks 16 and 20 and 2 to 3 times the ULN during the third trimester. ALP activity increases and persists until the onset of labor. Activity then returns to normal within 3 to 6 days.[22] Elevations also may be seen in complications of pregnancy such as hypertension, preeclampsia, and eclampsia, as well as in threatened abortion.

ALP levels are significantly decreased in the inherited condition of hypophosphatasia. Subnormal activity is a result of the absence of the bone isoenzyme, which results in inadequate bone calcification.

ALP exists as multiple isoenzymes, which have been studied by a variety of techniques. The major isoenzymes, which are found in the plasma and have been most extensively studied, are those isolated from the liver, bone, intestine, and placenta.[23]

The liver isoenzyme can be divided into two fractions—the major liver band and a smaller fraction called *fast liver*, or α_1 liver, which migrates anodal to the major band and corresponds to the α_1 fraction of protein electrophoresis. When total ALP levels are increased, the major liver fraction is the most frequently elevated. Many hepatobiliary conditions cause elevations of this fraction, usually early in the course of the disease. The fast-liver fraction has been reported in metastatic carcinoma of the liver, as well as in other hepatobiliary diseases. Its presence is regarded as a valuable indicator of obstructive liver disease. However, it is occasionally present in the absence of any detectable disease state.

The bone isoenzyme increases due to osteoblastic activity and is normally elevated in children during periods of growth and in adults older than age 50. In these cases, an elevated ALP level may be difficult to interpret.[24]

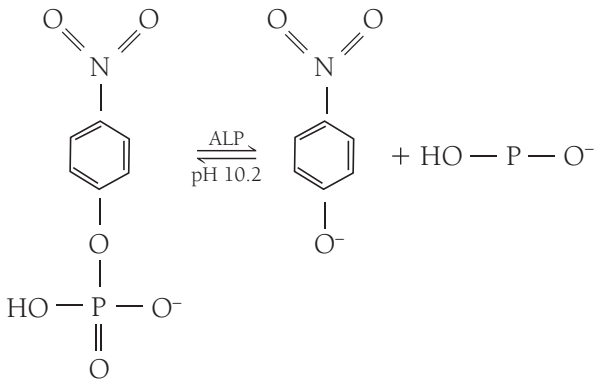
The presence of intestinal ALP isoenzyme in plasma depends on the blood group and secretor status of the individual. Individuals who are blood types B or O and are secretors are more likely to have this fraction. Apparently, intestinal ALP is bound by erythrocytes of group A. Furthermore, in B or O secretors, plasma concentration of intestinal ALP can increase up to 30 IU/L following the consumption of a meal. Intestinal ALP may increase in several disorders, such as diseases of the digestive tract and cirrhosis. Increased levels are also found in patients undergoing chronic hemodialysis.

In addition to the four major ALP isoenzyme fractions, certain abnormal fractions are associated with neoplasms. The most frequently seen are the Regan and Nagao isoenzymes. They have been referred to as *carcinoplacental alkaline phosphatases* because of their similarities to the placental isoenzyme. The frequency of occurrence ranges from 3% to 15% in cancer patients. The Regan isoenzyme has been characterized as an example of an ectopic production of an enzyme by malignant tissue. It has been detected in various carcinomas, such as the lung, breast, ovarian, and colon, with the highest incidences in ovarian and gynecologic cancers. Because of its low incidence in cancer patients, diagnosis of malignancy is rarely based on its presence. It is, however, useful in monitoring the effects of therapy because it will disappear on successful treatment.

The Regan isoenzyme migrates to the same position as the bone fraction and is the most heat stable of all ALP isoenzymes, resisting denaturation at 65°C for 30 minutes. Its activity is inhibited by phenylalanine. The Nagao isoenzyme may be considered a variant of the Regan isoenzyme. Its electrophoretic, heat stability, and phenylalanine inhibition properties are identical to those of the Regan fraction. However, Nagao also can be inhibited by L-leucine. Its presence has been detected in metastatic carcinoma of pleural surfaces and in adenocarcinoma of the pancreas and bile duct.

Assay for Enzyme Activity

Because of the relative nonspecificity of ALP, a variety of methodologies for its analysis have been proposed and are still in use today. The major differences between these relate to the concentration and types of substrate and buffer used and the pH of the reaction. A continuous monitoring technique based on a method devised by Bowers and McComb allows calculation of ALP activity based on the molar absorptivity of *p*-nitrophenol. The reaction proceeds according to **Equation 8.14**:



p-Nitrophenylphosphate | *p*-Nitrophenol | Phosphate ion

(**Eq. 8.14**)

p-Nitrophenylphosphate (colorless) is hydrolyzed to *p*-nitrophenol (yellow), and the increase in absorbance at 405 nm, which is directly proportional to ALP activity, is measured.

Several techniques have been used to determine ALP plasma activity with the different isoenzymes. These have included: (1) separation of the isoenzymes by electrophoresis, (2) stability to inactivation by heat or chemicals, (3) sensitivity to various inhibitors, (4) affinity to specific lectins, and (5) immunochemical assays.

Electrophoresis is considered the most useful single technique for distinguishing between ALP isoenzymes. However, because there may still be some degree of overlap between the fractions,

electrophoresis in combination with another separation technique may provide the most reliable information. The liver fraction migrates the fastest, followed by bone, placental, and intestinal fractions. Because of the similarity between liver and bone phosphatases, there often is not a clear separation between them. Quantification with use of a densitometer is sometimes difficult because of the overlap between the two peaks.

Difference in heat stability has been used as an approach to identify the isoenzyme source of an elevated ALP. The liver isoenzyme is more heat stable than the bone isoenzyme. Heating plasma to 56°C for 10 minutes will inactivate most of the bone isoenzyme, while the liver isoenzyme will be relatively stable. Placental ALP is the most heat stable of the four major fractions. It resists heat denaturation at 65°C for 30 minutes. Of the remaining fractions, the intestinal isoenzyme is the second most heat stable, followed by the liver, and with the bone isoenzyme being the most heat-labile. Heat inactivation is an imprecise method for differentiation because inactivation depends on many factors, such as correct temperature control, timing, and analytic methods sensitive enough to detect small amounts of residual ALP activity. In addition, there is some degree of overlap between heat inactivation of liver and bone fractions in both liver and bone diseases.

A third approach to identification of ALP isoenzymes is based on selective chemical inhibition. Phenylalanine is one of several inhibitors that have been used as it inhibits intestinal and placental ALP to a much greater extent than liver and bone ALP. With phenylalanine use, however, it is impossible to differentiate placental from intestinal ALP or liver from bone ALP.

Immunochemical methods for the measurement of bone-related ALP have been developed. However, these assays tend to lack specificity for the bone isoenzyme, with significant negative (22%)[25] and positive (27%)[26] bias having been reported.

Source of Error

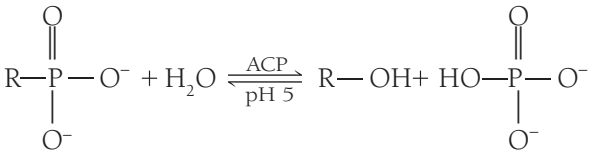
Hemolysis may cause elevations of ALP because ALP is approximately six times more concentrated in erythrocytes than in serum. ALP assays should be run as soon as possible after collection. Activity in serum increases approximately 3% to 10% on standing at 25°C or 4°C for several hours. Diet may induce elevations in ALP activity of blood group B and O individuals who are secretors. Values may be 25% higher following ingestion of a high-fat meal.

Reference Range

Gender	Age	Reference Range
ALP (total), (37°C)		
Males–Females	4–15 y	54–369 U/L
Males	20–50 y	53–128 U/L
	≥60 y	56–119 U/L
Females	20–50 y	42–98 U/L
	≥60 y	53–141 U/L

Acid Phosphatase

Acid phosphatase (ACP) includes a group of several phosphatases with an optimal activity at a pH below 7.0. ACP belongs to the same group of phosphatase enzymes as ALP and is a **hydrolase** that catalyzes the same type of reactions. The major difference between ACP and ALP is the pH of the reaction. Plasma ACP is stabilized by acidification to pH below 6.5 with an optimal pH of approximately 5.0. **Equation 8.15** outlines the reaction sequence:



Phosphomonoester Alcohol Phosphate ion

(**Eq. 8.15**)

Tissue Source

ACP activity is found in the prostate, bone, liver, spleen, kidney, erythrocytes, and platelets. The prostate is the richest source, with many times the activity found in other tissues.

Diagnostic Significance

Historically, ACP measurement was used as an aid in the detection of prostatic carcinoma, particularly metastatic carcinoma of the prostate. Total ACP determinations used relatively insensitive techniques and detected elevated ACP levels resulting from prostatic carcinoma only when the tumor had metastasized. Newer markers, such as prostate-specific antigen (PSA), are more useful screening and diagnostic tools. PSA is more likely than ACP to be elevated at each stage of prostatic carcinoma, even though a normal PSA level may be found in prostatic cancer that has spread to other parts of the body. PSA is particularly useful to monitor the success of

treatment; however, PSA is controversial as a screening test for prostatic malignancy because PSA elevation may occur in conditions other than prostatic carcinoma, such as benign prostatic hypertrophy and prostatitis.[27–29]

Because of the multiple tissues containing ACP, methods are needed to distinguish the prostatic form. One approach is the use of inhibitors such as tartrate. Serum and substrate are incubated both with and without the addition of L-tartrate. ACP activity remaining after inhibition by tartrate is subtracted from total ACP activity determined without inhibition, and the difference represents the prostatic portion:

$$\text{Total ACP} - \text{ACP after tartrate inhibition} = \text{prostatic ACP} \quad \textbf{(Eq. 8.16)}$$

The reaction is not entirely specific for prostatic ACP, with lysosomal ACP also being inhibited by tartrate. However, other tissue sources are largely uninhibited.

One technique with improved sensitivity over conventional ACP assays is an immunologic approach using antibodies that are specific for the prostatic form of ACP. Immunochemical techniques, however, are not of sufficient value as screening tests for prostatic carcinoma. Other prostatic conditions in which ACP elevations have been reported include hyperplasia of the prostate and prostatic surgery. There are conflicting reports of elevations following rectal examination and prostate massage. Certain studies have reported ACP elevations; others have indicated no detectable change. When elevations are found, levels usually return to normal within 24 hours.[30]

Plasma ACP activity also may frequently be elevated in bone disease, but this form is distinguishable because it is a tartrate-resistant form of ACP. Activity has been shown to be associated with the osteoclasts.[31] Elevations have been noted in Paget's disease, in breast cancer with bone metastases, and in Gaucher's disease, in which there is an infiltration of bone marrow and other tissue by Gaucher cells rich in ACP activity. Because of ACP activity in platelets, elevations are observed when platelet damage occurs, as in the thrombocytopenia resulting from excessive platelet destruction from idiopathic thrombocytopenic purpura.

ACP assays have proved useful in forensic clinical chemistry, particularly in the investigation of rape. Vaginal washings are examined for seminal fluid–ACP activity, which can persist for up to 4 days.[32] Elevated activity is presumptive evidence of rape in such cases.

Assay for Enzyme Activity

Simple assay procedures for total ACP use the same techniques as in ALP assays but are performed at an acid pH:

$$p\text{-Nitrophenolphosphate} \xrightarrow[\text{pH5}]{\text{ACP}} p\text{-Nitrophenol} + \text{phosphate ion} \quad \textbf{(Eq. 8.17)}$$

The reaction products are colorless at the acidic and require the addition of alkali, which stops the reaction by raising the pH out of the acidic range. The nitrophenol product at the alkaline pH is yellow and serves as a chromophore that can be measured spectrophotometrically.

Immunochemical techniques for prostatic ACP use several approaches, including enzyme-linked immunosorbent assay (ELISA) and immunoselective enzyme immunoassay (ISEA) methods. Also, an immune-enzymatic assay may include incubation with an antibody to prostatic ACP followed by washing and incubation with *p*-nitrophenylphosphate. The *p*-nitrophenol product formed is measured photometrically and reflects the amount of the prostatic ACP in the sample.

Source of Error

Plasma should be separated from the red cells to prevent leakage of ACP from erythrocytes and platelets. Hemolysis should be avoided because of contamination from erythrocyte ACP.

Plasma activity decreases within 1 to 2 hours if the sample is left at room temperature without the addition of a preservative. Decreased activity is a result of a loss of carbon dioxide from the plasma with a resultant increase in pH of the sample. If not assayed immediately, plasma should be frozen or acidified to a pH lower than 6.5. With acidification, ACP becomes stable for 2 days at room temperature, up to 1 week at 4°C, and for 4 months at –20°C.

Reference Range

Prostatic ACP, 0–3.5 ng/mL
Total ACP
Adults: 1.5–4.5 U/L (37°C)
Children: 3.5–9.0 U/L (37°C)

γ-Glutamyltransferase

γ-Glutamyltransferase (GGT) is an enzyme involved in the transfer of the γ-glutamyl residue from γ-glutamyl peptides to amino acids, H_2O, and other small peptides. In most biologic systems, glutathione serves as the γ-glutamyl donor. **Equation 8.18** outlines the reaction sequence:

$$\text{Glutathione} + \text{amino acid} \underset{}{\overset{\text{AST}}{\rightleftharpoons}} \text{Glutamyl peptide} + \text{L-cysteinylglycine} \quad \text{(Eq. 8.18)}$$

The specific physiologic function of GGT has not been clearly established, but it is suggested that GGT is involved in the maintenance of tissue glutathione levels and the transport of amino acids across cell membranes.[33]

Tissue Source

GGT activity is found in highest concentration in the kidney, and also in the brain, prostate, pancreas, and liver. Clinical applications of assay, however, are confined mainly to evaluation of liver and biliary system disorders because those are the primary sources of plasma GGT.

Diagnostic Significance

In the liver, GGT is located in the canaliculi of the hepatic cells and particularly in the epithelial cells lining the biliary ductules. Because of these locations, GGT is elevated in virtually all hepatobiliary disorders, making it one of the most sensitive of enzyme assays in these conditions. Higher elevations are generally observed in biliary tract obstruction, with plasma GGT activity increasing 5 to 30 times the ULN.

Within the hepatic parenchyma, GGT exists to a large extent in the smooth endoplasmic reticulum and is, therefore, subject to hepatic microsomal induction. Therefore, GGT levels will be increased in patients receiving enzyme-inducing drugs such as warfarin, phenobarbital, and phenytoin. Enzyme elevations may reach levels four times ULN.

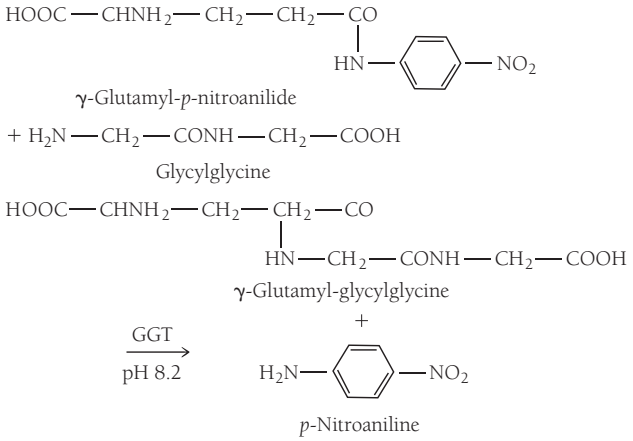
Elevated plasma GGT activity has been found in patients with high alcohol consumption, particularly chronic alcoholism. Generally, enzyme elevations in persons who are alcoholics or heavy drinkers range from two to three times ULN, although higher levels have been observed. GGT assays are useful in monitoring the effects of abstention from alcohol and are used as such by alcohol treatment centers. Levels usually return to normal within 2 to 3 weeks after cessation but can rise again if alcohol consumption is resumed. Because of the susceptibility to enzyme induction, any interpretation of GGT levels must be done with consideration of the consequent effects of drugs and alcohol.

GGT levels are also elevated in other conditions, such as acute pancreatitis, diabetes mellitus, and myocardial infarction. The source of elevation in pancreatitis and diabetes is probably due to changes in the pancreas, but the source of GGT in MI is unknown. GGT assays are of limited value in the diagnosis of these conditions and are not routinely requested.

GGT activity is useful in differentiating the source of an elevated plasma ALP level because GGT levels are normal in skeletal disorders and during pregnancy. It is particularly useful in evaluating hepatobiliary involvement in adolescents because ALP activity will be elevated as a result of bone growth.

Assay for Enzyme Activity

The most widely accepted substrate for use in GGT analysis is γ-glutamyl-*p*-nitroanilide. The γ-glutamyl residue is transferred to glycylglycine, releasing *p*-nitroaniline, a yellow chromogenic product with a strong absorbance at 405 to 420 nm. The reaction, which can be used as a continuous monitoring or fixed-point method, is outlined in **Equation 8.19**:



(**Eq. 8.19**)

Source of Error

GGT activity is stable with no loss of activity for 1 week at 4°C. Hemolysis does not interfere with GGT levels because the enzyme is not found in erythrocytes.

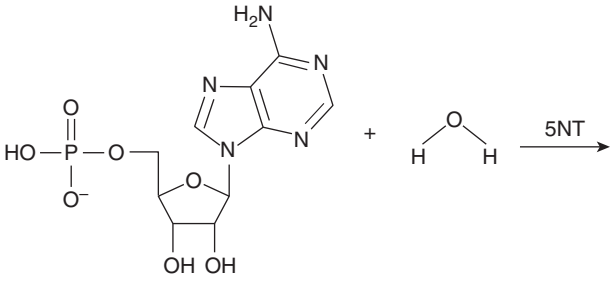
Reference Range

GGT: male, 6–55 U/L (37°C); female, 5–38 U/L (37°C)

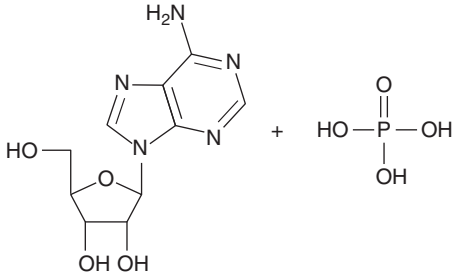
Values are lower in females, presumably because of suppression of enzyme activity resulting from estrogenic or progestational hormones.

5′-Nucleotidase

5′-Nucleotidase (5NT) is an enzyme belonging to the class of hydrolases. The enzyme requires a divalent cofactor, which is thought to be zinc. The preferred substrates for 5NT are 5′-nucleotide monophosphates used to produce the nucleoside and free phosphate, such as the catalysis of adenosine-5′-monophosphate (AMP) to adenosine and inorganic phosphate shown in **Equation 8.20**:



5′-Adenosine-Monophosphate H_2O



Adenosine Phosphate

(**Eq. 8.20**)

Tissue Source

5NT is widely distributed in tissues, but its clinical utility is in the diagnosis of hepatobiliary disorders. Within the cell, it is primarily found bound to the plasma membrane, although is also present in lower concentrations in the mitochondria, microsomes, Golgi apparatus, and the cytosol.[34]

Diagnostic Significance

Despite of its wide tissue distribution, increased plasma levels of 5NT are almost always an indication of hepatobiliary disease and is more commonly increased in biliary obstruction. Any increase in plasma 5NT levels more than a slight increase is highly specific for hepatobiliary disease,[35] even in patients with normal plasma levels of ALP. Like GGT, 5NT activity is useful in differentiating the source of elevated plasma ALP activity because 5NT levels are normal to slightly elevated in skeletal and bone disorders. A slight increase in 5NT also occurs in the second and third trimesters of pregnancy.

Assay for Enzyme Activity

The analysis for 5NT most often utilizes inosine-monophosphate or adenosine monophosphate. In an enzyme-coupled reaction, 5NT catalyzes the hydrolysis of inosine monophosphate to inosine and free inorganic phosphate. Using purine nucleoside phosphorylase as an auxiliary enzyme, the inosine is then converted to hypoxanthin. The enzyme xanthine oxidase then converts the hypoxanthine to urate and two molecules of hydrogen peroxide for every molecule of hypoxanthine. A peroxidase, such as horseradish peroxidase, acts as the indicator enzyme and acts on the peroxide and a chromogen to produce a colored compound that can be measured spectrophotometrically.

$$\text{Inosine Monophosphate} + H_2O \xrightarrow{\text{5NT}} \text{Inosine} + \text{Phosphate}$$

$$\text{Inosine} \xrightarrow{\text{Purine nucleoside Phosphorylase}} \text{Hypoxanthine} + \text{Ribose}$$

$$\text{Hypoxanthine} \xrightarrow{\text{Xanthine Oxidase}} \text{Uric Acid} + 2H_2O_2$$

(**Eq. 8.21**)

The difficulty with these assays is the fact that plasma and other body fluids will likely contain other nonspecific phosphatases, such as ALP, that may be present in higher concentrations. This disadvantage is corrected for in one of two ways. First, because the other phosphatases are not specific, the presence of excess nonnucleotide substrates will compete in the reaction with the nucleotide monophosphates with the nonspecific phosphatases, but not with the more specific 5NT. This allows the conversion of nucleotides to nucleosides to be the result of catalysis by 5NT. The second method used to minimize the reaction by ALP is to add an inhibitor such as β-phosphoglycerate.

Source of Error

High concentrations of nonspecific phosphatases may cause falsely increased results. EDTA will inhibit enzyme activity, likely due to the chelation of zinc.

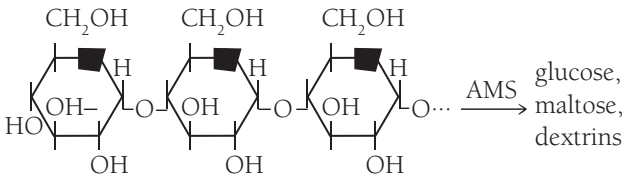
Reference Range

5NT: 3–9 U/L (37°C)

Plasma 5NT activity is generally low in children, begins to rise in adolescence, and plateaus at the age of 40 after levels increase significantly.[36]

Amylase

Amylase (AMY) is an enzyme belonging to the class of hydrolases that catalyze the breakdown of starch and glycogen. Starch consists of both amylose and amylopectin. Amylose is a long, unbranched chain of glucose molecules linked by α-1,4-glycosidic bonds; amylopectin is a branched-chain polysaccharide of α-1,4-glycosidic bonds with α-1–6 linkages at branch points. The structure of glycogen is similar to that of amylopectin but contains more α-1–6 linkage branch points. α-AMY attacks only the α-1,4-glycosidic bonds to produce degradation products consisting of glucose, maltose, and intermediate chains, called *dextrins*, which contain α-1–6 branching linkages. Cellulose and other structural polysaccharides contain β-linkages are not susceptible to α-AMY. AMY is therefore an important enzyme in the physiologic digestion of starches. The reaction proceeds according to **Equation 8.22**:



(**Eq. 8.22**)

AMY requires calcium and chloride ions for its activation.

Tissue Source

The acinar cells of the pancreas and the salivary glands are the major tissue sources of serum AMY. Lesser concentrations are found in skeletal muscle and the small intestine and fallopian tubes. AMY is a smaller enzyme, with a molecular weight of 50,000 to 55,000. Because of its small size, it is readily filtered by the renal glomerulus and appears in the urine.

Digestion of starches begins in the mouth with the hydrolytic action of salivary AMY. Salivary AMY activity, however, is of short duration because upon swallowing, it is inactivated by the acidity of the gastric contents. Pancreatic AMY then performs the major digestive action of starches once the polysaccharides reach the intestine.

Diagnostic Significance

The diagnostic significance of serum and urine AMY measurements is in the diagnosis of acute pancreatitis.[37] In acute pancreatitis, plasma AMY levels begin to rise 5 to 8 hours after the onset of an attack, peak at 24 hours, and return to normal levels within 3 to 5 days. Values usually exhibit a four- to six-fold increase above the ULN, but the amount of the increase in plasma enzyme activity is not related to the severity of the pancreatitis.

Disorders of tissues other than the pancreas can also produce elevations in AMY levels, so an elevated AMY level is a nonspecific finding. However, the degree of elevation of AMY is helpful, to some extent, in the differential diagnosis of acute pancreatitis. In addition, other laboratory tests (e.g., measurements of urinary AMY levels, AMY clearance studies, AMY isoenzyme studies, and measurements of plasma lipase [LPS] levels), when used in conjunction with plasma AMY measurement, increase the specificity of AMY measurements in the diagnosis of acute pancreatitis.

Other disorders causing an elevated serum AMY level include salivary gland lesions, such as mumps and parotitis, and other intra-abdominal diseases, such as perforated peptic ulcer, intestinal obstruction, cholecystitis, ruptured ectopic pregnancy, mesenteric infarction, and acute appendicitis. In addition, elevations have been reported in renal insufficiency and diabetic ketoacidosis. A decrease in urinary amylase activity has been found to be a useful early indicator of possible rejection of a transplanted pancreas.[38]

An apparently asymptomatic condition of hyperamylasemia has been noted in approximately 1% to 2% of the population. Hyperamylasemia can occur in neoplastic diseases with elevated results as high as 50 times ULN. *Macroamylasemia* is a condition that results when the AMY molecule combines with immunoglobulins to form a complex that is too large to be filtered across the glomerulus. Plasma AMY activity then increases because of the reduction in normal renal clearance of the enzyme and, consequently, the urinary excretion of AMY is abnormally low. The diagnostic significance of macroamylasemia lies in the need to differentiate it from other causes of increases in plasma amylase activity.

Much interest has been focused on the possible diagnostic use of AMY isoenzyme measurements.[37,39] Plasma AMY is a mixture of isoenzymes that can be separated, most notably by electrophoresis, although chromatography and isoelectric focusing also have been used for separation. In normal human serum, two major bands and as many as four minor bands may be seen. The bands are designated as P-type and S-type isoamylases, with varieties of each of these isoforms. P isoamylase is derived from pancreatic tissue, while S isoamylase is derived from salivary gland tissue, as well as the fallopian tube and lung.

The isoenzymes of salivary origin (S1, S2, S3) migrate most quickly, whereas those of pancreatic origin (P1, P2, P3) are slower. In normal human plasma, the isoamylases migrate in regions corresponding to the β- to α-globulin regions of protein electrophoresis. The most commonly observed fractions are P2, S1, and S2.

In acute pancreatitis, there is typically an increase in P-type activity, with P3 being the most predominant isoenzyme. However, P3 also has been detected in cases of renal failure and, therefore, is not entirely specific for acute pancreatitis. S-type isoamylase represents approximately two-thirds of AMY activity of normal plasma, whereas P-type predominates the isoamylase found in normal urine.

Assay for Enzyme Activity

Historically, methods measuring the disappearance of starch or the appearance of product were the methods of choice for the determination of amylase activity. However, smaller oligosaccharides have been found to be better, more stable substrates than starch because they provide more consistent hydrolysis products.

Glucose polymers have been synthesized containing a covalently bonded chemical tag that can be used in automated methods. As amylase acts on the larger dye-tagged polymer, smaller, labeled oligosaccharides are released. Using slide technology, the large size of the reagent polymer prevents it from traveling through the reagent layer. However, the smaller, labeled oligosaccharides are small enough to diffuse through the reagent layer where they can be quantified spectrophotometrically. The amount of color development is proportional to the activity of the amylase.

Fluorescence polarization methods use fluorescent tags bound to a large, rigid polymers. When stimulated with polarized light, the fluorescent tag is also polarized. As the polymer is hydrolyzed, smaller, fluorescent-tagged fragments are released. These free fragments are more able to move and rotate than the large, rigid-tagged polymer. This results in a decrease in fluorescence due to the movement of the smaller tagged fragments.

Coupled enzyme systems have been used to determine AMY activity by a continuous monitoring technique in which the change in absorbance of NAD^+ at 340 nm is measured. **Equation 8.23** is an example of a continuous monitoring method. For AMY activity, the optimal pH is 6.9.

$$\text{Maltopentose} \xrightleftharpoons{\text{AMS}} \text{Maltrotriose} + \text{maltose}$$

$$\text{Maltrotriose} + \text{maltose} \xrightleftharpoons{\alpha\text{-glucosidase}} 5\text{-Glucose}$$

$$5\text{-Glucose} + 5\ \text{ATP} \xrightleftharpoons{\text{Hexokinase}} 5\text{-Glucose-6-phosphate} + 5\ \text{ADP}$$

$$5\text{-Glucose-6-phosphate} + 5\ \text{NAD} + \xrightleftharpoons{\text{G-6-PD}} 5,6\text{-Phosphogluconolactone} + 5\ \text{NADH}$$

(**Eq. 8.23**)

Methods for the measurement of amylase activity using synthetic substrates also have been developed. One assay utilizes maltotetraose, a short oligomer of four glucose units, as the substrate for amylase. The maltotetraose is hydrolyzed to two maltose disaccharide molecules that are then treated with maltose phosphorylase to produce glucose-1-phosphate and inorganic phosphate. Glucose-1-phosphate is converted to glucose-6-phosphate by phosphoglucomutase, and the amount of glucose-6-phosphate is then analyzed using glucose-6-phosphate dehydrogenase assays. Use of this assay reduces potential interference from native glucose present in the sample, but the assay is not completely free of interference because the maltotetraose can be cleaved to maltotriose and glucose. Another assay using a synthetic substrate has been developed that uses an artificial substrate in which the chromophoric *para*-nitrophenol (pNP) is covalently linked to the reducing end of an oligosaccharide. The oligosaccharide, when cleaved by amylase, produces shorter pNP-labeled oligosaccharide fragments that can be treated with α-glucosidase to release the chromophore pNP for detection by colorimetry; the reaction may require the addition of alkali to stop the reaction and enhance the measurement of pNP.

Specific immunoassays are also available for measuring isoenzymes of AMY.

Source of Error

AMY in serum and urine is stable. Little loss of activity occurs at room temperature for 1 week or at 4°C for 2 months. Because plasma triglycerides suppress or inhibit serum AMY activity, AMY values may be normal in acute pancreatitis with hyperlipemia.

The administration of morphine and other opiates for pain relief before blood sampling will lead to falsely elevated serum AMY levels. The drugs may cause constriction of the sphincter of Oddi and of the pancreatic ducts, with consequent elevation of inarticulate pressure causing the regurgitation of AMY into

the plasma resulting in an artificial increase in plasma AMY levels.

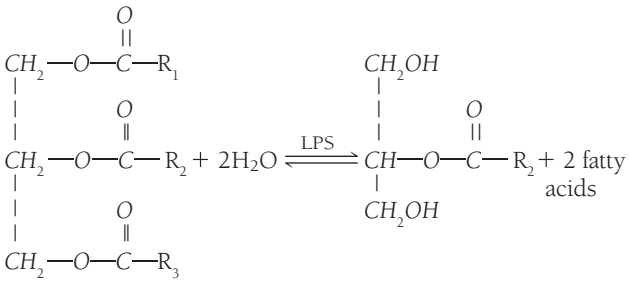
Reference Range

AMY: serum, 30–220 (37°C) U/L; urine, 6.5–48.1 IU/h

Because of the various AMY procedures currently in use, activity is expressed according to each procedure. There is no uniform expression of AMY activity, although Somogyi units are frequently used. The approximate conversion factor between Somogyi units and International Units is 1.85.

Lipase

Lipase (LPS) is an enzyme that hydrolyzes the ester linkages of fats to produce alcohols and fatty acids from triglycerides. The action of LPS preferentially catalyzes the partial hydrolysis of dietary triglycerides in the intestine producing the 2-monoglyceride intermediate, with the production of long-chain fatty acids. This reaction proceeds according to **Equation 8.24**:



(**Eq. 8.24**)

The enzymatic activity of pancreatic LPS shows preference for the removal of the fatty acids from positions 1 and 3 of the triglyceride molecule, but substrate must be present in an emulsion for activity to occur. The reaction rate is accelerated by the presence of colipase and a bile salt. Under alkali conditions, the lipase can show activity toward all three positions, such that glycerol with three free hydroxyl groups can be produced.

Tissue Source

LPS concentration is found primarily in the pancreas, although it is also present in the stomach and small intestine.

Diagnostic Significance

Clinical assays of serum LPS measurements are confined almost exclusively to the diagnosis of acute pancreatitis. Plasma LPS activity increases 4 to 8 hours after an attack of acute pancreatitis, with concentrations peaking at 24 hours and then decreasing within 8 to 14 days. LPS is similar in this respect to AMY measurements but is considered more specific for pancreatic disorders than AMY measurement. Both AMY and LPS levels rise quickly, but LPS elevations persist for approximately 8 days in acute pancreatitis, whereas AMY elevations persist for only 2 to 3 days; this is due to the fact that LIP is not cleared by the kidneys as AMY is, and so levels remain elevated longer. The extent of elevations does not correlate with severity of disease. Elevated LPS levels also may be found in other intra-abdominal conditions but with less frequency than elevations of serum AMY. Elevations of LPS have been reported in cases of penetrating duodenal ulcers and perforated peptic ulcers, intestinal obstruction, and acute cholecystitis. In contrast to AMY levels, LPS levels remain normal in conditions of salivary gland involvement. Therefore, LPS levels are useful in differentiating plasma AMY elevation as a result of pancreatic versus salivary involvement. Three lipase isoenzymes are known; forms designated L1 and L2 are ascribed as pancreatic lipase and L3 the more generic carboxyl-ester lipase. Isoforms L1 and L2 can be separated by electrophoresis. Changes in levels of isoenzyme L2 seem to have the most significant diagnostic value when combined with the analysis of AMY levels.

Assay for Enzyme Activity

Procedures used to measure LPS activity include estimation of liberated fatty acids via titrimetric and turbidimetric methods. The reaction is outlined in **Equation 8.25**:

$$\text{Triglyceride} + 2\,H_2O \xrightleftharpoons[(pH = 8.6\text{–}9.0)]{\text{LPS}} 2\text{-Monoglyceride} + 2 \text{ fatty acids}$$

(**Eq. 8.25**)

Titrimetric analyses utilize the effective titration of a product yielded by the lipase reaction that quantitatively measures the net reaction of titrant used to react with the product. Turbidimetric methods are simpler and more rapid than titrimetric assays. Fats in solution create a cloudy emulsion. As the fats are hydrolyzed by LPS, the particles disperse, and the rate of clearing can be measured as an estimation of LPS activity. Colorimetric methods are also available based on coupled reactions to directly measure the production of glycerol allowing for the continuous monitoring of the enzyme activity.

Source of Error

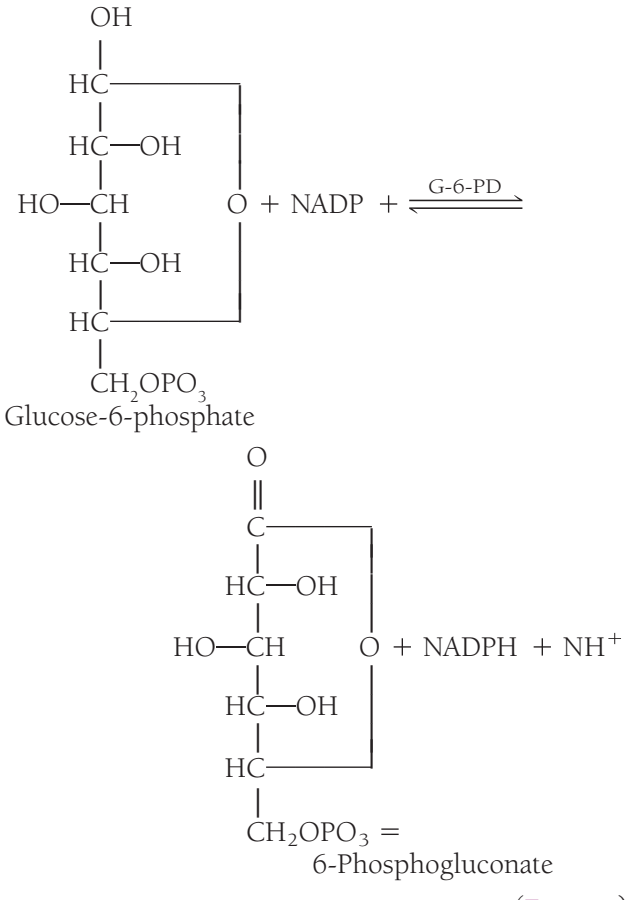
LPS is stable in plasma, with negligible loss in activity at room temperature for 1 week or for 3 weeks at 4°C. Hemolysis should be avoided because hemoglobin inhibits the activity of serum LPS, causing falsely low values.

Reference Range

LPS, < 38 U/L (37°C)

Glucose-6-Phosphate Dehydrogenase

Glucose-6-phosphate dehydrogenase (G-6-PD) is an **oxidoreductase** that catalyzes the oxidation of glucose-6-phosphate to 6-phosphogluconate or the corresponding lactone. The reaction is important as the first step in the pentose-phosphate shunt of glucose metabolism with the ultimate production of NADPH. The reaction is outlined in **Equation 8.26**:



(**Eq. 8.26**)

Tissue Source

Sources of G-6-PD include the adrenal cortex, spleen, thymus, lymph nodes, lactating mammary gland, and erythrocytes. Little activity is found in normal plasma.

Diagnostic Significance

Most of the interest of G-6-PD focuses on its role in the erythrocyte. Here, it functions to maintain NADPH in reduced form. An adequate concentration of NADPH is required for anabolic pathways of metabolism and the maintenance of glutathione levels from the oxidized to the reduced state. Glutathione in the reduced form, in turn, protects hemoglobin from oxidation by agents that may be present in the cell. A deficiency of G-6-PD results in an inadequate supply of NADPH and, ultimately, in the inability to maintain reduced glutathione levels. When erythrocytes are exposed to oxidizing agents, hemolysis occurs because of oxidation of hemoglobin and damage of the cell membrane.

G-6-PD deficiency is an inherited sex-linked trait. The disorder can result in several different clinical manifestations, including drug-induced hemolytic anemia. When exposed to an oxidant drug such as the antimalarial drug primaquine, affected individuals experience a hemolytic episode. The severity of the hemolysis is related to the drug concentration. G-6-PD deficiency is found in all races and ethnic backgrounds, but is most common in those of African, Asian, and Mediterranean ethnicity. Increased levels of G-6-PD in the serum have been reported in MI and megaloblastic anemias. No elevations are seen in hepatic disorders, however, as G-6-PD levels are not routinely performed as diagnostic aids in these conditions.

Assay for Enzyme Activity

The assay procedure for G-6-PD activity is outlined in **Equation 8.27**:

$$\text{Glucose-6-phosphate} + \text{NADPH}^{+} \xrightleftharpoons{\text{G-6-PD}} \text{6-Phosphogluconate} + \text{NADPH} + \text{H}^{+}$$

(**Eq. 8.27**)

A red cell hemolysate is used to assay for deficiency of the enzyme; serum is used for evaluation of enzyme elevations.

Reference Range

G-6-PD, 7.9–16.3 U/g Hgb

Macroenzymes

Macroenzymes are forms of the plasma enzymes (ACP, ALP, ALT, AMY, AST, CK, GGT, LD, LPS) that are bound to either an immunoglobulin (macroenzyme

type 1) or a nonimmunoglobulin substances (macroenzyme type 2), such as lipoproteins or self-polymerization. Macroenzymes are usually found in patients who have an unexplained persistent increase of enzyme concentrations in serum.[40] The presence of macroenzymes can also increase with increasing age.[41] Enzymes can bind to immunoglobulins in a nonspecific manner, but there is also evidence that the enzyme–immunoglobulin complex can be formed by specific interactions between circulating autoantibodies and serum enzymes.[40] The reason for the formation of antienzyme antibodies is not known, but there are two theories to explain their formation. According to the "antigen-driven theory," the self-antigen becomes immunogenic by being altered or released from a sequestered site and reacts with an antibody that is initially formed against a foreign antigen.[42] The dysregulation of immune tolerance theory explains the formation of enzymes with autoantibodies in patients with autoimmune disorders.[42] To date, no strong correlation has been found between the presence of antienzyme antibodies and the pathogenesis of disease.[40] However, the presence of macroenzymes should be documented in the patient's medical records because macroenzymes can persist for long periods.[40]

Macroenzymes accumulate in the plasma because their high molecular masses prevent them from being filtered out of the plasma by the kidneys. The detection of macroenzymes is clinically significant because the presence of macroenzymes can cause difficulty in the interpretation of diagnostic enzyme results. The formation of high-molecular-weight enzyme complexes can cause false elevations in plasma enzymes, or they can falsely decrease the activity of the enzyme by blocking the activity of the bound enzyme.[40]

The principal method to identify enzymes that are bound to immunoglobulins and nonimmunoglobulins is protein electrophoresis. The binding of enzymes to high-molecular-weight complexes can alter the normal electrophoretic pattern of enzymes (see Figure 8.5 for an example). Antienzyme antibodies can cause the formation of new enzyme bands on a gel, they can alter the intensity of enzyme bands, and they can cause band broadening on the gel.[40] Other test methods used to determine the presence of macroenzymes include gel filtration, immunoprecipitation, immunoelectrophoresis, counterimmunoelectrophoresis, and immunofixation. Last, the immunoinhibition test can also be used to determine the presence of macro-CK.

Drug-Metabolizing Enzymes

Drug-metabolizing enzymes function primarily to transform xenobiotics into inactive, water-soluble compounds for excretion through the kidneys. Metabolic enzymes can also transform inactive prodrugs into active drugs, convert xenobiotics into toxic compounds, or prolong the elimination half-life. Drug-metabolizing enzymes act as either phase I or phase II reactions. The enzymes involved in phase I reactions catalyze addition or removal of functional groups through hydroxylation, oxidation, dealkylation, dehydrogenation, reduction, deamination, and desulfonylation reactions and are often mediated by cytochrome P-450 (CYP 450) enzymes. Xenobiotics can also become transformed into more polar compounds through enzyme-mediated conjugation reactions, also known as phase II reactions, in which xenobiotics are conjugated with glucuronide (UDP-glucuronosyltransferase 1A1 [UGT1A1]), acetate (*N*-acetyltransferase [NAT]), glutathione (glutathione-*S*-transferase [GST]), sulfate (sulfotransferase), and methionine groups.[43]

CYP 450 enzymes are a superfamily of isoenzymes that are involved in the metabolism of more than 50% of all drugs. These enzymes contain heme molecules, and they are given the name CYP 450 because they absorb the maximum amount of light at 450 nm. More than 500 CYP 450 enzymes that have been identified, and they are classified into families according to their homology to other enzymes.[44] There are at least four CYP 450 (CYP 1, 2, 3, and 4) families that are expressed primarily in the liver, but some isoforms are also expressed in extrahepatic tissues such as the lung, kidneys, gastrointestinal tract, skin, and placenta.[43] The specific isozyme is classified by not only its family number but also by a subfamily letter, a number for an individual isozyme within the subfamily, and, if applicable, an asterisk followed by a number for each genetic (allelic) variant. Genetic variants have been identified that lead to complete enzyme deficiency (e.g., a frame shift, splice variant, stop codon, or a complete gene deletion), reduced enzyme function or expression, or enhanced enzyme function or expression. Recognition of genetic variants can explain interindividual differences in drug response and pharmacokinetics.[44–47] For example, four phenotypes are recognized for CYP 2D6: ultrametabolizers, extensive metabolizers, intermediate metabolizers, and poor metabolizers. Patients who are poor metabolizers for the CYP 2D6 enzyme are at risk for therapeutic failure when inactive prodrugs

such as tamoxifen require CYP 2D6 for drug activation. Tricyclic antidepressants such as nortriptyline require CYP 2D6 for inactivation. Thus, CYP 2D6 poor metabolizers may require lower doses than patients with extensive ("normal") metabolism and may be at high risk for adverse drug reactions.[43,46]

In addition to xenobiotic metabolism, CYP 450 enzymes are also involved in the biosynthesis of endogenous compounds. The CYP 5 family consists of thromboxane synthases that catalyze the reaction that leads to platelet aggregation. CYP 7 and CYP 27 families catalyze the hydroxylation of cholesterol for the biosynthesis of bile acids. The CYP 24 family catalyzes the hydroxylation and inactivation vitamin D_3. CYP 450 enzymes are also found in steroid-producing tissues and function to synthesize steroid hormones from cholesterol (CYP 11, 17, 19, and 21).[44]

Genetic variants that affect drug-metabolizing enzyme function and expression are recognized for other enzymes such as NAT, UGT1A1, GST, and TPMT. These variants are associated with distinct extensive (fast), intermediate, or poor (slow) metabolizer phenotypes, which could lead to adverse drug reactions or therapeutic failure. For example, two phenotypes—fast and slow acetylators—are recognized for *N*-acetyltransferase 2 (NAT2). NAT2 is the primary enzyme involved in the acetylation of isoniazid, a drug used to treat tuberculosis. Acetylation is the primary mechanism for the elimination of isoniazid, and therefore, patients with low NAT2 activity will not be able to inactivate isoniazid, putting those patients at increased risk for adverse drug reactions.[43] UGT1A1 has polymorphisms that can lead to a nonfunctioning enzyme. UGT1A1 is responsible for the metabolism of bilirubin; patients with nonfunctioning UGT1A1 are at risk for hyperbilirubinema.[47] Last, thiopurine methyltransferase (TPMT) is an enzyme that can be found in bone marrow and erythrocytes and functions to inactivate chemotherapeutic thiopurine drugs like azathioprine and 6-mercaptopurine. The TPMT enzyme has genetic polymorphisms, which causes variable responses (normal, intermediate, and low activity) to thiopurine metabolism. Patients with low TPMT activity are at risk of developing severe bone marrow toxicity when the standard-dose therapy for thiopurine drugs is administered; thus, genetic testing is essential for identifying patients with metabolizing enzyme polymorphisms.[47]

Pharmacogenetic testing is often used prior to drug therapy to assist clinicians in identifying patients with genetic polymorphisms, to guide drug and dose selection. Pharmacogenetic testing can be performed through phenotype tests that measure metabolic enzyme activity, through administration of a probe drug and subsequent evaluation of metabolic ratios, or through genotype testing that identifies clinically significant genetic variants.

The activity of drug-metabolizing enzymes can also be altered by food, nutritional supplements, or other drugs. Compounds that stimulate an increase in the synthesis CYP 450 enzymes are called *inducers*. Inducers will increase the metabolism of drugs and reduce the bioavailability of the parent compound. Compounds that reduce the expression or activity of a drug-metabolizing enzyme are referred to as *inhibitors*. For example, inhibitors can compete with substrates for the active site of the CYP 450 and thereby decrease the metabolism of drugs and increase the bioavailability of the parent compound, or block activity or expression through noncompetitive means.[43] **Table 8.5** lists the common families of CYP 450 enzymes, along with some of their substrates and drugs that can induce or inhibit enzyme activity.

Table 8.5 Common Substrates for Drug-Metabolizing Enzymes

Enzyme	Substrates	Inducers	Inhibitors
CYP 1A1	(*R*)-Warfarin	Omeprazole	
		TCDD	
		Benzo[a]pyrene, 3MC	
CYP 1A2	Acetaminophen	Insulin	Ciprofloxacin
	Caffeine	Tobacco	Cimetidine
	(*R*)-Warfarin	Polycyclic aromatic hydrocarbons	Amiodarone
	Estradiol		Fluoroquinolones
	Theophylline		

CYP 2A6	Cyclophosphamide	Dexamethasone	Pilocarpine
	Halothane		Coumarin
	Zidovudine		
	Coumarin		
CYP 2B6	Cyclophosphamide	Phenobarbital	Ticlopidine
	Diazepam	Rifampin	
	Bupropion		
CYP 2C9	(S)-Warfarin	Rifampin	Fluconazole
	Ibuprofen	Secobarbital	Amiodarone
	Tolbutamide		Isoniazid
	Diclofenac		Probenecid
	Losartan		Sertraline
	Phenytoin		Sulfamethoxazole
CYP 2C19	Diazepam	Barbiturate	Omeprazole
	Omeprazole	Phenytoin	Chloramphenicol
	Clomipramine		Cimetidine
	Indomethacin		Ketoconazole
			Indomethacin
CYP 2D6	Carvedilol	Dexamethasone	Bupropion
	Amitriptyline	Rifampin	Fluoxetine
	Haloperidol		Quinidine
	Amphetamine		Amiodarone
	Chlorpromazine		Sertraline
	Dextromethorphan		Celecoxib
	Codeine		Chlorpromazine
CYP 2E1	Acetaminophen	Ethanol	Disulfiram
	Chlorzoxazone	Isoniazid	Diethyldithiocarbamate
	Halothane		
	Ethanol		
CYP 3A4	Erythromycin	HIV antivirals	HIV antivirals
	Quinidine	Barbiturates	Ketoconazole
	Diazepam	Carbamazepine	Erythromycin
	Cortisol	Phenobarbital	Grapefruit juice
	Cyclosporine	Phenytoin	Cimetidine
	Indinavir	Rifampin	Chloramphenicol
	Chlorpheniramine	St. John's wort	
	Nifedipine	Troglitazone	

(continues)

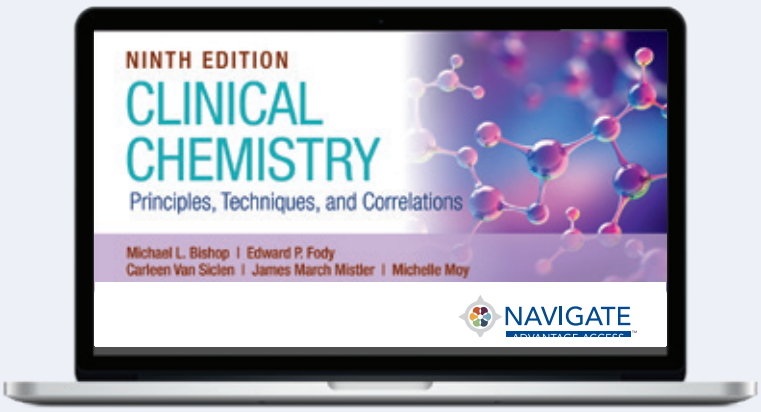
Table 8.5 Common Substrates for Drug-Metabolizing Enzymes *(continued)*

Enzyme	Substrates	Inducers	Inhibitors
	Lovastatin		
	Testosterone		
	Cocaine		
	Fentanyl		
	Tamoxifen		
TPMT	Azathioprine	Naproxen	TPMT
	6-Mercaptopurine	Furosemide	

Data from Rendic S, Di Carlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metab Rev.* 1997;29:413–580.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Enzyme Nomenclature 1978. *Recommendations of the Nomenclature Committee of the International Union of Biochemistry on the Nomenclature and Classification of Enzymes*. New York, NY: Academic Press; 1979.
2. Baron DN, Moss DW, Walker PG, Wilkinson JH. Abbreviations for names of enzymes of diagnostic importance. *J Clin Pathol.* 1971;24:656–657.
3. Baron DN, Moss DW, Walker PG, Wilkinson JH. Revised list of abbreviations for names of enzymes of diagnostic importance. *J Clin Pathol.* 1975;28:592–593.
4. Rej R. Accurate enzyme measurements. *Arch Pathol Lab Med.* 1998;117:352–364.
5. Spitzer M, Pinto A, Dasgupta R, Benjamin D. Early diagnosis of ectopic pregnancy: can we do it accurately using a chemical profile? *J Womens Health Gend Based Med.* 2000;9(5):537–544.
6. Na Kafusa J, Misago N, Miura Y, Kayaba M, Tanaka T, Narisawa Y. The importance of serum creatine phosphokinase levels in the early diagnosis, and as a prognostic factor, of *Vibrio vulnificus* infection. *Br Med J.* 2001;145(2):280–284.
7. Galen RS. The enzyme diagnosis of myocardial infarction. *Hum Pathol.* 1975;6(2):141–155.
8. Wilkinson JH. *The Principles and Practice of Diagnostic Enzymology*. Chicago, IL: Year Book Medical Publishers; 1976:395.
9. Silverman LM, Dermer GB, Zweig MH, Van Steirteghem AC, Tókés ZA. Creatine kinase BB: a new tumor-associated marker. *Clin Chem.* 1979;25(8):1432–1435.
10. Lang H, Ed. *Creatine Kinase Isoenzymes: Pathophysiology and Clinical Application*. Berlin, Germany: Springer-Verlag; 1981.
11. Irvin RG, Cobb FR, Roe CR. Acute myocardial infarction and MB creatine phosphokinase: relationship between onset of symptoms of infarction and appearance and disappearance of enzyme. *Arch Intern Med.* 1980;140:329–334.
12. Bark CJ. Mitochondrial creatine kinase—a poor prognostic sign. *JAMA.* 1980;243:2058–2060.

13. Bais R, Edwards JB. Creatine kinase. *Crit Rev Clin Lab Sci.* 1982;16:291–335.
14. Lang H, Wurzburg U. Creatine kinase, an enzyme of many forms. *Clin Chem.* 1982;28:1439–1447.
15. Lott JA. Serum enzyme determinations in the diagnosis of acute myocardial infarction: an update. Hum Pathol. 1984 Aug;15(8):706-16. https://pubmed.ncbi.nlm.nih.gov/6745913/.
16. Pesce MA. The CK isoenzymes: findings and their meaning. *Lab Mgmt.* 1982.
17. Roche Diagnostics. *Isomune-CK Package Insert*. Nutley, NJ: Hoffman-La Roche; 1979.
18. Faulkner WR, Meites S, Eds. *Selected Methods for the Small Clinical Chemistry Laboratory: Selected Methods of Clinical Chemistry*. Vol. 9. Washington, DC: American Association for Clinical Chemistry; 1982:475.
19. Jialal I, Sokoll LJ. Clinical utility of lactate dehydrogenase. *Am J Clin Pathol.* 2015;143:158–159.
20. Martins JT, Li DJ, Jialal I, Keffer JH. Comparison of cardiac troponin I and lactate dehydrogenase isoenzymes for the late diagnosis of myocardial injury. *Am J Clin Pathol.* 1996;106:705–708.
21. Giannini EG, Testa R Savarino V. Liver enzyme alteration: a guide for clinicians. *Can Med Assoc J.* 2005;172:367–379.
22. Posen S, Doherty E. The measurement of serum alkaline phosphatase in clinical medicine. *Adv Clin Chem.* 1981;22: 165–245.
23. Warren BM. *The Isoenzymes of Alkaline Phosphatase*. Beaumont, TX: Helena Laboratories; 1981.
24. Fleisher GA, Eickelberg ES, Elveback LR. Alkaline phosphatase activity in the plasma of children and adolescents. *Clin Chem.* 1977;23:469–472.
25. Milincovic N, Saric-Matutinovic M, Pejanovic S, Ignjatovic S. Comparison between bone alkaline phosphatase immunoassay and electrophoresis technique in hemodialysis patients. *J Med Biochem.* 2020;39:178–183.
26. Ahmed F, Gibbons SM. Bone-specific alkaline phosphatase by immunoassay or electrophoresis: their use in clinical practice. *J Clin Pathol.* 2015; 68:246–248.
27. Gittes RF. Carcinoma of the prostate. *N Engl J Med.* 1991;324:236–245.
28. Gittes RF. Prostate-specific antigen. *N Engl J Med.* 1987;318:954–955.
29. Oesterling JE. Prostate specific antigen: a critical assessment of the most useful tumor marker for adenocarcinoma of the prostate. *J Urol.* 1991;145:907–923.
30. Griffiths JC. The laboratory diagnosis of prostatic adenocarcinoma. *Crit Rev Clin Lab Sci.* 1983;19:187–204.
31. Yam LT. Clinical significance of the human acid phosphatases: a review. *Am J Med.* 1974;56:604–616.
32. Lantz RK, Berg MJ. From clinic to court: acid phosphatase testing. *Diagn Med.* 1981;Mar/Apr:55.
33. Rosalki SB. Gamma-glutamyl transpeptidase. *Adv Clin Chem.* 1975;17:53–107.
34. Sunderman FW. The clinical biochemistry of 5′-nucleotidase. *Ann Clin Lab Sci.* 1990;123–139.
35. Hyder MA, Hasan M Mohieldein A. Comparative study of 5′-nucleotidase test in various liver diseases. *J Clin Diag Res* 2016;10:BC01–BC03.
36. Moses GC, Tukerman JF, Henderson AR. Biological variance of cholinesterase and 5′-nucleotidase in serum of healthy persons. *Clin Chem* 1986; 32:175–177.
37. Salt WB II, Schenker S. Amylase: its clinical significance. A review of the literature. *Medicine.* 1976;4:269.
38. Han DJ, Sutherland DER. Pancreas transplantation. *Gut and Liver* 2010; 4:450–465.
39. Murthy U, DeGregorio F, Oates RP, Blair DC. Hyperamylasemia in patients with acquired immunodeficiency syndrome. *Am J Gastroenterol.* 1992;87:332–336.
40. Remaley AT, Wilding P. Macroenzymes: biochemical characterization, clinical significance, and laboratory detection. *Clin Chem.* 1989;35:2261–2270.
41. Turecky L. Macroenzymes and their clinical significance. *Bratisl Lek Listy.* 2004;105:260–263.
42. Theofilopoulous AN. Autoimmunity. In: Stites DP, Stobo JD, Wells JV, eds. *Basic and Clinical Immunology*. 6th ed. Norwalk, CT: Appleton & Lange; 1987:128–158.
43. Hardman JG, Limbird LE, eds. *Goodman and Gilman's The Pharmacological Basis of Therapeutics*. 9th ed. New York, NY: McGraw-Hill; 1996:11–16.
44. Rendic S, Di Carlo FJ. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metab Rev.* 1997;29:413–580.
45. Wilkinson GR. Drug metabolism and variability among patients in drug response. *N Engl J Med.* 2005;352:2211–2221.
46. Nebert DW, Dieter MZ. The evolution of drug metabolism. *Pharmacology.* 2000;61:124–135.
47. McMillin GA. *Pharmacogenetics Fundamental of Molecular Diagnostics*. St. Louis, MO: W.B. Saunders; 2007:197–215.

CHAPTER 9

Carbohydrates

Vicki S. Freeman

CHAPTER OUTLINE

KEY TERMS

Albuminuria
Carbohydrates
Diabetes mellitus
Disaccharides
Embden-Meyerhof pathway
Glucagon
Gluconeogenesis
Glucose
Glycogen
Glycogenesis
Glycogenolysis
Glycolysis
Glycosylated hemoglobin
Hemoglobin A1c
Hyperglycemic
Hypoglycemic
Insulin
Ketone
Monosaccharides
Oligosaccharides
Polysaccharides
Trioses

CHAPTER OBJECTIVES

At the end of this unit of study, the clinical laboratorian should be able to:

- Classify carbohydrates into their respective groups.
- Discuss the metabolism of carbohydrates in the body and the mode of action of hormones in carbohydrate metabolism.
- Differentiate the types of diabetes by clinical symptoms and laboratory findings according to the American Diabetes Association.
- Explain the clinical significance of the three ketone bodies.
- Relate expected laboratory results and clinical symptoms to the following metabolic complications of diabetes:
  - Ketoacidosis
  - Hyperosmolar coma
- Distinguish between reactive and spontaneous hypoglycemia.

- State the principle, specimen of choice, and the advantages and disadvantages of the glucose analysis methods.
- List the three commonly encountered methods of glycosylated hemoglobin testing, specimen(s) of choice, and sources of error.
- Interpret glycosylated hemoglobin values used for long-term monitoring of diabetes.
- Specify the methods of analysis and the advantages and disadvantages of ketone bodies.

CASE STUDY 9.1, PART 1

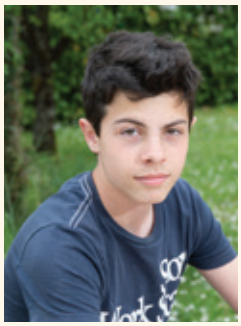
Dexter is a 47-year-old African American male with longstanding obesity (BMI 32 kg/m^2). His BP is 160/90 mm Hg. He states that he has a desk job and gets very little physical activity, and he notes that he often wakes up in the night needing to urinate. His father has diabetes mellitus. His physician suggests that he be screened for diabetes.

© sirtravelalot/Shutterstock.

1. Consider what tests the physician should order to determine if Dexter has diabetes.

CASE STUDY 9.2, PART 1

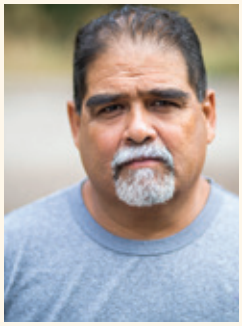
John, an 18-year-old male high school student and athlete who had a 4-year history of diabetes mellitus, was brought to the emergency department by his parents because of excessive drowsiness, vomiting, and diarrhea. His diabetes had been well controlled with approximately 45–60 units per day of insulin lispro (short-acting insulin) infused via an insulin pump. This regimen has been effective at maintaining stable blood sugar levels until several days ago, when John developed headaches, myalgia, and a low-grade fever. He has complained of excessive thirst and polyuria for the past 2 days; he has also had diarrhea, and vomiting began this morning.



© sylv1rob1/Shutterstock.

CASE STUDY 9.3, PART 1

Michael, a 58-year-old obese man, complains of frequent urination to his primary care physician.



© adamkaz/E+/Getty Images.

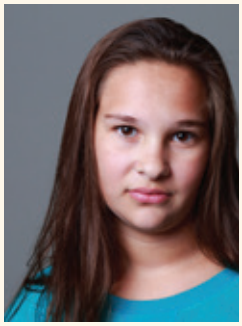
CASE STUDY 9.4, PART 1

James, a 14-year-old male student, was seen by his physician. His chief complaints are fatigue, weight loss, and increases in appetite, thirst, and frequency of urination.

© Tony Anderson/Stone/Getty Images.

CASE STUDY 9.5, PART 1

Myra, a 13-year-old girl, collapsed while playing dodgeball during recess at school.



© Kyle Lee/Shutterstock.

CASE STUDY 9.6, PART 1

Jo Ann, a 28-year-old woman, delivered a 9.5-lb infant daughter early this morning. The infant, Martha, was above the 95th percentile for weight and length.

© yasindmrblk/iStock/Getty Images Plus/Getty Images.

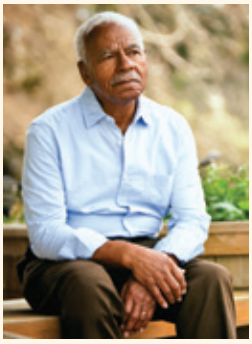
CASE STUDY 9.7, PART 1

Emily is a slender, 50-year-old White woman who reported no particular health issues during her annual physical examination.

© goodluz/Shutterstock.

CASE STUDY 9.8, PART 1

Alfred, a 63-year-old male, has been previously diagnosed with type 2 diabetes. He admits he does not always follow his physician's advice on diabetes management.



© Erickson Stock/Shutterstock.

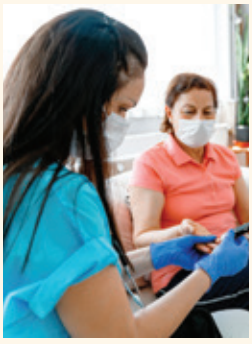
CASE STUDY 9.9, PART 1

Angela, a healthy 25-year-old female, complains of dizziness and shaking 1 hour after eating a large, high-carbohydrate meal with her friends.

© Tuzemka/Shutterstock.

CASE STUDY 9.10, PART 1

Shirley is a hospice nurse caring for patients with diabetes. She performed a fingerstick glucose test on Janet using an Accu-Chek glucometer.



© Phynart Studio/E+/Getty Images.

Organisms rely on the oxidation of complex organic compounds to obtain energy. Three general types of such compounds are carbohydrates, amino acids, and lipids. Although all three are used as a source of energy, carbohydrates are the primary source for brain, erythrocytes, and retinal cells in humans. **Carbohydrates** are the major food source and energy supply for the body and are stored primarily in the liver and muscle as glycogen. Disease states involving carbohydrates are split into two groups—hyperglycemia (which can lead to diabetes mellitus) and hypoglycemia. Early detection of diabetes mellitus is the aim of the American Diabetes Association (ADA) guidelines. Acute and chronic complications may be avoided with proper diagnosis, monitoring, and treatment. The laboratory plays an important role through periodic measurements of glycosylated hemoglobin and albuminuria.

General Description of Carbohydrates

Carbohydrates are compounds containing C, H, and O. The general formula for a carbohydrate is $C_x(H_2O)_y$. All carbohydrates contain C=O and –OH functional groups. There are some deviations from this basic formula because carbohydrate derivatives can be formed by the addition of other chemical groups, such as phosphates, sulfates, and amines. The classification of carbohydrates is based on four structural properties: (1) the size of the base carbon chain, (2) the location of the CO function group, (3) the number of sugar units, and (4) the stereochemistry of the compound.

Classification of Carbohydrates

Carbohydrates can be grouped into generic classifications based on the number of carbons in the molecule. For example, **trioses** contain three carbons, tetroses contain four, pentoses contain five, and hexoses contain six. In actual practice, the smallest carbohydrate is glyceraldehyde, a three-carbon compound.

Carbohydrates are hydrates of aldehyde or ketone derivatives based on the location of the CO functional group (**Figure 9.1**). The two forms of carbohydrates are aldose and ketose (**Figure 9.2**). The aldose form has a terminal carbonyl group (O=CH–) called an aldehyde group, whereas the ketose form has a carbonyl group (O=C) in the middle linked to two other carbon atoms (called a ketone group).

Figure 9.1 Carbohydrate derivatives in glucose metabolism.

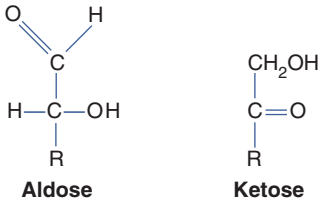


Figure 9.2 Two forms of carbohydrates.

Several models are used to represent carbohydrates. The Fisher projection of a carbohydrate has the aldehyde or ketone at the top of the drawing. The carbons are numbered starting at the aldehyde or ketone end. The compound can be represented as a straight chain or might be linked to show a representation of the cyclic, hemiacetal form (**Figure 9.3**). The Haworth projection represents the compound in the cyclic form that is more representative of the actual structure. This structure is formed when the functional (carbonyl) group (ketone or aldehyde) reacts with an alcohol group on the same sugar to form a ring called either a hemiketal or a hemiacetal ring, respectively (**Figure 9.4**).

Stereoisomers

The central carbons of a carbohydrate are asymmetric (chiral)—four different groups are attached to the carbon atoms. This allows for various spatial arrangements around each asymmetric carbon (also called stereogenic centers) forming molecules called *stereoisomers*. Stereoisomers have the same order and types of bonds but different spatial arrangements and different properties. For each asymmetric carbon, there

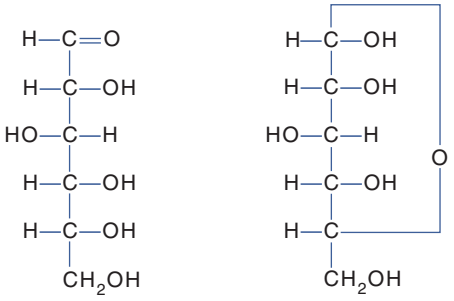


Figure 9.3 Fisher projection of glucose. **Left:** Open-chain Fisher projections. **Right:** Cyclic Fisher projection.

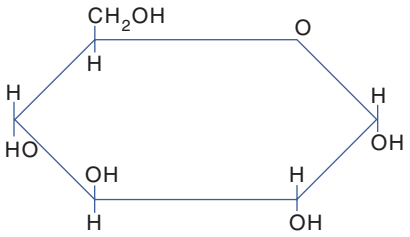


Figure 9.4 Haworth projection of glucose.

are 2^n possible isomers; therefore, there are 2^1, or two, forms of glyceraldehyde. Because an aldohexose contains four asymmetric carbons, there are 2^4, or 16, possible isomers. A monosaccharide is assigned to the D or the L series according to the configuration at the highest numbered asymmetric carbon. This asymmetrically substituted carbon atom is called the "configurational atom" or chiral center. Thus, if the hydroxyl group (or the oxygen bridge of the ring form) projects to the right in the Fisher projection, the sugar belongs to the D series and receives the prefix d-, and if it projects to the left, then it belongs to the L series and receives the prefix l-. These stereoisomers, called enantiomers, are images that cannot be overlapped and are nonsuperimposable. In **Figure 9.5**, d-glucose is represented in the Fisher projection with the hydroxyl group on carbon number 5 positioned on the right, whereas l-glucose has the hydroxyl group of carbon number 5 positioned on the left. Most naturally occurring sugars in humans are in the d-form.

Monosaccharides, Disaccharides, and Polysaccharides

Another classification of carbohydrates is based on the number of sugar units in the chain: monosaccharides (one), disaccharides (two), oligosaccharides (three to ten), and polysaccharides (ten or more). This chaining

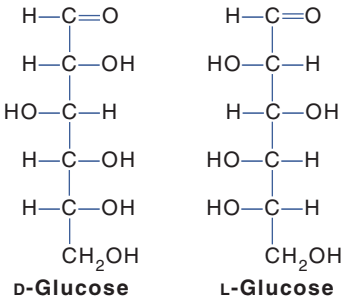


Figure 9.5 Stereoisomers of glucose.

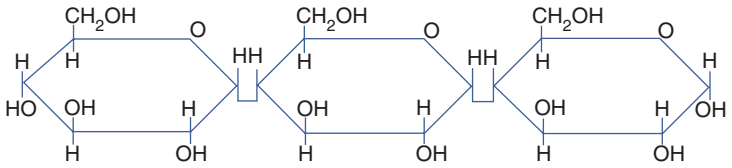


Figure 9.6 Linkage of monosaccharides.

of sugars relies on the formation of glycoside bonds that are bridges of oxygen atoms. When two carbohydrate molecules join, a water molecule is released. When they split, one molecule of water is consumed to form the individual sugar compounds. This reaction is called *hydrolysis*. The glycoside linkages of carbohydrate can involve any number of carbons; however, certain carbons are favored, depending on the carbohydrate. **Monosaccharides** are simple sugars that cannot be hydrolyzed to a simpler form; there is one sugar molecule. These sugars can contain three, four, five, or six or more carbon atoms (known as trioses, tetroses, pentoses, and hexoses, respectively). The most common hexose monosaccharides include glucose, fructose, and galactose.

Disaccharides are formed when two monosaccharide units are joined by a glycosidic linkage. On hydrolysis, disaccharides will be split into two monosaccharides by disaccharide enzymes (e.g., lactase) located on the microvilli of the intestine. These monosaccharides are then actively absorbed. The most common disaccharides are maltose (comprising two d-glucose molecules in a 1 → 4 linkage), lactose, and sucrose.

Oligosaccharides are the chaining of 3 to 10 sugar units, whereas **polysaccharides** are formed by the linkage of many monosaccharide units. On hydrolysis, polysaccharides will yield 10 or more monosaccharides. Amylase, an enzyme found in saliva and in the stomach, hydrolyzes starch to disaccharides in the duodenum. The most common polysaccharides are starch (plant based glucose molecules) and glycogen (animal based glucose molecules) (**Figure 9.6**).

Chemical Properties of Carbohydrates

Some carbohydrates are reducing substances; these carbohydrates can reduce other compounds while they themselves are oxidized. To be a reducing substance, the carbohydrate must contain an active (available) ketone or an aldehyde group. This property was used in many past laboratory methods for the determination of carbohydrates.

Carbohydrates can form glycosidic bonds with other carbohydrates and with noncarbohydrates. Two sugar molecules can be joined in tandem, forming a glycosidic bond between the hemiacetal group of one molecule and the hydroxyl group on the other molecule. In forming the glycosidic bond, an acetal is generated on one sugar (at carbon 1) in place of the hemiacetal. If the bond forms with one of the other carbons on the carbohydrate other than the anomeric (reducing) carbon, the anomeric carbon is unaltered, and the resulting compound remains a reducing substance. Examples of common-reducing sugars include glucose, maltose, fructose, lactose, and galactose. If a glycosidic bond is formed with the anomeric carbon on the other carbohydrate, the resulting compound is no longer a reducing substance. Nonreducing carbohydrates *do not* have an active ketone or aldehyde group and therefore will not reduce other compounds. The most common nonreducing sugar is sucrose—table sugar (**Figure 9.7**).

All monosaccharides and many disaccharides are reducing agents. This is because a free aldehyde or ketone group (the open chain form) can be oxidized under the proper conditions. A disaccharide remains a reducing agent when the hemiacetal or ketal hydroxyl group is not linked to another molecule; both maltose and lactose are reducing agents, whereas sucrose is not.

Glucose Metabolism

Glucose is a primary source of energy for humans. The nervous system, including the brain, totally depends on glucose drawn from the surrounding extracellular fluid (ECF) for energy. Nervous tissue

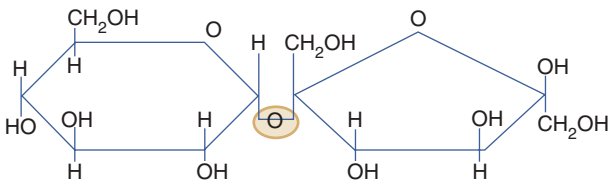


Figure 9.7 Haworth projection of sucrose.

cannot concentrate or store carbohydrates; therefore, it is critical to maintain a steady supply of glucose to the tissue. For this reason, the concentration of glucose in the ECF must be maintained in a narrow range. When the concentration falls below a certain level, the nervous tissue loses the primary energy source and is incapable of maintaining normal function.

Fate of Glucose

Most of our ingested carbohydrates are polymers, such as starch and **glycogen**. Salivary amylase and pancreatic amylase are responsible for the digestion of these nonabsorbable polymers to dextrins and disaccharides, which are further hydrolyzed to monosaccharides by maltase, an enzyme released by the intestinal mucosa. Sucrase and lactase are two other important gut-derived enzymes that hydrolyze sucrose to glucose and fructose and lactose to glucose and galactose.

When disaccharides are converted to monosaccharides, they are absorbed by the gut and transported to the liver by the hepatic portal venous blood supply. Glucose is the only carbohydrate that can be directly used for energy or stored as glycogen. Galactose and fructose must be converted to glucose before they can be used as energy. After glucose enters the cell, it is quickly shunted into one of three possible metabolic pathways, depending on the availability of substrates or the nutritional status of the cell. The ultimate goal of the cell is to convert glucose to carbon dioxide and water. During this process, the cell obtains the high-energy molecule adenosine triphosphate (ATP) from inorganic phosphate and adenosine diphosphate. The cell requires oxygen for the final steps in the electron transport chain (ETC). Nicotinamide adenine dinucleotide (NAD) in its reduced form (NADH) will act as an intermediate to couple glucose oxidation to the ETC in the mitochondria where much of the ATP is gained.

The first step for all three pathways requires glucose to be converted to glucose-6-phosphate using the high-energy molecule, ATP. This reaction is catalyzed by the enzyme hexokinase (**Figure 9.8**). Glucose-6-phosphate can enter the **Embden-Meyerhof pathway**, the *hexose monophosphate pathway* (HMP), or it can be converted to glycogen (Figure 9.8). The first two pathways are important for the generation of energy from glucose; the conversion to glycogen is important for the storage of glucose.

In the Embden-Meyerhof pathway, a form of glycolysis, glucose is broken down into two three-carbon molecules of pyruvic acid in the cell cytosol. Two molecules of ATP are needed to start the pathway, but 4 ATP molecules are created by the end, for a net gain of 2 ATP molecules for the cell. Pyruvic acid can then enter the *tricarboxylic acid* (TCA) *cycle* after conversion to acetyl-coenzyme A (acetyl-CoA). This pathway takes place in the mitochondria, requires oxygen, and is called the *aerobic pathway* (Figure 9.8). Other substrates can enter the pathway at several points. For example, glycerol released from the hydrolysis of triglycerides can enter at 3-phosphoglycerate, and fatty acids, ketones, and some amino acids can be converted or catabolized to acetyl-CoA before entering the TCA cycle. Other amino acids enter the pathway as pyruvate or as deaminated α-ketoacids and α-oxoacids. The conversion of amino acids by the liver and other specialized tissue, such as the kidney, to substrates that can be converted to glucose is called **gluconeogenesis**. Gluconeogenesis also encompasses the conversion of glycerol, lactate, and pyruvate to glucose.

Anaerobic **glycolysis** is important for tissue such as muscle, which often has important energy requirements without an adequate oxygen supply. These tissues can derive ATP from glucose in an oxygen-deficient environment by converting pyruvic acid into lactic acid. The lactic acid diffuses from the muscle cell, enters the systemic circulation, and is then taken up and used by the liver (Figure 9.8). For anaerobic glycolysis to occur, 2 mol of ATP must be consumed for each mol of glucose; however, 4 mol of ATP are directly produced, resulting in a net gain of 2 mol of ATP. Further gains of ATP result from the introduction of pyruvate into the TCA cycle and NADH into the ETC.

The second energy pathway is the HMP *shunt*, which is actually a detour of glucose-6-phosphate from the glycolytic pathway to become 6-phosphogluconic acid. This oxidized product permits the formation of ribose-5-phosphate and NADP in its reduced form (NADPH). NADPH is important to erythrocytes that lack mitochondria and are therefore incapable of the TCA cycle. The reducing power of NADPH is required for the protection of the cell from oxidative and free radical damage. Without NADPH, the lipid bilayer membrane of the cell and critical enzymes would eventually be destroyed, resulting in cell death. The HMP shunt also permits pentoses, such as ribose, to enter the glycolytic pathway.

When the cell's energy requirements are being met, glucose can be stored as glycogen. This third

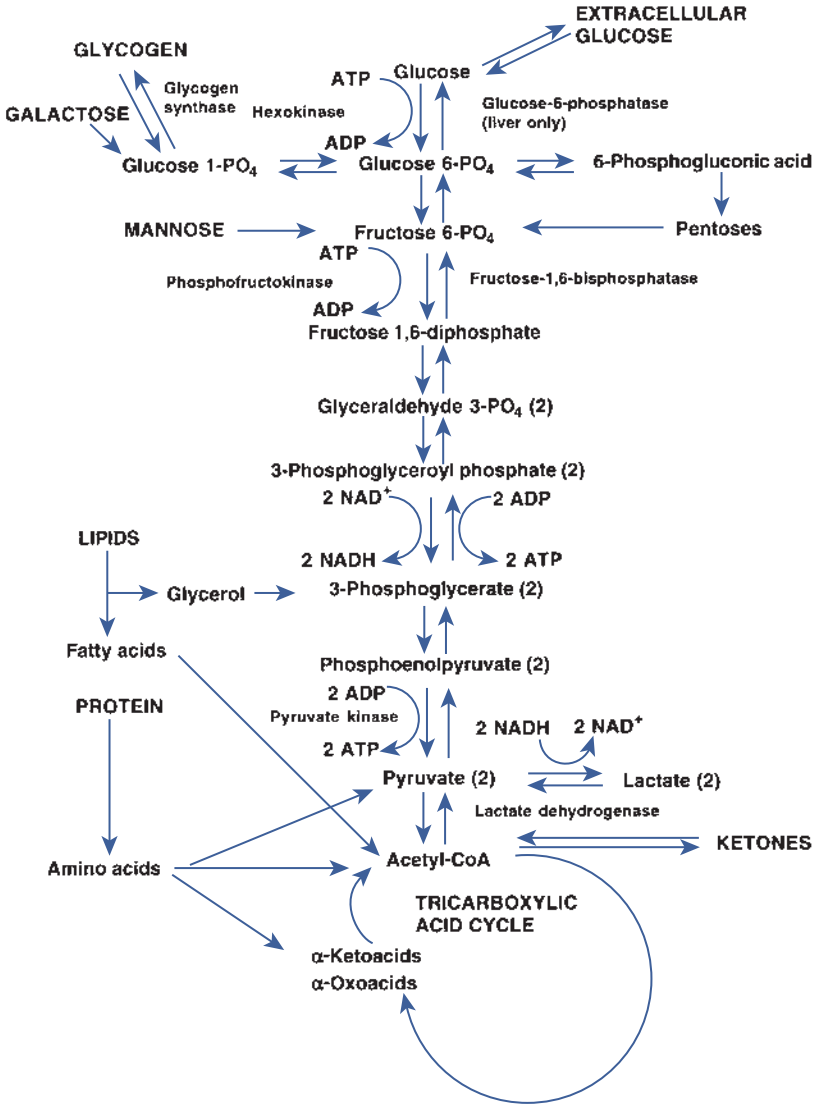


Figure 9.8 The Embden-Meyerhof pathway for anaerobic glycolysis.

pathway, which is called **glycogenesis**, is relatively straightforward. Glucose-6-phosphate is converted to glucose-1-phosphate, which is then converted to uridine diphosphoglucose, and then to glycogen, by glycogen synthase. Several tissues are capable of the synthesis of glycogen, especially the liver and muscles. Hepatocytes are capable of releasing glucose from glycogen or other sources to maintain the blood glucose concentration. This is because the liver synthesizes the enzyme glucose-6-phosphatase. Without this enzyme, glucose is trapped in the glycolytic pathway. Muscle cells do not synthesize glucose-6-phosphatase, and therefore, they are incapable of dephosphorylating glucose. Once glucose enters a muscle cell, it remains as glycogen unless it is catabolized. **Glycogenolysis** is the process by which glycogen is converted back to glucose-6-phosphate for entry into the glycolytic pathway. **Table 9.1** outlines the major energy pathways involved either directly or indirectly with glucose metabolism.

Overall, dietary glucose and other carbohydrates either can be used by the liver and other cells for energy or can be stored as glycogen for later use. When the supply of glucose is low, the liver will use glycogen and other substrates to elevate the blood glucose concentration. These substrates include glycerol from triglycerides, lactic acid from skin and muscles, and amino acids. If the lipolysis of triglycerides is unregulated, it results in the formation of ketone bodies, which the brain can use as a source of energy through the TCA cycle. The synthesis of glucose from amino acids is another form of gluconeogenesis. This process is used in conjunction with the formation of ketone bodies when glycogen

Table 9.1 Pathways in Glucose Metabolism

Glycolysis	Metabolism of glucose molecule to pyruvate or lactate for production of energy
Gluconeogenesis	Formation of glucose-6-phosphate from noncarbohydrate sources
Glycogenolysis	Breakdown of glycogen to glucose for use as energy
Glycogenesis	Conversion of glucose to glycogen for storage
Lipogenesis	Conversion of carbohydrates to fatty acids
Lipolysis	Decomposition of fat

Table 9.2 The Action of Hormones

Action of Insulin
Increases glycogenesis and glycolysis: glucose → glycogen → pyruvate → acetyl-CoA
Increases lipogenesis
Decreases glycogenolysis
Action of Glucagon
Increases glycogenolysis: glycogen → glucose
Increases gluconeogenesis: fatty acids → acetyl-CoA → ketone, proteins → amino acids

stores are depleted, which is associated with starvation. The principal pathway for glucose oxidation is through the Embden-Meyerhof pathway. NADPH can be synthesized through the HMP shunt, which is a side pathway from the anaerobic glycolytic pathway (Figure 9.8).

Regulation of Carbohydrate Metabolism

The liver, pancreas, and other endocrine glands are all involved in controlling the blood glucose concentrations within a narrow range. During a brief fast, glucose is supplied to the ECF from the liver through glycogenolysis. When the fasting period is longer than 1 day, glucose is synthesized from other sources through gluconeogenesis. Control of blood glucose is under two major hormones: insulin and glucagon, both produced by the pancreas. Their actions oppose each other. Other hormones and neuroendocrine substances also exert some control over blood glucose concentrations, permitting the body to respond to increased demands for glucose or to survive prolonged fasts. It also permits the conservation of energy as lipids when excess substrates are ingested.

The β-cells of the pancreatic islets synthesize pro-insulin, which is then converted to insulin and C-peptide after proteolytic cleavage. **Insulin** is the primary hormone responsible for the entry of glucose into the cell and therefore reducing blood glucose levels. It is synthesized by the β-cells of islets of Langerhans in the pancreas. When these cells detect an increase in blood glucose, they release insulin. The release of insulin causes an increased movement of glucose into the cells and increased glucose metabolism. Insulin is normally released when glucose levels are high and is *not* released when glucose levels are decreased. It decreases plasma glucose levels by increasing the transport entry of glucose in muscle and adipose tissue by way of nonspecific receptors. It also regulates glucose by increasing glycogenesis, lipogenesis, and glycolysis and inhibiting glycogenolysis. Insulin is the only hormone that decreases glucose levels and can be referred to as a **hypoglycemic** agent (**Table 9.2**).

Glucagon is the primary hormone responsible for increasing blood glucose levels. It is synthesized by the α-cells of islets of Langerhans in the pancreas and released during stress and fasting states. When these cells detect a decrease in blood glucose, they release glucagon. Glucagon acts by increasing blood glucose levels by glycogenolysis in the liver and an increase in gluconeogenesis. It can be referred to as a **hyperglycemic** agent (Table 9.2).

Two hormones produced by the adrenal gland affect carbohydrate metabolism. *Epinephrine*, produced by the adrenal medulla, increases plasma glucose by inhibiting insulin secretion, increasing glycogenolysis, and promoting lipolysis. Epinephrine is released during times of stress. *Glucocorticoids*, primarily cortisol, are released from the adrenal cortex on stimulation by adrenocorticotropic hormone (ACTH). Cortisol increases plasma glucose by decreasing intestinal entry into the cell and increasing gluconeogenesis and lipolysis, while decreasing glycogenesis.

Two anterior pituitary hormones, growth hormone and ACTH, also promote increased plasma glucose. *Growth hormone* increases plasma glucose by decreasing the entry of glucose into the cells and increasing glycogenolysis. Its release from the

pituitary is stimulated by decreased glucose levels and inhibited by increased glucose. Decreased levels of cortisol stimulate the anterior pituitary to release ACTH. ACTH, in turn, stimulates the adrenal cortex to release cortisol and increases plasma glucose levels by converting liver glycogen to glucose and promoting gluconeogenesis.

Two other hormones affect glucose levels: thyroxine and somatostatin. The thyroid gland is stimulated by the production of thyroid-stimulating hormone to release *thyroxine* that increases plasma glucose levels by increasing glycogenolysis, gluconeogenesis, and intestinal absorption of glucose. *Somatostatin*, produced by the δ-cells of the islets of Langerhans of the pancreas, increases plasma glucose levels by the inhibition of insulin, glucagon, growth hormone, and other endocrine hormones.

Hyperglycemia

In healthy patients, during a hyperglycemia state, insulin is secreted by the β-cells of the pancreatic islets of Langerhans. Insulin enhances membrane permeability to cells in the liver, muscle, and adipose tissue. It also alters the glucose metabolic pathways. A fasting plasma glucose of 100 mg/dL has been designated as the upper limit of normal blood glucose by the ADA; in a non-fasting state, levels of ≤140 mg/dL are considered normal.[1] *Hyperglycemia*, or increased plasma glucose levels, is caused by an imbalance of hormones.

Diabetes Mellitus

Diabetes mellitus is actually a group of metabolic diseases characterized by hyperglycemia resulting from defects in insulin secretion, insulin action, or both. In 1979, the National Diabetes Data Group developed a classification and diagnosis scheme for diabetes mellitus.[2] This scheme included dividing diabetes into two broad categories: type 1, previously called insulin-dependent diabetes mellitus (IDDM), and type 2, previously called non-insulin-dependent diabetes mellitus (NIDDM).

Established in 1995, the International Expert Committee on the Diagnosis and Classification of Diabetes Mellitus, working under the sponsorship of the ADA, was given the task of updating the 1979 classification system. The proposed changes included eliminating the older terms of IDDM and NIDDM. The categories of type 1 and type 2 were retained, with the adoption of Arabic numerals instead of Roman numerals (**Table 9.3**).[3]

Therefore, the ADA/World Health Organization (WHO) guidelines recommend the following categories of diabetes[1]:

- Type 1 diabetes
- Type 2 diabetes
- Other specific types of diabetes
- Gestational diabetes mellitus (GDM) (diagnosed in the second or third trimester of pregnancy that is not clearly overt diabetes)

Type 1 diabetes mellitus is characterized by hyperglycemia, primarily as a result of pancreatic islet

Table 9.3 Classification of Diabetes Mellitus

Classification	Pathogenesis
Type 1	β-Cell destruction leading to absolute insulin deficiency and development of autoantibodies (includes latent autoimmune diabetes of adulthood) ■ Islet cell autoantibodies ■ Insulin autoantibodies ■ Glutamic acid decarboxylase autoantibodies ■ Tyrosine phosphatase IA-2 and IA-2B autoantibodies
Type 2	■ Insulin resistance with an insulin secretory defect (due to progressive loss of adequate β-cell insulin secretion) ■ Relative insulin deficiency due to insulin resistance
Other	■ Monogenic diabetes syndromes (neonatal diabetes mellitus and maturity-onset diabetes of the young) ■ Diseases of the exocrine pancreas ■ Drug or chemical induced
Gestational	Glucose intolerance during pregnancy diagnosed in the second or third trimester of pregnancy

CASE STUDY 9.1, PART 2

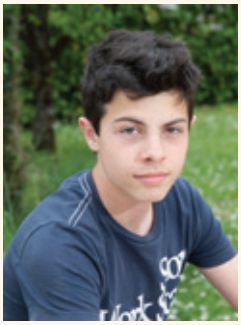
Dexter's physician ordered a fasting plasma glucose and HbA1c.
The results were as follows: FPG: 113 mg/dL (6.3 mmol/L) and a HbA1c: 6.7%.

© sirtravelalot/Shutterstock.

2. Based on these results, what would you tell the physician to do next? Do these results indicate that Dexter has type 2 diabetes?
3. If the tests are inconclusive, which test would you repeat to confirm or refute the diagnosis of type 2 diabetes?

CASE STUDY 9.2, PART 2

Recall John the high school athlete. His physician ordered urinalysis and chemistry profile. The results are shown in the table below. Discuss John's symptoms and laboratory results with his physician.



© sylv1rob1/Shutterstock.

Urinalysis Results		Chemistry Test Results	
Specific gravity	1.012	Sodium	126 mmol/L
pH	5.0	Potassium	6.1 mmol/L
Glucose	4+	Chloride	87 mmol/L
Ketones	Large	Bicarbonate	6 mmol/L
		Plasma glucose	600 mg/dL
		BUN	48 mg/dL
		Creatinine	2.0 mg/dL

1. What is the probable diagnosis of this patient based on the data presented?
2. What does the serum bicarbonate level tell the clinician? What do the sodium and chloride levels tell you about the cause of the patient's acidosis?
3. Why are the urine ketones positive?
4. What methods are used to quantitate urine ketones? What ketone(s) do they detect?

β-cell destruction, and a tendency to develop ketoacidosis. *Type 2 diabetes*, in contrast, includes hyperglycemia cases that result from insulin resistance combined with an insulin secretory defect. An intermediate stage, in which the fasting glucose is increased above normal limits but not to the level of diabetes, has been named *impaired fasting glucose* (IFG) or *prediabetes*. Use of the term *impaired glucose tolerance* to indicate glucose tolerance values above normal but below diabetes levels was retained. Also, the term GDM was retained for women who develop glucose intolerance during pregnancy.

Type 1 Diabetes, LADA, and Idiopathic Type 1

Type 1 diabetes constitutes 5% to 10% of all cases of diabetes and most commonly occurs in childhood and adolescence. It is a result of rapid, T-cell-mediated autoimmune destruction of the β-cells of the pancreas, causing an absolute deficiency of insulin secretion. This disease is usually initiated by an environmental factor or infection (usually a virus) in individuals with a genetic predisposition and causes the immune destruction of the β-cells of the pancreas and, therefore, a decreased production of insulin. Characteristics of type 1 diabetes include abrupt onset, insulin dependence, and a tendency for ketoacidosis. In recent years, a relationship has been identified between type 1 diabetes and mutations in a number of specific genes.[4] In addition, one or more of the following markers are found in 85% to 90% of individuals with fasting hyperglycemia: islet cell autoantibodies, insulin autoantibodies, glutamic acid decarboxylase (GAD) autoantibodies, tyrosine phosphatase IA-2 and IA-2β autoantibodies, and/or zinc transporter 8 antibody (ZnT8).[1] A similar condition called *latent autoimmune diabetes of adulthood (LADA)*, sometimes referred to as "Type 1a," is a slower autoimmune β-cell destruction that can occur in adults,

leading to a long duration of marginal insulin secretory capacity.

Signs and symptoms of type 1 include polydipsia (excessive thirst), polyphagia (increased food intake), polyuria (excessive urine production), rapid weight loss, hyperventilation, mental confusion, and possible loss of consciousness (due to increased glucose in the brain). Complications include microvascular problems such as nephropathy, neuropathy, and retinopathy. Increased incidence of heart disease is also found in patients with diabetes. **Table 9.4** lists the laboratory findings in hyperglycemia.

Idiopathic type 1 diabetes is a form of type 1 diabetes that has no known etiology, is strongly inherited, and does not have β-cell autoimmunity. Individuals with this form of diabetes have episodic requirements for insulin replacement.

Type 2 Diabetes

Type 2 diabetes mellitus is characterized by hyperglycemia that is caused by an individual's resistance to insulin combined with an insulin secretory defect. This is due to a progressive loss of adequate β-cell insulin secretion. This resistance results in a relative, not an absolute, insulin deficiency. Type 2 is the most common form of diabetes mellitus. Most patients affected by type 2 diabetes are obese or have an increased percentage of body fat distribution in the abdominal region. This type of diabetes often goes undiagnosed for many years and is associated with a strong genetic predisposition, with patients at increased risk with an increase in age, obesity, and lack of physical exercise. Characteristics usually include adult onset of the disease and milder symptoms than in type 1, with ketoacidosis seldom occurring. However, patients with type 2 diabetes are more likely to go into a hyperosmolar coma and are at an increased risk of developing macrovascular and microvascular complications.

Table 9.4 Laboratory Findings in Hyperglycemia

Decreased or absent insulin
Increased glucose in plasma and urine
Increased urine-specific gravity
Increased serum and urine osmolality
Ketones in serum and urine (ketonemia and ketonuria, respectively)
Decreased blood and urine pH (acidosis)
Electrolyte imbalance

Other Specific Types of Diabetes

Other specific types of diabetes are associated with (secondary to) certain conditions, including genetic defects of β-cell function or insulin action, pancreatic disease or injury, diseases of endocrine origin, drug or chemical-induced insulin receptor abnormalities, and certain genetic syndromes. The characteristics and prognosis of these forms of diabetes depend on the primary disorder. Monogenic diabetes are forms caused by a single gene mutation; they include maturity-onset diabetes of youth (MODY) and neonatal diabetes. Both are rare forms of diabetes that are inherited in an autosomal dominant fashion.[1]

Gestational Diabetes Mellitus (GDM)

GDM has been defined as any degree of glucose intolerance with onset or first recognition during the second or third trimester of pregnancy. According to ADA guidelines, women at high risk and who have diabetes at their initial prenatal visit, using standard criteria (**Table 9.5**), should be diagnosed with diabetes mellitus, not gestational diabetes.[1] Women identified through the oral glucose tolerance, listed in Table 9.8, should receive a diagnosis of GDM. Causes of GDM include metabolic and hormonal changes that frequently return to normal postpartum. However, this disease is associated with increased perinatal complications and an increased risk for the development of diabetes in later years. Infants born to mothers with diabetes are at increased risk for respiratory distress syndrome, hypocalcemia, and hyperbilirubinemia.

Table 9.5 Diagnostic Criteria for Diabetes Mellitus

1. HbA1c ≥6.5% using a method that is NGSP certified and standardized to the DCCT assay[a]
2. Fasting plasma glucose ≥126 mg/dL (≥7.0 mmol/L)[a]
3. 2-h plasma glucose ≥200 mg/dL (≥11.1 mmol/L) during an OGTT[a]
4. Random plasma glucose ≥200 mg/dL (≥11.1 mmol/L) plus symptoms of diabetes[a]

DCCT, Diabetes Control and Complications Trial; HbA1c, hemoglobin A1c; NGSP, National Glycohemoglobin Standardization Program; OGTT, oral glucose tolerance test.

[a]In the absence of unequivocal hyperglycemia, diagnosis requires two abnormal test results from the same sample or in two separate test samples.

Fetal insulin secretion is stimulated in the neonate of a mother with diabetes, which can lead to severe hypoglycemia when the infant is born and the umbilical cord is severed, abruptly terminating the infant's oversupply of glucose.

Pathophysiology of Diabetes Mellitus

In both type 1 and type 2 diabetes, the individual will be hyperglycemic, which can be severe. Glucosuria can also occur after the renal tubular transporter system for glucose becomes saturated. This happens when the glucose concentration of plasma exceeds roughly 180 mg/dL in an individual with normal renal function and urine output. As hepatic glucose overproduction continues, the plasma glucose concentration reaches a plateau, around 300 to 500 mg/dL (17 to 28 mmol/L). Provided renal output is maintained, glucose excretion will match the overproduction, causing the plateau.

The individual with type 1 diabetes has a higher tendency to produce ketones. Patients with type 2 diabetes seldom generate ketones but instead have a greater tendency to develop hyperosmolar nonketotic states. The difference in glucagon and insulin concentrations in these two groups appears to be responsible for the generation of ketones through increased β-oxidation. In type 1, there is an absence of insulin with an excess of glucagon. This permits gluconeogenesis and lipolysis to occur. In type 2, insulin is not absent and may, in fact, present as hyperinsulinemia at times; therefore, glucagon is reduced. Fatty acid oxidation is inhibited in type 2. This causes fatty acids to be incorporated into triglycerides for release as very low density lipoproteins.

The laboratory findings of a patient with diabetes with ketoacidosis tend to reflect dehydration, electrolyte disturbances, and metabolic acidosis. Acetoacetate, β-hydroxybutyrate, and acetone (ketone bodies) are produced from the oxidation of fatty acids. The two former ketone bodies contribute to the acidosis. Lactate, fatty acids, and other organic acids can also contribute to a lesser degree. Bicarbonate and total carbon dioxide are usually decreased due to Kussmaul-Kien respiration (deep respirations). This is a compensatory mechanism to blow off carbon dioxide and remove hydrogen ions in the process. The anion gap in this acidosis can exceed 16 mmol/L. Serum osmolality is high as a result of hyperglycemia and water loss; sodium concentrations tend to be lower due in part to losses (polyuria) and in part to a shift of water from cells because of the hyperglycemia. The sodium value should not be falsely underestimated because of hypertriglyceridemia. Grossly elevated triglycerides will displace plasma volume and give the appearance of decreased electrolytes when flame photometry or prediluted, ion-specific electrodes are used for sodium determinations. Hyperkalemia is almost always present as a result of the displacement of potassium from cells in acidosis. This is somewhat misleading because the patient's total body potassium is usually decreased.

More typical of the untreated patient with type 2 diabetes is the nonketotic hyperosmolar state. The individual presenting with this syndrome has an overproduction of glucose; however, there appears to be an imbalance between production and elimination in urine. Often, this state is precipitated by heart disease, stroke, or pancreatitis. Glucose concentrations exceed 300 to 500 mg/dL (17 to 28 mmol/L), and severe dehydration is present. The severe dehydration contributes to the inability to excrete glucose in the urine. Mortality is high with this condition. Ketones are not observed because the severe hyperosmolar state inhibits the ability of glucagon to stimulate lipolysis. The laboratory findings of nonketotic hyperosmolar coma include plasma glucose values exceeding 1000 mg/dL (55 mmol/L), normal or elevated plasma sodium and potassium, slightly decreased bicarbonate, elevated blood urea nitrogen (BUN) and creatinine, and an elevated osmolality (>320 mOsm/dL). The gross elevation in glucose and osmolality, the elevation in BUN, and the absence of ketones distinguish this condition from diabetic ketoacidosis.

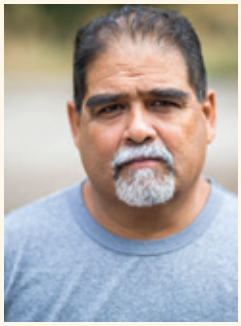
Other forms of impaired glucose metabolism that do not meet the criteria for diabetes mellitus include impaired fasting glucose and impaired glucose tolerance. These forms are discussed in the following section.

Criteria for Testing for Prediabetes and Diabetes

The testing criteria for asymptomatic adults for type 2 diabetes mellitus were modified by the ADA Expert Committee to allow for earlier detection of the disease. According to the ADA recommendations, all adults beginning at the age of 45 years should be tested for diabetes every 3 years using the **hemoglobin A1c (HbA1c)**, fasting plasma glucose, or a 2-hour 75 g oral glucose tolerance test (OGTT) unless the individual has otherwise been diagnosed with diabetes.[1] Testing should be carried out at an earlier age or more frequently in individuals who display overweight tendencies, that is,

CASE STUDY 9.3, PART 2

Recall Michael? The laboratory work was performed after his first visit, and the following results were obtained:



© adamkaz/E+/Getty Images.

Non-Fasting Plasma Glucose		225 mg/dL	
Urinalysis Results			
Color and appearance	Pale/clear	Blood	Negative
pH	6.0	Bilirubin	Negative
Specific	1.025	Urobilinogen	Negative
Glucose	2+	Nitrites	Negative
Ketones	Negative	Leukocyte esterase	Negative

1. What is the probable diagnosis of this patient?
2. What other test(s) should be performed to confirm this? Which is the preferred test?
3. What laboratory values from Question 2 would confirm the diagnosis of diabetes?
4. After diagnosis, what test(s) should be performed to monitor his condition?

BMI ≥25 kg/m^2 (at-risk BMI may be lower in some ethnic groups, i.e., Asian Americans ≥23 kg/m^2), and have additional risk factors, as follows:

- Habitually physically inactive
- Family history of diabetes in a first-degree relative
- In a high-risk minority population (e.g., African American, Latino, Native American, Asian American, and Pacific Islander)
- History of GDM or delivering a baby weighing more than 9 lb. (4.1 kg)
- Hypertension (blood pressure ≥140/90 mm Hg)
- Low high-density lipoprotein (HDL) cholesterol concentrations (<35 mg/dL [0.90 mmol/L])
- Elevated triglyceride concentrations > 250 mg/dL (2.82 mmol/L)
- HbA1c ≥5.7% (33 mmol/mol), IGT, or IFG on previous testing
- History of impaired fasting glucose/impaired glucose tolerance
- Women with polycystic ovarian syndrome (PCOS)
- Other clinical conditions associated with insulin resistance (e.g., severe obesity and acanthosis nigricans)
- History of cardiovascular disease

In the absence of the above criteria, testing for prediabetes and diabetes should begin at the age of 45 years. If results are normal, testing should be repeated at least at 3-year intervals, with consideration of more frequent testing depending on initial results and risk status.

As the incidence of adolescent type 2 diabetes has risen dramatically in the past few years, criteria for the testing for type 2 diabetes in asymptomatic children have been developed. These criteria include initiation of testing at the age 10 years or at the onset of puberty, if puberty occurs at a younger age, with follow-up testing every 2 years. Testing should be carried out on children who display the following characteristics: overweight (BMI >85th percentile for age and sex, weight for height >85th percentile, or weight >120% of ideal for height) plus any two of the following risk factors:

- Family history of type 2 diabetes in first- or second-degree relative
- High-risk minority population (e.g., Native American, African American, Latino, Asian American, and Pacific Islander)
- Signs of insulin resistance or conditions associated with insulin resistance (e.g., acanthosis nigricans, hypertension, dyslipidemia, and PCOS)
- Maternal history of diabetes or GDM

Criteria for the Diagnosis of Diabetes Mellitus

Four methods of diagnosis are suggested: (1) HbA1c ≥6.5% using a National Glycohemoglobin Standardization Program (NGSP)-certified method, (2) a fasting plasma glucose ≥126 mg/dL (7.0 mmol/L), or (3) an OGTT with a 2-hour post load (75 g glucose load) level ≥200 mg/dL (11.1 mmol/L), and

CASE STUDY 9.4, PART 2

Recall James, a 14-year-old male with complaints of fatigue, weight loss, and increases in appetite, thirst, and frequency of urination. For the past 3 to 4 weeks, he had been excessively thirsty and had to urinate every few hours. He began to get up three to four times a night to urinate. His grandparents had a history of diabetes mellitus. James's laboratory results are as follows:

© Tony Anderson/Stone/Getty Images.

Laboratory Data		
Fasting plasma glucose	160 mg/dL	
Urinalysis	Specific gravity	1.040
	Glucose	4+
	Ketones	Moderate

1. According to the ADA, what criteria are required for the diagnosis of diabetes?
2. Based on the preceding information, can this patient be diagnosed with diabetes?
3. What further tests might be performed to confirm the diagnosis?
4. Assuming this patient has diabetes, which type would be diagnosed?

(4) symptoms of diabetes plus a random plasma glucose level ≥200 mg/dL (11.1 mmol/L), each of which should be confirmed on a subsequent day by any one of the first three methods (**Tables 9.5–9.7**). Any of the first three methods are considered appropriate for the diagnosis of diabetes. The decision on which method to use is the decision of the healthcare provider depending on various patient factors. Point-of-care (POC) assay methods for either plasma glucose or HbA1c are not recommended for diagnosis.

An intermediate group of individuals who do not meet the criteria of diabetes mellitus but who have glucose levels above normal can be placed into three categories based on their risk of developing diabetes. First, those individuals with fasting glucose levels ≥100 mg/dL but less than 126 mg/dL are placed in the impaired fasting glucose category. Individuals who have 2-hour OGTT levels ≥140 mg/dL but less than 200 mg/dL are placed in the impaired glucose tolerance category. Additionally, individuals with a HbA1c of 5.7% to 6.4% are placed in an at-risk category. Individuals in these three categories are referred to as having "prediabetes" indicating the relatively high risk for the development of type 2 diabetes in these patients.

Criteria for the Testing and Diagnosis of GDM

The diagnostic criteria for gestational diabetes were revised by the International Association of the Diabetes and Pregnancy Study Groups. The revised criteria recommend that all nondiabetic pregnant women should be screened for GDM at 24 to 28 weeks of gestation.

The approach for screening and diagnosis is either a one-step or two-step approach. The one-step approach is the performance of a 2-hour OGTT using

Table 9.6 Categories of Fasting Plasma Glucose

Normal fasting glucose	FPG 70–99 mg/dL (3.9–5.5 mmol/L)
Impaired fasting glucose	FPG 100–125 mg/dL (5.6–6.9 mmol/L)
Provisional diabetes diagnosis	FPG ≥126 mg/dL (≥7.0 mmol/l)[a]

FPG, fasting plasma glucose.

[a]Must be confirmed.

Table 9.7 Categories of Oral Glucose Tolerance

Normal glucose tolerance	2-h PG ≤140 mg/dL (≤7.8 mmol/L)
Impaired glucose tolerance	2-h PG140–199 mg/dL (7.8–11.1 mmol/l)
Provisional diabetes diagnosis	2-h PG ≥200 mg/dL (≥11.1 mmol/L)[a]

PG, plasma glucose.

[a]Must be confirmed.

CASE STUDY 9.1, PART 3

Recall Dexter. His physician ordered a HbA1c to confirm the first result 10 days later.

The repeat HbA1c was performed, and the result was 6.6%. Diagnosis of type 2 diabetes is confirmed.

4. Now that Dexter has been diagnosed with diabetes, for what risks should he be monitored?

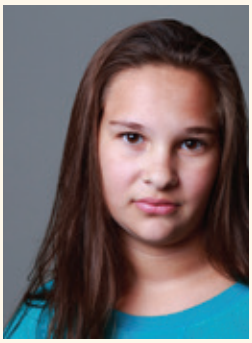
© sirtravelalot/Shutterstock.

a 75-g glucose load at 24 to 28 weeks' gestation. Glucose specimens are collected at fasting, 1 hour, and 2 hours. A fasting plasma glucose value ≥ 92 mg/dL, or a 1-hour value ≥ 180 mg/dL, or a 2-hour glucose value ≥ 153 mg/dL is diagnostic of GDM. If any one of the three criteria is met, then a diagnosis of GDM is rendered. The OGTT test should be performed in the morning after an overnight fast of at least 8 hours (**Table 9.8**).

In the two-step approach, an initial measurement of plasma glucose at 1-hour post load (50-g glucose load) is performed. A plasma glucose value ≥140 mg/dL (≥7.8 mmol/L) indicates the need to perform a 3-hour OGTT using a 100 g glucose load. GDM is diagnosed

CASE STUDY 9.5, PART 2

Recall Myra. Her mother was contacted by the EMTs, and she mentioned that her daughter had been losing weight and making frequent trips to the bathroom at night. The emergency squad noticed a fruity odor on Myra's breath. The decision was made to transport Myra to the emergency department. On arrival, her vital signs were as follows:



© Kyle Lee/Shutterstock.

Blood pressure	**98/50 mm Hg**
Respirations	Rapid
Temperature	99 °F

Stat laboratory testing was ordered and the results are reported:

Random Urine		**Serum Chemistries**	
pH	5.5	Glucose	500 mg/dL
Protein	Negative	Ketones	Positive
Glucose	4+	BUN	6 mg/dL
Ketones	Moderate	Creatinine	0.4 mg/dL
Blood	Negative		

Assist the physician's in explaining Myra's diagnosis to her mother.

1. Based on the preceding information, can Myra be diagnosed with diabetes?
2. Which type of diabetes is most likely to be diagnosed?
3. What is the cause of the fruity breath?
4. Why is Myra's respiration elevated?

Table 9.8 Diagnostic Criteria for Gestational Diabetes

One-Step Approach	
Fasting plasma glucose	≥92 mg/dL (5.1 mmol/L)
1-h plasma glucose	≥180 mg/dL (10 mmol/L)
2-h plasma glucose	≥153 mg/dL (8.5 mmol/L)
Two-Step Approach	
Fasting	≥95 mg/dL (5.3 mmol/L)
1-h plasma glucose	≥180 mg/dL (10.0 mmol/L)
2-h plasma glucose	≥155 mg/dL (8.6 mmol/L)
3-h plasma glucose	≥140 mg/dL (7.8 mmol/L)

when any two of the following four values are met or exceeded: fasting, >95 mg/dL (5.3 mmol/L); 1 hour, ≥180 mg/dL (10.0 mmol/L); 2 hours, ≥155 mg/dL (8.6 mmol/L); or 3 hours, ≥140 mg/dL (7.8 mmol/L). This test should be performed in the morning after an overnight fast of between 8 and 14 h, after at least 3 days of unrestricted diet (≥150 g carbohydrate per day) and unlimited physical activity. Note: The National Diabetes Data Group's levels are slightly higher than those listed above.

Hypoglycemia

Hypoglycemia involves decreased plasma glucose levels and can have many causes—some are transient and relatively insignificant, but others can be life threatening. Hypoglycemia causes brain fuel deprivation, which can result in impaired judgment and behavior, seizures, comas, functional brain failure, and death. Hypoglycemia is the result of an imbalance in the rate of glucose appearance and disappearance from the circulation. This imbalance may be caused by treatment, such as oral antiglycemic drugs or injected insulin, or by biological factors. The plasma glucose concentration at which glucagon and other glycemic factors are released is between 65 and 70 mg/dL (3.6 to 3.9 mmol/L); at about 50 to 55 mg/dL (2.8 to 3.1 mmol/L), observable symptoms of hypoglycemia appear. The warning signs and symptoms of hypoglycemia are all related to the central nervous system. The release of epinephrine into the systemic circulation and of norepinephrine at nerve endings of specific neurons acts in unison with glucagon to increase plasma glucose. Glucagon, which inhibits insulin, is released from the islet cells of the pancreas. Epinephrine released from the adrenal gland both increases blood glucose metabolism and inhibits insulin. In addition, cortisol and growth hormone are released, both of which increase blood glucose levels.

Historically, hypoglycemia was classified as postabsorptive (fasting) and postprandial (reactive) hypoglycemia. *Postprandial hypoglycemia* describes hypoglycemia that occurs within 4 hours after a meal. Current approaches suggest classifying postprandial hypoglycemia based on the severity of symptoms and measured plasma glucose levels. This approach is especially important for individuals with diabetes who are a high risk for hypoglycemic episodes (**Table 9.9**).[5] The ADA and the Endocrine Society recommend that a plasma concentration of less than or equal to 70 mg/dL (3.9 mmol/L) be used as cutoff as well as an alert value to prevent a clinical hypoglycemic episode. This value also allows for a margin of error for self-monitoring glucose devices. Patients with diabetes who present with iatrogenic

Table 9.9 Classification of Hypoglycemia

Level	Glycemic Criteria	Description
Level 1 Glucose alert value	<70 mg/dL	Sufficiently low for treatment with fast-acting carbohydrate and dose adjustment of glucose-lowering therapy
Level 2 Clinically significant hypoglycemia	<54 mg/dL (3.0 mmol/L)	Sufficiently low to indicate serious, clinically important hypoglycemia
Level 3 Severe hypoglycemia	No specific glucose threshold	Hypoglycemia associated with severe cognitive impairment requiring external assistance for recovery

Data from International Hypoglycaemia Study Group. Glucose concentrations of less than 3.0 mmol/L (54 mg/dL) should be reported in clinical trials: a joint position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2017; 40:155–157.

hypoglycemia may potentially be harmed by the low glucose level.

Hypoglycemia is rare in individuals with normal glucose metabolism. In individuals without diabetes, a diagnosis of hypoglycemia should be made only in those who demonstrate the Whipple triad: (1) hypoglycemic symptoms; (2) plasma glucose concentration is low (<50 mg/dL) when the symptoms are present; and (3) symptoms are relieved by correction of the hypoglycemia when administered glucose or glucagon.[6] Symptoms of hypoglycemia include increased hunger, sweating, nausea and vomiting, dizziness, nervousness and shaking, blurring of speech and sight, and mental confusion. When hypoglycemia symptoms present in individuals in a postabsorptive (fasting) state, an insulinoma (pancreatic β-cell tumor) might be suspected. Laboratory findings include decreased plasma glucose levels during a hypoglycemic episode and extremely elevated insulin levels in patients with an insulinoma. To investigate an insulinoma, the patient is required to fast under controlled conditions. Men and women have different metabolic patterns in prolonged fasts. The healthy male will maintain plasma glucose of 55 to 60 mg/dL for several days. Healthy females will produce ketones more readily and permit plasma glucose to decrease to 40 mg/dL or lower. Diagnostic criteria for an insulinoma include a change in glucose level ≥25 mg/dL coincident with an insulin level ≥6 μU/mL, C-peptide levels ≥0.2 nmol/L, proinsulin levels ≥5 pmol/L, and/or β-hydroxybutyrate levels less than or equal to 2.7 mmol/L.[7]

Genetic Defects in Carbohydrate Metabolism

Glycogen storage diseases are the result of the deficiency of a specific enzyme that causes an alternation of glycogen metabolism. The most common congenital form of glycogen storage disease is glucose-6-phosphatase deficiency type 1, also called von Gierke disease, which is an autosomal recessive disease. This disease is characterized by severe hypoglycemia that coincides with metabolic acidosis, ketonemia, and elevated lactate and alanine. Hypoglycemia occurs because glycogen cannot be converted back to glucose by way of hepatic glycogenolysis. A glycogen buildup is found in the liver, causing hepatomegaly. The patients usually have severe hypoglycemia, hyperlipidemia, uricemia, and growth inhibition. A liver biopsy will show a positive glycogen stain. Although the glycogen accumulation is irreversible, the disease can be kept under control by avoiding the development of hypoglycemia. Liver transplantation corrects the hypoglycemic condition. Other enzyme defects or deficiencies that cause hypoglycemia include glycogen synthase, fructose-1,

CASE STUDY 9.6, PART 2

Recall Jo Ann, a 28-year-old woman who delivered her daughter, Martha , early this morning. The mother's history was incomplete; she claimed to have had no medical care through her pregnancy. Several hours after birth, Martha became lethargic and flaccid. A whole blood glucose and ionized calcium were ordered and performed in the nursery with the following results:

Whole blood glucose	25 mg/dL
Ionized calcium	4.9 mg/dL
Plasma glucose was drawn and analyzed in the main laboratory to confirm the whole blood findings.	
Plasma glucose	33 mg/dL
An intravenous glucose solution was started, and whole blood glucose was measured hourly.	

© yasindmrblk/iStock/Getty Images Plus/ Getty Images.

1. Give the possible explanation for the Martha's large birth weight and size.
2. If Jo Ann has gestational diabetes, why is her baby hypoglycemic?
3. Explain the discrepancy between the whole blood glucose concentration and the plasma glucose concentration in the neonate.
4. If Jo Ann had received appropriate prenatal care, what laboratory tests should have been performed, and what criteria would have indicated that she had gestational diabetes?

6-bisphosphatase, phosphoenolpyruvate carboxykinase, and pyruvate carboxylase. Glycogen debrancher enzyme deficiency does not cause hypoglycemia but does cause hepatomegaly.

Galactosemia, a cause of failure to thrive syndrome in infants, is a congenital deficiency of one of three enzymes involved in galactose metabolism, resulting in increased levels of galactose in plasma. The most common enzyme deficiency is galactose-1-phosphate uridyltransferase. Galactosemia occurs because of the inhibition of glycogenolysis and is accompanied by diarrhea and vomiting. Galactose must be removed from the diet to prevent the development of irreversible complications. If left untreated, the patient will develop intellectual deficiencies and cataracts. The disorder can be identified by measuring erythrocyte galactose-1-phosphate uridyltransferase activity. Laboratory findings include hypoglycemia, hyperbilirubinemia, and galactose accumulation in the blood, tissue, and urine following milk ingestion. Another enzyme deficiency, fructose-1-phosphate aldolase deficiency, causes nausea and hypoglycemia after fructose ingestion.

Specific inborn errors of amino-acid metabolism and long-chain fatty acid oxidation are also responsible for hypoglycemia. There are also alimentary and idiopathic hypoglycemias. Alimentary hypoglycemia appears to be caused by an increase in the release of insulin in response to rapid absorption of nutrients after a meal or the rapid secretion of insulin-releasing gastric factors.

Role of the Laboratory in Differential Diagnosis and Management of Patients with Glucose Metabolic Alterations

The demonstration of hyperglycemia or hypoglycemia under specific conditions is used to diagnose diabetes mellitus and hypoglycemic conditions. Other laboratory tests have been developed to identify insulinomas and to monitor glycemic control and the development of renal complications.

Methods of Glucose Measurement

Glucose can be measured from serum, plasma, or whole blood. Today, most glucose measurements are performed on serum or plasma. The glucose concentration in whole blood is approximately 11% lower than the glucose concentration in plasma. Serum or plasma must be refrigerated and separated from the cells within 1 hour to prevent substantial loss of glucose by the cellular fraction, particularly if the white blood cell count is elevated. Sodium fluoride ions (gray-top tubes) are often used as an anticoagulant and preservative of whole blood, particularly if analysis is delayed. The fluoride inhibits glycolytic enzymes. However, although fluoride maintains

CASE STUDY 9.7, PART 2

Recall Emily J., who reported no particular health issues during her annual physical examination. Her physician received her pre-physical laboratory test. Emily indicated she had no family history of diabetes, nor any history of elevated glucose levels during her two pregnancies. The results were as follows:

Laboratory Results	
Fasting blood glucose (FBG)	90 mg/dL
Cholesterol	140 mg/dL
HDL	40 mg/dL
Triglycerides	90 mg/dL

© goodluz/Shutterstock.

1. What is the probable diagnosis of this patient?
2. Describe the proper follow-up for this patient.
3. What are the appropriate screening tests for diabetes in nonpregnant adults?
4. What risk factors would indicate a potential for this patient to develop diabetes?

long-term glucose stability, the rates of decline of glucose in the first hour after sample collection in tubes with and without fluoride are virtually identical. Therefore, the plasma should be separated from the cells as soon as possible. FBG should be obtained in the morning after an approximately 8- to 10-hour fast (not longer than 16 hours). Fasting plasma glucose values have a diurnal variation with the mean FBG higher in the morning than in the afternoon.[8] Diabetes in patients tested in the afternoon may be missed because of this variation. Cerebrospinal fluid and urine can also be analyzed. Urine glucose measurement is not used in diabetes diagnosis; however, some patients use this measurement for monitoring purposes.

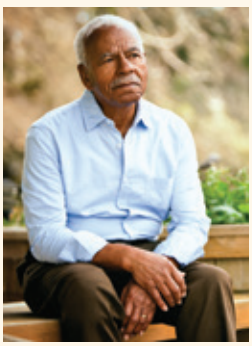
The ability of glucose to function as a reducing agent has been useful in the detection and quantitation of carbohydrates in body fluids. Glucose and other carbohydrates are capable of converting cupric ions in alkaline solution to cuprous ions. The solution loses its deep-blue color and a red precipitate of cuprous oxide forms. Benedict's and Fehling's reagents, which contain an alkaline solution of cupric ions stabilized by citrate or tartrate, respectively, have been used to detect reducing agents in urine and other body fluids. Another chemical characteristic that used to be exploited to quantitate carbohydrates is the ability of these molecules to form Schiff bases with aromatic amines. *O*-toluidine in a hot acidic solution will yield a colored compound with an absorbance maximum at 630 nm. Galactose, an aldohexose, and mannose, an aldopentose, will also react with *O*-toluidine and produce a colored compound that can interfere with the reaction. The Schiff base reaction with *O*-toluidine is of historical interest only and has been replaced by more specific enzymatic methods, which are discussed in the following section.

The most common methods of glucose analysis use the enzyme glucose oxidase or hexokinase (**Table 9.10**). Glucose oxidase is the most specific enzyme reacting with only β-D-glucose. Glucose oxidase converts β-d-glucose to gluconic acid. Mutarotase may be added to the reaction to facilitate the conversion of α-d-glucose to β-d-glucose. Oxygen is consumed and hydrogen peroxide (H_2O_2) is produced. The reaction can be monitored polarographically either by measuring the rate of disappearance of oxygen using an oxygen electrode or by consuming H_2O_2 in a side reaction. Horseradish peroxidase is used to catalyze the second reaction, and the H_2O_2 is used to oxidize a dye compound. Two commonly used chromogens are 3-methyl-2-benzothiazolinone hydrazone and *N*,*N*-dimethylaniline. The shift in absorbance can be monitored spectrophotometrically and is proportional to the amount of glucose present in the specimen. This coupled reaction is known as the Trinder reaction. However, the peroxidase coupling reaction used in the glucose oxidase method is subject to positive and negative interference. Increased levels of uric acid, bilirubin, and ascorbic acid can cause falsely decreased values as a result of these substances being oxidized by peroxidase, which then prevents the oxidation and detection of the chromogen. Strong oxidizing substances, such as bleach, can cause falsely increased values. An oxygen consumption electrode can be used to perform the direct measurement of oxygen by the polarographic technique, which avoids this interference. Oxygen depletion is measured and is proportional to the amount of glucose present. Polarographic glucose

CASE STUDY 9.8, PART 2

Recall Alfred. He admits he does not always eat a healthy diet, especially when he travels with his family. For three consecutive quarters, a fasting glucose and glycosylated hemoglobin were performed to monitor his disease. The results are as follows:

	Quarter 1	Quarter 2	Quarter 3
Plasma glucose, fasting	280 mg/dL	85 mg/dL	91 mg/dL
Glycosylated hemoglobin	7.8%	15.3%	8.5%



© Erickson Stock/Shutterstock.

1. In which quarter was the patient's glucose best controlled? The least controlled?
2. Do the fasting plasma glucose and glycosylated hemoglobin match? Why or why not?
3. What methods are used to measure glycosylated hemoglobin?
4. What potential conditions might cause erroneous results?

Table 9.10 Methods of Glucose Measurement

Glucose oxidase	$Glucose + O_2 + H_2O \xrightarrow{glucose\ oxidase} gluconic\ acid + H_2O_2$
	$H_2O_2 + reduced\ chromogen \xrightarrow{peroxidase} oxidized\ chromogen + H_2O$
Hexokinase	$Glucose + ATP \xrightarrow{hexokinase} glucose\text{-}6\text{-}PO_4 + ADP$
	$Glucose\text{-}6\text{-}PO_4 + NADP^+ \xrightarrow{G\text{-}6\text{-}PD} NADPH + 6\text{-}phosphogluconate$
Clinitest	$Cu^{2+} \xrightarrow{Reducing\ substance} Cu^{1+}O$

analyzers measure the rate of oxygen consumption because glucose is oxidized under first-order conditions using glucose oxidase reagent. The H_2O_2 formed must be eliminated in a side reaction to prevent the reaction from reversing. Molybdate can be used to catalyze the oxidation of iodide to iodine by H_2O_2 or catalase can be used to catalyze oxidation of ethanol by H_2O_2, forming acetaldehyde and H_2O.

The hexokinase method is considered more accurate than the glucose oxidase methods because the coupling reaction using glucose-6-phosphate dehydrogenase is highly specific; therefore, it has less interference than the coupled glucose oxidase procedure. Hexokinase in the presence of ATP converts glucose to glucose-6-phosphate. Glucose-6-phosphate and the cofactor $NADP^+$ are converted to 6-phosphogluconate and NADPH by glucose-6-phosphate dehydrogenase. NADPH has a strong absorbance maximum at 340 nm, and the rate of appearance of NADPH can be monitored spectrophotometrically and is proportional to the amount of glucose present in the sample. Generally accepted as the reference method, this method is not affected by ascorbic acid or uric acid. Gross hemolysis and extremely elevated bilirubin may cause a false decrease in results. The hexokinase method may be performed on serum or plasma collected using heparin, ethylenediaminetetraacetic acid (EDTA), fluoride, oxalate, or citrate. The method can also be used for urine, cerebrospinal fluid, and serous fluids.

Nonspecific methods of measuring glucose are still used in the urinalysis section of the laboratory primarily to detect reducing substances other than glucose, specifically in the newborn population.

Self-Monitoring of Blood Glucose

The ADA has recommended that individuals with diabetes monitor their blood glucose levels in an effort to maintain levels as close to normal as possible. For persons with type 1 diabetes, the recommendation is three to four times per day; for persons with type 2 diabetes, the optimal frequency is unknown. Recommended glycemic target for many nonpregnant adult is a preprandial capillary plasma glucose of 80 to 130 mg/dL. This recommendation has been determined to correlate with the achievement of an HbA1c of less than 7%. Peak postprandial, 1 to 2 hours after the beginning of a meal, capillary plasma glucose of less than 180 mg/dL may be suggested for individuals who are not achieving the HbA1c target. It is important that patients be taught how to use control solutions and calibrators to ensure the accuracy of their results.[9] Self-monitoring of blood glucose is performed using fingerstick glucometers or continuous glucose monitors. Patients with type 1 diabetes and those with GDM may also use blood or urine ketone meters to monitor for ketoacidosis.

Glucose Tolerance and 2-Hour Postprandial Tests

Guidelines for the performance and interpretation of the 2-hour postprandial test were set by the Expert Committee. A variation of this test is to use a standardized load of glucose. A solution containing 75 g of glucose is administered, and a specimen for plasma glucose measurement is drawn 2 hours later. Under this criterion, the patient drinks a standardized (75 g) glucose load, and a glucose measurement is taken 2 hours later. If that level is ≥200 mg/dL and is confirmed on a subsequent day by either an increased random or fasting glucose level, the patient is diagnosed with diabetes (see earlier discussion).

The OGTT is not recommended for routine use under the ADA guidelines. This procedure is inconvenient to patients and is not being used by physicians for diagnosing diabetes. However, if OGTT is used, the WHO recommends the criteria listed in Table 9.7. It is important that proper patient

preparation be given before this test is performed. The patient should be ambulatory and on a normal-to-high carbohydrate intake for 3 days before the test. The patient should be fasting for at least 10 hours and not longer than 16 hours, and the test should be performed in the morning because of the hormonal diurnal effect on glucose. Just before tolerance and while the test is in progress, patients should refrain from exercise, eating, drinking (except that the patient may drink water), and smoking. Factors that affect the tolerance results include medications such as large doses of salicylates, diuretics, anticonvulsants, oral contraceptives, and corticosteroids. Also, gastrointestinal problems, including malabsorption problems, gastrointestinal surgery, and vomiting and endocrine dysfunctions, can affect the OGTT results. The guidelines recommend that only the fasting and the 2-hour sample be measured, except when the patient is pregnant. The adult dose of glucose solution (Glucola) is 75 g; children receive 1.75 g/kg of glucose to a maximum dose of 75 g.

Glycosylated Hemoglobin/HbA1c

The aim of diabetic management is to maintain the blood glucose concentration within or near the nondiabetic range with a minimal number of fluctuations. Serum or plasma glucose concentrations can be measured by laboratories in addition to patient self-monitoring of whole blood glucose concentrations. Long-term blood glucose regulation can be followed by measurement of glycosylated hemoglobin.

Glycosylated hemoglobin is the term used to describe the formation of a hemoglobin compound produced when glucose (a reducing sugar) reacts with the amino group of hemoglobin (a protein). The glucose molecule attaches nonenzymatically to the hemoglobin molecule to form a ketoamine. The rate of formation is directly proportional to the plasma glucose concentrations. Because the average red blood cell lives approximately 120 days, the glycosylated hemoglobin level at any one time reflects the average blood glucose level over the previous 2 to 3 months. Therefore, measuring the glycosylated hemoglobin provides the clinician with a time-averaged picture of the patient's blood glucose concentration over the past 3 months.

HbA1c, the most commonly measured glycosylated hemoglobin, is a glucose molecule attached to one or both N-terminal valines of the β-polypeptide chains of normal adult hemoglobin.[10] HbA1c is a more reliable method of monitoring long-term diabetes control than random plasma glucose. Normal values range from 4.0% to 6.0%. Studies have shown that there is a strong linear relationship between average blood glucose and HbA1c. Using a linear regression model, Nathan et al.[11] determined that an estimated average glucose (eAG) can be calculated from the HbA1c reported value using the equation eAG (mg/dL) = 28.7 × HbA1c – 46.7 (**Table 9.11**). However, this information needs to be used carefully, as another study has shown that the relationship between average plasma glucose and HbA1c can differ substantially depending on the glycemic control of the population studied.[12] It is also important to remember that two factors determine the glycosylated hemoglobin levels: the average glucose concentration and the red blood cell life span. If the red blood cell life span is decreased because of another disease state, such as hemoglobinopathies, the hemoglobin will have less time to become glycosylated, and the glycosylated hemoglobin level will be lower. Other conditions associated with increased red cell turnover include pregnancy (second and

CASE STUDY 9.9, PART 2

Recall Angela, a 25-year-old healthy female. Her friends insisted she go to urgent care. On arrival, the nurse performed a random glucose test performed via fingerstick, with a result of 60 mg/dL.

1. Which of Angela's symptoms are characteristic of hypoglycemia?
2. What test(s) should be performed next to determine this young woman's problem?
3. To which category of hypoglycemia would this individual belong?
4. What criteria would be used to diagnose a potential insulinoma?

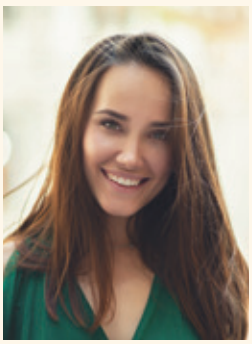


Table 9.11 Translation of HbA1c Assay into eAG Values and A1c Levels

Estimated Average Glucose (eAG)	HbA1c	
	NGSP (%)	IFCC (mmol/mol)
97 mg/dL (5.4 mmol/L)	5	31
126 mg/dL (7.0 mmol/L)	6	42
140 mg/dL (7.7 mmol/L)	6.5	48
154 mg/dL (8.6 mmol/L)	7	53
183 mg/dL (10.2 mmol/L)	8	64
212 mg/dL (11.8 mmol/L)	9	75
240 mg/dL (13.4 mmol/L)	10	86
269 mg/dL (14.9 mmol/L)	11	97
298 mg/dL (16.5 mmol/L)	12	108

Data from Nathan DM, Kuenen J, Borg R, et al. Translating the A1C assay into estimated average glucose values. *Diabetes Care.* 2008;31: 1473–1478 and Kilpatrick ES, Rigby AS, Atkin SL. Variability in the relationship between mean plasma glucose and HbA1c: implications for the assessment of glycemic control. *Clin Chem.* 2007;53:897–901.

third trimesters), recent blood loss or transfusion, erythropoietin therapy, or hemolysis.

Current ADA guidelines recommend that an HbA1c test be performed at least two times per year with patients who are meeting treatment goals and who have stable glycemic control. For patients whose therapy has changed or who are not meeting glycemic goals, a quarterly HbA1c test is recommended. The use of POC testing for HbA1c allows for more timely decisions on therapy changes and has been shown to result in tighter glycemic control.[13,14] However, using POC assays for diagnostic purposes is not recommended, since proficiency testing of personnel is not required.[2] Lowering HbA1c to an average of less than 7% has clearly been shown to reduce the microvascular, retinopathic, and neuropathic complications of diabetes. Therefore, the HbA1c goal for nonpregnant adults in general is less than 7%. Further studies have shown a small benefit to lowering HbA1c to less than 6%, making this a goal for selected individual patients if it is possible without significant hypoglycemia. More (<6.5%) or less (<8%) stringent goals might be selected for individual patients based on their medical history. For pediatric populations, the recommended target HbA1c is less than 7.5%.

The specimen requirement for HbA1c measurement is an EDTA whole blood sample. Before analysis, a hemolysate must be prepared. The methods of measurement are grouped into two major categories: (1) based on charge differences between glycosylated and nonglycosylated hemoglobin (cation-exchange chromatography, electrophoresis, and isoelectric focusing), and (2) structural characteristics of glycogroups on hemoglobin (affinity chromatography and immunoassay).

In the clinical laboratory, affinity chromatography is the preferred method of measurement. In this method, the glycosylated hemoglobin attaches to the boronate group of the resin and is selectively eluted from the resin bed using a buffer. This method is not temperature dependent and not affected by hemoglobin F, S, or C. Another method of measurement uses cation-exchange chromatography in which the negatively charged hemoglobins attach to the positively charged resin bed. The glycosylated hemoglobin is selectively eluted from the resin bed using a buffer of specific pH in which the glycohemoglobins are the most negatively charged and elute first from the column. However, this method is highly temperature dependent and affected by hemoglobinopathies. The presence of hemoglobin F yields false increased levels, and the presence of hemoglobins S and C yields false decreased levels. A common POC instrument HbA1c assay is based on a latex immunoagglutination inhibition methodology. In this method, both the concentration of HbA1c specifically and the concentration of total hemoglobin are measured, and the ratio is reported as percent HbA1c. In this method, glycated hemoglobin F is not measured, so at a very high level of hemoglobin F (>10%), the amount of HbA1c will be lower than expected because a greater proportion of the glycated hemoglobin will be in the form of glycated hemoglobin F. High-performance liquid chromatography (HPLC) and electrophoresis methods are also used to separate the various forms of hemoglobin. With HPLC, all forms of glycosylated hemoglobin—A1a, A1b, and A1c—can be separated.

Standardization of glycosylated hemoglobin has been a continuing problem; there was no consensus on the reference method, and no single standard is available to be used in the assays. Because of this, HbA1c values vary with the method and laboratory performing them (**Table 9.12**). The International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) developed a common definition for HbA1c and a reference method that specifically measures the concentration of only one molecular species of glycated A1c, the glycated N-terminal residue of the β-chain of hemoglobin. This method,

Table 9.12 Methods of Glycated Hemoglobin Measurement

Methods Based on Structural Differences		
Immunoassays	Polyclonal or monoclonal antibodies toward the glycated N-terminal group of the β-chain of hemoglobin	
Affinity chromatography	Separates based on chemical structure using borate to bind glycosylated proteins	Not temperature dependent Not affected by other hemoglobins
Methods Based on Charge Differences		
Ion-exchange chromatography	Positive-charge resin bed	Highly temperature dependent Affected by hemoglobinopathies
Electrophoresis	Separation is based on differences in charge	Hemoglobin F values > 7% interfere
Isoelectric focusing	Type of electrophoresis using isoelectric point to separate	Pre-HbA1c interferes
High-pressure liquid chromatography	A form of ion-exchange chromatography	Separates all forms of glyco-Hb: A1a, A1b, and A1c

using either HPLC/electrospray mass spectrometry or HPLC/capillary electrophoresis, is only used to standardize A1c assays and cannot be used for the clinical measurement of HbA1c.[15] International consensus groups have determined that HbA1c results will be represented worldwide in IFCC units (mmol HbA1c/mol Hb) and derived NGSP units (%) using the IFCC–NGSP master equation.[16,17] The master equation for this conversion is NGSP = [0.09148 × IFCC] + 2.152. HbA1c reagent and instrument manufacturers are required to document their traceability to the IFCC reference system with both the IFCC units (mmol/mol) and derived NGSP units (%). In the United States, the NGSP, with the Diabetes Control and Complications Trial (DCCT) HPLC method, is used as a primary reference method. An allowable error of 5 mmol HbA1c/mol Hb (0.46%) has been determined as acceptable.[18–20]

With these developments, HbA1c measurement has become a more reliable indicator of the long-term patient blood glucose regulation with more consistency in the results from laboratory to laboratory.

Ketones

Ketone bodies are produced by the liver through metabolism of fatty acids to provide a ready energy source from stored lipids at times of low carbohydrate availability. The three ketone bodies are acetone (2%), acetoacetic acid (20%), and 3-β-hydroxybutyric acid (78%). Low levels of ketone bodies are normally present in the body. However, in cases of carbohydrate deprivation or decreased carbohydrate use, such as occurs in starvation/fasting, high-fat diets, diabetes mellitus, prolonged vomiting, or glycogen storage disease, blood levels of ketones increase to meet the energy needs. The term *ketonemia* refers to the accumulation of ketones in blood, and the term *ketonuria* refers to the accumulation of ketones in urine (**Figure 9.9**). The measurement of ketones is recommended for patients with type 1 diabetes during acute illness, stress, pregnancy, or elevated blood glucose levels above 300 mg/dL, or when the patient has signs of ketoacidosis.

The specimen requirement is fresh serum or urine; the sample should be tightly stoppered and analyzed

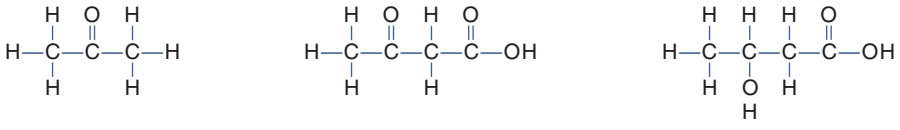


Figure 9.9 The three ketone bodies.

Table 9.13 Methods of Ketone Measurement

Nitroprusside	Acetoacetic acid + nitroprusside $\xrightarrow{\text{alkaline pH}}$ purple color
Enzymatic	3-β-hydroxybutyrate + NAD^+ $\xrightarrow{\text{3-Hbd}}$ acetoacetate + H^+ + NADH

immediately. No method used for the determination of ketones reacts with all three ketone bodies. The historical test (Gerhardt's) that used ferric chloride reacted with acetoacetic acid to produce a red color. The procedure had many interfering substances, including salicylates. A more common method using sodium nitroprusside ($NaFe[CN]_5NO$) reacts with acetoacetic acid in an alkaline pH to form a purple color. If the reagent contains glycerin, then acetone is also detected. This method is used with the urine reagent strip test and Acetest tablets. A newer enzymatic method adapted to some automated instruments uses the enzyme 3-hydroxybutyrate dehydrogenase to detect either 3-β-hydroxybutyric acid or acetoacetic acid, depending on the pH of the solution. A pH of 7.0 causes the reaction to proceed to the right (increasing absorbance) and a pH of 8.5 to 9.5 causes the reaction to proceed to left (decreasing absorbance; **Table 9.13**).

Albuminuria

Diabetes mellitus causes progressive changes to the kidneys and ultimately results in diabetic renal nephropathy, now referred to as diabetic kidney disease. This complication progresses over years and may be delayed by aggressive glycemic control. An early sign that diabetic kidney disease is occurring is an increase in urinary albumin. In the latest ADA guidelines, the term microalbumin is no longer used "since albuminuria occurs on a continuum."[21] **Albuminuria** measurements are useful to assist in diagnosis at an early stage and before the development of proteinuria (**Table 9.14**). An annual assessment of kidney function by the determination of urinary albumin excretion and estimated glomerular filtration rate is recommended for diabetic patients. Persistent albuminuria, defined as urinary albumin to creatinine ratio of 30 to 299 mg/g creatinine in two out of three urine collections over a 3- to 6-month period, is an early indicator of diabetic kidney disease. Factors that may elevate the urinary excretion of albumin include exercise within 24 hours, infection, fever, congestive heart failure, marked hyperglycemia, and marked hypertension.[22]

Table 9.14 Reference Ranges for the Nondiabetic Patient

Glucose, plasma, or serum, fasting	70–99 mg/dL
HbA1c	4.0-5.6%
Microalbumin, Urine (Random)	<25 mg/g creatinine
Spot collection	<30 mg/g Cr
Timed collection	<20 μg/min
24-h collection	<30 mg/24 h

The use of a random spot collection for the measurement of the albumin–creatinine ratio is the preferred method.[23] Using the spot method, without the simultaneous creatinine measurement, may result in false-positive and false-negative results because of variation in urine concentration. The two other alternatives, a 24-hour collection and a timed 4-hour overnight collection, which are more burdensome to the patient and add little to prediction or accuracy, are seldom required.

Islet Autoantibody, Insulin Testing, and C-Peptide Testing

The presence of autoantibodies to the β-islet cells of the pancreas is characteristic of type 1 diabetes (**Table 9.15**). However, islet autoantibody testing is only recommended in the setting of a research trial or as an option for first-degree family members with type 1 diabetes.[24] In the future, this testing might identify at-risk, prediabetic patients. Insulin measurements are not required for the diagnosis of diabetes mellitus, but in certain hypoglycemic states, it is important to know the concentration of insulin in relation to the plasma glucose concentration. Although C-peptide has no known biological activity, its levels serve as an indication of β-cell function, decreasing with loss of β-cell mass in type 1

CASE STUDY 9.10, PART 2

Recall Shirley, the hospice nurse caring for patients with diabetes. Janet's value from the Accu-Chek glucose monitor is 200 mg/dL. The Accu-Chek glucose monitor had acceptable QC results several hours previously. A plasma sample collected 5 minutes later by a phlebotomist and processed by the laboratory, resulted in a glucose value of 225 mg/dL.

5. Are these two glucose results significantly different?
6. Explain.

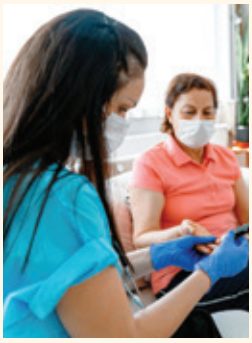


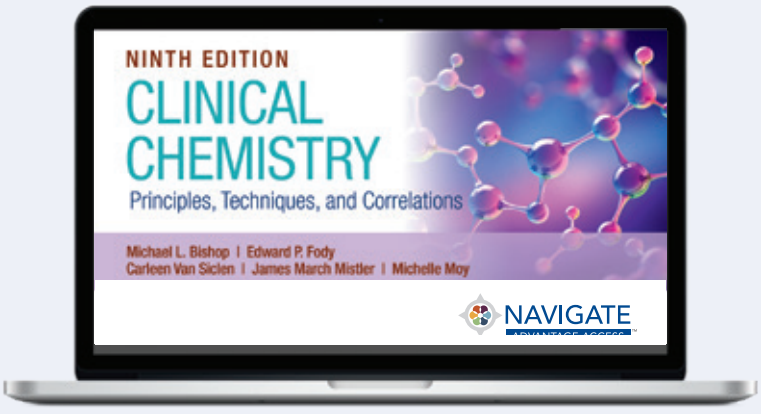
Table 9.15 Autoantibody Markers

Test	Abbreviation	Description
Islet Cell Cytoplasmic Autoantibodies	ICA	Measures a group of islet cell autoantibodies targeted against a variety of islet cell proteins
Glutamic Acid Decarboxylase Autoantibodies	GADA	Tests for autoantibodies directed against β cell protein but is not specific to β cells
Insulinoma-Associated-2 Autoantibodies	IA-2A	Tests for autoantibodies directed against β cell antigens but is nonspecific
Insulin Autoantibodies	IAA	Tests for autoantibodies targeted to insulin; insulin is the only antigen thought to be highly specific for β cells.
Zinc transporter 8	ZnT8	Autoantibodies targeted against the zinc transporter

diabetes, or increasing with insulin resistance-associated hyperinsulinemia. Using a combination of urinary C-peptide/creatinine ratio and antibody screening as markers may aid in determining the diagnosis of MODY and LADA, or when there is uncertainty about a type 1 versus a type 2 diagnosis.[25]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. American Diabetes Association. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes 2021. *Diabetes Care* 2021;44(suppl. 1):S15–S33.
2. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose tolerance. *Diabetes*. 1979;28:1039–1057.
3. American Diabetes Association. Classification and diagnosis of diabetes. Sec. 2. In Standards of Medical Care in Diabetes. 2016. *Diabetes Care*. 2015;39(suppl. 1):S13–S22.
4. Redondo MJ, Steck AK, Pugliese A. Genetics of type 1 diabetes. *Pediatr Diabetes*. 2018;19(3):346–353.
5. International Hypoglycaemia Study Group. Glucose concentrations of less than 3.0 mmol/L (54 mg/dL) should be reported in clinical trials: a joint position statement of the American Diabetes Association and the European Association for the Study of Diabetes. *Diabetes Care* 2017;40:155–157.
6. Cryer PE, Axelrod L, Grossman AB, et al.; Endocrine Society. Evaluation and management of adult hypoglycemic disorders: an Endocrine Society Clinical Practice Guideline. *J Clin Endocrinol Metab*. 2009;94:709–728.
7. Service FJ. Medical progress: hypoglycemic disorders. *N Engl J Med*. 1995;332(17):1144–1152.
8. Troisi RJ, Cowie CC, Harris MI. Diurnal variation in fasting plasma glucose: implications for diagnosis of diabetes in patients examined in the afternoon. *JAMA*. 2000;284: 3157–3159.
9. American Diabetes Association. Tests of glycemia in diabetes. *Diabetes Care*. 2004;27(suppl. 1):S91–S93.
10. Eckfeldt JH, Bruns DE. Another step towards standardization of methods for measuring hemoglobin A1c. *Clin Chem*. 1997;43(10):1811–1813.
11. Nathan DM, Kuenen J, Borg R, et al. Translating the A1c assay into estimated average glucose values. *Diabetes Care*. 2008;31:1473–1478.
12. Kilpatrick ES, Rigby AS, Atkin SL. Variability in the relationship between mean plasma glucose and HbA1c: implications for the assessment of glycemic control. *Clin Chem*. 2007;53:897–901.
13. American Diabetes Association. Standards of medical care in diabetes—2016. *Diabetes Care*. 2016;39(suppl. 1):S1–S109.
14. American Diabetes Association. Glycemic targets: Standards of medical care in diabetes 2021. *Diabetes Care* 2021; 44(suppl. 1):S73–S84.
15. Jeppsson JO, Kobold U, Barr J, et al. International Federation of Clinical Chemistry and Laboratory Medicine (IFCC): Approved IFCC reference method for the measurement of HbA1c in human blood. *Clin Chem Lab Med*. 2002;40:78–89.
16. Sacks DB. Measurement of hemoglobin A1c. *Diabetes Care* 2012;35(12):2674–2680.
17. Geistanger A, Arends S, Berding C, et al. Statistical methods for monitoring the relationship between the IFCC reference measurement procedure for hemoglobin A1c and the designated comparison methods in the United States, Japan, and Sweden. *Clin Chem*. 2008;54:1379–1385.
18. Weykamp CW, Mosca A, Gillery P, et al. The analytical goals for hemoglobin A1c measurement in IFCC units and National Glycohemoglobin Standardization Program units are different. *Clin Chem*. 2011;57:1204–1206.
19. Weykamp C, John G, Gillery P, et al. Investigation of 2 models to set and evaluate quality targets for HbA1c: biological variation and sigmametrics. *Clin Chem*. 2015;61:752–759.
20. Braga F, Panteghini M. Standardization and analytical goals for glycated hemoglobin measurement. *Clin Chem Lab Med*. 2013;51(9):1719–1726.
21. American Diabetes Association. Microvascular complications and foot care. Sec. 9. In Standards of Medical Care in Diabetes–2016. *Diabetes Care*. 2016;39(suppl. 1): S72–S66, S80.
22. Sacks DB, Arnold M, Bakris GL, et al. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Clin Chem*. 2011;57: p.e1–p.e47. Published May 26, 2011.
23. National Kidney Foundation. KDOQI Clinical Practice Guideline for Diabetes and CKD: 2012 update. *Am J Kidney Dis*. 2012;60(5):850–886.
24. Winter WE, Schatz DA. Autoimmune markers in diabetes. *Clin Chem*. 2011;57(2):168–175.
25. Leighton E, Sainsbury CA, Jones GC. A practical review of C-peptide testing in diabetes. *Diabetes Ther*. 2017;8(3): 475–487.

CHAPTER 10

Lipids and Lipoproteins

Kathleen M. Kenwright

CHAPTER OUTLINE

KEY TERMS

Arteriosclerosis
Atherosclerosis
Cholesterol
Chylomicrons
Endogenous pathway
Exogenous pathway
Fatty acids
Friedewald equation
HDL
LDL
Lipoprotein
Lipoprotein a (LP[a])
Phospholipids
Triglycerides
VLDL
Xanthomas

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Explain lipoprotein physiology and metabolism.
- Describe the structure of fatty acids, phospholipids, triglycerides, cholesterol, and the various types of lipoprotein particles discussed.
- State the laboratory tests used to assess lipids and lipoproteins, including principles and procedures.
- Correlate common lipid disorders with clinical and laboratory data.
- Discuss the incidence and types of lipid and lipoprotein abnormalities.
- Identify the reference ranges for the major serum lipids.
- Interpret the clinical significance of lipid and lipoprotein values in the assessment of coronary heart disease.
- Illustrate the role of standardization in the measurement of lipids and lipoproteins.

Lipoproteins constitute the body's "petroleum industry." Like the great oil tankers that travel the oceans of the world transporting petroleum for fuel needs, **chylomicrons** are large, lipid-rich transport vessels that ferry their cargo of dietary triglycerides, the main oil in the body, throughout the circulatory system to peripheral tissues, finally docking when nearly empty at the liver as chylomicron remnants. The very-low-density lipoproteins (**VLDLs**) are like tanker trucks, redistributing dietary and hepatic-synthesized triglycerides to peripheral cells mostly during fasting for energy needs or storage as fat. The low-density lipoproteins (**LDLs**), rich in cholesterol, which start out as VLDLs, are like nearly empty tankers that just deliver cholesterol to peripheral cells and return to the liver after their main cargo, triglycerides, have been off-loaded. The high-density lipoproteins (**HDLs**) are the cleanup crew, gathering up excess cholesterol for transport back to the liver. Cholesterol is used by the body for such useful functions as facilitating triglyceride transport by lipoproteins and maintaining the normal structure and integrity of cell membranes, and as a precursor for steroid hormone synthesis, but when in excess, it can lead to cardiovascular disease.

Lipids and lipoproteins, which are central to the energy metabolism of the body, have become increasingly important in clinical practice, primarily because of their association with coronary heart disease (CHD). Numerous epidemiologic studies have demonstrated that, especially in affluent countries with high fat consumption, there is a clear association between the blood lipid levels and the development of **atherosclerosis** (hardening and narrowing of large arteries). Decades of basic research have also contributed to knowledge about the nature of the lipoproteins and their lipid and protein constituents, as well as their role in the pathogenesis of the atherosclerotic process.

The accurate measurement of the various lipid and lipoprotein parameters is critical in the diagnosis and treatment of patients with **dyslipidemia** (disease states associated with abnormal serum lipids). International efforts to reduce the impact of CHD on public health have focused attention on improving the reliability and convenience of the

CASE STUDY 10.1, PART 1

Maggie is a 13-year-old White female with no preexisting medical conditions. When she visits her pediatrician for an annual well-child visit, her mother tells the doctor that Maggie is an active child who enjoys playing volleyball, roller blading, and riding her bicycle. In addition, Maggie has recently discovered that she loves baking and decorating cakes.

lipid and lipoprotein assays. Expert panels have developed guidelines for the detection and treatment of high cholesterol, as well as laboratory performance goals of accuracy and precision for the measurement of the lipid and lipoprotein analytes. This chapter begins with a review of lipid chemistry and lipoprotein metabolism, followed by the diagnosis and treatment of dyslipidemia. Finally, the clinical laboratory measurement of lipids and lipoproteins will be discussed in the context of various national and international guidelines, such as the National Cholesterol Education Program (NCEP).

Lipid Chemistry

Lipids, commonly referred to as fats, are ubiquitous constituents of all living cells and have a dual role. First, because they are composed of mostly carbon–hydrogen (C–H) bonds, they are a rich source of energy and an efficient way for the body to store excess calories. Because of their unique physical properties, lipids are also an integral part of cell membranes, and therefore also play an important structural role in cells. Lipids are also precursors for the steroid hormones, prostaglandins, leukotrienes, and lipoxins. The lipids transported by lipoproteins, namely, triglycerides, **phospholipids**, cholesterol, and cholesteryl esters, are also the principal lipids found in cell membranes and intracellular lipid droplets, and the main focus of this section.

Fatty Acids

Fatty acids, as seen in the structure shown in **Figure 10.1**, are simply linear chains of C–H bonds that terminate with a carboxyl group (–COOH). In plasma, only relatively small amounts of fatty acids exist in the free or unesterified form, most of which are noncovalently bound to albumin. The majority of plasma fatty acids are instead found as a constituent of triglycerides or phospholipids (Figure 10.1). Fatty acids are covalently attached to the glycerol backbone of triglycerides and phospholipids by an ester bond that forms between the carboxyl group on the fatty acid and the hydroxyl group (–OH) on glycerol. Fatty acids are variable in length and can be classified as short-chain (4 to 6 carbon atoms), medium-chain

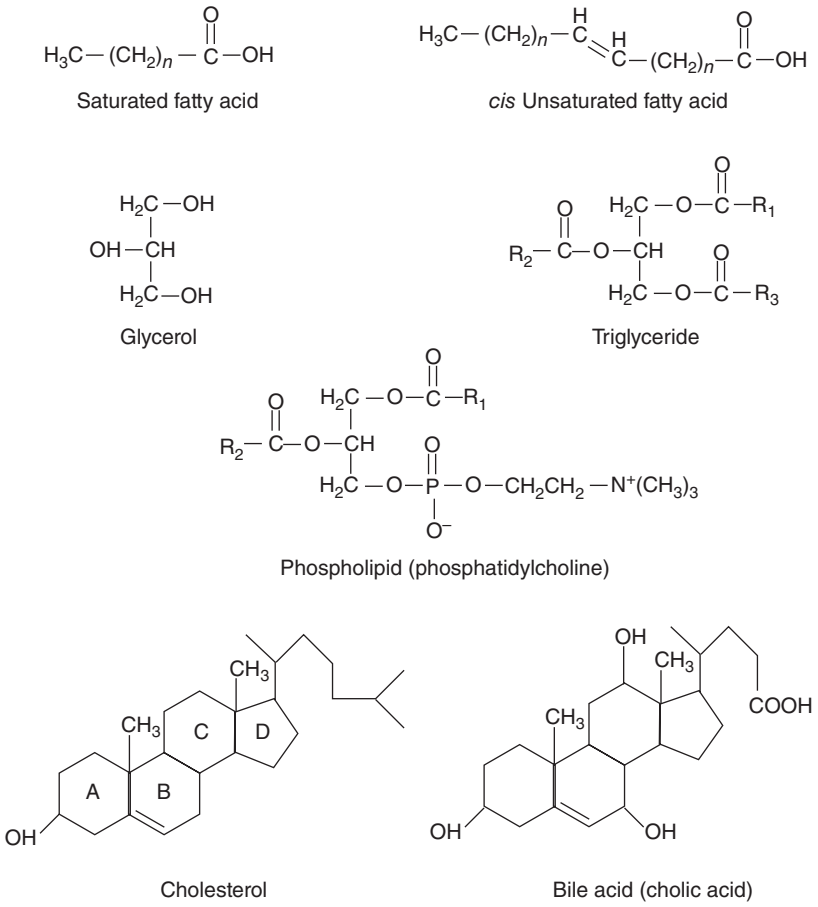


Figure 10.1 Chemical structures of lipids. Fatty acids are abbreviated as (R) for triglycerides and phospholipids.

(8 to 12 carbon atoms), long-hain (12 to 18 carbon atoms), or very-long-chain (>20 carbon atoms) fatty acids. Most fatty acids in our diet are of the long-chain variety and contain an even number of carbon atoms, such as palmitic acid (16 carbons) and stearic, oleic, linoleic, and linolenic acids (18 carbons). Not all carbon atoms in fatty acids are fully saturated—or, in other words, bonded with hydrogen atoms; some of them may instead form unsaturated carbon–carbon (C=C) double bonds. Depending on the number of C=C double bonds, fatty acids can be classified as being saturated (no double bonds) like palmitic acid, monounsaturated (one double bond) like oleic acid, or polyunsaturated (two or more double bonds) like linoleic and linolenic acids. The C=C double bonds of unsaturated fatty acids are typically arranged in the cis configuration, with both hydrogen atoms on the same side of the C=C double bond, which causes a bend in their molecular structure (Figure 10.1). These bends increase the space that unsaturated fatty acids require when packed in a lipid layer, and as a result, these fatty acids are more fluid because they do not as readily self-associate and pack together tightly. Fatty acid C=C double bonds can also occur in the trans configuration, with both hydrogen atoms on opposite sides of the C=C double bond. Because of the spatial orientation of their double bonds, trans fatty acids do not bend and have physical properties more like saturated fatty acids. The trans fatty acids are not commonly found in nature; however, they are commonly found in our diet because of the chemical hydrogenation treatment used in food processing. This hydrogenation increases the viscosity of oils (hardening fat) by converting polyunsaturated plant oils into solid margarine, which introduces trans double bonds. The major dietary trans fatty acid is elaidic acid, an 18-carbon fatty acid with one double bond. Both metabolic and epidemiological studies have shown that the consumption of trans fatty acids increases the risk of CHD. Therefore, the Food and Drug Administration (FDA) passed a rule that bans companies from adding trans fat to their food products. They had until 2020 to meet compliance, with only a few exceptions.[1] The adverse effects of trans fatty acids are mostly attributed to an elevation of LDL and a decrease in HDL.[2]

Most fatty acids are synthesized in the body from carbohydrate precursors, except linoleic and linolenic acids, which are referred to as essential fatty acids.[3] These two fatty acids are found in plants and must be ingested in the diet. Both these essential fatty acids are important for growth, maintenance, and proper functioning of many physiological processes.

The polyunsaturated fatty acids are classified into omega-3, omega-6, and omega-9 families, depending on the position of the first double bond from the terminal (omega) methyl group of the fatty acid chain. Polyunsaturated fatty acids are precursors for the synthesis of eicosanoids that include prostaglandins, thromboxanes, prostacyclins, and leukotrienes. These fatty acids play a vital role in the structure and function of most biological membranes, and some fatty acids like omega-3 polyunsaturated fatty acids found in fish oils are beneficial in lowering the risk of cardiovascular disease.

Triglycerides

As can be inferred from the name, **triglycerides** contain three fatty acid molecules attached to one molecule of glycerol by ester bonds in one of three stereochemically distinct bonding positions, referred to as sn-1, sn-2, and sn-3 (Figure 10.1). Each fatty acid in the triglyceride molecule can potentially be different in structure, thus producing many possible types of triglycerides. Triglycerides containing saturated fatty acids, which do not have bends in their structure, pack together more closely and tend to be solid at room temperature. In contrast, triglycerides, containing cis unsaturated fatty acids (Figure 10.1), typically form oils at room temperature. Most triglycerides from plant sources, such as corn, sunflower seeds, and safflower seeds, are rich in polyunsaturated fatty acids and are oils at room temperature, whereas triglycerides from animal sources contain mostly saturated fatty acids and are usually solid at room temperature. As can be seen by inspecting the structure of triglycerides (Figure 10.1), there are no charged or polar hydrophilic groups, making it very hydrophobic and virtually water insoluble. Because it is uncharged, a triglyceride is classified as a neutral lipid.

Phospholipids

Phospholipids are similar in structure to triglycerides except that they only have two esterified fatty acids (Figure 10.1). The third position on the glycerol backbone instead contains a phospholipid head group. There are several types of phospholipid head groups, such as choline, inositol, inositol phosphates, glycerol, serine, and ethanolamine, which are all hydrophilic in nature. The various types of phospholipids are named based on the type of phospholipid head group present. Phosphatidylcholine (often referred as lecithin) (Figure 10.1) has a choline head group and is the most common phospholipid found on lipoproteins and in cell membranes. The two asymmetrically

positioned fatty acids in phospholipids are typically 14 to 24 carbon atoms long, with one fatty acid commonly saturated and the other unsaturated, most commonly in the sn-2 position.

Because phospholipids contain both hydrophobic fatty acid C–H chains and a hydrophilic head group, they are by definition amphipathic lipid molecules and, as such, are found on the surface of lipid layers or on the surface of lipoprotein particles. The polar hydrophilic head group faces outward toward the aqueous environment, whereas the fatty acid chains face inward away from the water in a perpendicular orientation with respect to the lipid surface. Phospholipids are synthesized in the cytosolic compartment of all organs of the body, especially in the liver, with phosphatidylcholine and phosphatidylethanolamine being the most abundant phospholipids in the body.

Cholesterol

Cholesterol is an unsaturated steroid alcohol containing four rings (A, B, C, and D), and it has a single C–H side chain tail similar to a fatty acid in its physical properties (Figure 10.1). The only hydrophilic part of cholesterol is the hydroxyl group in the A-ring. Cholesterol is, therefore, also an amphipathic lipid and is found on the surface of lipid layers along with phospholipids. Cholesterol is oriented in lipid layers so that the four rings and the side chain tail are buried in the membrane in a parallel orientation to the fatty acid acyl chains on adjacent phospholipid molecules. The polar hydroxyl group on the cholesterol A-ring faces outward, away from the lipid layer, allowing it to interact with water by noncovalent hydrogen bonding or with the polar head groups of phospholipids.

Cholesterol can also exist in an esterified form called cholesteryl ester, with the hydroxyl group conjugated by an ester bond to a fatty acid, in the same way as in triglycerides. In contrast to free cholesterol, there are no polar groups on cholesteryl esters, making them very hydrophobic. Because it is not charged, cholesteryl esters are also classified as a neutral lipid and are not found on the surface of lipid layers but instead are located in the center of intracellular lipid droplets or in the hydrophobic core of lipoproteins, along with triglycerides.

Cholesterol is almost exclusively synthesized by animals, but plants do contain other sterols (called phytosterols) similar in structure to cholesterol. Dietary phytosterols are known to lower plasma total cholesterol and LDL cholesterol (LDL-C) and raise HDL cholesterol (HDL-C) levels, most likely by competing with the intestinal absorption of cholesterol.

Cholesterol is synthesized in most tissues of the body from acetyl coenzyme A (acetyl-CoA) in the microsomal and cytosolic compartments of the cell. More than 20 enzymes are involved in the formation of cholesterol from acetyl-CoA. The principal steps include the conversion of acetyl-CoA, derived from β-oxidation of fatty acids and dicarboxylic acids, to β-hydroxy β-methyl glutaryl CoA (HMG-CoA).[4] HMG-CoA is then converted to melovanate by the enzyme HMG-CoA reductase (HMHCR), the rate-limiting enzyme in cholesterol biosynthesis.[4] Mevalonic acid is phosphorylated, isomerized, and converted to geranyl and farnesyl pyrophosphates, which, in turn, form squalene.[4] Cyclization of squalene occurs in the endoplasmic reticulum with oxidation, methyl group transfer reactions, saturation of the side chain, and double bond shifting to produce cholesterol.[4]

Cholesterol is also unique in that, unlike other lipids, it is not readily catabolized by most cells and, therefore, does not serve as a source of fuel. Cholesterol can, however, be converted in the liver to primary bile acids, such as cholic acid (Figure 10.1) and chenodeoxycholic acid, which promote fat absorption in the intestine by acting as detergents. A small amount of cholesterol can also be converted by some tissues, such as the adrenal gland, testis, and ovary, to steroid hormones, such as glucocorticoids, mineralocorticoids, and estrogens. Finally, a small amount of cholesterol, after first being converted to 7-dehydrocholesterol, can also be transformed to vitamin D_3 in the skin by irradiation from sunlight.

General Lipoprotein Structure

The prototypical structure of a lipoprotein particle is shown in **Figure 10.2**. Lipoproteins are typically spherical in shape and range in size from as small as 10 nm to more than 1 μm (**Table 10.1**) As the name implies, lipoproteins are composed of both lipids and proteins, called apolipoproteins. The amphipathic cholesterol and phospholipid molecules are primarily found on the surface of lipoproteins as a single monolayer, whereas the hydrophobic and neutral triglyceride and cholesteryl ester molecules are found in the central or core region and thus are micelles (Figure 10.2). Because the main role of lipoproteins is the delivery of fuel to peripheral cells, the core of the lipoprotein particle essentially represents the

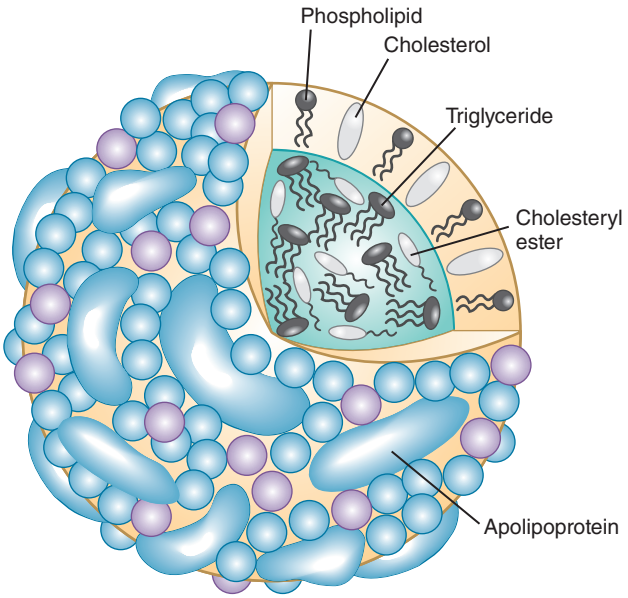


Figure 10.2 Model of lipoprotein structure.

cargo that is being transported by lipoproteins. The size of the lipoprotein particle correlates with its core neutral lipid content. The larger lipoprotein particles have correspondingly larger core regions and, therefore, contain relatively more triglyceride and cholesteryl ester. The larger lipoprotein particles also contain more lipid relative to protein and thus are lighter in density. The various lipoprotein particles were originally separated by ultracentrifugation into different density fractions (chylomicrons [chylos], VLDL, LDL, and HDL), which still form the basis for the most commonly used lipoprotein classification system (Table 10.1).

Apolipoproteins are primarily located on the surface of lipoprotein particles (**Table 10.2**). They help maintain the structural integrity of lipoproteins and also serve as ligands for cell receptors and as activators and inhibitors of the various enzymes that modify lipoprotein particles (Table 10.2). Apolipoproteins contain a structural motif called an amphipathic α helix, which accounts for the ability of these proteins to bind to lipids. Amphipathic helices are protein segments arranged in coils so that the hydrophobic amino acid residues interact with lipids, whereas the part of the helix containing hydrophilic amino acids faces away from the lipids and toward the aqueous environment.

Apolipoprotein (Apo) A1, the major protein of HDL, is frequently used as a measure of the amount of the antiatherogenic HDL present in plasma.[5] Apo B is a large protein with a molecular weight of approximately 500 kD; it is the principal protein of LDL, VLDL, and chylomicrons.[6] The Apo B gene (*APOB*) is found on chromosome 2 and produces two forms of Apo B: Apo B100 and Apo B48. Apo B100 is found in LDL and VLDL and is a ligand for the LDL receptor; it is, therefore, critical in the uptake of LDL by cells.[7] Apo B48, exclusively found in chylomicrons, is produced in the intestine. Apo B48 is essentially the first 48% of the Apo B molecule, whereas full-length Apo B is produced in the liver.[7] Apo B also contains amphipathic α helices but, unlike other apolipoproteins, also contains amphipathic β sheets that avidly bind to lipids. Apo B100 can also be found covalently linked to Apo(a),[8] a plasminogen-like protein that is found in a proatherogenic lipoprotein particle called **lipoprotein (a) [Lp(a)]**. Apo E, another important apolipoprotein found in many types of lipoproteins (LDL, VLDL, and HDL), also serves as a ligand for the LDL receptor and the chylomicron remnant receptor. There are three major isoforms of Apo E: Apo E2, E3, and E4. The Apo E isoforms affect lipoprotein metabolism because they differ in their ability to interact with the LDL receptor. For example, patients who are homozygous for the Apo E2 isoform have low affinity

Table 10.1 Characteristics of the Major Human Lipoproteins

Characteristics	Chylomicrons	VLDL	LDL	HDL
Density (g/mL)	<0.93	0.93–1.006	1.019–1.063	1.063–1.21
Molecular weight (kD)	(0.4–30) × 10^9	(10–80) × 10^6	2.75 × 10^6	(1.75–3.6) × 10^5
Diameter (nm)	80–1200	30–80	18–30	5–12
Total lipid (% by weight)	98	89–96	77	50
Triglycerides (% by weight)	84	44–60	11	3
Total cholesterol (% by weight)	7	16–22	62	19

Chylos, chylomicrons; HDL, high-density lipoprotein; LDL, low-density lipoprotein; VLDL, very-low-density lipoprotein.

Table 10.2 Characteristics of the Major Human Apolipoproteins

Apolipoprotein	Molecular Weight (kD)	Plasma Concentration (mg/dL)	Major Lipoprotein Location	Function
Apo A1	28,000	100–200	HDL	Structural, LCAT activator, ABCA1 lipid acceptor
Apo A2	17,400	20–50	HDL	Structural
Apo A4	44,000	10–20	Chylos, VLDL, HDL	Structural
Apo B100	5.4×10^5	70–125	LDL, VLDL	Structural, LDL receptor ligand
Apo B48	2.6×10^5	<5	Chylos	Structural, remnant receptor ligand
Apo C1	6630	5–8	Chylos, VLDL, HDL	Structural
Apo C2	8900	3–7	Chylos, VLDL, HDL	Structural, LPL cofactor
Apo C3	9400	10–12	Chylos, VLDL, HDL	Structural, LPL inhibitor
Apo E	34,400	3–15	VLDL, HDL	Structural, LDL receptor ligand
Apo(a)	$(3–7) \times 10^5$	<30	Lp(a)	Structural, plasminogen inhibitor

Chylos, chylomicrons; HDL, high-density lipoprotein; LCAT, lecithin–cholesterol acyltransferase; LDL, low-density lipoprotein; LPL, lipoprotein lipase; VLDL, very-low-density lipoprotein.

for the LDL receptor and are at an increased risk for developing hyperlipoproteinemia type III (dysbetalipoproteinemia or broad β disease), a rare inborn error of metabolism. The connection with lipid metabolism is not completely understood, but individuals with the Apo E4 isoform have higher affinity for the LDL receptor and have been shown to have an increased risk of developing Alzheimer's disease.[9]

Chylomicrons

Chylomicrons, which contain Apo B48, are the largest and the least dense of the lipoprotein particles, having diameters as large as 1200 nm (Table 10.1).[10] Because of their large size, they scatter light, which accounts for the turbidity or milky appearance of postprandial plasma specimens. Because they are so light, they also readily float to the top of plasma and form a creamy layer when stored for hours or overnight at 4°C. Chylomicrons are produced by the intestine, where they are packaged with absorbed dietary lipids and apolipoproteins. Once they enter the circulation, triglycerides and cholesteryl esters in chylomicrons are rapidly hydrolyzed by lipases, such as lipoprotein lipase (LPL), and within a few hours are transformed into chylomicron remnant particles, which are recognized by proteoglycans and remnant receptors in the liver, facilitating their uptake.[10] The principal role of chylomicrons is the delivery of dietary lipids to hepatic and peripheral cells.

Very-Low-Density Lipoprotein

VLDL is produced primarily by the liver and contains Apo B100, the main apolipoprotein, Apo E, and Apo Cs; like chylomicrons, they are also rich in triglycerides. They are the major carriers of endogenous (hepatic-derived) triglycerides and shuttle triglycerides from the liver to peripheral tissue, mostly during the fasting state, for energy utilization and storage. Like chylomicrons, they also reflect light and account for most of the turbidity observed in fasting hyperlipidemic plasma specimens, although they do not form a creamy top layer like chylomicrons because they are smaller and less buoyant (Table 10.1). Excess dietary intake of carbohydrates, saturated fatty acids, and trans fatty acids enhances the hepatic synthesis of triglycerides, which in turn increases VLDL production. Free fatty acids released into the circulation by adipocytes are avidly taken up by the liver, which also stimulates VLDL secretion.

Intermediate-Density Lipoproteins

Intermediate-density lipoproteins (IDLs), also referred to as VLDL remnants, normally only exist transiently during the conversion of VLDL to LDL. The triglyceride and cholesterol contents of IDL are intermediate between those of VLDL and LDL. Normally, the conversion of VLDL to IDL proceeds so efficiently that

appreciable quantities of IDL usually do not accumulate in the plasma after an overnight fast; thus, IDLs are not typically present in high quantities in normal plasma. In patients with hyperlipoproteinemia type III, elevated levels of IDLs can be found in plasma. This defect is due to an abnormal form of Apo E that delays the clearance of IDL. Individuals with this disorder are at a significant risk for peripheral vascular disease (PVD) and coronary artery disease (CAD), presumably because these lipoprotein particles are damaging to blood vessel walls.

Low-Density Lipoproteins

LDL primarily contains Apo B100 and is more cholesterol rich than other Apo B–containing lipoproteins (Table 10.1). They form as a consequence of the lipolysis of VLDL. LDL is readily taken up by cells via the LDL receptor in the liver and peripheral cells. In addition, because LDL particles are significantly smaller than VLDL particles and chylomicrons, they can infiltrate into the extracellular space of the vessel wall, where they can be oxidized and taken up by macrophages through various scavenger receptors.[11] Macrophages that take up too much lipid become filled with intracellular lipid drops and turn into foam cells, which is the predominant cell type of fatty streaks, and an early precursor of atherosclerotic plaques.

LDL particles can exist in various sizes and compositions and have been separated into as many as eight subclasses through density ultracentrifugation or gradient gel electrophoresis.[12] The LDL subclasses differ largely in their content of core lipids; the smaller particles are denser and have relatively more triglyceride than cholesteryl esters. Recently, there has been great interest in measuring LDL subfractions, because small, dense, LDL particles have been shown to be more proatherogenic and may be a better marker for CHD risk.

Lipoprotein (a)

Lp(a) particles are LDL-like particles that contain one molecule of Apo(a) linked to Apo B100 by a single disulfide bond. Lp(a) particles are heterogeneous in both size and density as a result of a differing number of repeating peptide sequences, called *kringles*, in the Apo(a) portion of the molecule. Lp(a) is larger than LDL and has a higher lipid content and a slightly lower density. The concentration of Lp(a) is inversely related to the size of the isoform; the larger size isoforms are not as efficiently secreted from the liver. Plasma levels of Lp(a) vary widely among individuals in the general population but remain relatively constant within an individual. Lp(a) appears to be poorly cleared by the LDL receptor, but the kidney has been postulated as the site of removal.[13]

Elevated levels of Lp(a) (>30 mg/dL) are now known to increase the risk of premature CHD and stroke. Because the kringle domains of Lp(a) have a high level of homology with plasminogen, a precursor of plasmin that promotes clot lysis via fibrin cleavage, it has been proposed that Lp(a) may compete with plasminogen for binding sites on endothelium and on fibrin, thereby promoting clot formation. Clinical studies have demonstrated increased risk of both myocardial infarction and stroke with increasing Lp(a) concentration; however, the measurement of Lp(a) is often underutilized in clinical practice. Part of the reason is because the accurate measurement of Lp(a) is difficult, and specific therapies for reducing its concentration in blood are limited, although several are in development. Measuring Lp(a) is recommended for patients with a strong family history of CHD, particularly in the absence of other known risk factors such as increased LDL-C, for patients who develop CHD on statin therapy, or for patients with premature aortic stenosis, which has been shown to be caused by high Lp(a) levels.

High-Density Lipoproteins

HDL, the smallest and most dense lipoprotein particle, is synthesized by both the liver and the intestine (Table 10.1). HDL can exist as either disk-shaped particles or, more commonly, spherical particles.[5] Discoidal HDL typically contains two molecules of Apo A1, which form a ring around a central lipid bilayer of phospholipid and cholesterol. Discoidal HDL is believed to represent nascent or newly secreted HDL and is the most active form in removing excess cholesterol from peripheral cells. The ability of HDL to remove cholesterol from cells, called reverse cholesterol transport, is one of the main mechanisms proposed to explain the antiatherogenic property of HDL. When discoidal HDL has acquired an additional lipid, cholesteryl esters and triglycerides form a core region between its phospholipid bilayer, which transforms discoidal HDL into spherical HDL. HDL is highly heterogeneous in size and lipid and protein composition and is separable into as many as 13 or 14 different subfractions. There are two major types of spherical HDL based on density differences: HDL_2 (1.063 to 1.125 g/mL) and HDL_3 (1.125 to 1.21 g/mL). HDL_2 particles are larger in size, less dense, and richer in lipid than HDL_3 and may be more efficient in the delivery of lipids to the liver.

Lipoprotein X

Lipoprotein X (LpX) is an abnormal lipoprotein only produced in patients with cholestatic liver disease or in patients with mutations or deficiencies of lecithin–cholesterol acyltransferase (LCAT), the enzyme that esterifies cholesterol.[14] The mechanism of LpX formation in these disease states is currently unknown. LpX is different from other lipoproteins in the endogenous pathway in that it is composed of phospholipids, free cholesterol, Apo A1, and albumin but lacks Apo B100.[14] Phospholipids and cholesterol are its main lipid components (~90% by weight), and albumin and Apo A1 are the main protein components (<10% by weight). LpX is mainly removed by the reticuloendothelial system of the liver and the spleen. Other organs, such as the kidney, also actively clear lipoprotein X from the plasma, which may account for renal disease in patients with genetic LCAT deficiency.[14] Lipoprotein X is similar in density to LDL-C, which prevents differentiation of the two by common methods including the direct homogenous method and the estimated LDL-C method as calculated by the **Friedewald equation**; this may cause false elevations of LDL-C. LpX can be quantified after lipoprotein electrophoresis using filipin, a special stain.

Lipoprotein Physiology and Metabolism

The four major pathways involved in lipoprotein metabolism are shown in **Figure 10.3** and discussed in the following sections. The lipid absorption pathway, the **exogenous pathway**, and the **endogenous pathway** all depend on Apo B–containing lipoprotein particles and can be viewed as the process to transport dietary lipid and hepatic-derived lipid to peripheral cells. In terms of energy metabolism, these three pathways are critical in the transport to peripheral cells of fatty acids, which are generated during the lipolysis of triglycerides and, to a lesser degree, cholesteryl esters on lipoproteins. In regard to the pathogenesis of atherosclerosis, the net result of these three pathways is the net delivery of cholesterol to peripheral cells, which can lead to atherosclerosis when the cells in the vessel wall accumulate too much cholesterol. Peripheral cells are prone to accumulating cholesterol because they also synthesize their own cholesterol, and, unlike liver cells, they do not have the enzymatic pathways to catabolize cholesterol. Furthermore, cholesterol is relatively water insoluble and cannot readily diffuse away from its site of deposition or synthesis.

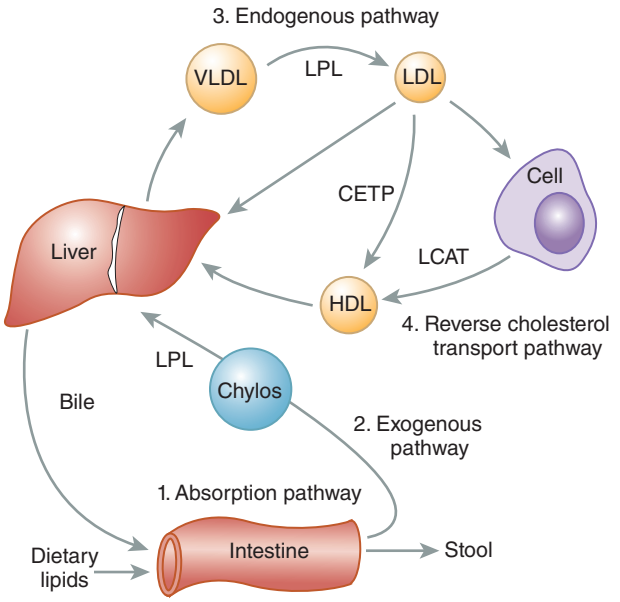


Figure 10.3 Diagram of major lipoprotein metabolism pathways.

The principal way that peripheral cells maintain their cholesterol equilibrium is the reverse cholesterol transport pathway (Figure 10.3), which is mediated by HDL. In this pathway, excess cholesterol from peripheral cells is transported back to the liver, where it can be excreted into the bile as free cholesterol or converted to bile acids. The liver is, therefore, involved in both forward and reverse cholesterol transport pathways and, in many ways, acts as a buffer in helping the body maintain its overall cholesterol homeostasis. There are several genetic defects in the genes that encode for proteins in the forward and reverse cholesterol transport pathways, which result in a predisposition for atherosclerosis. Most individuals with CAD, however, do not have a clear, single, genetic defect but instead have multiple genetic variations or gene polymorphisms that most likely interact with various lifestyle habits, such as exercise, diet, and smoking, leading to a predisposition for disease.

Lipid Absorption

Because fats are water insoluble, special mechanisms are required to facilitate the intestinal absorption of the fat consumed per day; in a typical Western diet, this could be upwards of 60 to 130 g. During the process of digestion, pancreatic lipase first converts dietary lipids into more polar compounds with amphipathic properties by cleaving off fatty acids. Thus, triglycerides are transformed into monoglycerides and diglycerides; cholesterol esters are transformed into free cholesterol; and phospholipids are

transformed into lysophospholipids. These amphipathic lipids in the intestinal lumen form large aggregates with bile acids called micelles. Lipid absorption occurs when the micelles come in contact with the microvillus membranes of the intestinal mucosal cells. Absorption of some of these lipids may occur via a passive transfer process; however, recent evidence suggests that in some cases, it might also be facilitated by specific transporters. Short-chain free fatty acids, with 10 or fewer carbon atoms, can readily pass directly into the portal circulation and are carried by albumin to the liver. The absorbed long-chain fatty acids, monoglycerides, and diglycerides are re-esterified in intestinal cells to form triglycerides and cholesteryl esters. The newly formed triglycerides and cholesteryl esters are then packaged into chylomicrons by the microsomal transfer protein, along with Apo B48.

Triglyceride absorption is extremely efficient; greater than 90% of dietary triglycerides are taken up by the intestine. In contrast, only about half of the 500 mg of cholesterol in the typical diet is absorbed each day. An even smaller fraction of plant sterols is absorbed by the intestine.

Exogenous Pathway

The newly synthesized chylomicrons in the intestine (Figure 10.3) are initially secreted into the lacteals (small intestine lymphatic vessels) and then pass into the lymphatic ducts, eventually entering the circulation by way of the thoracic duct. After entering the circulation, chylomicrons interact with proteoglycans, such as heparin sulfate, on the luminal surface of capillaries in various tissues, such as skeletal muscle, heart, and adipose tissue. The proteoglycans along with a specific protein also promote the binding of LPL,[15] which hydrolyzes triglycerides on chylomicrons. The free fatty acids and glycerol generated by the hydrolysis of triglycerides by LPL can then be taken up by cells and used as a source of energy. Excess fatty acids, particularly in fat cells (adipocytes), are re-esterified into triglycerides for long-term energy storage in intracellular lipid drops. A key protein in triglyceride metabolism is Apo C2, which is found on VLDL and is critical in the activation of LPL. Hormone-sensitive lipase, another lipase that is found inside adipose cells, releases free fatty acids from triglycerides in stored fat during fasting states when energy sources from carbohydrates are insufficient for the body's energy needs. The hormones epinephrine and cortisol play a key role in the mobilization and hydrolysis of triglycerides from adipocytes, whereas insulin prevents lipolysis by adipocytes and promotes fat storage and glucose utilization.

During lipolysis of chylomicrons, there is a transfer of lipid (mainly triglyceride) and apolipoproteins (Apo A1 and C2) onto HDL, and chylomicrons are converted within a few hours after a meal into chylomicron remnant particles. Chylomicron remnants are rapidly taken up by the liver through interaction of Apo E with the LDL receptor and another receptor on the surface of liver cells called LDL-related receptor protein.[16] Once in the liver, lysosomal enzymes break down the remnant particles to release free fatty acids, free cholesterol, and amino acids. Some cholesterol is converted to bile acids. Both bile acids and free cholesterol are directly excreted into the bile, but not all the excreted cholesterol and bile salts exit the body. As previously described, approximately half of the excreted biliary cholesterol is reabsorbed by the intestine, with the remainder appearing in the stool as fecal neutral steroids. In the case of bile acids, almost all bile acids are reabsorbed and reused by the liver for bile production.

Endogenous Pathway

Most triglycerides in the liver that are packaged into VLDL are derived from the diet after recirculation from adipose tissue. Normally, only a small fraction is synthesized de novo in the liver from dietary carbohydrate. VLDL particles, once secreted into the circulation, undergo a lipolytic process similar to that of chylomicrons (Figure 10.3). VLDL loses core lipids, causing dissociation and transfer of apolipoproteins and phospholipids to other lipoprotein particles like HDL, primarily by the action of LPL. This results in the conversion of VLDL to denser particles, called IDLs or VLDL remnants. IDL persists for short periods of time and receives cholesterol esters from HDL in exchange for triglycerides via a cholesteryl ester transport protein. As with chylomicron remnants, IDL is taken up by the liver via Apo E and the LDL receptor, and the triglycerides in IDL are removed by hepatic triglyceride lipase, located on hepatic endothelial cells, ultimately producing LDL. About half of VLDL is eventually completely converted to LDL, and the remainder is taken up as VLDL remnants by the liver remnant receptors.

LDL particles are the major lipoproteins responsible for the delivery of exogenous cholesterol to peripheral cells due to the efficient uptake of LDL by the LDL receptors.[13] LDL can pass between capillary endothelial cells and bind to LDL receptors on cell membranes that recognize Apo B100. Once

bound to LDL receptors, they are endocytosed by cells and transported to the lysosome, where they are degraded. The triglycerides in LDL are converted by acid lipase into free fatty acids and glycerol, and further metabolized by the cell for energy. They can also be re-esterified and stored in lipid drops for later use. Free cholesterol derived from degraded LDL can be used for membrane biosynthesis, and excess cholesterol is converted by acyl-CoA:cholesterol acyltransferase (ACAT) into cholesteryl esters and stored in intracellular lipid drops. The regulation of cellular cholesterol biosynthesis is, in part, coordinated by the availability of cholesterol delivered by the LDL receptor. Many enzymes in the cholesterol biosynthetic pathway (e.g., HMG-CoA reductase) are down-regulated, along with the LDL receptor, when there is excess cellular cholesterol.

Abnormalities in LDL receptor function result in elevation of LDL in the circulation and lead to hypercholesterolemia and premature atherosclerosis. Patients who are heterozygous for a disease called familial hypercholesterolemia (FH) have only approximately half the normal LDL receptors. FH occurs in approximately 1:500 and results in decreased hepatic uptake of LDL by the liver and increased hepatic cholesterol biosynthesis. The LDL that accumulates in the plasma of these individuals often leads to the development of CHD by mid-adulthood, and even earlier for homozygous FH individuals.

Reverse Cholesterol Transport Pathway

As previously described, one of the major roles of HDL is to maintain the equilibrium of cholesterol in peripheral cells by the reverse cholesterol transport pathway (Figure 10.3). HDL is believed to remove excess cholesterol from cells by multiple pathways. In the aqueous diffusion pathway, HDL acts as a sink for the small amount of cholesterol that can diffuse away from the cells. Although cholesterol is relatively water insoluble, because it is an amphipathic lipid, it is soluble in plasma in micromolar amounts and can spontaneously dissociate from the surface of cell membranes to enter the extracellular fluid. Some free cholesterol will then bind to nascent HDL in the extracellular space, and once bound, it becomes trapped in lipoproteins after conversion to cholesteryl ester by the enzyme LCAT; LCAT resides on nascent HDL and is activated by its cofactor, Apo A1. The nascent HDL is first converted to HDL_3, then to HDL_2 after absorbing additional free cholesterol and Apo C1, C2, C3, E and phospholipids. HDL_2 can then directly deliver cholesterol to the liver. HDL particles can also transfer cholesteryl esters to chylomicrons and VLDL remnants to be transported to the liver. The cholesterol on HDL is transferred to LDL by the cholesteryl ester transfer protein (CETP), allowing for approximately half of the cholesterol on HDL to return to the liver. This connects the forward and reverse cholesterol transport pathways (Figure 10.3). Cholesterol that reaches the liver is then directly excreted into the bile or first converted to a bile acid before excretion.

Another pathway in which HDL mediates the removal of cholesterol from cells involves the ABCA1 transporter. The ABCA1 transporter is a member of the ATP-binding cassette transporter family,[17] which pumps various ligands across the plasma membrane. The exact substrate for the ABCA1 transporter is not known, but it is believed that the transporter modifies the plasma membrane by transferring a lipid that then enables Apo A1 that has dissociated from HDL to bind to the cell membrane. In a detergent-like extraction mechanism, Apo A1 then removes excess cholesterol and phospholipid from the plasma membrane of cells to form a discoidal-shaped HDL particle. The newly formed HDL is then competent to accept additional cholesterol by the aqueous diffusion pathway and is eventually converted into spherical HDL by the action of LCAT (Figure 10.3). Defects in the gene for the ABCA1 transporter lead to Tangier disease, a disorder associated with low HDL and a predisposition to premature CHD.[18]

Lipid and Lipoprotein Population Distributions

Serum lipoprotein concentrations differ between adult men and women, primarily because of differences in sex hormone levels, with women having, on average, higher HDL-C levels and lower total cholesterol and triglyceride levels than men. The difference in total cholesterol, however, disappears after menopause as estrogen decreases. Men and women both show a tendency toward increased total cholesterol, LDL-C, and triglyceride concentrations with increased age. HDL-C concentrations generally remain stable after the onset of puberty and do not drop in women with the onset of menopause. General adult reference ranges are shown in **Table 10.3**.

Circulating levels of total cholesterol, LDL-C, and triglycerides in young children are generally much lower than those seen in adults.[19] In addition, concentrations do not significantly differ between boys and girls. HDL-C levels for both boys and girls are

Table 10.3 Adult Reference Ranges For Lipids

Analyte	Reference Range
Total cholesterol	140–200 mg/dL (3.6–5.2 mmol/L)
HDL-C	40–75 mg/dL (1.0–2.0 mmol/L)
LDL-C	50–130 mg/dL (1.3–3.4 mmol/L)
Triglycerides	60–150 mg/dL (0.7–1.7 mmol/L)

HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol.

comparable to those of adult women. At the onset of puberty, however, HDL-C concentrations in boys fall to adult male levels, a drop of approximately 20%, whereas those of girls do not significantly change. It is the lower concentration of HDL-C in men, combined with their higher LDL-C and triglyceride levels, that accounts for much of the increased risk of premature heart disease in men.

The incidence of heart disease is strongly associated with serum cholesterol concentration, particularly LDL-C. Comparisons across different societies and geographic locations show that eating less animal fat and more whole grains, fruits, and vegetables is associated with lower LDL-C and lower rates of heart disease compared with societies that ingest more animal fat and less fruits, vegetables, and whole grains. These differences can be attributed to both genetic and lifestyle factors in the various countries and ethnic groups. The importance of diet was clearly shown in a study that compared the dietary patterns and heart disease rates in Japanese men living in Japan, Hawaii, and California.[20] In this study, as dietary intake became more Westernized with increased consumption of fat and cholesterol, the LDL-C concentrations increased, as did the rates of heart disease. Japanese men living in California were found to have much higher rates of heart disease than Japanese men living in Japan; those in Hawaii were intermediate. Within societies in which diet tends to be more homogeneous, LDL-C levels become somewhat less discriminatory as a risk factor, and HDL-C levels become more important as a negative risk factor.

National Cholesterol Education Program

The National Cholesterol Education Program (NCEP) was formed to alert the American population to the risk factors associated with heart disease. The NCEP has used panels of experts, including Adult Treatment Panels (ATPs), the Children and Adolescents Treatment Panel, and the Laboratory Standardization Panel, to produce various recommendations.[21] In 1988, the first NCEP ATP developed a list of heart disease risk factors. These guidelines were most recently updated by ATP IV in 2013.[22] In general, the ATP guidelines and other similar guidelines are based on the central tenet that lifestyle changes, such as a low-fat diet, increased physical activity, and weight control, should be the first-line treatment for all patients at risk for cardiovascular disease (CVD), but that more aggressive drug therapy should also be considered for the highest risk individuals and/or for those who do not respond to lifestyle changes. According to the more recent guidelines by the American Heart Association (AHA) and the American College of Cardiology (ACC),[22] patients at high risk for CVD (i.e., those with diabetes and preexisting CVD) or those with an LDL-C greater than 190 mg/dL are advised to initiate simultaneous lifestyle changes along with drug therapy, defined as high-dose statin, in order to reduce LDL-C by 50% from baseline.[22] Lower-risk patients between age 40 and 75 should have their 10-year CVD risk calculated and be treated with high- or moderate-dose statin if their risk score is greater than 7.5%, and moderate-dose statin when their risk score is between 5% and 7.5%. For those patients whose 10-year risk is $< 5\%$ or who fall outside the age range for the risk calculator ($<$40 years or $>$75 years), only lifestyle changes would typically be recommended, but statins could be considered if they have other extenuating circumstances, such as a strong family history or are at risk based on another risk marker like high-sensitivity C-reactive protein (hsCRP). Repeat laboratory testing is recommended to monitor compliance and to determine if a minmum $\geq$50% reduction in LDL-C is achieved, but there are no longer any specific treatment target goals for LDL-C as in past guidelines. The current list of risk factors used in calculating the 10-year risk of CVD is shown in **Table 10.4**. It is recommended that all adults (20 years and older) have a fasting lipoprotein profile performed (total cholesterol, LDL-C, and HDL-C and triglycerides) once every 5 years. As described below, the Children and Adolescents Treatment Panel has developed similar guidelines for CVD risk assessment in the pediatric population.[19]

Dyslipidemia and Children

It is well recognized and documented that the atherosclerotic process begins in childhood, with dyslipidemia as one of the main risk factors for CVD. An analysis of National Health and Nutrition

Table 10.4 Coronary Heart Disease Risk Factors Determined by the NCEP Adult Treatment Panels

Positive Risk Factors
■ Age: ≥45 y for men; ≥55 y or premature menopause for women
■ Family history of premature CHD
■ Current cigarette smoking
■ Hypertension (blood pressure ≥140/90 mm Hg or taking antihypertensive medication)
■ LDL-C concentration ≥160 mg/dL (≥4.1 mmol/L), with ≤1 risk factor
■ LDL-C concentration ≥130 mg/dL (3.4 mmol/L), with ≥2 risk factors
■ LDL-C concentration ≥100 mg/dL (2.6 mmol/L), with CHD or risk equivalent
■ HDL-C concentration < 40 mg/dL (<1.0 mmol/L)
■ Diabetes mellitus = CHD risk equivalent
■ Metabolic syndrome (multiple metabolic risk factors)
Negative Risk Factors
■ HDL-C concentration ≥60 mg/dL (≥1.6 mmol/L)
■ LDL-C concentration <100 mg/dL (<2.6 mmol/L)

CHD, coronary heart disease; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol.

Examination Survey data collected between 1998 and 2010 estimated that 24.6% of US children aged 9 to 11 had either abnormally low HDL cholesterol or elevated non-HDL cholesterol.[23] Elevated total cholesterol, LDL-C, VLDL-C, and decreased HDL-C concentrations are significantly associated with early atherosclerotic lesions in adolescent and young adults.[24] Thus, early identification and treatment of children with hyperlipidemia is likely to delay the atherosclerotic process and reduce CVD risk in adulthood. Guidelines for the assessment and treatment of CVD risk in children and adolescents have been published by the National Heart, Lung, and Blood Institute (NHLBI).[19] The NCEP published the first set of pediatric guidelines for dyslipidemia in 1992; however, some considered the reliance of these guidelines on family history to be too restrictive. For example, the Coronary Artery Risk Detection in Appalachian Communities (CARDIAC) Project, which utilized universal screening, found that 36% of children with severe dyslipidemia would have been overlooked.[25]

National Heart, Lung, and Blood Institute

With childhood obesity on the rise and almost half of obese children having at least one abnormal lipid level, the NHLBI Expert Panel in 2011 published a comprehensive updated evidence-based guideline for the identification, management, and treatment of all major CV risk factors (including diabetes, hypertension, cigarette smoking, and obesity) in pediatric patients.[19] Universal lipid screening of all children from 9 to 11 years and again at 17 to 21 years with a nonfasting, non-HDL-C levels and targeted lipid screening are both supported by the NHLBI guidelines and endorsed by the American Academy of Pediatrics.[19] The 9 to 11 age range is important because that is when the atherosclerotic process appears to accelerate but before the decline in LDL-C seen with puberty. The NHLBI universal screening recommendations for lipid abnormalities in children and adolescents are controversial, and some health care providers have resisted.[26] Acceptable lipid levels for nonfasting blood specimens for youth and young adults are shown in **Table 10.5**.[19] If a nonfasting lipid profile reveals a non-HDL-C greater than 145 mg/dL, then two separate fasting lipid profiles should be done at least 2 weeks apart but within 3 months of each other and the results averaged to confirm the abnormality. If an initial fasting lipid profile reveals an LDL-C ≥130 mg/dL, then another fasting lipid profile should be obtained at least 2 weeks later but within 3 months and the results averaged. The NHLBI recommends that children identified with dyslipidemia should begin with diet and lifestyle changes along with patient education and progress to more stringent diet and lifestyle changes; however, even high-risk children are not being screened in accordance with recommendations.[19,27,28] If dyslipidemia persists, then pharmacological treatment and/or referral to a lipid specialist may be recommended for children aged ≥10 years.[19,28] Likewise, children with a genetic dyslipidemia, such as FH, should also be referred to a lipid specialist for more aggressive type therapy.[19,28] Future studies are needed, however, to firmly establish if treatment of hyperlipidemia early in life prevents CHD in adulthood.

The NCEP Laboratory Standardization Panel and its successor, the Lipoprotein Measurement Working Group,[29] set laboratory guidelines for acceptable precision and accuracy when measuring total

Table 10.5 Lipid Cut-Off Points (mg/dL)

	For Youth <17 y			For Young Adults 17–21 y		
Category	**Acceptable**	**Borderline***	**High****	**Acceptable[a]**	**Borderline[a]**	**High[a]**
TC	<170	170–199	≥200	>190	190–225	≥200
LDL-C	<110	110–129	≥130	<120	120–160	≥130
Non-HDL-C	<120	120–144	≥145	<150	150–190	≥145
Triglycerides						
0–9 y	<75	75–99	≥100			
10–19 y	<90	90–129	≥130	<115	115–150	≥150
Category	**Acceptable**	**Borderline**	**Low*****	**Acceptable**	**Borderline**	**Low**
HDL-C	>45	45–40	<40	>45	45–40	<40

Values represent approximately the *75th and **95th percentiles and approximately the ***10th percentile for low HDL-C.
[a]The cut points for TC, HDL-C, and non-HDL-C represent the 95th percentile for 20- to 24-year-old subjects; acceptable values are at the <75th percentile. For HDL-C, low is <25th percentile and acceptable values are >50th percentile.
HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol.

Data from Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents, National Heart, Lung, and Blood Institute. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics*. 2011;128(suppl 5):S213–S256.

cholesterol, triglycerides, and lipoprotein cholesterol (HDL-C and LDL-C) (**Table 10.6**).

The best way to reduce the prevalence of heart disease is through prevention. Lifestyle changes that include aerobic exercise, a healthy diet, normal blood pressure, and not smoking are key to avoiding CHD and stroke. Lipoprotein profile measurements provide a method of identifying individuals who may have levels that put them at risk so that they can receive appropriate treatment. Treatment of other diseases that may affect lipoproteins, such as diabetes mellitus, hypothyroidism, and renal disease, is also important. A diet low in fat, with a caloric intake adjusted to meet and maintain ideal body weight, along with regular exercise, can significantly reduce the risk of heart disease, stroke, diabetes, and cancer.[28] Additionally, saturated fat is more atherogenic than unsaturated fat.[30] The American Heart Association has recommended dietary guidelines for the intake of fat and cholesterol for most adult Americans (**Table 10.7**).

Table 10.6 NCEP Analytic Performance Goals

	Precision	Bias	Total Error
Total cholesterol	3% CV	±3%	±8.9%
HDL-C			
≥42 mg/dL	4% CV	±5%	±12.8%
<42 mg/dL	SD	<1.7 mg/dL	
LDL-C	4% CV	±4%	±11.8%
Triglycerides	5% CV	±5%	±14.8%

CV, coefficient of variation; HDL-C, high-density lipoprotein cholesterol; SD, standard deviation; LDL-C, low-density lipoprotein cholesterol.

Table 10.7 Composition of the Therapeutic Lifestyle Changes Diet Recommended by the NCEP Adult Treatment Panel III (Compared with the Average American Diet)

Dietary Nutrient	TLC Diet	Average American Diet
Total fat	25%–35%	36% (% total calories)
Saturated	<7%	15%
Monounsaturated	≤20%	15%
Polyunsaturated	≤10%	6%
Cholesterol		>400 mg/d
Carbohydrate	50%–60%	
Fiber	20–30 g/d	
Protein	<15%	

TLC, therapeutic lifestyle changes.

CASE STUDY 10.1, PART 2

Recall Maggie. Her pediatrician does an extensive family history, which reveals that Maggie's mother has high cholesterol and her maternal grandparents were diagnosed with coronary artery disease at an early age.

1. According to NCEP pediatric guidelines, which lipid tests should be performed at this visit?
2. Maggie has a fasting lipid profile performed, and the LDL-C is 135 mg/dL. What is the next step?

© Monkey Business Images/Shutterstock.

Diagnosis and Treatment of Lipid Disorders

Diseases associated with abnormal lipid concentrations are referred to as *dyslipidemias*. They can be caused directly by genetic abnormalities or through environmental/lifestyle imbalances, or they can develop secondarily, as a consequence of other diseases. Dyslipidemias are generally defined by the clinical characteristics of patients and the results of laboratory tests and are not necessarily defined by the specific genetic defect associated with the abnormality. Many, but not all, dyslipidemias, regardless of etiology, are associated with CHD or arteriosclerosis.

Arteriosclerosis

In the United States and many other developed countries, **arteriosclerosis** (the thickening of blood vessels due to buildup of cholesterol plaques) and atherosclerosis (a type of arteriosclerosis, in this type, the large arteries are hardened and narrowed) are the most important underlying causes of death and disability. The mortality rate has leveled off and started to decrease in the United States in the past three decades, partly as a result of advances in diagnosis and treatment but also because of lifestyle changes in the American population. This increased awareness of the importance of diet and exercise in preventing CHD has resulted in an overall decrease in the average serum cholesterol concentration and in a lower prevalence of heart disease; however, it still exceeds all other causes of death. Although as many women as men now develop arteriosclerosis, they typically develop it 10 years later than men; unfortunately, it sometimes goes unrecognized due to a of lack of awareness, and the symptoms of myocardial infarction are subtler and often different in women.

The relationship between heart disease and dyslipidemias stems from the deposition of lipids, mainly in the form of esterified cholesterol, in artery walls. This lipid deposition first results in fatty streaks, which are thin streaks of excess fat in macrophages in the subendothelial space. Studies have shown that fatty streaks occur early in life.[31] Fatty streaks can develop over time into plaques that contain an increased number of smooth muscle cells, extracellular lipid, calcification, and fibrous tissue, which can partially block or occlude blood flow. Also, established plaque for unknown reasons can become vulnerable to rupture or erosion, triggering a thrombosis that can block circulation. When plaque develops in arteries of the arms or legs, it is called peripheral vascular disease (PVD); when it develops in the heart, it is referred to as coronary artery disease (CAD); and, when it develops in the vessels of the brain, it is called cerebrovascular disease. CAD is associated with angina and myocardial infarction, and cerebrovascular disease is associated with stroke. Many genetic and acquired dyslipidemias may also lead to lipid deposits in the liver and kidney, resulting in impaired function of these vital organs. Lipid deposits in skin form nodules called **xanthomas**, which are often a clue to the presence of an underlying genetic abnormality.

Plaque formation involves repeated cycles of cell injury, followed by infiltration and cell proliferation to repair the site. LDL is believed to play a central role in initiating and promoting plaque formation. It is deposited into the subendothelial space, where it

is taken up by various cells, including macrophages. This alters the gene and protein expression pattern of these cells and can promote an inflammatory response, particularly when LDL becomes oxidized.[32] Injury signals from the evolving plaque trigger the expression of adhesion proteins on endothelial cells and the production of soluble chemotactic proteins from resident macrophages, which promotes the attachment and infiltration of additional macrophages, lymphocytes, and platelets to the plaque. Continual injury and repair lead to additional narrowing of the vessel opening, or lumen, causing the blood to circulate in a nonlaminar manner under greater and greater pressure, which further aggravates plaque formation. The final event leading to complete occlusion of blood flow occurs when there is a hemorrhage into the plaque, which results in the formation of a thrombus that blocks blood flow and precipitates a myocardial infarction.

In some individuals, high levels of blood cholesterol or triglycerides are caused by genetic abnormalities in which either too much is synthesized or too little is removed. High levels of cholesterol and/or triglycerides in most people, however, are a result of increased consumption of foods rich in fat and cholesterol, smoking, and lack of exercise or a result of other disorders or disease states that affect lipid metabolism, such as diabetes, hypertension, hypothyroidism, obesity, liver and kidney diseases, and alcoholism. Low levels of HDL-C are also associated with increased risk of heart disease, but there are currently limited ways to pharmacologically raise HDL-C levels; it is unclear to what degree low HDL is causal in the pathogenesis of atherosclerosis.

Laboratory analyses are an important adjunct to managing patients with dyslipidemia because accurate measurement of total cholesterol, HDL-C, LDL-C, and triglyceride levels is needed to determine the most appropriate diet or diet and drug therapy. According to the ATP III guidelines (**Table 10.8**), individuals on a low-fat diet who continue to have LDL-C levels of 190 mg/dL (4.9 mmol/L) or higher on repeated measurement will likely benefit from drug intervention. If they have two or more CAD risk factors and continue to have LDL-C levels of 160 mg/dL (4.1 mmol/L) or higher, they also would benefit from drug therapy. And if they have already been previously diagnosed with heart disease, drug therapy should be considered when the LDL-C level is 130 mg/dL (3.4 mmol/L) or higher. The average of at least two measurements, taken 1 to 8 weeks apart, should be used to determine the best treatment approach.[21]

The most effective class of drugs for managing patients with dyslipidemia are the HMG-CoA reductase drug inhibitors, such as lovastatin, atorvastatin, and rosuvastatin. These drugs, commonly known as statins, block intracellular cholesterol synthesis by inhibiting HMG-CoA reductase, a rate-limiting enzyme in cholesterol biosynthesis. The reduced level of cholesterol in hepatocytes increases the expression of the LDL receptor, which removes LDL from the circulation, thus reducing the deposition of LDL into vessels and the formation of plaques. All statins are probably identical in their main mechanism but differ in dose–response curves. The major safety issues with statins are myositis and hepatotoxic effects; however, patient monitoring in clinical trials has shown that fewer than 2% of patients have sustained increases in liver enzymes. Routine monitoring of serum transaminases and creatine kinase is recommended for patients complaining of muscle-related symptoms. These drugs typically reduce LDL-C by as much as 20% to 40%, raise HDL-C by 5% to 10%, and can lower triglyceride by 7% to 43%, depending on the initial triglyceride level.[33] Statins are generally well tolerated, with maximum effects observed after 4 to 6 weeks.[33] Clinical trials have shown that statins are beneficial for both the primary and secondary prevention of CHD.[21]

Niacin or nicotinic acid at high doses is also a potent drug for reducing LDL-C levels and is the most widely used drug at this time for significantly raising HDL-C levels.[34] Niacin can also lower triglyceride levels; however, it causes unpleasant side effects in some patients.

Ezetimibe is a drug that inhibits the absorption of cholesterol by inhibiting the Niemann-Pick C1-like 1 (NPC1-L1) transporter in the intestine without impacting the absorption of fat-soluble nutrients. Ezetimibe therapy has been shown to decrease LDL-C concentrations by approximately 20%.[35] Ezetimibe in conjunction with statins results in greater reductions in LDL-C concentration than with statins alone. As with the use of statin drugs, liver function tests should be performed on individuals taking ezetimibe.[35] Fish oil products that contain omega-3 fatty acids, eicosapentaenoic and docosahexaenoic acids, have been shown to reduce the risk of cardiovascular events.[36] Fish oils also

Table 10.8 Treatment Guidelines Established by the NCEP Adult Treatment Panels (Initial Testing Should Consist of Fasting for ⩾12 Hours)

Risk Category and Action
■ TC, <200 mg/dL (<5.2 mmol/L); TG, <150 mg/dL (<1.7 mmol/L); LDL-C, <130 mg/dL (<3.4 mmol/L); HDL-C, ≥40 mg/dL (≥1.0 mmol/L) ■ Repeat within 5 y ■ Provide risk reduction information ■ TC, 200–239 mg/dL (5.2–6.2 mmol/L); TG, 150–199 mg/dL (1.7–2.2 mmol/L); LDL-C, 130–159 mg/dL (3.4–4.1 mmol/L); HDL-C, ≥40 mg/dL (≥1.0 mmol/L); and 0–1 risk factors ■ Provide therapeutic lifestyle changes diet and physical activity information and reevaluate in 1 y ■ Provide risk reduction information ■ TC, ≥200 mg/dL (≥5.2 mmol/L); TG, ≥200 mg/dL (≥2.2 mmol/L); LDL-C, 130–159 mg/dL (3.4–4.1 mmol/L); HDL-C, <40 mg/dL (<1.0 mmol/L); and ≥2 risk factors ■ Do clinical evaluation, including family history ■ Start dietary therapy (see below) ■ TC, ≥240 mg/dL (6.2 mmol/L) ■ Perform lipoprotein (LDL-C) analysis (see below)

Treatment Decisions

Risk Category	Action Level	Goal
Dietary Therapy		
No CHD; 0–1 risk factors	≥160 mg/dL (4.1 mmol/L)	<160 mg/dL (4.1 mmol/L)
No CHD; ≥2 risk factors (10-y risk, ≥ 20%)	≥130 mg/dL (3.4 mmol/L)	<130 mg/dL (3.4 mmol/L)
CHD; CHD risk equivalent (10-y risk, >20%)	≥100 mg/dL (2.6 mmol/L)	<100 mg/dL (2.6 mmol/L)
Drug Therapy		
No CHD; 0–1 risk factors	≥190 mg/dL (4.9 mmol/L)	<160 mg/dL (4.1 mmol/L)
No CHD; 2 risk factors (10-y risk, ≤10%)	≥160 mg/dL (4.1 mmol/L)	<130 mg/dL (3.4 mmol/L)
No CHD; ≥2 risk factors (10-y risk, 10%–20%)	≥130 mg/dL (3.4 mmol/L)	<100 mg/dL (3.4 mmol/L)
CHD; CHD risk equivalent	≥130 mg/dL (3.4 mmol/L)	<100 mg/dL (2.6 mmol/L)

CHD, coronary heart disease; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglycerides.

enhance fatty acid β-oxidation by stimulating peroxisome proliferator–activated receptors.

Hyperlipoproteinemia

Disease states associated with abnormal serum lipids are generally caused by malfunctions in the synthesis, transport, or catabolism of lipoproteins. Dyslipidemias can be subdivided into two major categories: *hyperlipoproteinemias*, which are diseases associated with elevated lipoprotein levels, and *hypolipoproteinemias*, which are associated with decreased lipoprotein levels. The hyperlipoproteinemias can be subdivided into hypercholesterolemia, hypertriglyceridemia, and combined hyperlipidemia, with elevations of both cholesterol and triglycerides.

Hypercholesterolemia

Studies continue to show a strong correlation with hypercholesterolemia and coronary heart disease mortality. One form of the disease, which is associated with genetic abnormalities that predispose affected individuals to elevated cholesterol levels, is called familial hypercholesterolemia (FH). Homozygotes for FH fortunately are rare (1:1 million in the population) and have total cholesterol concentrations ≥ 500 mg/dL (≥ 13 mmol/L). These patients can have their first heart attack when still in their teenage years.[37] Heterozygotes for the disease are seen much more frequently (1:310 in the population), because it is an autosomal codominant disorder; a defect in just one of the two copies of the LDL receptor can

CASE STUDY 10.1, PART 3

Recall Maggie. Ten years later, Maggie is enrolled in graduate school and has little time for exercise. She still enjoys baking and has gained weight from eating her delicious cakes. Maggie has missed her last few annual physical exams; however, she participated in a health fair and discovered that her glucose is elevated. She schedules a physical and has a fasting lipid profile as well as a complete metabolic panel performed. Her abnormal results are shown below:

© Monkey Business Images/Shutterstock.

Laboratory Results		
Analyte	**Patient Value**	**Reference Range**
Total Cholesterol	295	140–200 mg/dL
HDL-C	35	40–75 mg/dL
LDL-C	calculate	50–130 mg/dL
Triglycerides	200	60–150 mg/dL
Glucose	140	75–100 mg/dL

3. What is Maggie's calculated LDL-C?
4. What other laboratory tests should be performed?

adversely affect lipid levels. Heterozygotes tend to have total cholesterol concentrations in the range of 200 to 550 mg/dL (5 to 14 mmol/L) and, if not treated, become symptomatic for heart disease in before age 55.[38] Approximately 5% of patients younger than age 50 with CAD are FH heterozygotes. Other symptoms associated with FH include tendinous and tuberous xanthomas, which are cholesterol deposits in tendons and under the skin, respectively, and arcus senilis, which are cholesterol deposits in the cornea.[37] In both homozygotes and heterozygotes, the cholesterol elevation is primarily associated with an increase in LDL-C. These individuals synthesize intracellular cholesterol normally but lack, or are deficient in, active LDL receptors. There are several classes of defects in the LDL receptor gene that are associated with FH. Consequently, LDL builds up in the circulation, because there are insufficient receptors to bind the LDL and transfer the cholesterol into the cells. Cells that require cholesterol for use in cell membrane and hormone production, however, synthesize cholesterol intracellularly at an increased rate to compensate for the lack of cholesterol from the receptor-mediated mechanism.

In FH heterozygotes and other forms of hypercholesterolemia, reduction in the rate of internal cholesterol synthesis by inhibition of HMG-CoA reductase with statin drugs stimulates the production of additional LDL receptors, particularly in the liver, which removes LDL from the circulation. FH homozygotes, however, do not usually fully benefit as much from this type of therapy because they do not have enough functional receptors to stimulate. Homozygotes can be treated by a technique called LDL apheresis, a method similar to dialysis treatment, in which blood is periodically drawn from the patient, processed to remove LDL, and returned to the patient.[38]

Most individuals with elevated LDL-C levels do not have FH but are still at increased risk for premature CHD[25] and should be maintained on a low-fat diet and receive statin treatment when necessary. Regular physical activity should also be incorporated, with drug therapy.

PCSK9

The principal route of clearance of circulating LDL-C from the blood is hepatocyte endocytosis, a process mediated by binding of LDL-C to LDL-R on the hepatocyte cell membrane.[39] Proprotein convertase subtilisin/kexin type 9 (PCSK9) is a 72-kDa serine protease expressed in the liver, kidney, and intestine that plays an important role in LDL metabolism. PCSK9 binds to the LDL-R on the cell surface and is subsequently escorted to lysosomes for degradation of the LDL-R. The binding of PCSK9 thus prevents the receptor from recycling to the plasma membrane, which can occur up to 150 times before it is normally

inactivated or degraded. Thus, because of its role as major inhibitor of LDL-R, PCSK9 gene inactivation is an attractive target for lipid-lowering therapy.[40]

Hypertriglyceridemia

The NCEP ATP III has identified borderline high triglycerides as levels of 150 to 200 mg/dL (1.7 to 2.3 mmol/L), high as 200 to 500 mg/dL (2.3 to 5.6 mmol/L), and very high as greater than 500 mg/dL (>5.6 mmol/L).[21] Hypertriglyceridemia can be a consequence of genetic abnormalities, called FH, or the result of secondary causes, such as hormonal abnormalities associated with the pancreas, adrenal glands, and pituitary, or of diabetes mellitus or nephrosis. Diabetes mellitus leads to increased shunting of glucose into the pentose pathway, causing increased fatty acid synthesis. Renal failure depresses the removal of large-molecular-weight constituents like triglycerides, causing increased serum levels. Hypertriglyceridemia is generally a result of an imbalance between synthesis and clearance of VLDL in the circulation. Hypertriglyceridemia is now generally considered an important and potentially treatable risk factor for CHD and ischemic stroke.[41]

Triglycerides are influenced by a number of hormones, such as insulin, glucagon, pituitary growth hormone, adrenocorticotropic hormone (ACTH), thyrotropin, and epinephrine and norepinephrine. Epinephrine and norepinephrine influence serum triglyceride levels by triggering production of hormone-sensitive lipase, which is found in adipose tissue.[42] Other body processes that trigger hormone-sensitive lipase activity are cell growth (growth hormone), adrenal stimulation (ACTH), thyroid stimulation (thyrotropin), and fasting (glucagon). Each process, through its action on hormone-sensitive lipase, results in an increase in serum triglyceride values.

Severe hypertriglyceridemia (>880 mg/dL [>10 mmol/L]) is a potentially life-threatening abnormality because it can cause acute and recurrent pancreatitis.[43] In addition, it is associated with atherosclerotic CVD when smaller triglyceride-rich lipoproteins are present. It is, therefore, imperative that these patients be diagnosed and treated with low-fat diets and triglyceride-lowering medication. Severe hypertriglyceridemia is generally caused by a deficiency of LPL, Apo C2, or one of the other genes involved in triglyceride metabolism.[44] Normally, LPL hydrolyzes triglycerides carried in chylomicrons and VLDL to provide cells with free fatty acids for energy from exogenous and endogenous triglyceride sources. A deficiency in LPL or Apo C2 activity keeps chylomicrons from being cleared, and serum triglycerides remain extremely elevated, even when the patient has fasted for longer than 12 to 14 hours.

Treatment of hypertriglyceridemia consists of dietary modifications, fish oil, and/or triglyceride-lowering drugs (primarily fibric acid derivatives) in cases of severe hypertriglyceridemia or when accompanied with low HDL-C.[44] It is possible that certain subspecies of chylomicrons and VLDL are atherogenic. Remnants of chylomicrons and VLDL represent subspecies that have been partially hydrolyzed by lipases and are thought to be potentially atherogenic.

Combined Hyperlipidemia

Combined hyperlipidemia is generally defined as the presence of elevated levels of serum total cholesterol and triglycerides. Individuals presenting with this syndrome are considered at increased risk for CHD. In one genetic form of this condition, called familial combined hyperlipidemia (FCH), individuals from an affected family may only have elevated cholesterol, whereas others only have elevated triglycerides, and yet others, elevations of both. FCH is due in part to excessive hepatic synthesis of apoprotein B, leading to increased VLDL secretion and increased production of LDL from VLDL. These patients may have eruptive xanthomas and are at high risk for developing CHD. Another rare genetic form of combined hyperlipidemia is called familial dysbetalipoproteinemia or hyperlipidemia type III. The disease results from an accumulation of cholesterol-rich VLDL and chylomicron remnants as a result of defective catabolism of those particles. Individuals with hyperlipidemia type III will frequently have total cholesterol values of 200 to 300 mg/dL (5 to 8 mmol/L) and triglycerides of 300 to 600 mg/dL (3 to 7 mmol/L). This disorder is associated with an increased risk of PVD and coronary disease.[45] Patients often have palmar xanthomas and tuberoeruptive xanthomas.[46] To distinguish them from other forms of combined hyperlipidemias, it is first necessary to isolate the VLDL fraction with ultracentrifugation. A ratio derived from the cholesterol concentration in VLDL to total serum triglycerides will be greater than 0.30 in the presence of hyperlipidemia type III. If the VLDL fraction is analyzed by agarose electrophoresis, the particles will migrate in a broad β region rather than in the normal pre-β region. Definitive diagnosis requires a determination of Apo E isoforms by isoelectric focusing or DNA typing, resulting in either Apo E2/2 homozygosity or, rarely, Apo E mutation or

CASE STUDY 10.1, PART 4

Recall Maggie. She is now in her late 40s and is a working mother with a stressful career. Although she tried to watch her carbohydrate intake and get more exercise, she has developed type 2 diabetes mellitus. Her most recent fasting blood work shows a glucose of 205 mg/dL, triglycerides of 350 mg/dL, and a total cholesterol of 295 mg/dL. In addition, Maggie has noticed yellow nodules on her skin.

5. What is a possible secondary cause of Maggie's high triglycerides?
6. Should Maggie be concerned about her children's lipid profiles?
7. What additional blood work should be performed on Maggie?
8. What are the yellow nodules on Maggie's skin?

deficiency. As with other dyslipidemias, these patients can be treated with a combination of drugs, such as niacin, gemfibrozil, HMG-CoA reductase inhibitors, and low-fat diets. Because of the cholesterol-enriched composition of these particles, use of the Friedewald equation[47] to calculate LDL-C levels will result in an underestimation of VLDL-C and, therefore, an overestimation of LDL-C, compared with β-quantification ultracentrifugation procedure.[48]

Lp(a) Elevation

Elevations in the serum concentration of Lp(a), especially in conjunction with elevations of LDL, increase the risk of CHD and CVD.[49] Higher Lp(a) levels have been observed in studies more frequently in patients with CHD than in normal control subjects. Lp(a) are variants of LDL with an extra apolipoprotein, called Apo(a); the size and serum concentrations of Lp(a) are largely genetically determined. Because Apo(a) has a high degree of homology with the coagulation factor plasminogen,[50] it has been proposed that it competes with plasminogen for fibrin binding sites, thus increasing plaque formation.[50] Most LDL-lowering drugs have no effect on Lp(a) concentration, even when LDL-C is significantly lowered. The two drugs shown to have some effects are niacin and estrogen replacement in postmenopausal women, but the value of specifically treating patients for high Lp(a) is still not clear, and instead, lowering LDL-C remains the primary the first goal.

Non-HDL Cholesterol

Although LDL-C is widely recognized as an established risk marker for CVD, many studies have demonstrated that LDL-C alone does not provide a sufficient measure of atherogenesis, especially in hypertriglyceridemic patients. To address this issue, the calculation of non-HDL-C has been employed. Non-HDL-C reflects total cholesterol minus HDL-C and encompasses all cholesterol present in potentially atherogenic, Apo B–containing lipoproteins [LDL, VLDL, IDL, and Lp(a)]. Unlike LDL-C, which can be incorrectly calculated using the Friedewald equation in the presence of postprandial hypertriglyceridemia, non-HDL-C is reliable when measured in the nonfasting state.[51] On average, non-HDL-C levels are approximately 30 mg/dL higher than LDL-C levels.[25] Recent studies have shown that elevated levels of non-HDL-C are associated with increased CVD risk, even if the LDL-C levels are normal.[52] In clinical studies, non-HDL-C has been found to be an independent predictor of CVD and for diabetes patients; it may be a stronger predictor than LDL-C and triglycerides.[52]

Hypobetalipoproteinemia

Hypobetalipoproteinemia is associated with isolated low levels of LDL-C as a result of a defect in the Apo B gene, but because it is not generally associated with CHD, it is not discussed further here. Abetalipoproteinemia, which is due to a defect in the microsomal transfer protein used in the synthesis and secretion of VLDL, can also present with low LDL and Apo B–like hypobetalipoproteinemia. It is an autosomal recessive disorder and like hypobetalipoproteinemia patients, they are not at an increased risk of cardiovascular disease but can develop several neurologic and ophthalmologic problems from fat-soluble vitamin deficiencies.

Hypoalphalipoproteinemia

Hypoalphalipoproteinemia indicates an isolated decrease in circulating HDL, typically defined as an HDL-C concentration less than 40 mg/dL

(1.0 mmol/L), without the presence of hypertriglyceridemia. The "alpha" in the term denotes the region in which HDL migrates on agarose electrophoresis. There are several defects, often genetically determined, which are associated with hypoalphalipoproteinemia, such as LCAT, Apo A1, and ABCA1 transporter gene mutations.[53] Most of these defects are associated with a somewhat increased risk of premature CHD, but the association is not as strong as seen with genetic disorders that result in increased LDL-C. An extreme form of hypoalphalipoproteinemia, Tangier disease, is associated with HDL-C concentrations as low as 1 to 2 mg/dL (0.03 to 0.05 mmol/L) in homozygotes, accompanied by total cholesterol concentrations of 50 to 80 mg/dL (1.3 to 2.1 mmol/L).

Treatment options of individuals with isolated decreases of HDL-C are limited and the clinical utility is not clear. Niacin is somewhat effective in raising HDL-C but can have adverse effects, such as flushing or even hepatotoxicity, although newer, timed-release preparations may ameliorate those effects.[54] Lifestyle modifications, which can raise HDL-C, and treatment of any coexisting disorders that increase CHD risk are likely to be beneficial in these patients.

Acute, transitory hypoalphalipoproteinemia can also be seen in cases of severe physiologic stress, such as acute infections (primarily viral), other acute illnesses, and surgical procedures. HDL-C concentrations, as well as total cholesterol, can be significantly reduced under these conditions but will return to normal levels as recovery proceeds. For this reason, lipoprotein concentrations drawn during hospitalization or with a known disease state should be reassessed when the patient is in a healthy, nonhospitalized state before intervention is considered.

Lipid and Lipoprotein Analyses

Lipid Measurement

Lipids and lipoproteins are important indicators of CHD risk, which is a major reason for their measurement in research, as well as in clinical practice. Decision cut points used to characterize CHD risk have been developed by various guidelines based on consideration of population distributions from large epidemiologic studies, intervention studies that demonstrated the efficacy of treatment regimens, and cost-effectiveness.[21] To improve the reliability of the analytic measurements, standardization programs have been implemented for the research laboratories performing the population and intervention studies, which helped to make results comparable among laboratories and over time.[55] More recently, standardization programs have been extended to diagnostic manufacturers and routine clinical laboratories to facilitate reliable classification of patients using the national decision cut points. Thus, accuracy and standardization of results are especially important with the lipid and lipoprotein analytes.

Cholesterol Measurement

Serum or plasma specimens collected in blood collection tubes from patients who have fasted for at least 12 hours are usually preferred for total cholesterol testing but do not usually significantly change after a meal. If analysis is delayed, the serum/plasma specimen can be refrigerated at 4°C for several days. The lipid workup traditionally has begun with measurement of total serum cholesterol. The reference

CASE STUDY 10.1, PART 5

Recall Maggie. She is now 56 years old and continues to struggle with elevated cholesterol and diabetes. She refuses to increase her dose of statin drugs or try new medications. Her lab values have not changed significantly over the years. In addition, she has developed hypertension.

9. What disease is Maggie at risk of developing?
10. How many known risk CHD risk factors does Maggie have?

© Monkey Business Images/Shutterstock.

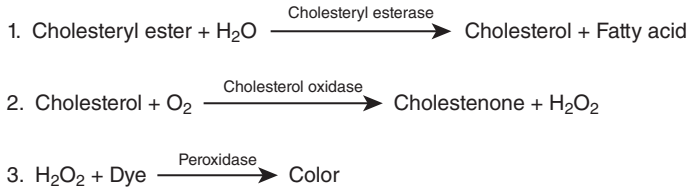


Figure 10.4 Enzymatic assay sequence—cholesterol.

method is a gas chromatography-mass spectrometry (GC-MS) method that specifically measures cholesterol and does not detect related sterols. This method shows good agreement with the gold standard method developed and applied at the U.S. National Institute of Standards and Technology, the so-called Definitive Method, using isotope dilution mass spectrometry (IDMS).[56]

The lipoproteins HDL and LDL are generally quantified based on their cholesterol content. Thus, the standard lipid panel, including measurements of total cholesterol, LDL-C, and HDL-C, together with triglycerides, can be completed routinely using chemistry analyzers.

Although several enzymatic reaction sequences have been described, one sequence (**Figure 10.4**) is most common for measuring cholesterol. The enzyme cholesteryl ester hydrolase cleaves the fatty acid residue from cholesteryl esters, which comprise about two-thirds of circulating cholesterol, converting them to unesterified or free cholesterol. The free cholesterol is reacted by the second enzyme, cholesterol oxidase, producing hydrogen peroxide (H_2O_2), a substrate for a common enzymatic color reaction using horseradish peroxidase to couple two colorless chemicals into a colored compound. The intensity of the resulting color, proportional to the amount of cholesterol, can be measured by a spectrophotometer, usually at a wavelength around 500 nm. Enzymes and reagents have improved so that most appropriately calibrated commercial reagents can be expected to give reliable results. This reaction sequence is generally used on serum without an extraction step but can be subject to interference. For example, vitamin C and bilirubin are reducing agents that could interfere with the peroxidase-catalyzed color reaction, unless appropriate additional enzyme systems are added to eliminate the interference.

Triglyceride Measurement

Measurement of serum triglycerides in conjunction with cholesterol is useful in detecting certain genetic and other types of metabolic disorders, as well as in characterizing the risk of CVD. The triglyceride value is also commonly used in the estimation of LDL-C by the Friedewald equation and is usually done on a fasting sample because triglycerides can significantly increase after a meal. Several enzymatic reaction sequences are available for triglyceride measurement, all including lipases to cleave fatty acids from the glycerol backbone.[57] The freed glycerol participates in any one of several enzymatic sequences. A second sequence (**Figure 10.5**), involves glycerol kinase and glycerophosphate oxidase, coupled to the same peroxidase color reaction described for cholesterol.

The enzymatic triglyceride reaction sequences also react with any endogenous free glycerol, which is universally present in serum and can be a significant source of interference.[57] In most specimens, the endogenous free glycerol contributes a 10 to 20 mg/dL (0.11 to 0.023 mmol/L) overestimation of triglycerides. About 20% of specimens will have higher glycerol, with levels increased in certain

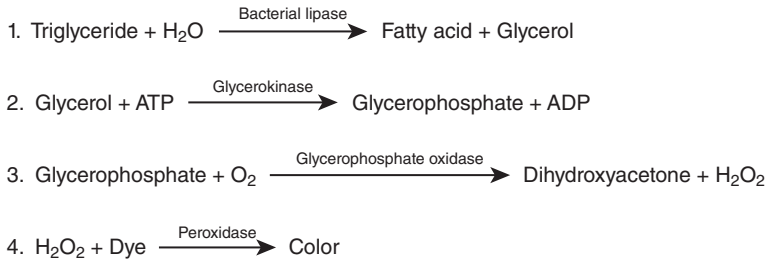


Figure 10.5 Enzymatic assay sequence—triglycerides.

conditions, such as diabetes and liver disease, or from glycerol-containing medications. Reagents are available that correct for endogenous free glycerol and are used by many research laboratories, but such methods are less efficient and therefore uncommon in clinical laboratories.

Until recently, the triglyceride reference method involved alkaline hydrolysis, solvent extraction, and a color reaction with chromotropic acid, an assay that is tedious and poorly characterized and was largely only done by the lipid standardization laboratory at the Centers for Disease Control and Prevention (CDC). Like the reference method for cholesterol, the CDC has recently switched to a GC-MS method that involves the hydrolysis of fatty acids on triglycerides and the measurement of glycerol. However, it must be noted that accuracy in triglyceride measurements for clinical purposes might be considered less relevant than that for cholesterol because the physiologic variation is so large, with the coefficient of variation (CV) in the range of 25% to 30%, making the contribution of analytic variation relatively insignificant.

Lipoprotein Methods

Various methods have been used for the separation and quantitation of serum lipoproteins, taking advantage of physical properties, such as density, size, charge, and apolipoprotein content. The range in density observed among the lipoprotein classes is a function of the relative lipid and protein content and enables fractionation by density, using ultracentrifugation. Electrophoretic separations take advantage of differences in charge and size. Chemical precipitation methods, once common in clinical laboratories and now primarily used in research laboratories, depend on particle size, charge, and differences in the apolipoprotein content. Antibodies specific to apolipoproteins can be used to bind and separate lipoprotein classes. Chromatographic methods take advantage of size differences in molecular sieving methods or composition in affinity methods, using, for example, heparin sepharose. Most common in clinical laboratories are direct homogeneous reagents designed for fully automated use with chemistry analyzers, using combinations of detergents and, in some cases, antibodies to selectively assay cholesterol in lipoprotein classes.

Many ultracentrifugation methods have been used in the research laboratory, but ultracentrifugation is uncommon in the clinical laboratory. The most common approach, called preparative ultracentrifugation, uses sequential density adjustments of serum to fractionate major and minor lipoprotein classes. Density gradient methods, either nonequilibrium techniques in which separations are based on the rate of flotation or equilibrium techniques in which the lipoproteins separate based on their density, permit fractionation of several or all classes in a single run.[58] Ultracentrifugation is also used in the reference methods for lipoprotein quantitation because lipoproteins are classically defined in terms of hydrated density.

Electrophoretic methods allow separation and quantitation of major lipoprotein classes, as well as subclasses, and provide a visual display useful in detecting unusual or variant patterns.[59] Agarose gel has been the most common medium for separation of intact lipoproteins, providing a clear background and convenience in use.[60] Commercial automated electrophoretic systems' lipoprotein determinations can be precise and accurate but are no longer widely used.[61]

Electrophoresis in polyacrylamide gels is used for separation of lipoprotein classes, subclasses, and the apolipoproteins.[62] Of particular interest are methods that fractionate LDL subclasses to characterize the more atherogenic smaller, denser, lipid-depleted fractions versus the larger, lighter subclasses.

Chemical precipitation, usually with polyanions, such as heparin and dextran sulfate, together with divalent cations, such as manganese or magnesium, can be used to separate any of the lipoproteins but are most common for HDL.[63] Apo B in VLDL and LDL is rich in positively charged amino acids, which

CASE STUDY 10.1, PART 6

Recall Maggie. She participates in a lipid study at a local university where the research lab uses EDTA for their cholesterol measurements. Maggie notices that her cholesterol results from the research lab are lower than the results from her physician's office lab.

11. What is a possible explanation for the lower results?

© Monkey Business Images/Shutterstock.

preferentially form complexes with polyanions. The addition of divalent cations neutralizes the charged groups on the lipoproteins, making them aggregate and become insoluble, resulting in their precipitation leaving HDL in solution. At appropriate concentrations of polyanion and divalent cation, the separation is reasonably specific.

Newer technologies for lipid biomarkers use liquid chromatography mass spectrometry (LC-MS). This method is not used in routine clinical laboratories: however, research with LC-MS has identified more lipid metabolites and may lead to better outcomes for patients.[64]

HDL Methods

For routine diagnostic purposes, HDL for many years was separated almost exclusively by chemical precipitation, involving a two-step procedure with manual pretreatment. A precipitation reagent added to serum or plasma aggregated non-HDLs, which were sedimented by centrifugation, at forces of approximately 1500 *g* (gravity) with lengthy centrifugation times of 10 to 30 minutes, or higher forces of 10,000 to 15,000 *g* and decreasing the centrifugation times to 3 minutes. HDL is then quantified as cholesterol in the supernate, usually by one of the enzymatic assays modified for the lower HDL-C range.

Older precipitation methods for HDL have largely been replaced with the development of a new class of direct, sometimes termed *homogeneous*, methods, which automate the HDL quantification, making them better suited for the modern chemistry laboratory. Specific polymers, detergents, and even modified enzymes are used to suppress the enzymatic cholesterol reaction in lipoproteins other than HDL.[65] In general, a first reagent is added to "block" non-HDLs, followed by a second reagent with the enzymes to quantify the accessible HDL cholesterol. Homogeneous assays, which are highly precise and reasonably accurate, have generally replaced pretreatment methods in the routine laboratory.[65] However, the methods have been shown to lack specificity for HDL in unusual specimens, for example, from patients with liver or kidney conditions. Also, the reagents have been subject to frequent modifications by the manufacturers in an effort to improve performance, which can affect results in long-term studies. For these reasons, precipitation-based methods are still used by some reference labs and are still performed by research laboratories.

The accepted reference method for HDL-C is a three-step procedure developed at the CDC. However, because this method is tedious and expensive, a simpler, direct precipitation method has been validated by the CDC Network Laboratory Group as a designated comparison method, using direct dextran sulfate (50 kD) precipitation of serum with Abell-Kendall cholesterol analysis.[55]

LDL Methods

LDL-C, well validated as a treatable risk factor for CHD, is the primary basis for treatment decisions in all guidelines. The most common research method for LDL-C quantitation and the basis for the reference method has been designated β-quantification, in which β designation refers to the electrophoretic term for LDL. β-quantification combines ultracentrifugation and chemical precipitation.[66] Ultracentrifugation of serum at the native density of 1.006 g/mL is used to float VLDL and any chylomicrons for separation. The fractions are recovered by pipetting after separating the fractions by slicing the tube. Ultracentrifugation has been preferred for VLDL separation because other methods, such as precipitation, are not as specific for VLDL and may be subject to interference from chylomicrons. In general, ultracentrifugation is a robust but tedious technique that can give reliable results, provided the technique is meticulous.

In a separate step, chemical precipitation is used to separate HDL from either the whole serum or the infranate (layer just below the supernate) obtained from ultracentrifugation. Cholesterol is quantified in serum, in the 1.006 g/mL infranate, and in the HDL supernate by enzymatic or other assay methods. LDL-C is calculated as the difference between cholesterol measured in the infranate and in the HDL fraction. VLDL-C is usually calculated as the difference between that in whole serum and the amount in the infranate fraction. The requirement for the need of an ultracentrifuge has generally limited β-quantification to lipid specialty laboratories. β-quantification is also the basis for the accepted reference method for LDL. The method is the same as that described for HDL above, with the measured HDL-C subtracted from that in the bottom fraction to obtain LDL-C.

A more common approach, bypassing ultracentrifugation and commonly used in routine and sometimes research laboratories, is calculating LDL-C

by the Friedewald calculation.[47] HDL-C is quantified either after precipitation or using one of the direct methods, and total cholesterol and triglycerides are measured in the serum. VLDL-C is estimated as the triglyceride level divided by 5 (when using mg/dL units), an approximation that works reasonably well in most normolipemic specimens. The presence of elevated triglycerides (400 mg/dL [4.52 mmol/L] is the accepted limit), chylomicrons, and β-VLDL characteristic of the rare hyperlipoproteinemia type III precludes this estimation. The estimated VLDL-C and measured HDL-C are subtracted from total serum cholesterol to estimate or derive LDL-C. Thus,

$$\text{LDL-C} = \text{total cholesterol} - \text{HDL} - \text{Trig}/5$$

This method, commonly performed as the lipid panel, is widely used in estimating LDL-C in routine clinical practice, having been recommended in the NCEP guidelines. Investigations in lipid specialty laboratories have suggested that the method is reasonably reliable for patient classification, provided the underlying measurements are made with appropriate accuracy and precision.

Like for HDL, direct LDL-C methods have been developed or refined for general use, similar to the homogeneous assays for HDL-C.[67] Besides achieving full automation of the challenging LDL-C separation, these assays have the potential to streamline the measurement while improving precision. However, separating LDL with adequate specificity from other lipoproteins in this manner is more challenging even than that for HDL.[68] Because of the extra cost compared to Friedewald calculations and analytical concerns, direct LDL-C assays are not as widely used as direct HDL-C assays, although they have the advantage of being used on nonfasting samples.

Apolipoprotein Methods

Lipids by nature tend to be insoluble in the aqueous environment of the circulation; hence, the lipoproteins include various protein constituents, designated apolipoproteins, which enhance solubility, as well as playing other functional roles (Table 10.2). Apolipoproteins are commonly measured in research and some specialty laboratories supporting cardiovascular practices, or clinical studies measure them routinely in addition to the lipoproteins. For clinical diagnostic purposes, three apolipoproteins in particular have been of interest. Apo B, the major protein of LDL and VLDL, is an indicator of combined LDL and VLDL concentration that can be measured directly in serum by immunoassay. Some studies suggest that Apo B may be a better indicator of atherogenic particles than LDL-C.[69] Apo A1, as the major protein of HDL, could be measured directly in serum in place of separation and analysis of HDL-C; however, because quantification in terms of cholesterol content is more common, the latter practice has prevailed. Lp(a), the variant of LDL shown to be an independent indicator of CHD risk, is sometimes determined for managing patients. Measurement of these three apolipoproteins, considered to be emerging markers, can be useful in patient management by experienced practitioners.

Lp(a) has pre-β mobility on agarose electrophoresis and can be quantified by this technique. However, the apolipoproteins are commonly measured by immunoassays of various types, with several commercial kit methods available.[70] Most common in routine laboratories are turbidimetric assays for chemistry analyzers or nephelometric assays for dedicated nephelometers. Especially for ApoB and Lp(a), these light-scattering assays may be subject to interference from the larger triglyceride-rich lipoproteins (chylomicrons) and VLDL. Enzyme-linked immunosorbent assay (ELISA), radial immunodiffusion (RID), and radioimmunoassay (RIA) methods have also been available, but the latter two methods are becoming less common. Antibodies used in the immunoassays may be polyclonal or monoclonal. International efforts to develop reference materials and standardization programs for the assays are in progress. Because Lp(a) is genetically heterogeneous and the levels and CHD risk correlate with the isoform size, qualitative assessment of isoform distribution may also be useful.[71]

Phospholipid Measurement

Quantitative measurements of phospholipids are primarily done for research but not as part of routine laboratory assessments of lipids and lipoproteins because they have not been shown to add any value for cardiovascular risk prediction. The choline-containing phospholipids lecithin, lysolecithin, and sphingomyelin, which account for at least 95% of total phospholipids in serum, can be measured by an enzymatic reaction sequence using phospholipase D, choline oxidase, and horseradish peroxidase.[72]

Fatty Acid Measurement

Although studies suggest that fatty acids have potential in assessing CHD risk (e.g., the n-3 fatty acids), analysis is also primarily used in research laboratories for studies of diet. Less common is their measurement in the diagnosis of rare genetic conditions. Individual fatty acids can be analyzed by gas–liquid chromatography (after extraction, alkaline hydrolysis, and conversion to methyl esters of diazomethane),[73] but enzymatic methods exist for measuring total plasma free fatty acids.[74]

Standardization of Lipid and Lipoprotein Assays

Precision

Precision is a prerequisite for accuracy; a method may have no overall systematic error or bias, but if it is imprecise, it will still be inaccurate on some individual measurements. With modern automated analyzers, analytic variation has generally become less of a concern than biologic and other sources of preanalytic variation. Cholesterol levels are affected by many factors that can be categorized into biologic, clinical, and sampling sources.[75] Changes in lifestyle that affect diet, exercise, weight, and smoking patterns can result in fluctuations in the observed cholesterol and triglyceride values and the distribution of the lipoproteins. Similarly, the presence of clinical conditions, various diseases, or the medications used in their treatment can affect the circulating lipoproteins. Conditions present during blood collection, such as fasting status, posture, the choice of anticoagulant in the collection tube, and storage conditions, can alter the measurements. Typical observed biologic variation for more than 1 year for total cholesterol averages approximately 6.1% CV. In the average patient, measurements made over the course of a year would fall 66% of the time within ±6.1% of the mean cholesterol concentration and 95% of the time within twice this range. Some patients may exhibit substantially more biologic variation. Thus, preanalytic variation is generally relatively large in relation to the usual analytic variation, which is typically less than 3% CV, and must be considered in interpreting cholesterol results. Some factors, such as posture and blood collection, can be standardized to minimize the variation. The NCEP guidelines recommend averaging at least two successive measurements to reduce the effects of both preanalytic and analytic sources.[21] The use of stepped cut points also reduces the practical effect of variation.

Accuracy

Accuracy or trueness is ensured by demonstrating traceability or agreement through calibration to the respective "gold standard" reference system. With cholesterol, the reference system is advanced and complete, having served as a model for standardization of other laboratory analytes. The definitive method at the National Institute of Standards and Technology provides the ultimate accuracy target but is too expensive and complicated for frequent use.[56] A reference method has been made conveniently accessible through a network of standardized laboratories, the Cholesterol Reference Method Laboratory Network. This network was established in the United States and other countries to extend standardization to manufacturers and clinical laboratories.[55] The network provides accuracy comparisons leading to certification of performance using fresh native serum specimens, necessary for reliable accuracy transfer because of analyte–matrix interaction problems on processed reference materials.[76]

Matrix Interactions

In the early stages of cholesterol standardization, which were directed toward diagnostic manufacturers and routine laboratories, commercial lyophilized or freeze-dried materials were used. These materials, made in large quantities, often with spiking or artificial addition of analytes, were assayed by the definitive and/or reference methods and distributed widely for accuracy transfer. Subsequently, biases were observed with some systems on fresh patient specimens even though they appeared to be accurate on the reference materials. Although such manufactured reference materials are convenient, stable, and amenable to shipment at ambient temperatures, the manufacturing process, especially spiking and lyophilization, altered the measurement properties in enzymatic assays such that results were not representative of those on patient specimens. To achieve reliable feedback on accuracy and facilitate transfer of the accuracy base, direct comparisons with the reference methods on actual patient specimens were determined to be necessary.[55] In response to these problems, the College of American Pathologists, a major provider of proficiency testing, now offers Accuracy-Based Lipid survey, which is made

of fresh frozen serum and is commutable and without any matrix problems. Lipid and lipoprotein values are assigned to the material using reference methods.

CDC Cholesterol Reference Method Laboratory Network

The CDC cholesterol reference method laboratory network program was organized to improve the accuracy of lipid and lipoprotein testing. (Information is available at http://www.cdc.gov/labstandards/crmln.html.) The network offers formal certification programs for total cholesterol, HDL-C, and LDL-C and triglycerides, whereby laboratories and manufacturers can document traceability to the national reference systems.[55] Through this program, clinical laboratories are able to identify certified commercial methods. Certification does not ensure all aspects of quality in a reagent system but primarily ensures that the accuracy is traceable to reference methods within accepted limits and that precision can meet the NCEP targets. The certification process is somewhat tedious and, thus, most efficient through manufacturers, but individual laboratories desiring to confirm the performance of their systems can complete a scaled-down certification protocol for cholesterol.

Analytic Performance Goals

The NCEP laboratory panels have established analytic performance goals based on clinical needs for routine measurements (Table 10.6).[77] For analysis of total cholesterol, the performance goal for total error is 8.9%. That is, the overall error should be such that each individual cholesterol measurement falls within ±8.9% of the reference method value. Because the goals are based on 95% certainty, 95 of 100 measurements should fall within the total error limit. One can assay a specimen many times and calculate the mean to determine the usual value or the central tendency. The scatter or random variation around the mean is described by the standard deviation, an interval around the mean that includes, by definition, two-thirds of the observations. In the laboratory, because the scatter or imprecision is often proportional to the concentration, random variation is usually specified in relative terms as CV—the coefficient of variation or relative standard deviation, calculated as the standard deviation divided by the mean. Overall accuracy or systematic error is described as *bias* or *trueness*—the difference between the mean and the true value. Bias is primarily a function of the method's calibration and may vary by concentration. Of greatest concern in this context is bias at the NCEP decision cut points. The bias and CV targets presented in Table 10.6 are representative of performance that will meet the NCEP goals for total error.

Quality Control

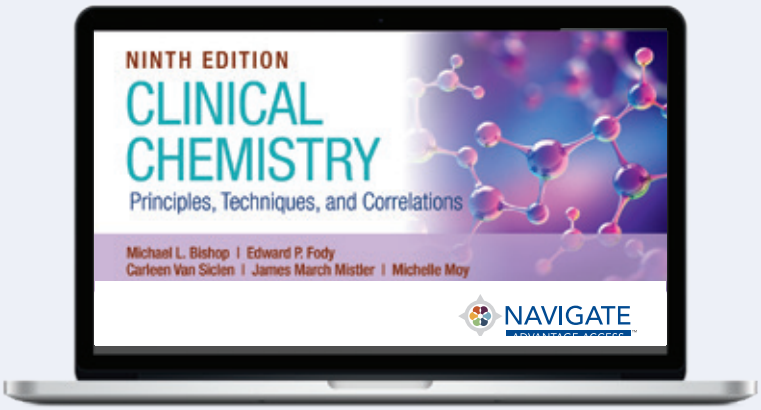
Achieving acceptable analytic performance requires the use of reliable quality control materials, which should preferably closely emulate actual patient specimens. Commercial control materials have improved in recent years but may not approximate results with patients' specimens. Control materials can also be prepared in-house from freshly collected patient serum, aliquoted into securely sealed vials, quick frozen, and stored at –70°C. Such pools of fresh frozen serum are less subject to matrix interactions than the usual commercial materials, which is most important in monitoring accuracy in lipoprotein separation and analysis and preferable for monitoring cholesterol and other lipid measurements. At least two pools should be analyzed, preferably with levels at or near decision points for each analyte.

Specimen Collection

Serum is the sample of choice for lipoprotein measurement in the routine clinical laboratory. Ethylenediaminetetraacetic acid (EDTA) plasma was the traditional choice in lipid research laboratories, especially for lipoprotein separations, because the anticoagulant was thought to enhance stability by chelating metal ions. EDTA, however, has potential disadvantages that discourage routine use. Microclots, which can form in plasma during storage, could plug the sampling probes on the modern chemistry analyzers. EDTA also osmotically draws water from red cells, diluting the plasma constituents, and the dilution effect can vary depending on such factors as fill volume, the analyte being measured, and the extent of mixing. Because the NCEP cut points are based on serum values, cholesterol measurements made on EDTA plasma require correction by the factor of 1.03.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. U.S. Food and Drug Administration (FDA). Trans fats. 2018. Retrieved from: https://www.fda.gov/food/food-additives-petitions/trans-fat
2. Muller H, Kirkhus B, Pedersen JI. Serum cholesterol predictive equations with special emphasis on trans and saturated fatty acids. an analysis from designed controlled studies. *Lipids.* 2001;36(8):783–791.
3. Cunnane SC. Problems with essential fatty acids: time for a new paradigm? *Prog Lipid Res.* 2003;42(6):544–568.
4. Charles KN, Shackelford JE, Faust PL, Fliesler SJ, Stangl H, Kovacs WJ. Functional peroxisomes are essential for efficient cholesterol sensing and synthesis. *Front Cell Dev Biol.* 2020;8:560266. doi:10.3389/fcell.2020.560266
5. Segrest JP, Li L, Anantharamaiah GM, Harvey SC, Liadaki KN, Zannis V. Structure and function of apolipoprotein A-I and high-density lipoprotein. *Curr Opin Lipidol.* 2000;11(2):105–115.
6. Burnett JR, Barrett PH. Apolipoprotein B metabolism: tracer kinetics, models, and metabolic studies. *Crit Rev Clin Lab Sci.* 2002;39(2):89–137.
7. Whitfield AJ, Barrett PHR, Van Bockxmeer FM, Burnett, JR. Lipid disorders and mutations in the *APOB* gene. *Clin Chem.* 2004;50(10):1725–1732. https://doi.org/10.1373/clinchem.2004.038026
8. Kostner KM, Kostner GM. Lipoprotein(a): still an enigma? *Curr Opin Lipidol.* 2002;13(4):391–396.
9. Lane RM, Farlow MR. Lipid homeostasis and apolipoprotein E in the development and progression of Alzheimer's disease. *J Lipid Res.* 2005;46(5):949–968.
10. Cooper AD. Hepatic uptake of chylomicron remnants. *J Lipid Res.* 1997;38(11):2173–2192.
11. Kruth HS. Sequestration of aggregated low-density lipoproteins by macrophages. *Curr Opin Lipidol.* 2002;13(5):483–488.
12. Warnick GR, McNamara JR, Boggess CN, Clendenen F, Williams PT, Landolt CC. Polyacrylamide gradient gel electrophoresis of lipoprotein subclasses. *Clin Lab Med.* 2006;26(4):803–846.
13. Albers JJ, Koschinsky ML, Marcovina SM. Evidence mounts for a role of the kidney in lipoprotein(a) catabolism. *Kidney Int.* 2007;71(10):961–962.
14. Nishiwaki M, Ikewaki K, Bader G, et al. Human lecithin:cholesterol acyltransferase deficiency: in vivo kinetics of low-density lipoprotein and lipoprotein-X. *Arterioscler Thromb Vasc Biol.* 2006;26(6):1370–1375.
15. Mead JR, Irvine SA, Ramji DP. Lipoprotein lipase: structure, function, regulation, and role in disease. *J Mol Med (Berl).* 2002;80(12):753–769.
16. Vaziri ND. Causes of dysregulation of lipid metabolism in chronic renal failure. *Semin Dial.* 2009;22(6):644–651.
17. Dean M, Hamon Y, Chimini G. The human ATP-binding cassette (ABC) transporter superfamily. *J Lipid Res.* 2001;42(7):1007–1017.
18. Nofer JR, Remaley AT. Tangier disease: still more questions than answers. *Cell Mol Life Sci.* 2005;62(19–20):2150–2160.
19. National Heart, Lung, and Blood Institute Expert Panel on Integrated Guidelines for Cardiovascular Health and Risk Reduction in Children and Adolescents. Expert panel on integrated guidelines for cardiovascular health and risk reduction in children and adolescents: summary report. *Pediatrics.* 2011;128(suppl 5):S213–S256.
20. Kato H, Tillotson J, Nichaman MZ, Rhoads GG, Hamilton HB. Epidemiologic studies of coronary heart disease and stroke in Japanese men living in Japan, Hawaii and California. *Am J Epidemiol.* 1973;97(6):372–385.
21. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III. Third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation.* 2002;106(25):3143–3421.

22. Jensen MD, Ryan DH, Apovian CM, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *Circulation.* 2014;129(25 suppl 2):S102–S138.
23. Kit BK, Carroll MD, Lacher DA, Sorlie PD, DeJesus JM, Ogden C. Trends in serum lipids among US youths aged 6 to 19 years, 1988–2010. *JAMA.* 2012;308(6):591–600.
24. McMahan CA, Gidding SS, Malcom GT, et al. Pathobiological determinants of atherosclerosis in youth risk scores are associated with early and advanced atherosclerosis. *Pediatrics.* 2006;118(4):1447–1455.
25. Ritchie SK, et al. Universal versus targeted blood cholesterol screening among youth: the CARDIAC project. *Pediatrics.* 2010;126(2):260–265.
26. Newman TB, Pletcher MJ, Hulley SB. Overly aggressive new guidelines for lipid screening in children: evidence of a broken process. *Pediatrics.* 2012;130(2):349–352.
27. Berger JH, Chen F, Faerber JA, O'Byrne ML, Brothers JA. Adherence with lipid screening guidelines in standard- and high-risk children and adolescents. *Am Heart J.* 2021;232:39–46.
28. Niebauer J, Hambrecht R, Velich T, et al. Attenuated progression of coronary artery disease after 6 years of multifactorial risk intervention: role of physical exercise. *Circulation.* 1997;96(8):2534–2541.
29. Warnick GR, Wood PD. National Cholesterol Education Program recommendations for measurement of high density lipoprotein cholesterol: executive summary. The National Cholesterol Education Program Working Group on Lipoprotein Measurement. *Clin Chem.* 1995;41(10):1427–1433.
30. Islam M, Amin M, Siddiqui S, Hossain M, Sultana F, Kabir M. Trans fatty acids and lipid profile: A serious risk factor to cardiovascular disease, cancer and diabetes. *Diabetes Metab Syndrome Clin Res Rev*. 2019;13(2):1643–1647.
31. Abrignani MG, Lucà F, Favilli S, et al. Lifestyles and cardiovascular prevention in childhood and adolescence. *Pediatr Cardiol*. 2019;40(6):1113–1125. doi:10.1007/s00246-019-02152-w
32. Keaney JF Jr. Atherosclerosis: from lesion formation to plaque activation and endothelial dysfunction. *Mol Aspects Med.* 2000;21(4–5):99–166.
33. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med.* 1995;333(20):1301–1307.
34. Mani P, Rohatgi A. Niacin therapy, HDL cholesterol, and cardiovascular disease: is the HDL hypothesis defunct? *Curr Atheroscler Rep*. 2015;17(8):43. doi:10.1007/s11883-015-0521-x
35. Kosoglou T, Statkevich P, Johnson-Levonas AO, Paolini JF, Bergman AJ, Alton KB. Ezetimibe: a review of its metabolism, pharmacokinetics and drug interactions. *Clin Pharmacokinet.* 2005;44(5):467–494.
36. Kłosiewicz-Latoszek L, Cybulska B, Tyszko P. Current state-of-the-art knowledge on the role of omega-3 fatty acids in the prevention of cardiovascular disease. *Ann Agric Env Med.* 2020;27(4):519–525. doi:10.26444/aaem/126674
37. Bartuli A, Bucci M, Buonuomo PS, et al. Homozygous familial hypercholesterolemia in Italy: Clinical and molecular features. *Atherosclerosis*. 2020;312:72–78. doi:10.1016/j.atherosclerosis.2020.08.027
38. Palcoux JB, Meyer M, Jouanel P, Vanlieferinghen P, Malpuech G. Comparison of different treatment regimens in a case of homozygous familial hypercholesterolemia. *Ther Apher.* 2002;6(2):136–139.
39. Ling H, Burns TL, Hilleman DE. An update on the clinical development of proprotein convertase subtilisin kexin 9 inhibitors, novel therapeutic agents for lowering low-density lipoprotein cholesterol. *Cardiovasc Ther.* 2014;32(2):82–88.
40. Giugliano RP, Sabatine MS. Are PCSK9 inhibitors the next breakthrough in the cardiovascular field? *J Am Coll Cardiol.* 2015;65(24):2638–2651.
41. Saeed A, Feofanova EV, Yu B, et al. Remnant-like particle cholesterol, low-density lipoprotein triglycerides, and incident cardiovascular disease. *JACC*. 2018;72(2):156-169. doi:10.1016/j.jacc.2018.04.050
42. Recazens E, Mouisel E, Langin D. Hormone-sensitive lipase: sixty years later. *Prog Lipid Res*. 2021;82:101084. doi:10.1016/j.plipres.2020.101084
43. Gill PK, Dron JS, Dilliott AA, et al. Ancestry-specific profiles of genetic determinants of severe hypertriglyceridemia. *J Clin Lipidol*. 2021;15(1):88–96. doi:10.1016/j.jacl.2020.11.007
44. Esan O, Wierzbicki AS. Volanesorsen in the treatment of familial chylomicronemia syndrome or hypertriglyceridaemia: design, development and place in therapy. *Drug Des Devel Ther*. 2020;14:2623–2636. Accessed March 3, 2021.
45. Civeira F, Pocoví M, Cenarro A, et al. Apo E variants in patients with type III hyperlipoproteinemia. *Atherosclerosis.* 1996;127(2):273–282.
46. Sijbrands EJ, Hoffer MJ, Meinders AE, et al. Severe hyperlipidemia in apolipoprotein E2 homozygotes due to a combined effect of hyperinsulinemia and an SstI polymorphism. *Arterioscler Thromb Vasc Biol.* 1999;19(11):2722–2729.
47. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem.* 1972;18(6):499–502.
48. Jialal I, Hirany SV, Devaraj S, Sherwood TA. Comparison of an immunoprecipitation method for direct measurement of LDL-cholesterol with beta-quantification (ultracentrifugation). *Am J Clin Pathol.* 1995;104(1):76–81.
49. Paquette M, Bernard S, Thanassoulis G, Baass A. LPA genotype is associated with premature cardiovascular disease in familial hypercholesterolemia. *J Clin Lipidol*. 2019;13(4):627. doi:10.1016/j.jacl.2019.04.006
50. Petersen NH, Schmied AB, Zeller JA, Plendl H, Deuschl G, Zunker P. Lp(a) lipoprotein and plasminogen activity in patients with different etiology of ischemic stroke. *Cerebrovasc Dis*. 2007;23(2–3):188–193. doi:10.1159/000097640
51. Bittner V. Non-high-density lipoprotein cholesterol and cardiovascular disease. *Curr Opin Lipidol.* 2003;14(4):367–371.
52. Charlton-Menys V, Betteridge DJ, Colhoun H, et al. Targets of statin therapy: LDL cholesterol, non-HDL cholesterol, and apolipoprotein B in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS). *Clin Chem.* 2009;55(3):473–480.
53. Baldassarre D, Amato M, Pustina L, et al. Increased carotid artery intima-media thickness in subjects with primary

hypoalphalipoproteinemia. *Arterioscler Thromb Vasc Biol.* 2002;22:317–322.
54. McCormack PL, Keating GM. Prolonged-release nicotinic acid: a review of its use in the treatment of dyslipidaemia. *Drugs.* 2005;65(18):2719–2740.
55. Myers GL, Cooper GR, Greenberg N, et al. Standardization of lipid and lipoprotein measurements. In: Rifai N, Warnick GR, Dominczak MH, eds. *Handbook of Lipoprotein Testing*. 2nd ed. Washington, DC: AACC Press; 2000.
56. Cohen A, Hertz HS, Mandel J, et al. Total serum cholesterol by isotope dilution/mass spectrometry: a candidate definitive method. *Clin Chem.* 1980;26(7):854–860.
57. Warnick GR, Nakajima K. Fasting versus nonfasting triglycerides: implications for laboratory measurements. *Clin Chem.* 2008;54(1):14–16. doi:10.1373/clinchem.2007.098863
58. Patsch JR, Patsch W. Zonal ultracentrifugation. *Methods Enzymol.* 1986;129:3–26.
59. Schmitz G, Boettcher A, Barlage S. New approaches to the use of lipoprotein electrophoresis in the clinical laboratory. In: Rifai N, Warnick GR, Dominczak MH, eds. *Handbook of Lipoprotein Testing*. 2nd ed. Washington, DC: AACC Press; 2000.
60. Greenspan P, Mao FW, Ryu BH, Gutman RL. Advances in agarose gel electrophoresis of serum lipoproteins. *J Chromatogr A.* 1995;698(1–2):333–339.
61. Warnick GR, Leary ET, Goetsch J. Electrophoretic quantification of LDL-cholesterol using the Helena REP. *Clin Chem.* 1993;39:1122.
62. Li Z, McNamara J, Ordovas J, Schaefer E. et al. Analysis of high density lipoproteins by a modified gradient gel electrophoresis method. *J Lipid Res.* 1994;35(9):1698–1711.
63. Burstein M, Legmann P. Lipoprotein precipitation. *Monogr Atheroscler.* 1982;11:1–131.
64. Cai F, Ren F, Zhang Y, et al. Screening of lipid metabolism biomarkers in patients with coronary heart disease via ultra-performance liquid chromatography-high resolution mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci.* 2021;1169:122603. doi:10.1016/j.jchromb.2021.122603
65. Matsushima-Nagata K, Sugiuchi H, Anraku K, et al. A homogeneous assay to determine high-density lipoprotein subclass cholesterol in serum. *Anal Biochem.* 2021;613:114019. doi:10.1016/j.ab.2020.114019
66. Bachorik PS. Lipid and lipoprotein analysis with desktop analyzers. In: Rifai N, Warnick GR, Dominczak MH, eds. *Handbook of Lipoprotein Testing*. 2nd ed. Washington, DC: AACC Press; 2000.
67. Sugiuchi H, Irie T, Uji Y, et al. Homogeneous assay for measuring low-density lipoprotein cholesterol in serum with triblock copolymer and alpha-cyclodextrin sulfate. *Clin Chem.* 1998;44(3):522–531.
68. Rifai N, Iannotti E, DeAngelis K, Law T. Analytical and clinical performance of a homogeneous enzymatic LDL-cholesterol assay compared with the ultracentrifugation-dextran sulfate-Mg^{2+} method. *Clin Chem.* 1998;44(6 Pt 1):1242–1250.
69. Barter PJ, Ballantyne CM, Carmena R, et al. Apo B versus cholesterol in estimating cardiovascular risk and in guiding therapy: report of the thirtyperson/ten-country panel. *J Intern Med.* 2006;259(3):247–258.
70. Bhatnager D, Durrington PN. Measurement and clinical significance of apolipoproteins A-1 and B. In: Rifai N, Warnick GR, Dominczak MH, eds. *Handbook of Lipoprotein Testing*. 2nd ed. Washington, DC: AACC Press; 2000.
71. Marcovina SM, Koschinsky ML. Lipoprotein (a): structure, measurement, and clinical significance. In: Rifai N, Warnick GR, Dominczak MH, eds. *Handbook of Lipoprotein Testing*. 2nd ed. Washington, DC: AACC Press; 2000.
72. McGowan MW, Artiss JD, Zak B. A procedure for the determination of high-density lipoprotein choline-containing phospholipids. *J Clin Chem Clin Biochem.* 1982;20(11):807–812.
73. Hellerstein MK. Methods for measurement of fatty acid and cholesterol metabolism. *Curr Opin Lipidol.* 1995;6(3):172–181.
74. Husek P, Simek P, Tvrzická E. Simple and rapid procedure for the determination of individual free fatty acids in serum. *Anal Chim Acta.* 2002;465:433–439.
75. Cooper GR, Myers GL, Smith SJ, Schlant RC. Blood lipid measurements. Variations and practical utility. *JAMA.* 1992;267(12):1652–1660.
76. Bernert JT Jr, Akins JR, Cooper GR, Poulose AK, Myers GL, Sampson EJ. Factors influencing the accuracy of the national reference system total cholesterol reference method. *Clin Chem.* 1991;37(12):2053–2061.
77. Report from the Laboratory Standardization Panel of the National Cholesterol Education Program. *Recommendations for Improving Cholesterol Measurement*. Bethesda, MD: National Institutes of Health; 1990; (No. 902964).

CHAPTER 11

Electrolytes

James March Mistler

CHAPTER OUTLINE

KEY TERMS

Active transport
Anion
Anion gap (AG)
Arginine vasopressin hormone (AVP)
Atrial natriuretic peptide (ANP)
Cation
Diabetes insipidus
Diffusion
Electrolyte
Extracellular fluid (ECF)
Hypercalcemia
Hyperchloremia
Hyperkalemia
Hypermagnesemia
Hypernatremia
Hyperphosphatemia
Hypocalcemia
Hypochloremia
Hypokalemia
Hypomagnesemia
Hyponatremia
Hypophosphatemia
Hypovolemia
Interstitial fluid
Intracellular fluid (ICF)
Intravascular fluid (IVF)
Osmolal gap
Osmolality
Osmolarity
Osmometer
Polydipsia
Renin–angiotensin–aldosterone system (RAAS)
Resting membrane potential (RMP)
Syndrome of inappropriate antidiuretic hormone (SIADH)
Tetany
Transcellular fluid

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define electrolyte, osmolality, anion gap, anion, and cation.
- Describe the physiology of each electrolyte discussed in the chapter.
- State the clinical significance of each of the electrolytes.
- Calculate osmolality, osmolal gap, and anion gap and discuss the clinical usefulness of each.
- Discuss the analytic techniques used to assess electrolyte concentrations.
- Correlate the information with disease state, given patient data.
- Identify the reference ranges for sodium, potassium, chloride, bicarbonate, magnesium, and calcium.
- State the specimen of choice for the major electrolytes.
- Explain the role of the kidney in electrolyte excretion and conservation.
- Discuss the usefulness of urine electrolyte results: sodium, potassium, calcium, and osmolality.

CASE STUDY 11.1, PART 1

Alyssa is a 34-year-old female who lives alone. Alyssa was found by her partner after she failed to respond to text messages or phone calls for 2 days. Alyssa was found on the floor of the bathroom, severely dehydrated and barely conscious. Her partner brought her to the emergency department and tells the staff he thinks she was drinking previously. The ED provider notices dry mucous membranes and orders several laboratory tests.

CASE STUDY 11.2, PART 1

Chris, a 15-year-old non-binary teenager who has type 1 diabetes (insulin dependent), was brought to the emergency department in a coma by their parents. Chris' parents stated that the teen has had some recent episodes of hyperglycemia and ketosis because Chris is often "too busy" to test and take their insulin. Chris's mother is hopeful that the new insurance she got through her job will cover a continuous glucose monitor to avoid these dangerous episodes.

Electrolytes are ions (minerals) capable of carrying an electric charge. They are classified as either anions or cations based on the type of charge they carry as well as how the ion migrates in an electric field; those electrolytes with a positive charge are **cations** that move toward the cathode, while those with a negative charge are **anions** that move toward the anode.

Electrolytes are an essential component in numerous processes, including volume and osmotic regulation (sodium [Na^+], chloride [Cl^-], potassium [K^+]); myocardial rhythm and contractility (K^+, magnesium [Mg^{2+}], calcium [Ca^{2+}]); cofactors in enzyme activation (e.g., Mg^{2+}, Ca^{2+}, zinc [Zn^{2+}]); regulation of adenosine triphosphatase (ATPase) ion pumps (Mg^{2+}); acid–base balance (bicarbonate [HCO_3^-], K^+, Cl^-); blood coagulation (Ca^{2+}, Mg^{2+}); neuromuscular excitability (K^+, Ca^{2+}, Mg^{2+}); and the production and use of ATP from glucose (e.g., Mg^{2+}, phosphate PO_4^-). Because many of these functions require electrolyte concentrations to be held within very narrow ranges, the body has complex systems for monitoring and maintaining electrolyte concentrations.

This chapter explores both the metabolic physiology and regulation of each electrolyte, and relates these factors to the clinical significance of electrolyte measurements. In addition, methodologies used in determining concentrations of the individual analytes are discussed.

Water

The average water content of the human body varies from 40% to 75% of total body weight, with values declining with age and especially with obesity. Women have lower average water content than do men as a result of a higher fat content. Water is the solvent for all processes in the human body as it is responsible for transporting nutrients to cells, determining cell volume by its transport into and out of cells, removal of waste products by way of urine, and acting as the body's coolant by way of sweating. Water is located in both intracellular and extracellular compartments. **Intracellular fluid (ICF)** is the fluid inside the cells and accounts for about two-thirds of total body water. **Extracellular fluid (ECF)** accounts for the other one-third of total body water and can be further subdivided into the **intravascular fluid** (IVF, plasma), **interstitial fluid** that surrounds the cells in the tissue, and **transcellular fluid** such as cerebral spinal fluid. Normal plasma is about 93% water, with the remaining volume occupied by lipids and proteins. The concentrations of ions within cells and in plasma are maintained both by energy-consuming active transport processes and by diffusion or passive transport processes.

Active transport is a mechanism that requires energy to move ions across cellular membranes. For example, maintaining a high intracellular concentration of K^+ and a high extracellular (plasma) concentration of Na^+ requires use of energy, which comes from the breakdown of ATP by ATPase-dependent ion pumps. **Diffusion** is the passive movement of ions (no energy consumed) across a membrane and depends on both the size and charge of the ion being transported and on the nature of the membrane through which it is passing. The rate of diffusion of various ions also may be altered by physiologic and hormonal processes.

Distribution of water in the various body fluid compartments is controlled by maintaining the

concentration of electrolytes and proteins in the individual compartments. Because most biologic membranes are freely permeable to water but not to ions or proteins, the concentration of ions and proteins on either side of the membrane will influence the flow of water across a membrane (known as an osmoregulator). In addition to the osmotic effects of Na^+, other ions, proteins, and blood pressure influence the flow of water across a membrane.

Osmolality

Osmolality is a physical property of a solution that is based on the concentration of solutes (expressed as millimoles) per kilogram of solvent (w/w). Osmolality is related to several changes in the properties of a solution relative to pure water, such as freezing point depression and vapor pressure decrease. These colligative properties are the basis for routine measurements of osmolality in the laboratory (see Chapter 4, *Analytic Techniques*). The term **osmolarity** is still occasionally used, with results reported in milliosmoles per liter (w/v), but it is inaccurate in cases of hyperlipidemia or hyperproteinemia; for urine specimens; or in the presence of certain osmotically active substances, such as alcohol or mannitol. Both the sensation of thirst and **arginine vasopressin hormone (AVP)**, formerly called antidiuretic hormone (ADH), are stimulated by the hypothalamus in response to an increased osmolality of blood. The natural response to the thirst sensation is to consume more fluids, increasing the water content of the ECF, therefore diluting the elevated solute (usually Na^+) levels and decreasing the osmolality of the plasma. Thirst, therefore, is important in mediating fluid intake. The other means of controlling osmolality is by secretion of AVP. This hormone is secreted by the posterior pituitary gland and acts on the collecting ducts in the kidneys to increase water reabsorption. This is an example of a negative feedback loop; as water is conserved, the osmolality decreases, which in turn shuts off AVP secretion (feedback loops are discussed more in Chapter 13, *Basic Endocrinology*).[1]

Clinical Significance of Osmolality

Osmolality in plasma is important because it is the parameter to which the hypothalamus responds. The regulation of osmolality primarily affects the Na^+ concentration in plasma because Na^+ and its associated anions account for approximately 90% of the osmotic activity in plasma, whereas urea and glucose account for the other 10%. Another important process affecting the Na^+ concentration in blood is the regulation of blood volume. As discussed later, although osmolality and volume are regulated by separate mechanisms (except for AVP and thirst), they are closely related because osmolality (Na^+) is regulated by changes in water balance, whereas volume is regulated by changes in Na^+ balance.[1]

To maintain a normal plasma osmolality (275 to 295 mOsm/kg of plasma H_2O), osmoreceptors in the hypothalamus respond quickly to small changes in osmolality. A 1% to 2% increase in osmolality causes a fourfold increase in the circulating concentration of AVP, and a 1% to 2% decrease in osmolality shuts off AVP production. AVP acts by increasing the reabsorption of water in the cortical and medullary collecting tubules but has a half-life in the circulation of only 15 to 20 minutes.

Renal water regulation by AVP and thirst play important roles in regulating plasma osmolality. Renal water excretion is more important in controlling water excess, whereas thirst is more important in preventing water deficit (dehydration). Consider what happens in the following conditions.

Water Load. As excess intake of water, sometimes caused by **polydipsia** (increased thirst), begins to lower plasma osmolality, both AVP and thirst are suppressed. In the absence of AVP, water is not reabsorbed, causing a large volume of dilute urine to be excreted; as much as 3 to 20 L of water daily can be excreted, well above any normal intake of water. Therefore, hypoosmolality and hyponatremia usually occur only in patients with impaired renal excretion of water.[1,2]

Water Deficit. As the deficit of water begins to increase plasma osmolality, both AVP secretion and thirst are activated. Although AVP contributes by minimizing renal water loss, thirst is the major defense against hyperosmolality and hypernatremia. Although hypernatremia rarely occurs in a person with a normal thirst mechanism and access to water, it becomes a concern in infants, unconscious patients, or anyone who is unable to either drink or ask for water. Osmotic stimulation of thirst progressively diminishes in people who are older than age 60. In the older patient with illness and diminished mental status, dehydration becomes increasingly likely. An example of the effectiveness of thirst in preventing dehydration can be seen in patients with **diabetes insipidus**, a disease of water and salt imbalance. In diabetes insipidus, there is a deficiency of AVP, causing increased urinary excretion, which may or may not interfere with Na^+ concentration.

Refer to Chapter 15, *Hypothalamic and Pituitary Function*, for more information on diabetes insipidus.

Regulation of Blood Volume

Adequate blood volume is essential to maintain blood pressure and ensure good perfusion to all tissues and organs. Regulation of both Na^+ and water is interrelated in controlling blood volume. The **Renin–Angiotensin–Aldosterone System (RAAS)** responds primarily to a decreased blood volume. Renin is secreted near the renal glomeruli in response to decreased renal blood flow (decreased blood volume or blood pressure). Renin converts the hormone angiotensinogen to angiotensin I, which then becomes angiotensin II. Angiotensin II causes both vasoconstriction, which quickly increases blood pressure, and secretion of aldosterone, an adrenal hormone that acts on the collecting ducts to increase reabsorption of Na^+, thus causing water to follow. The overall RAAS and the effects of blood volume and osmolality on Na^+ and water metabolism are shown in **Figure 11.1**. Changes in blood volume (which is actually a change in pressure) are initially detected by a series of stretch receptors located in areas such as the cardiopulmonary circulation, carotid sinus, aortic arch, and glomerular arterioles. These receptors then activate a series of responses (effectors) that restore volume by appropriately varying vascular resistance, cardiac output, and renal Na^+ and water retention.[1]

Four other factors affect blood volume: (1) **atrial natriuretic peptide (ANP)**, released from the myocardial atria in response to volume expansion, promotes Na^+ excretion in the kidney (B-type natriuretic peptide and ANP act together in regulating blood pressure and fluid balance); (2) volume receptors independent of osmolality stimulate the release of AVP, which conserves water by renal reabsorption; (3) glomerular filtration rate (GFR) increases with volume expansion and decreases with volume depletion; and (4) all other things equal, an increased plasma Na^+ will increase urinary Na^+ excretion (and therefore water excretion) and vice versa. The reabsorption of 98% to 99% of filtered Na^+ by the renal tubules conserves nearly all of the 150 L of glomerular filtrate produced daily, while a 1% to 2% reduction in tubular reabsorption of Na^+ can increase water loss by several liters per day.

Urine osmolality values may vary widely depending on water intake and the circumstances of collection. However, it is generally decreased in diabetes insipidus and excessive water intake, and it is increased in conditions such as the **syndrome of inappropriate antidiuretic hormone** (**SIADH**, increased AVP secretion) and **hypovolemia** (low blood volume), although urinary Na^+ is usually decreased.

Determination of Osmolality

Specimen. Osmolality may be measured in serum or random urine samples. Major electrolyte concentrations, mainly sodium, chloride, and bicarbonate, provide the largest contribution to the osmolality value of serum. Plasma use is not recommended because osmotically active substances may be

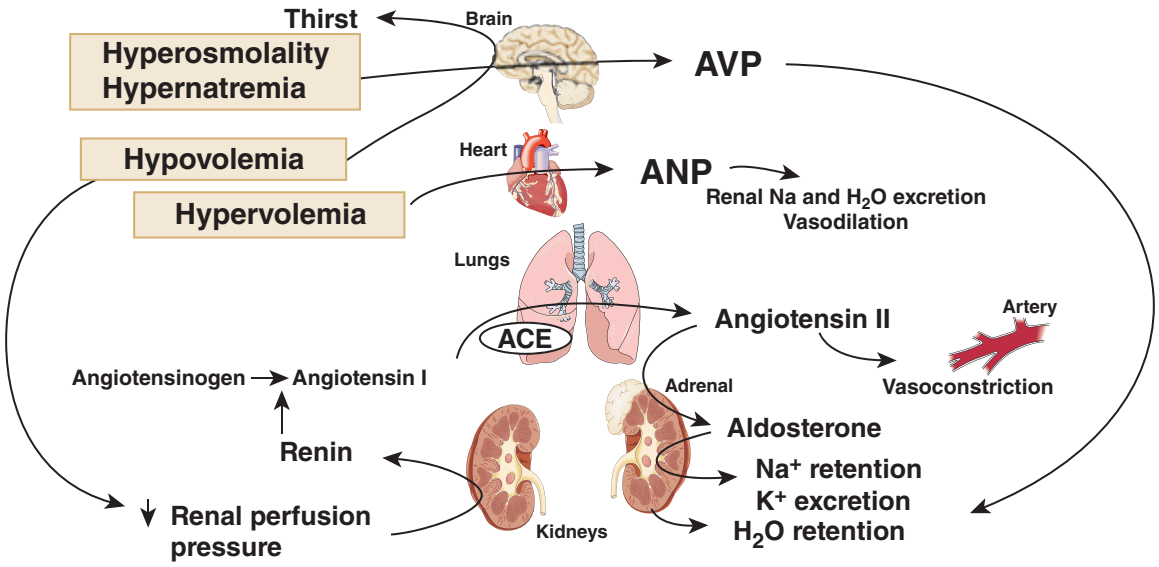


Figure 11.1 Responses to changes in blood osmolality and blood volume. ANP, atrial natriuretic peptide; AVP, arginine vasopressin; ACE, angiotensin-converting enzyme. The primary stimuli are shown in *boxes* (e.g., hypovolemia).

introduced into the specimen from the anticoagulant. Samples must be free of particulate matter to obtain accurate results. Turbid serum and urine samples should be centrifuged before analysis to remove any extraneous particles.

Discussion. The methods for determining osmolality are based on properties of a solution that are related to the number of molecules of solute per kilogram of solvent (colligative properties), such as changes in freezing point and vapor pressure. An increase in osmolality decreases both the freezing point temperature and the vapor pressure. Measurements of freezing point depression and vapor pressure decrease (actually, the dew point) are the two most frequently used methods of analysis. For detailed information on theory and methodology, consult Chapter 4, *Analytic Techniques* or the operator's manual of the instrument being used.

Osmometers that operate by freezing point depression are standardized using sodium chloride reference solutions. After calibration, the appropriate amount of sample is pipetted into the required cuvette or sample cup and placed in the analyzer. The sample is then supercooled to −7°C and seeded to initiate the freezing process. When temperature equilibrium has been reached, the freezing point is measured, with results for serum and urine osmolality reported as milliosmoles per kilogram.

Calculation of osmolality is useful either as an estimate of the true osmolality or to determine the **osmolal gap**, which is the difference between the measured osmolality and the calculated osmolality. The osmolal gap indirectly indicates the presence of osmotically active substances other than Na^+, urea, or glucose, such as ethanol, methanol, ethylene glycol, isopropanol, lactate, or β-hydroxybutyrate.

Two formulas are presented, each having theoretic advantages and disadvantages. Both are adequate for the purpose previously described. The numbers in the formula (1.86, 18, and 2.8) are the molecular weights of the respective compounds. In the first equation, these are rounded for simplicity; both are commonly used in osmolality formulas.

$$2\,Na + \frac{\text{glucose (mg / dL)}}{20} + \frac{\text{BUN (mg / dL)}}{3}$$

or

$$1.86\,Na + \frac{\text{glucose}}{18} + \frac{\text{BUN}}{2.8} + 9 \qquad \textbf{(Eq. 11.1)}$$

Reference Ranges

See **Table 11.1**.

Table 11.1 Reference Ranges for Osmolality

Serum	275–295 mOsm/kg
Urine (24 h)	300–900 mOsm/kg
Urine/serum ratio	1.0–3.0
Random urine	50–1,200 mOsm/kg
Osmolal gap	5–10 mOsm/kg

CASE STUDY 11.1, PART 2

Remember Alyssa, the woman who was found unconscious after 2 days. Some labs are now back.

Analyte	Patient Value	Reference Range
Serum osmolality	278 mOsm/kg	275–295 mOsm/kg
Glucose	75 mg/dL	70–100 mg/dL
BUN	10 mg/dL	6–20 mg/dL
Na^+	129 mmol/L	135–145 mmol/L
Urine Na^+	8 mmol/L (random)	>20 mmol/L

1. Calculate the osmolal gap. What does this indicate?

The Electrolytes

Sodium

Na^+ is the most abundant cation in the ECF, representing 90% of all extracellular cations, and largely determines the osmolality of the plasma. A normal plasma osmolality is approximately 295 mmol/L, with Na^+ and associated anions accounting for 270 mmol/L.

Na^+ concentration in the ECF is much larger than inside the cells. Because a small amount of Na^+ can diffuse through the cell membrane, the two sides would eventually reach equilibrium. To prevent equilibrium and to ensure a concentration gradient, active transport systems, such as ATPase ion pumps, are present in all cells. K^+ (see section "Potassium") is the major intracellular cation. Like Na^+, K^+ would eventually diffuse across the cell membrane until equilibrium is reached. The Na^+/K^+-ATPase ion pump moves three Na^+ ions out of the cell in exchange for two K^+ ions moving into the cell as ATP is converted to ADP. Because water follows electrolytes across cell membranes, the continual removal of Na^+ from the cell prevents osmotic rupture of the cell by also drawing water from the cell.

Regulation

The plasma Na^+ concentration depends greatly on the intake and excretion of water and, to a somewhat lesser degree, on the renal regulation of Na^+. Three processes are of primary importance: (1) the intake of water in response to thirst, as stimulated or suppressed by plasma osmolality; (2) the excretion of water, largely affected by AVP release in response to changes in either blood volume or osmolality; and (3) the blood volume status, which affects Na^+ excretion through aldosterone, angiotensin II, and ANP. The kidneys have the ability to conserve or excrete large amounts of Na^+, depending on the Na^+ content of the ECF and the blood volume. Normally, 60% to 75% of filtered Na^+ is reabsorbed in the proximal tubule; electroneutrality is maintained by either Cl^- reabsorption or hydrogen ion (H^+) secretion. Some Na^+ is also reabsorbed in the loop and distal tubule and (controlled by aldosterone) exchanged for K^+ in the connecting segment and cortical collecting tubule. The regulation of osmolality and volume has been summarized in Figure 11.1.

Clinical Applications

Hyponatremia. **Hyponatremia** is defined as a low serum/plasma Na^+ levels, usually less than 135 mmol/L, and levels below 120 mmol/L are clinically significant.[3] Hyponatremia is one of the most common electrolyte disorders in hospitalized and nonhospitalized patients and can be assessed either by the cause of the decrease or with the osmolality level.[3,4]

Decreased levels may be caused by increased Na^+ loss, increased water retention, or water imbalance (**Table 11.2**). Increased Na^+ loss in the urine can occur with decreased aldosterone, certain diuretics (thiazides), ketonuria (Na^+ lost with ketones), or a salt-losing nephropathy (with some renal tubular disorders). K^+ deficiency also causes Na^+ loss because of the inverse relationship of the two ions in the renal tubules; when serum K^+ levels are low, the renal tubules will conserve K^+ and excrete Na^+ in exchange for the loss of the monovalent cation. Each disorder results in an increased urine Na^+ level ($\geq$20 mmol/d), which exceeds the amount of water loss.

Prolonged vomiting, diarrhea, or severe burns can result in Na^+ loss. Urine Na^+ levels are usually less than 20 mmol/d in these disorders, which can be used to differentiate among causes for urinary loss.

Increased water retention causes dilution of serum/plasma Na^+ as with acute or chronic renal failure. In

Table 11.2 Causes of Hyponatremia

Increased Sodium Loss
Hypoadrenalism
Potassium deficiency
Diuretic use
Ketonuria
Salt-losing nephropathy
Prolonged vomiting or diarrhea
Severe burns
Increased Water Retention
Renal failure
Nephrotic syndrome
Hepatic cirrhosis
Congestive heart failure
Water Imbalance
Excess water intake
SIADH
Pseudohyponatremia

SIADH, syndrome of inappropriate antidiuretic hormone.

nephrotic syndrome and hepatic cirrhosis, plasma proteins are decreased, resulting in a decreased colloid osmotic pressure in which intravascular fluid migrates to the tissue (edema results). The low plasma volume causes AVP to be produced, causing fluid retention and resulting in dilution of Na^+. This compensatory mechanism is also seen with congestive heart failure (CHF) as a result of increased venous pressure. Urine Na^+ levels can be used to differentiate the cause for increased water retention; when urine Na^+ is $\geq$20 mmol/d, acute or chronic renal failure is the likely cause, and when urine levels are less than 20 mmol/d, water retention may be a result of nephrotic syndrome, hepatic cirrhosis, or CHF.

Water imbalance can occur as a result of excess water intake, as with polydipsia. The increased intake must be chronic before water imbalance occurs, which may cause mild or severe hyponatremia. In a normal individual, excess intake will not affect Na^+ levels. SIADH causes an increase in water retention because of increased AVP production. A defect in AVP regulation has been associated with pulmonary disease, malignancies, central nervous system (CNS) disorders, infections (e.g., *Pneumocystis jirovecii* pneumonia), or trauma.[5] Pseudohyponatremia can occur when Na^+ is measured using indirect ion-selective electrodes (ISEs) in a patient who is hyperproteinemic or hyperlipidemic. An indirect ISE dilutes the sample prior to analysis, and as a result of serum/plasma water displacement, the ion levels are falsely decreased. (For detailed information on the theory of water displacement with indirect ISEs, consult Chapter 4, *Analytic Techniques*.) Hyponatremia can also be classified according to serum osmolality (**Table 11.3**). Because Na^+ is a major contributor to osmolality, both levels can assist in identifying the cause of hyponatremia. There are three categories of hyponatremia: low osmolality, normal osmolality, or high osmolality.[3] Most instances of hyponatremia occur with decreased osmolality. This may be a result of Na^+ loss or water retention, as previously mentioned.

Hyponatremia with a normal osmolality may be a result of a high increase in non-sodium cations as listed in Table 11.3. In multiple myeloma, the cationic γ-globulins replace some Na^+ to maintain the electroneutrality; however, because it is a multivalent cation, it has little effect on osmolality.

Pseudohyponatremia, as mentioned earlier, may also be seen with in vitro hemolysis, considered the most common cause for a false decrease. When red blood cells (RBCs) lyse, Na^+, K^+, and water are released, and since Na^+ concentration is lower in RBCs, this results in a false decrease. Hyponatremia with a high osmolality is associated with hyperglycemia. The elevated levels of glucose increase the serum osmolality and cause a shift of water from the cells to the blood, resulting in a dilution of Na^+.

Table 11.3 Classification of Hyponatremia by Osmolality

With Low Osmolality
Increased sodium loss
Increased water retention
With Normal Osmolality
Increased non-sodium cations
Lithium excess
Increased γ-globulins—cationic (multiple myeloma)
Severe hyperkalemia
Severe hypermagnesemia
Severe hypercalcemia
Pseudohyponatremia
Hyperlipidemia
Hyperproteinemia
Pseudohyperkalemia as a result of in vitro hemolysis
With High Osmolality
Hyperglycemia
Mannitol infusion

Symptoms of hyponatremia. Symptoms depend on the serum level. Between 125 and 130 mmol/L, symptoms are primarily gastrointestinal (GI). More severe neuropsychiatric symptoms are seen below 125 mmol/L, including nausea and vomiting, muscular weakness, headache, lethargy, and ataxia. More severe symptoms also include seizures, coma, and respiratory depression.[5] A level below 120 mmol/L for 48 hours or less (acute hyponatremia) is considered a medical emergency.[3,4] Serum and urine electrolytes are monitored while the patient is treated to return Na^+ levels to normal.

Treatment of hyponatremia. Treatment is directed at correction of the condition that caused either water loss or Na^+ loss in excess of water loss. In addition, the onset of hyponatremia either acute or chronic (less than or more than 48 hours, respectively) and the

severity of hyponatremia are considered in treatment. Conventional treatment of hyponatremia involves fluid restriction and providing hypertonic saline and/or other pharmacologic agents that may take several days to reach the desired effect and may have deleterious side effects.[3,4] Correcting severe hyponatremia too rapidly can cause cerebral myelinolysis. On the other hand, correcting hyponatremia too slowly can cause cerebral edema.[3,4] Appropriate management of fluid administration is critical. Fluid administration and monitoring are required during treatment of the underlying cause of the hyponatremia.

AVP receptor antagonists (known as vaptans) have been found to be an effective treatment for euvolemic or hypervolemic hyponatremia. Conivaptan blocks the action of AVP in the collecting ducts of the nephron, thus decreasing water reabsorption and can be used to treat euvolemic and hypervolemic hyponatremia.[4] Euvolemic hyponatremia is associated with SIADH, adrenal insufficiency, and rarely hypothyroidism.[4] Hypervolemic hyponatremia is associated with liver cirrhosis with ascites, CHF, and overhydrated postoperative patients.[4] Conivaptan is not an effective treatment with hypovolemic hyponatremia because the increased water loss would accentuate the volume depletion problem.

Hypernatremia. **Hypernatremia** (increased serum Na^+ concentration) results from excess loss of water relative to Na^+ loss, decreased water intake, or increased Na^+ intake or retention. Hypernatremia is less commonly seen in hospitalized patients than is hyponatremia. Loss of hypotonic fluid may occur either by the kidney or through profuse sweating, diarrhea, or severe burns.

Hypernatremia may result from loss of water in diabetes insipidus, either because the kidney cannot respond to AVP (nephrogenic diabetes insipidus) or because AVP secretion is impaired (central diabetes insipidus). Diabetes insipidus is characterized by copious production of dilute urine (3 to 20 L/d). Because people with diabetes insipidus drink large volumes of water, hypernatremia usually does not occur unless the thirst mechanism is also impaired. Partial defects of either AVP release or the response to AVP may also occur. In such cases, urine is concentrated to a lesser extent than appropriate to correct the hypernatremia. Excess water loss may also occur in renal tubular disease, such as acute tubular necrosis, in which the tubules become unable to fully concentrate the urine. See **Table 11.4**.

The measurement of urine osmolality is necessary to evaluate the cause of hypernatremia. With renal loss of water, the urine osmolality is low or normal. With extrarenal fluid losses, the urine osmolality is increased. Interpretation of the urine osmolality in hypernatremia is shown in **Table 11.5**.

Water loss through the skin and by breathing (insensible loss) accounts for about 1 L of water loss per day in adults. Any condition that increases water loss, such as fever, burns, diarrhea, or exposure

Table 11.4 Causes of Hypernatremia

Excess Water Loss
Diabetes insipidus
Renal tubular disorder
Prolonged diarrhea
Profuse sweating
Severe burns
Decreased Water Intake
Older persons
Infants
Mental impairment
Increased Intake or Retention
Hyperaldosteronism
Sodium bicarbonate excess
Dialysis fluid excess

Table 11.5 Hypernatremia (150 mmol/L) Related to Urine Osmolality

Urine Osmolality <300 mOsm/kg
Diabetes insipidus (AVP secretion or kidney response)
Urine Osmolality 300–700 mOsm/kg
Partial defect in AVP release or response to AVP
Osmotic diuresis
Urine Osmolality >700 mOsm/kg
Loss of thirst
Insensible loss of water (breathing, skin)
GI loss of hypotonic fluid
Excess intake of sodium

AVP, arginine vasopressin hormone.

to heat, will increase the likelihood of developing hypernatremia. Commonly, hypernatremia occurs in those persons who may be thirsty but who are unable to ask for or obtain water, such as adults with altered mental status and infants. When urine cannot be fully concentrated (e.g., in neonates, young children, older persons, and certain patients with renal insufficiency), a relatively lower urine osmolality may occur.

Chronic hypernatremia in an alert patient is indicative of hypothalamic disease, usually with a defect in the osmoreceptors rather than from a true resetting of the osmostat. A reset osmostat may occur in primary hyperaldosteronism, in which excess aldosterone induces mild hypervolemia that inhibits AVP release, increasing plasma Na^+ upward by approximately 3 to 5 mmol/L.[1]

Hypernatremia may be from excess ingestion of salt or administration of hypertonic solutions of Na^+, such as sodium bicarbonate (used for heartburn/indigestion and correcting acidosis) or hypertonic dialysis solutions. Neonates are especially susceptible to hypernatremia from this cause. In these cases, AVP response is appropriate, resulting in urine osmolality of greater than 800 mOsm/kg (Table 11.5).

Symptoms of hypernatremia. Symptoms most commonly involve the CNS as a result of the hyperosmolar state. These symptoms include altered mental status, lethargy, irritability, restlessness, seizures, muscle twitching, hyper reflexes, fever, nausea or vomiting, difficult respiration, and increased thirst. Serum or plasma Na^+ levels >160 mmol/L are associated with a mortality rate of 60% to 75%.[5]

Treatment of hypernatremia. Treatment is directed at correction of the underlying condition that caused the water depletion or Na^+ retention. The speed of correction depends on the rate with which the condition developed. Hypernatremia must be corrected gradually because too rapid a correction of serious hypernatremia ($\geq$160 mmol/L) can induce cerebral edema and death; the maximal rate should be 1 mmol/L/h.[1]

Determination of Sodium

Specimen. Serum, plasma, and urine are all acceptable for Na^+ measurements. When plasma is used, lithium heparin, ammonium heparin, and lithium oxalate are suitable anticoagulants. Hemolysis does not cause a significant change in serum or plasma values as a result of decreased levels of intracellular Na^+. However, with marked hemolysis, levels may be decreased as a result of a dilutional effect.

Whole blood samples may be used with some analyzers; however, consult the instrument operation manual for acceptability. The specimen of choice in urine Na^+ analysis is a 24-hour collection. Sweat is also suitable for analysis. Sweat collection and analysis are discussed in Chapter 23, *Body Fluid Analysis*.

Methods. Through the years, Na^+ has been measured in various ways, including chemical methods, flame emission spectrophotometry, atomic absorption spectrophotometry (AAS), and ISEs. Chemical methods are outdated because of large sample volume requirements and lack of precision. ISEs are the most routinely used method in clinical laboratories today.

The ISE method uses a semipermeable membrane to develop an electrical potential produced by having different ion concentrations on either side of the membrane. In this type of system, two electrodes are used. One electrode has a constant potential, making it the reference electrode, and the difference in potential between the reference and measuring electrodes can be used to calculate the "concentration" of the ion in solution. However, it is the activity of the ion, not the concentration that is being measured (see Chapter 4, *Analytic Techniques*).

Ion-selective electrodes are selective for the ion but are not absolutely specific because other monovalent cations may react with the electrode but not in the physiologic range. Most analyzers use a glass ion-exchange membrane in its ISE system for Na^+ measurement (**Figure 11.2**).

There are two types of ISE measurement, based on sample preparation: direct and indirect. Direct measurement provides an undiluted sample to interact with the ISE membrane. With the indirect method, a diluted sample is used for measurement. Indirect ISE devices use diluted plasma or serum samples, whereas direct ISE devices use whole (undiluted) blood samples.

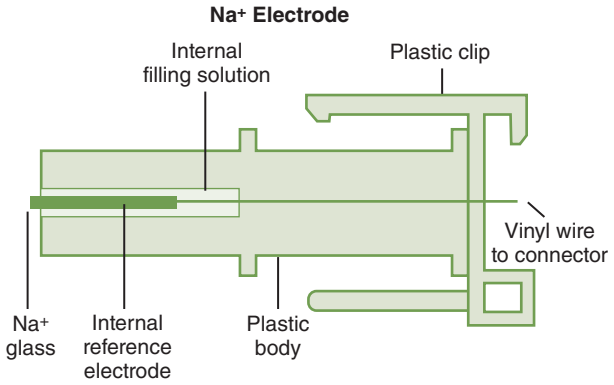


Figure 11.2 Diagram of sodium ISE with glass capillary membrane.

Courtesy of Nova Biomedical, Waltham, MA.

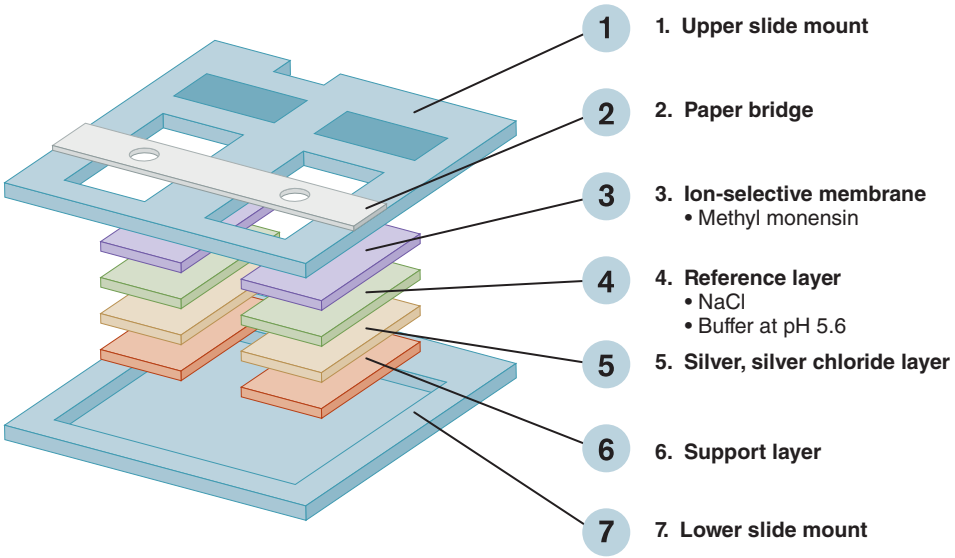


Figure 11.3 Schematic diagram of the ISE system for the potentiometric slide on the VITROS.

Courtesy of OCD, a Johnson & Johnson company, Rochester, NY.

Direct ISE measurement is used in blood gas analyzers whereas indirect ISE measurement is used in most automated chemistry analyzers. There is no significant difference in results, except when samples are hyperlipidemic or hyperproteinemic as discussed previously.

One source of error with ISEs is protein buildup on the membrane through continuous use. The protein-coated membranes cause poor selectivity, which results in poor reproducibility of results. A routine maintenance of these ISEs requires removal of this protein buildup to ensure quality results.

VITROS analyzers (Ortho-Clinical Diagnostics) use a single-use, direct ISE potentiometric system. Each disposable slide contains a reference and measuring electrode (**Figure 11.3**). A drop of sample fluid and a drop of reference fluid are simultaneously applied to the slide, and the potential difference between the two is measured, which is proportional to the Na^+ concentration.[6]

Reference Ranges

See **Table 11.6**.

Table 11.6 Reference Ranges for Sodium

Serum, plasma	135–145 mmol/L
Urine (24 h)	120–240 mmol/d, varies with diet
Cerebrospinal fluid	136–150 mmol/L

Potassium

Potassium (K^+) is the major intracellular cation in the body, with a concentration 20 times greater inside the cells than outside. Many cellular functions require that the body maintain a low ECF concentration of K^+ ions. As a result, only 2% of the body's total K^+ circulates in the plasma. Functions of K^+ in the body include regulation of neuromuscular excitability, contraction of the heart, ICF volume, and H^+ concentration.[7]

The K^+ concentration has a major effect on the contraction of skeletal and cardiac muscles. An elevated plasma K^+ decreases the **resting membrane potential** (**RMP**; ground state cellular response) of the cell, meaning the RMP is closer to zero. This decreases the net difference between the cell's resting potential and threshold (action) potential. A lower than normal difference increases cell overexcitability, leading to muscle weakness. **Hyperkalemia** (increase in K^+ levels) can ultimately cause a lack of muscle excitability as a result of a higher RMP than action potential, which may lead to paralysis or a fatal cardiac arrhythmia.[8] **Hypokalemia** (decreased K^+ levels) decreases cell excitability by increasing the RMP, often resulting in an arrhythmia or paralysis.[8] The heart may cease to contract in extreme cases of either hyperkalemia or hypokalemia.

K^+ concentration also affects the H^+ concentration in the blood. For example, in hypokalemia, as K^+ is lost from the body, Na^+ and H^+ move into the cell to ensure electroneutrality. The H^+ concentration is therefore decreased in the ECF, resulting in alkalosis.

Regulation

Renal function related to tubular reabsorption and secretion is important in the regulation of K^+ balance. Initially, the proximal tubules reabsorb nearly all the K^+. Then, under the influence of aldosterone, additional K^+ is secreted into the urine in exchange for Na^+ in both the distal tubules and the collecting ducts. Thus, the distal nephron is the principal determinant of urinary K^+ excretion. Most individuals consume far more K^+ than needed; the excess is excreted in the urine but may accumulate to toxic levels if renal failure occurs.

K^+ uptake from the ECF into the cells is important in normalizing an acute rise in plasma K^+ concentration due to an increased K^+ intake. Excess plasma K^+ rapidly enters the cells to normalize plasma K^+. As the cellular K^+ gradually returns to the plasma, it is removed by urinary excretion. Note that chronic loss of cellular K^+ may result in cellular depletion before there is an appreciable change in the plasma K^+ concentration because excess K^+ is normally excreted in the urine.

Three factors that influence the distribution of K^+ between cells and ECF are as follows: (1) K^+ loss frequently occurs whenever the Na^+, K^+-ATPase pump is inhibited by conditions such as hypoxia, hypomagnesemia, or digoxin overdose; (2) insulin promotes acute entry of K^+ into skeletal muscle and liver by increasing Na^+, K^+-ATPase activity; and (3) catecholamines, such as epinephrine (β_2-stimulator), promote cellular entry of K^+, whereas propranolol (β-blocker) impairs cellular entry of K^+. Dietary deficiency or excess is rarely a primary cause of hypokalemia or hyperkalemia. However, with a preexisting condition, dietary deficiency or excess can enhance the degree of hypokalemia or hyperkalemia.

Exercise. K^+ is released from muscle cells during exercise, which may increase plasma K^+ by 0.3 to 1.2 mmol/L with mild to moderate exercise and by as much as 2 to 3 mmol/L with exhaustive exercise. These changes are usually reversed after several minutes of rest. Forearm exercise and fist clenching during venipuncture can cause erroneously high plasma K^+ concentrations and should be avoided.

Hyperosmolality. Hyperosmolality, as with uncontrolled diabetes mellitus, causes water to diffuse from the cells, carrying K^+ with the water, which leads to gradual depletion of K^+ if kidney function is normal.

Cellular Breakdown. Cellular breakdown releases K^+ into the ECF. Examples are severe trauma, tumor lysis syndrome, rhabdomyolysis, and massive blood transfusions.

Clinical Applications

Hypokalemia. Hypokalemia is a plasma K^+ concentration below the lower limit of the reference range. Hypokalemia can occur with GI or urinary loss of K^+, or with increased cellular uptake of K^+. Common causes of hypokalemia are shown in **Table 11.7**. GI loss occurs when GI fluid is lost through vomiting, diarrhea, gastric suction, or discharge from an intestinal fistula. Increased K^+ loss in the stool also occurs with certain tumors, malabsorption, cancer therapy (chemotherapy or radiation therapy), and large doses of laxatives.

Table 11.7 Causes of Hypokalemia

Gastrointestinal Loss
Vomiting
Diarrhea
Gastric suction
Intestinal tumor
Malabsorption
Cancer therapy (chemotherapy, radiation therapy)
Large doses of laxatives
Renal Loss
Diuretics (thiazides, mineralocorticoids)
Nephritis
Renal tubular acidosis
Hyperaldosteronism
Cushing's syndrome
Hypomagnesemia
Acute leukemia
Cellular Shift
Alkalosis
Insulin overdose
Hydration
Decreased Intake

Renal loss of K^+ can result from kidney disorders such as K^+-losing nephritis and renal tubular acidosis (RTA). In RTA, as tubular excretion of H^+ decreases, K^+ excretion increases maintaining electroneutrality. Because aldosterone promotes Na^+ retention and K^+ loss, hyperaldosteronism can lead to hypokalemia and metabolic alkalosis.[7] Hypomagnesemia (decreased Mg^{2+} levels) can lead to hypokalemia by promoting urinary loss of K^+ to conserve Mg^{2+}. Mg^{2+} deficiency also diminishes the activity of Na^+, K^+-ATPase and enhances the secretion of aldosterone. Effective treatment may require supplementation with both Mg^{2+} and K^+.[8] Renal K^+ loss also occurs with acute myelogenous leukemia, acute myelomonocytic leukemia, and acute lymphocytic leukemia. Although reduced dietary intake of K^+ rarely causes hypokalemia in healthy persons, decreased intake may intensify hypokalemia caused by use of diuretics, for example.

Increased cellular uptake of potassium is encountered in alkalemia because alkalemia promotes intracellular loss of H^+ to minimize elevation of intracellular pH. As this happens, both K^+ and Na^+ enter cells to preserve electroneutrality, lowering the plasma concentration of K^+ and Na^+. Plasma K^+ decreases by about 0.4 mmol/L per 0.1 unit rise in pH.[8]

Insulin also promotes the entry of K^+ into skeletal muscle and liver cells. Because insulin therapy can sometimes uncover an underlying hypokalemic state, plasma K^+ should be monitored carefully whenever insulin is administered to susceptible patients.[8] A rare cause of hypokalemia is associated with a blood sample from a leukemic patient with a significantly elevated white blood cell count as the K^+ present in the sample is taken up by the white cells if the sample is left at room temperature for several hours.

Symptoms of hypokalemia. Symptoms (e.g., weakness, fatigue, and constipation) often become apparent as plasma K^+ decreases below 3 mmol/L. Hypokalemia can lead to muscle weakness or paralysis, which can interfere with breathing. The dangers of hypokalemia concern all patients but especially those with cardiovascular disorders because of an increased risk of arrhythmia, which may cause sudden death in certain patients. Mild hypokalemia (3.0 to 3.4 mmol/L) is usually asymptomatic.

Treatment of hypokalemia. Treatment typically includes oral KCl replacement over several days. In some instances, intravenous (IV) replacement may be indicated. In some cases, chronic mild hypokalemia may be corrected simply by including food in the diet with high K^+ content, such as dried fruits, nuts, bran cereals, bananas, and orange juice. Plasma electrolytes are monitored as treatment to return K^+ levels to normal occurs.

Hyperkalemia. The most common causes of hyperkalemia are shown in **Table 11.8**. Patients with hyperkalemia often have an underlying disorder, such as renal insufficiency, diabetes mellitus, or metabolic acidosis, that contributes to hyperkalemia.[8] For example, during administration of KCl, a person with renal insufficiency is far more likely to develop hyperkalemia than is a person with normal renal function. The most common cause of hyperkalemia in hospitalized patients is due to therapeutic K^+ administration. The risk is greatest with IV K^+ replacement.

In healthy persons, an acute oral load of K^+ will briefly increase plasma K^+ because most of the absorbed K^+ rapidly moves intracellularly. Normal

Table 11.8 Causes of Hyperkalemia

Decreased Renal Excretion
Acute or chronic renal failure (GFR < 20 mL/min)
Hypoaldosteronism
Addison's disease
Diuretics
Cellular Shift
Acidosis
Muscle/cellular injury
Chemotherapy
Leukemia
Hemolysis
Increased Intake
Oral or intravenous potassium replacement therapy
Artifactual
Sample hemolysis
Thrombocytosis
Prolonged tourniquet use or excessive fist clenching

cellular processes gradually release this excess K^+ back into the plasma, where it is normally removed by renal excretion. Impairment of urinary K^+ excretion is usually associated with chronic hyperkalemia.[8]

If a shift of K^+ from cells into plasma occurs too rapidly to be removed by renal excretion, acute hyperkalemia develops. In diabetes mellitus, insulin deficiency promotes cellular loss of K^+. Hyperglycemia also contributes by producing a hyperosmolar plasma that pulls water and K^+ from cells, promoting further loss of K^+ into the plasma.[8] In metabolic acidosis, as excess H^+ moves intracellularly to be buffered, K^+ leaves the cell to maintain electroneutrality. Plasma K^+ increases by 0.2 to 1.7 mmol/L for each 0.1 unit reduction of pH.[8] Because cellular K^+ often becomes depleted in cases of acidosis with hyperkalemia (including diabetic ketoacidosis), treatment with agents such as insulin and bicarbonate can cause a rapid intracellular movement of K^+, producing severe hypokalemia.

Various drugs may cause hyperkalemia, especially in patients with either renal insufficiency or diabetes mellitus. These drugs include captopril (inhibits ACE), nonsteroidal anti-inflammatory agents (inhibit aldosterone), spironolactone (K^+-sparing diuretic), digoxin (inhibits Na^+–K^+ pump), cyclosporine (inhibits renal response to aldosterone), and heparin therapy (inhibits aldosterone secretion).

Hyperkalemia may result when K^+ is released into the ECF during enhanced tissue breakdown or catabolism, especially if renal insufficiency is present. Increased cellular breakdown may be caused by trauma, administration of cytotoxic agents, hemolysis, rhabdomyolysis, tumor lysis syndrome, and multiple blood transfusions. In banked blood, K^+ is gradually released from erythrocytes during storage, often causing elevated K^+ concentration in plasma supernatant.

Patients on cardiac bypass may develop mild elevations in plasma K^+ during warming after surgery because warming causes cellular release of K^+, whereas hypothermia causes movement of K^+ into cells.

Symptoms of hyperkalemia. Hyperkalemia can cause muscle weakness, tingling, numbness, or mental confusion by altering neuromuscular conduction. Since hyperkalemia disturbs muscle conduction, this can lead to cardiac arrhythmias and possible cardiac arrest. Plasma K^+ concentrations of 6 to 7 mmol/L may alter the electrocardiogram (ECG), and concentrations more than 10 mmol/L may cause fatal cardiac arrest.[8]

Treatment of hyperkalemia. Treatment should be immediately initiated when serum K^+ is 6.0 to 6.5 mmol/L or greater or if there are ECG changes.[8] To offset the effect of K^+, which lowers the resting potential of myocardial cells, Ca^{2+} may be given to reduce the threshold potential of myocardial cells. Therefore, Ca^{2+} provides immediate but short-lived protection to the myocardium against the effects of hyperkalemia. Substances that acutely shift K^+ back into cells, such as sodium bicarbonate, glucose, or insulin, may also be administered. K^+ may be quickly removed from the body by use of loop diuretics or by sodium polystyrene sulfonate (Kayexalate) enemas, which bind to K^+ secreted in the colon. Hemodialysis can be used if other measures fail.[8] Patients treated with these agents must be monitored carefully to prevent hypokalemia as K^+ moves back into cells or is removed from the body.

Collection of Samples

Proper collection and handling of samples for K^+ analysis is extremely important because there are many causes of artifactual hyperkalemia. First, the coagulation process releases K^+ from platelets, so that serum K^+ may be 0.1 to 0.7 mmol/L higher than plasma K^+ concentrations. If the patient's platelet count is elevated (thrombocytosis), serum K^+ may be further elevated. Second, if a tourniquet is left on the arm too long during blood collection or if patients excessively clench their fists or otherwise exercise their forearms before venipuncture, cells may release K^+ into the plasma. The first situation may be avoided by using a heparinized tube to prevent clotting of the specimen and the second by using proper care in the drawing of blood. Third, because storing blood on wet ice promotes the release of K^+ from cells, whole blood samples for K^+ determinations should be stored at room temperature (never iced) and analyzed promptly or centrifuged to remove the cells. Fourth, if hemolysis occurs after the blood is drawn, K^+ may be falsely elevated—the most common cause of artifactual hyperkalemia. Slight hemolysis (≈50 mg/dL of hemoglobin) can cause an increase of approximately 3%, while gross hemolysis (>500 mg/dL of hemoglobin) can cause an increase of up to 30%.

Determination of Potassium

Specimen. Serum, plasma, and urine may be acceptable for analysis. Hemolysis must be avoided because of the high K^+ content of erythrocytes. Heparin is the anticoagulant of choice. Whereas serum

and plasma generally give similar K^+ levels, serum reference ranges tend to be slightly higher due to the process of blood clotting. Significantly elevated platelet counts may result in the release of K^+ during clotting from rupture of these cells, causing a spurious hyperkalemia. In this case, plasma is preferred. Whole blood samples may be used with some analyzers. Urine specimens should be collected over a 24-hour period to eliminate the influence of diurnal variation.

Methods. As with Na^+, the current method of choice is ISE. For ISE measurements, a valinomycin membrane is used to selectively bind K^+, causing an impedance change that can be correlated to K^+ concentration where KCl is the inner electrolyte solution.

Reference Ranges

See **Table 11.9**.

Table 11.9 Reference Ranges for Potassium

Serum	3.5–5.1 mmol/L
Urine (24 h)	33–86 mmol/d

Chloride

Chloride (Cl^-) is the major extracellular anion, and its precise function in the body is not well understood; however, it is involved in maintaining osmolality, blood volume, and electroneutrality. In most processes, Cl^- shifts secondarily to a movement of Na^+ or HCO_3^-.

Cl^- ingested in the diet is almost completely absorbed by the intestinal tract. Cl^- is then filtered out by the glomerulus and passively reabsorbed, in conjunction with Na^+ by the proximal tubules. Excess Cl^- is excreted in the urine and sweat. Excessive sweating stimulates aldosterone secretion, which also acts on the sweat glands to conserve Na^+ and Cl^-.

Cl^- maintains electroneutrality in two ways. First, Na^+ is reabsorbed along with Cl^- in the proximal tubules (Cl^- follows Na^+). In effect, Cl^- acts as the rate-limiting component, in that Na^+ reabsorption is limited by the amount of Cl^- available. Electroneutrality is also maintained by Cl^- through the *chloride shift*. In this process, CO_2 generated by cellular metabolism within the tissue diffuses out into both the plasma and the red cell. In the red cell, CO_2 with H_2O forms carbonic acid (H_2CO_3), which splits into H^+ and HCO_3^- (bicarbonate). Deoxyhemoglobin buffers H^+, whereas the HCO_3^- diffuses out into the plasma and Cl^- diffuses into the red cell to maintain the electric balance of the cell (**Figure 11.4**).

Clinical Applications

Cl^- disorders are often a result of the same causes that disturb Na^+ levels because Cl^- passively follows Na^+. There are a few exceptions. **Hyperchloremia** (increased Cl^- levels) may also occur when there is an excess loss of HCO_3^- as a result of GI losses, RTA, or metabolic acidosis. **Hypochloremia** (decreased Cl^- levels) may also occur with excessive loss of Cl^- from prolonged vomiting, diabetic ketoacidosis, aldosterone deficiency, or salt-losing renal diseases such as pyelonephritis. A low serum level of Cl^- may also be encountered in conditions associated with high serum HCO_3^- concentrations, such as compensated respiratory acidosis or metabolic alkalosis.

Determination of Chloride

Specimen. Serum or plasma may be used, with lithium heparin being the anticoagulant of choice. Hemolysis does not cause a significant change in serum or plasma values as a result of decreased levels of intracellular Cl^-. However, with marked hemolysis, levels may be decreased as a result of a dilutional effect. Whole blood samples may be used with some analyzers. The specimen of choice for urine Cl^- analysis is a 24-hour collection because of diurnal variation. Sweat is also suitable for analysis. Sweat collection and analysis are discussed in Chapter 23, *Body Fluid Analysis*.

Methods. There are several methodologies available for measuring Cl^-, including ISEs, amperometric-coulometric titration, mercurimetric titration, and colorimetry. The most commonly used is ISE. For ISE measurement, an ion-exchange membrane is used to selectively bind Cl^- ions.

Reference Ranges

See **Table 11.10**.

Table 11.10 Reference Ranges for Chloride

Plasma, serum	98–107 mmol/L
Urine (24 h)	110–250 mmol/d, varies with diet

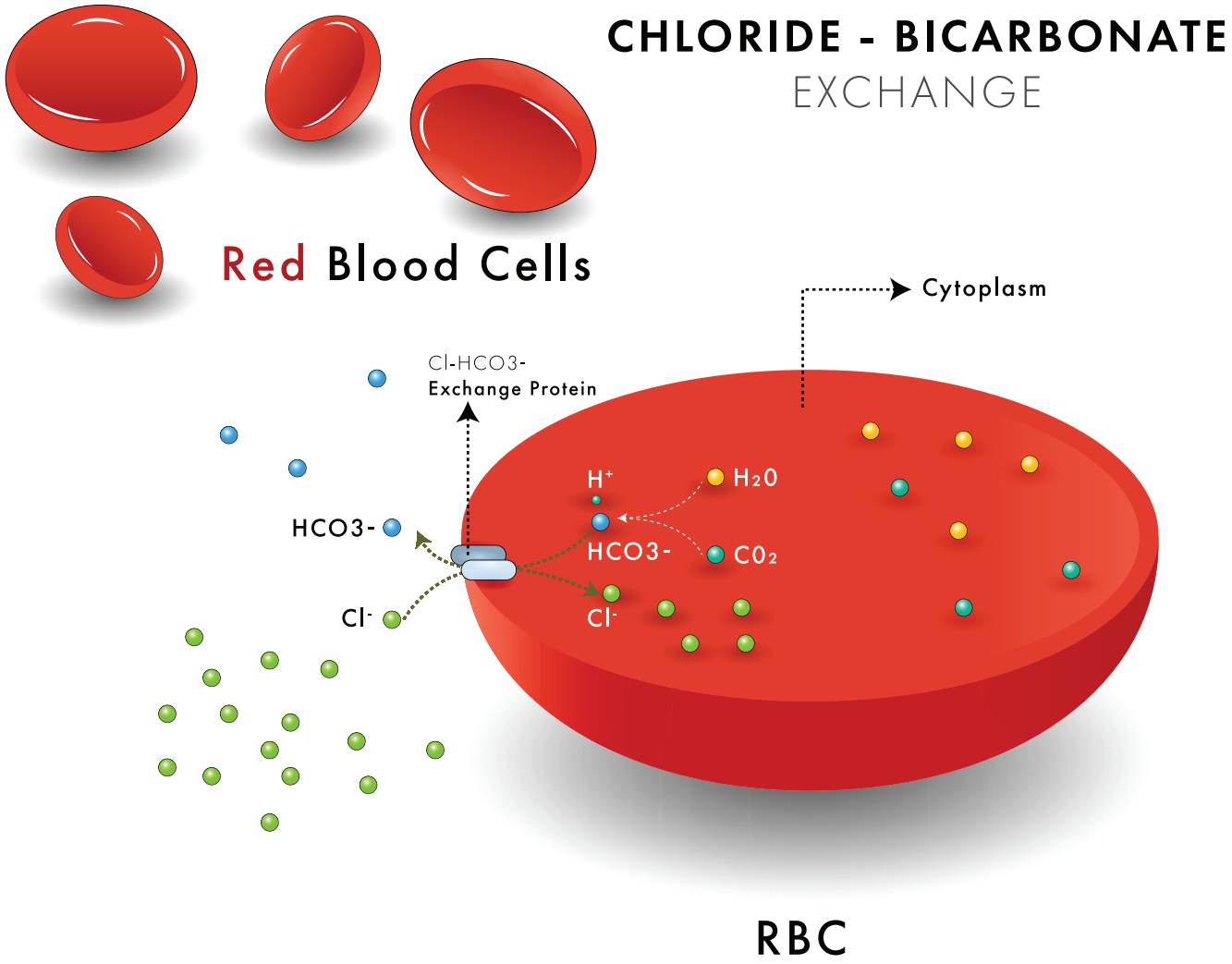


Figure 11.4 Chloride shift mechanism. See text for details.

Bicarbonate

Bicarbonate is the second most abundant anion in the ECF. Total CO_2 comprises the bicarbonate ion HCO_3^-, H_2CO_3, and dissolved CO_2, with HCO_3^- accounting for more than 90% of the total CO_2 at physiologic pH. Because HCO_3^- composes the largest fraction of total CO_2, total CO_2 measurement is indicative of HCO_3^- measurement.

HCO_3^- is the major component of the buffering system in the blood. Carbonic anhydrase in RBCs converts CO_2 and H_2O to H_2CO_3, which dissociates into H^+ and HCO_3^-.

$$CO_2 + H_2O \xleftrightarrow{CA} H^+ + HCO_3^- \qquad \text{(Eq. 11.2)}$$

where CA is carbonic anhydrase. HCO_3^- diffuses out of the cell in exchange for Cl^- to maintain ionic charge neutrality within the cell (chloride shift; see Figure 11.4). This process converts potentially toxic CO_2 in the plasma to an effective buffer: HCO_3^-. HCO_3^- buffers excess H^+ by combining with acid, then eventually dissociating into H_2O and CO_2 in the lungs where the acidic CO_2 gas is eliminated.

Regulation

Most of the HCO_3^- in the kidneys (85%) is reabsorbed by the proximal tubules, with 15% being reabsorbed by the distal tubules. Because tubules are only slightly permeable to HCO_3^-, it is usually reabsorbed as CO_2. This happens as HCO_3^-, after filtering into the tubules, combines with H^+ to form H_2CO_3, which then dissociates into H_2O and CO_2. The CO_2 readily diffuses back into the ECF. Normally, nearly all the HCO_3^- is reabsorbed from the tubules, with little lost in the urine. When HCO_3^- is filtered in excess of the available H^+, almost all excess HCO_3^- flows into the urine.

In alkalosis, with a relative increase in HCO_3^- compared with CO_2, the kidneys increase excretion of HCO_3^- into the urine, carrying along a cation such as Na^+. This loss of HCO_3^- from the body helps correct pH by increasing the ratio of H^+ (acid) to HCO_3^- (base).

During compensatory responses for acidosis, the body will reabsorb more than 90% of HCO_3^- in the proximal and distal tubules. As the kidneys reabsorb HCO_3^-, excretion of H^+ is increased to decrease the ratio of acid to base.

Clinical Applications

Acid–base imbalances cause changes in HCO_3^- and CO_2 levels. A decreased HCO_3^- may occur from metabolic acidosis as HCO_3^- combines with H^+ to produce CO_2, which is exhaled by the lungs. The typical response to metabolic acidosis is compensation by hyperventilation, which lowers pCO_2. Elevated total CO_2 concentrations occur in metabolic alkalosis as HCO_3^- is retained, often with increased pCO_2 due to compensation by hypoventilation. Typical causes of metabolic alkalosis include severe vomiting, hypokalemia, and excessive alkali intake.

Determination of CO_2

Specimen. This chapter deals specifically with venous serum or plasma determinations rather than arterial and whole blood pCO_2 measurements. For discussion of arterial and whole blood pCO_2 measurements, refer to Chapter 12, *Blood Gases, pH, and Buffer Systems*.

Serum or lithium heparin plasma is suitable for analysis. Although specimens should be anaerobic for the highest accuracy, many current analyzers (excluding blood gas analyzers) do not permit anaerobic sample handling. In most instances, the sample is capped until the serum or plasma is separated and the sample is analyzed immediately. If the sample is left uncapped before analysis, CO_2 escapes and levels decrease by approximately 6 mmol/L/h.

CO_2 measurements may be obtained in several ways; however, the actual portion of the total CO_2 being measured may vary with the method used. Two common methods are ISE and an enzymatic method. One type of ISE for measuring total CO_2 uses an acid reagent to convert all the forms of CO_2 to CO_2 gas and is measured by a pCO_2 electrode (see Chapter 4, *Analytic Techniques*). The enzyme method alkalinizes the sample to convert all forms of CO_2 to HCO_3^-. HCO_3^- is used to carboxylate phosphoenolpyruvate (PEP) in the presence of PEP carboxylase, which catalyzes the formation of oxaloacetate:

$$\text{Phosphoenolpyruvate} + HCO_3^- \xrightarrow{\text{PEP carboxylase}} \text{Oxalocetate} + H_2PO_4^- \quad \textbf{(Eq. 11.3)}$$

This is coupled to the following reaction, in which NADH is consumed as a result of the action of malate dehydrogenase:

$$\text{Oxaloacetate} + \text{NADH} + H^+ \xrightarrow{\text{MDH}} \text{Malate} + NAD^+ \quad \textbf{(Eq. 11.4)}$$

The rate of change in absorbance of NADH is proportional to the concentration of HCO_3^-.

Reference Ranges

CO_2, venous 22 to 28 mmol/L (plasma, serum).[2]

Calcium

Calcium Physiology

About 99% of Ca^{2+} in the body is part of bone, and the remaining 1% is mostly in the blood and other ECF, as little is in the cytosol of most cells. Bone contains about 1 kg of calcium and serves as a repository for calcium, phosphate, and magnesium. Ca^{2+} in blood is distributed among several forms. About 45% circulates as free Ca^{2+} ions (referred to as ionized Ca^{2+}); 40% is bound to protein, mostly albumin; and 15% is bound to anions, such as HCO_3^-, citrate, and lactate: however, this distribution can change in disease states.

The concentration of ionized Ca^{2+} in blood is 5000 to 10,000 times higher than in the cytosol of cardiac or smooth muscle cells. Maintenance of this large gradient is vital to maintain the essential rapid inward flux of Ca^{2+} during muscle contractility. Ionized Ca^{2+} is closely regulated and has a mean concentration in humans of about 1.18 mmol/L. Because decreased ionized Ca^{2+} impairs myocardial function, it is important to maintain ionized Ca^{2+} at a near-normal concentration during surgery and in critically ill patients. Decreased ionized Ca^{2+} concentrations in blood can cause neuromuscular irritability, which may become clinically apparent as **tetany** (irregular muscle spasms). It is noteworthy that concentrations of citrate, HCO_3^-, lactate, and albumin can change dramatically during surgery or critical care. This is why ionized Ca^{2+} cannot be reliably calculated from total Ca^{2+} measurements, especially in acutely ill individuals.

Regulation

Three hormones—parathyroid hormone (PTH), vitamin D, and calcitonin—are known to regulate serum Ca^{2+} by altering their secretion rate in response to changes in ionized Ca^{2+} and exert their effects on three organ systems: the gastrointestinal (GI) tract, the kidneys, and the bone. The actions of these hormones are shown in **Figure 11.5**.

PTH secretion from the parathyroid glands is stimulated by a decrease in ionized Ca^{2+}, and conversely, PTH secretion is stopped by an increase in ionized Ca^{2+}. PTH exerts three major effects on both bone and kidney. In the bone, PTH activates a process known as *bone resorption*, in which activated osteoclasts break down bone and subsequently release Ca^{2+} into the ECF. In the kidneys, PTH conserves Ca^{2+} by increasing tubular reabsorption of Ca^{2+} ions. PTH also stimulates renal production of active vitamin D.

Vitamin D_3, a cholecalciferol, is obtained from the diet (seafood, dairy, egg yolks) or exposure of skin to sunlight. Vitamin D_3 is then converted in the liver to 25-hydroxycholecalciferol (25-OH-D_3), still an inactive form of vitamin D. In the kidney, 25-OH-D_3 is hydroxylated again to form 1,25-dihydroxycholecalciferol (1,25-$[OH]_2$-D_3), the biologically active form. This active form of vitamin D increases Ca^{2+} absorption in the intestine and enhances the effect of PTH on bone resorption.

While both PTH and active 1,25$(OH)_2D_3$ raise calcium levels, only calcitonin, which originates in the medullary cells of the thyroid gland, is secreted to decrease calcium levels. Calcitonin exerts its Ca^{2+}-lowering effect by inhibiting the actions of both PTH and vitamin D. Although calcitonin is apparently not secreted during normal regulation of the ionized Ca^{2+} concentration in blood, it is secreted in response to a hypercalcemic stimulus.

Normal calcium homeostasis is achieved by a complex and regulated interplay between vitamin D, PTH, and the intestines, bone, and kidneys. Hormonal control, metabolism, and defects in the function and formation of hormones or organ system disorders can induce disease states. These are reviewed in more detail in Chapter 18, *Parathyroid Function*.

Clinical Applications

Tables 11.11 and **11.12** summarize causes of hypocalcemic and hypercalcemic disorders. Although both total Ca^{2+} and ionized Ca^{2+} measurements are available in many laboratories, ionized Ca^{2+} is usually a more sensitive and specific marker for Ca^{2+} disorders.

Hypocalcemia. **Hypocalcemia** is defined as a decreased calcium level in the blood below normal reference ranges. When there is a decrease or defect of PTH, there will be a decrease in calcium levels.

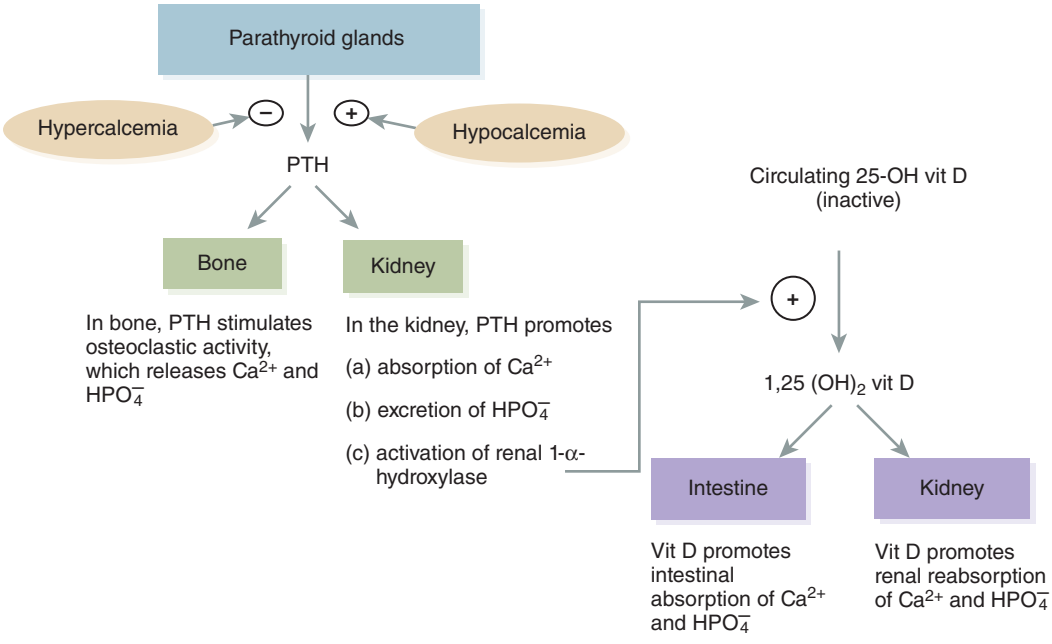


Figure 11.5 Hormonal response to hypercalcemia and hypocalcemia. PTH, parathyroid hormone; 25-OH vit D, 25-hydroxyvitamin D; 1,25$(OH)_2$ vit D, dihydroxyvitamin D.

Table 11.11 Causes of Hypocalcemia

Primary hypoparathyroidism—glandular aplasia, destruction, or removal
Vitamin D deficiency
Pseudohypoparathyroidism
Hypomagnesemia
Hypermagnesemia
Hypoalbuminemia (total only, ionized calcium not affected)—chronic liver disease, nephrotic syndrome, malnutrition
Acute pancreatitis
Renal disease
Rhabdomyolysis

Vitamin D deficiency and malabsorption can cause decreased absorption, which leads to increased PTH production or secondary hyperparathyroidism and hypocalcemia. Pseudohypoparathyroidism is a rare hereditary disorder where PTH response is decreased. These causes are discussed in more detail in Chapter 18, *Parathyroid Function*.

Because **hypomagnesemia** (decreased levels of magnesium) has become more frequent in hospitalized patients, chronic hypomagnesemia has also become recognized as a frequent cause of hypocalcemia. Hypomagnesemia may cause hypocalcemia by three mechanisms: (1) it inhibits the glandular secretion of PTH across the parathyroid gland membrane; (2) it impairs PTH action at its receptor site on bone; and (3) it causes vitamin D resistance.[9] Elevated Mg^{2+} levels (**hypermagnesemia**) may also inhibit PTH release and target tissue response, potentially leading to hypocalcemia and hypercalciuria.

Table 11.12 Causes of Hypercalcemia

Primary hyperparathyroidism—adenoma or glandular hyperplasia
Hyperthyroidism
Increased vitamin D
Benign familial hypocalciuria
Malignancy
Multiple myeloma
Milk alkali syndrome
Thiazide diuretics
Prolonged immobilization

When total Ca^{2+} is the only result reported, hypocalcemia can appear with hypoalbuminemia. Common causes are associated with chronic liver disease, nephrotic syndrome, and malnutrition. In general, for each 1 g/dL decrease in serum albumin, there is a 0.2 mmol/L (0.8 mg/dL) decrease in total Ca^{2+} levels.[10] About one-half of the patients with acute pancreatitis develop hypocalcemia. The most consistent cause appears to be a result of increased intestinal binding of Ca^{2+} as increased intestinal lipase activity occurs.

Patients with renal disease caused by glomerular failure often have altered concentrations of Ca^{2+}, PO_4^-, albumin, Mg^{2+}, and H^+ (pH). In chronic renal disease, secondary hyperparathyroidism frequently develops as the body tries to compensate for hypocalcemia caused either by hyperphosphatemia (PO_4^- binds and lowers ionized Ca^{2+}) or altered vitamin D metabolism. Monitoring and controlling ionized Ca^{2+} concentrations may avoid problems due to hypocalcemia, such as osteodystrophy, unstable cardiac output or blood pressure, or problems arising from hypercalcemia, such as renal stones and other calcifications. Rhabdomyolysis, as with major crush injury and muscle damage, may cause hypocalcemia as a result of increased PO_4^- release from cells, which bind to Ca^{2+} ions.[10]

Surgery and intensive care. Because appropriate Ca^{2+} concentrations promote good cardiac output and maintain adequate blood pressure, the maintenance of a normal ionized Ca^{2+} concentration in blood is beneficial to patients in either surgery or intensive care. Controlling Ca^{2+} concentrations may be critical in open heart surgery when the heart is restarted and during liver transplantation because large volumes of citrated blood are given.

Because these patients may receive large amounts of citrate, HCO_3^-, Ca^{2+} salts, or fluids, the greatest discrepancies between total Ca^{2+} and ionized Ca^{2+} concentrations may be seen during major surgical operations. Consequently, ionized Ca^{2+} measurements are the Ca^{2+} measurement of greatest clinical value.

Hypocalcemia occurs commonly in critically ill patients—that is, those with sepsis, thermal burns, renal failure, or cardiopulmonary insufficiency. These patients frequently have abnormalities of acid–base regulation and losses of protein and albumin, which are best suited to monitoring Ca^{2+} status by ionized

Ca^{2+} measurements. Normalization of ionized Ca^{2+} may have beneficial effects on cardiac output and blood pressure.

Neonatal monitoring. Typically, blood ionized Ca^{2+} concentrations in neonates are high at birth and then rapidly decline by 10% to 20% after 1 to 3 days. After about 1 week, ionized Ca^{2+} concentrations in the neonate stabilize at levels slightly higher than in adults.

The concentration of ionized Ca^{2+} may decrease rapidly in the early neonatal period because the infant may lose Ca^{2+} rapidly and not readily reabsorb it. Several possible etiologies have been suggested: abnormal PTH and vitamin D metabolism, hypercholesterolemia, hyperphosphatemia, and hypomagnesemia.

Symptoms of hypocalcemia. Neuromuscular irritability and cardiac irregularities are the primary groups of symptoms that occur with hypocalcemia. Neuromuscular symptoms include paresthesia, muscle cramps, tetany, and seizures. Cardiac symptoms may include arrhythmia or heart block. Symptoms usually occur with severe hypocalcemia, in which total Ca^{2+} levels are below 1.88 mmol/L (7.5 mg/dL).

Treatment of hypocalcemia. Oral or parenteral Ca^{2+} therapy may be used to treat hypocalcemia, depending on the severity of the decreased level and the cause. Vitamin D may sometimes be administered in addition to oral Ca^{2+} to increase absorption. If hypomagnesemia is a concurrent disorder, Mg^{2+} therapy should also be provided.

Hypercalcemia. **Hypercalcemia** is an abnormal increase of calcium in the bloodstream. Primary hyperparathyroidism (excess secretion of PTH) is the main cause of hypercalcemia and may show obvious clinical signs or may be asymptomatic.[10] Although either total or ionized Ca^{2+} measurements are elevated in serious cases, ionized Ca^{2+} is more frequently elevated in subclinical or asymptomatic hyperparathyroidism. The second leading cause of hypercalcemia is associated with various types of malignancy, with hypercalcemia sometimes being the sole biochemical marker for disease. A rare, benign form of hypercalcemia called familial hypocalciuric hypercalcemia has also been reported. These are discussed in more detail in Chapter 18, *Parathyroid Function*.

Because of the proximity of the parathyroid glands to the thyroid gland, hyperthyroidism can sometimes cause hyperparathyroidism. Due to the direct effects of thyroxine on the bones, there is increased resorption and hypercalcemia, leaving PTH levels low.

Milk alkali syndrome in its classic form described in the 1930s follows the oral administration of very high (>20 g) amount of calcium and large amounts of milk to control gastric acid levels in patients with peptic ulcer disease. The hallmark of this disorder includes hypercalcemia, alkalosis, and renal impairment. Serum phosphate is often elevated. It is very rare to see the typical form of MAS now because of the use of proton pump inhibitors and H_2 blockers.

Thiazide diuretics (such as hydrochlorothiazide, HCTZ) used to treat hypertension increase Ca^{2+} reabsorption but do not usually cause hypercalcemia. However, in subjects with other conditions, such as subclinical primary hyperparathyroidism, hypercalcemia can be uncovered. Lithium carbonate (lithium), when used to treat bipolar disorder, can cause hypercalcemia. Lithium reduces intracellular formation of inositol triphosphate levels altering the "set point" and increasing Ca^{2+} concentrations. Hypervitaminosis D is a condition that may result from excessive intake of vitamin D. It may also result from aberrant production of $1,25(OH)_2D_3$ as a result of extrarenal 1α-hydroxylation of 25-hydroxy vitamin D. High doses of vitamin A, or vitamin A analogs/metabolites in the retinoic acid family, may cause hypercalcemia. Vitamin A is believed to activate osteoclasts and enhance bone resorption, elevating blood calcium. In this condition, both PTH and $1,25(OH)_2D$ are suppressed.

Prolonged immobilization may cause increased bone resorption. Hypercalcemia associated with immobilization is further compounded by renal insufficiency.

Symptoms of hypercalcemia. A mild hypercalcemia (2.62 to 3.00 mmol/L [10.5 to 12 mg/dL]) is often asymptomatic. Moderate or severe Ca^{2+} elevations include neurologic, GI, and renal symptoms. Neurologic symptoms may include mild drowsiness or weakness, depression, lethargy, and coma. GI symptoms may include constipation, nausea, vomiting, anorexia, and peptic ulcer disease. Hypercalcemia may cause renal symptoms of nephrolithiasis and nephrocalcinosis. Hypercalciuria can result in nephrogenic diabetes insipidus, which causes polyuria that results in hypovolemia, which further aggravates the hypercalcemia. Hypercalcemia can also cause symptoms of digitalis toxicity.

Treatment of hypercalcemia. Treatment of hypercalcemia depends on the level of hypercalcemia and the cause. People with primary hyperparathyroidism often are asymptomatic. Estrogen deficiency in postmenopausal women can increased PTH secretion in older women.[10] Often, estrogen replacement therapy reduces Ca^{2+} levels. Parathyroidectomy may be necessary in some hyperparathyroid patients. Patients with moderate to severe hypercalcemia are treated to reduce Ca^{2+} levels. Salt and water intake is encouraged to increase Ca^{2+} excretion and avoid dehydration, which can compound the hypercalcemia. Thiazide diuretics should be discontinued. Bisphosphonates (a derivative of pyrophosphate) are the main drug class used to lower Ca^{2+} levels, achieved by its binding action to bone, which prevents bone resorption.

Determination of Calcium

Specimen. The preferred specimen for total Ca^{2+} determinations is either serum or lithium heparin plasma collected without venous stasis. Because anticoagulants such as ethylenediaminetetraacetic acid (EDTA) or oxalate bind Ca^{2+} tightly and interfere with measurement, they are unacceptable for use.

The proper collection of samples for ionized Ca^{2+} measurements require great care. Because loss of CO_2 will increase the pH of the sample, ionized Ca^{2+} levels will be affected. As pH increases, ionized Ca^{2+} levels decrease as more Ca^{2+} becomes bound to proteins (albumin). **Figure 11.6** shows the effects of pH on ionized calcium. Because of this, samples must be collected anaerobically. Although heparinized whole blood is the preferred sample, serum from sealed evacuated blood collection tubes may be used if clotting and centrifugation are done quickly (<30 minutes) and at room temperature. No liquid heparin products should be used. Most heparin anticoagulants (Na^+ and lithium) partially bind to Ca^{2+} and lower ionized Ca^{2+} concentrations. A heparin concentration of 25 IU/mL, for example, decreases ionized Ca^{2+} by about 3%. Dry heparin products available are titrated with small amounts of Ca^{2+} or Zn^{2+} or with small amounts of heparin dispersed in an inert "puff" that essentially eliminates the interference by heparin.

For analysis of Ca^{2+} in urine, an accurately timed urine collection is preferred. The urine should be acidified with 6 mol/L HCl, with approximately 1 mL of the acid added for each 100 mL of urine.

Methods. The two commonly used methods for total Ca^{2+} analysis use either orthocresolphthalein complexone (CPC) or arsenazo III dye to form a complex with Ca^{2+}. Prior to the dye-binding reaction, Ca^{2+} is released from its protein carrier and complexes by acidification of the sample. The CPC method uses 8-hydroxyquinoline to prevent Mg^{2+} interference. AAS remains the reference method for total Ca^{2+}, although it is rarely used in the clinical setting.

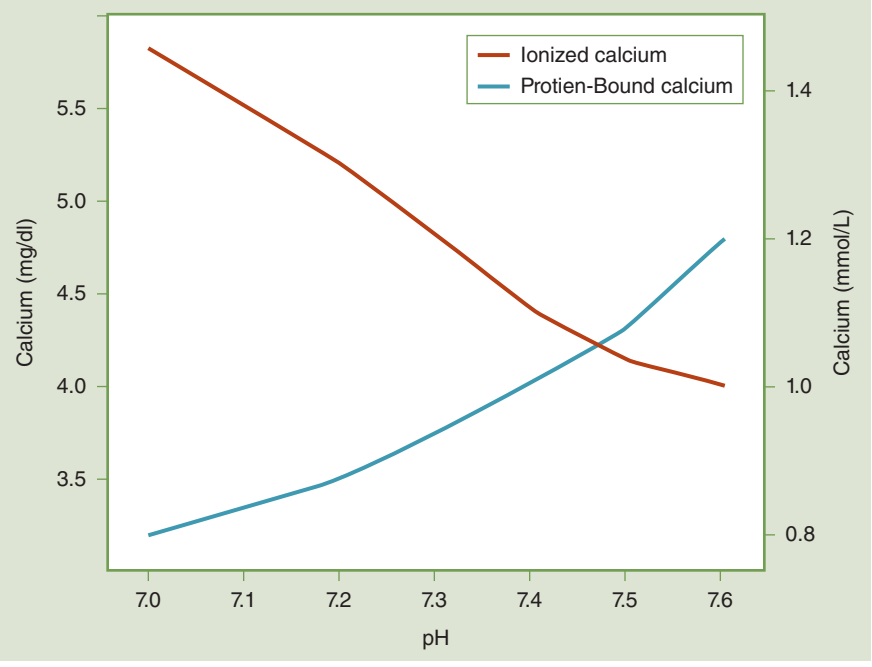


Figure 11.6 pH change of ionized calcium.

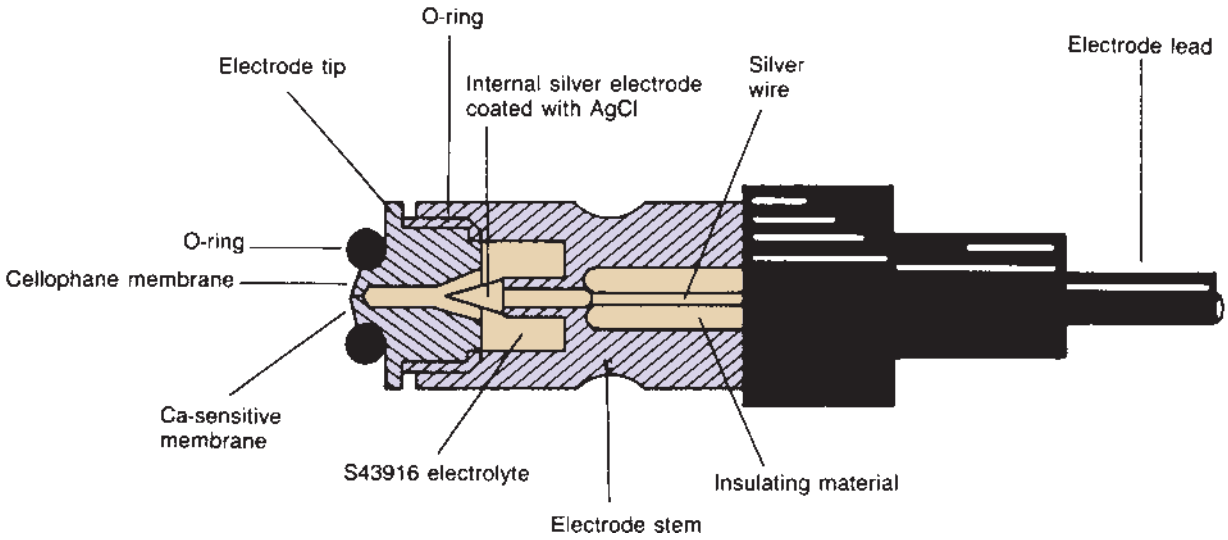


Figure 11.7 Diagram of ionized calcium electrode for the ionized calcium analyzer.

Current commercial analyzers that measure ionized/free Ca^{2+} use ISEs for this measurement. These systems may use membranes impregnated with special molecules that selectively, but reversibly, bind Ca^{2+} ions. As Ca^{2+} binds to these membranes, an electric potential develops across the membrane that is proportional to the ionized Ca^{2+} concentration. A diagram of one such electrode is shown in **Figure 11.7**.

Reference Ranges

For total Ca^{2+}, the reference range varies slightly with age. In general, Ca^{2+} concentrations are higher through adolescence when bone growth is most active. Ionized/free Ca^{2+} concentrations can change rapidly from day 1 to day 3 of life. Following this, they stabilize at relatively high levels, with a gradual decline through adolescence; see **Table 11.13**.

Table 11.13 Reference Ranges for Calcium

Total Calcium—Serum, Plasma	
Child, <3 y	2.13–2.63 mmol/L (8.5–10.5 mg/dL)
Adult	2.24–2.53 mmol/L (9.0–10.1 mg/dL)
Ionized Calcium—Serum	
Adult	1.15–1.33 mmol/L (4.6–5.3 mg/dL)
Ionized Calcium—Whole Blood	
Adult	1.15–1.27 mmol/L (4.6–5.1 mg/dL)

Magnesium

Magnesium Physiology

Magnesium (Mg^{2+}) is the fourth most abundant cation in the body and second most abundant intracellular ion after K^+. The average human body (70 kg) contains 1 mol (24 g) of Mg^{2+}. Approximately 53% of Mg^{2+} in the body is found in bone and 46% in muscle, other organs, and soft tissue, and less than 1% is present in serum and RBCs.[10] Of the Mg^{2+} present in serum, about one-third is bound to protein, primarily albumin. Of the remaining two-thirds, 61% exists in the free or ionized state, and about 5% is complexed with other ions, such as PO_4^- and citrate. Similar to Ca^{2+}, it is the free ion that is physiologically active in the body.[10]

The role of Mg^{2+} in the body is widespread. It is an essential cofactor of more than 300 enzymes, including those important in glycolysis; transcellular ion transport; neuromuscular transmission; synthesis of carbohydrates, proteins, lipids, and nucleic acids; and the release of and response to certain hormones.

The clinical usefulness of serum Mg^{2+} levels has greatly increased as more information about the analyte has been discovered. The most significant findings are the relationship between abnormal serum Mg^{2+} levels and cardiovascular, metabolic, and neuromuscular disorders. Although serum levels may not reflect total body stores of Mg^{2+}, they are useful in determining acute changes in the ion.

Regulation

Rich sources of Mg^{2+} in the diet include raw nuts, dry cereal, and "hard" drinking water; other sources include vegetables, meats, fish, and fruit. Processed foods, an ever-increasing part of the average U.S. diet, have low levels of Mg^{2+} that may cause an inadequate intake. This in turn may increase the likelihood of Mg^{2+} deficiency. The small intestine may absorb 20% to 65% of the dietary Mg^{2+}, depending on the need and intake.

The overall regulation of body Mg^{2+} is controlled largely by the kidney, which can reabsorb Mg^{2+} in deficiency states or readily excrete excess Mg^{2+} in overload states. Of the nonprotein-bound Mg^{2+} that gets filtered by the glomerulus, only 25% to 30% is reabsorbed by the proximal convoluted tubule (PCT), unlike Na^+, in which 60% to 75% is absorbed in the PCT. Henle's loop is the major renal regulatory site, where 50% to 60% of filtered Mg^{2+} is reabsorbed in the ascending limb. In addition, 2% to 5% is reabsorbed in the distal convoluted tubule.[10] The renal threshold for Mg^{2+} is approximately 0.60 to 0.85 mmol/L (≈1.46 to 2.07 mg/dL). Because this is close to normal serum concentration, slight excesses of Mg^{2+} in serum are rapidly excreted by the kidneys. Normally, only about 4% to 6% of filtered Mg^{2+} is excreted in the urine per day.[10]

Mg^{2+} regulation appears to be related to that of Ca^{2+} and Na^+. Parathyroid hormone (PTH) increases the renal reabsorption of Mg^{2+} and enhances the absorption of Mg^{2+} in the intestine. However, changes in ionized Ca^{2+} have a far greater effect on PTH secretion.

Clinical Applications

Hypomagnesemia. Hypomagnesemia, decreased levels of Mg^{2+}, is most frequently observed in hospitalized individuals in intensive care units (ICUs) or those receiving diuretic therapy or digitalis therapy (to treat CHF, atrial fibrillation). These patients most likely have an overall tissue depletion of Mg^{2+} as a result of severe illness or loss, which leads to low serum levels. Hypomagnesemia occurs in about 15% of nonhospitalized individuals.[10]

There are many causes of hypomagnesemia; however, it can be grouped into general categories (**Table 11.14**). Reduced intake is least likely to cause severe deficiencies in the United States. A Mg^{2+}-deficient diet as a result of starvation, chronic alcoholism, or Mg^{2+}-deficient IV therapy can cause a loss of the ion.

Table 11.14 Causes of Hypomagnesemia

Causes Related to Increased Excretion	
Renal excretion	Tubular disorder Glomerulonephritis Pyelonephritis
Endocrine excretion	Hyperparathyroidism Hyperaldosteronism Hyperthyroidism Hypercalcemia Diabetic ketoacidosis
Drug-induced excretion	Diuretics Antibiotics Cyclosporine Digitalis
Causes Related to Decreased Absorption	Malabsorption syndrome Surgical resection of small intestine Nasogastric suction Pancreatitis Vomiting Diarrhea Laxative abuse Neonatal Primary Congenital
Causes Related to Reduced Intake	Poor diet/starvation Prolonged magnesium-deficient intravenous therapy Chronic alcoholism
Miscellaneous Causes	Excess lactation Pregnancy

Modified from Polancic JE. Magnesium: metabolism, clinical importance, and analysis. *Clin Lab Sci.* 1991;4(2):105–109.

Various GI disorders may cause decreased absorption by the intestine, which can result in an excess loss of Mg^{2+} via the feces. Malabsorption syndromes; intestinal resection or bypass surgery; nasogastric suction; pancreatitis; and prolonged vomiting, diarrhea, or laxative use may lead to an Mg^{2+} deficiency. Neonatal hypomagnesemia has been reported as a result of various surgical procedures.

Mg^{2+} loss due to increased excretion by way of urine can occur as a result of various renal and endocrine disorders or the effects of certain drugs on the

kidneys. Renal tubular disorders and other select renal disorders may result in excess amounts of Mg^{2+} being lost through the urine because of decreased tubular reabsorption.

Several endocrine disorders can cause a loss of Mg^{2+}. Hyperparathyroidism and hypercalcemia may cause increased renal excretion of Mg^{2+} as a result of excess Ca^{2+} ions. Excess serum Na^{+} levels caused by hyperaldosteronism may also cause increased renal excretion of Mg^{2+}. In both of these situations, the body is trying to maintain electroneutrality and removes Mg^{2+} in order to maintain balance. A pseudohypomagnesemia may also be the result of hyperaldosteronism caused by increased water reabsorption (dilutional effect). Hyperthyroidism may result in an increased renal excretion of Mg^{2+} and may also cause an intracellular shift of the ion. In persons with diabetes, excess urinary loss of Mg^{2+} is associated with glycosuria. Hypomagnesemia can aggravate the neuromuscular and vascular complications commonly found in this disease. Some studies have shown a relationship between Mg^{2+} deficiency and insulin resistance. Mg^{2+} is not thought to play a role in the pathophysiology of diabetes mellitus; however, magnesium does show an inverse relationship with diabetes.[11]

Several drugs, including diuretics, gentamicin (an antibiotic), cisplatin (a chemotherapeutic agent), and cyclosporine (an immunosuppressant) increase renal loss of Mg^{2+} and frequently result in hypomagnesemia. The loop diuretics, such as furosemide, are especially effective in increasing renal loss of Mg^{2+}. Thiazide diuretics require a longer period of use to cause hypomagnesemia. Cisplatin has a nephrotoxic effect that inhibits the ability of the renal tubule to conserve Mg^{2+}. Cyclosporine severely inhibits the renal tubular reabsorption of Mg^{2+} and has many adverse effects, including nephrotoxicity, hypertension, hepatotoxicity, and neurologic symptoms such as seizures and tremors. Cardiac glycosides, such as digoxin and digitalis, can interfere with Mg^{2+} reabsorption; the resulting hypomagnesemia is a significant finding because the decreased level of Mg^{2+} can amplify the symptoms of digitalis toxicity.

Excess lactation has been associated with hypomagnesemia as a result of increased use and loss through milk production. Mild deficiencies have been reported in pregnancy, which may cause a hyperexcitable uterus, anxiety, and insomnia.

Symptoms of hypomagnesemia. A patient who has hypomagnesemia may be asymptomatic until serum levels fall below 0.5 mmol/L.[10] A variety of symptoms can occur, and the most frequent involve cardiovascular, neuromuscular, psychiatric, and metabolic abnormalities (**Table 11.15**). The cardiovascular and neuromuscular symptoms result primarily from the ATPase enzyme's requirement for Mg^{2+}. Mg^{2+} loss leads to decreased intracellular K^{+} levels because of a faulty Na^{+}/K^{+} pump (ATPase). This change in cellular RMP causes increased excitability that may lead to cardiac arrhythmias. This condition may also lead to digitalis toxicity.

Muscle contraction also requires Mg^{2+} and ATPase for normal Ca^{2+} uptake following contraction. Normal nerve and muscle cell stimulation requires Mg^{2+} to assist with the regulation of acetylcholine, a potent neurotransmitter. Hypomagnesemia can cause a variety of symptoms, from weakness to tremors, tetany,

Table 11.15 Symptoms of Hypomagnesemia

Cardiovascular
Arrhythmia
Hypertension
Digitalis toxicity
Neuromuscular
Weakness
Cramps
Ataxia
Tremor
Seizure
Tetany
Paralysis
Coma
Psychiatric
Depression
Agitation
Psychosis
Metabolic
Hypokalemia
Hypocalcemia
Hypophosphatemia
Hyponatremia

Data from Polancic JE. Magnesium: metabolism, clinical importance, and analysis. *Clin Lab Sci.* 1991;4(2):105–109.

paralysis, or coma. The CNS can also be affected, resulting in psychiatric disorders that range from subtle changes to depression or psychosis.

Metabolic disorders are also associated with hypomagnesemia. Studies have indicated that approximately 40% of hospitalized patients with hypokalemia also have hypomagnesemia.[10] In addition, 20% to 30% of patients with hyponatremia, hypocalcemia, or hypophosphatemia also have hypomagnesemia. Mg^{2+} deficiency can impair PTH release and target tissue response, resulting in hypocalcemia. Replenishing any of these deficient ions alone often does not remedy the disorder unless Mg^{2+} therapy is provided, whereas Mg^{2+} therapy alone may restore both ion levels to normal. However, serum levels of the ions must be monitored during treatment.

Treatment of hypomagnesemia. The preferred form of treatment is by oral intake using magnesium lactate, magnesium oxide, or magnesium chloride or an antacid that contains Mg^{2+}. In severely ill patients, an $MgSO_4$ solution is given parenterally. Before initiation of therapy, renal function must be evaluated to avoid inducing hypermagnesemia during treatment.[10]

Hypermagnesemia

Hypermagnesemia, increased levels of Mg^{2+}, is observed less frequently than is hypomagnesemia.[10] Causes for elevated serum Mg^{2+} levels are summarized in **Table 11.16**; the most common is renal failure (GFR < 15 mL/min). The most severe elevations are usually a result of the combined effects of decreased renal function and increased intake of commonly prescribed Mg^{2+}-containing medications, such as antacids, enemas, or laxatives. Nursing home patients are at greatest risk for this occurrence.

Table 11.16 Causes of Hypermagnesemia

Causes Related to Decreased Excretion	Acute or chronic renal failure Hypothyroidism Hypoaldosteronism Hypopituitarism (↓growth hormone)
Causes Related to Increased Intake	Antacids Enemas Cathartics Therapeutic—eclampsia, cardiac arrhythmia
Miscellaneous Causes	Dehydration Bone carcinoma Bone metastases

Data from Polancic JE. Magnesium: metabolism, clinical importance, and analysis. *Clin Lab Sci.* 1991;4(2):105–109.

Hypermagnesemia has been associated with several endocrine disorders. Thyroxine and growth hormone cause a decrease in tubular reabsorption of Mg^{2+}, and a deficiency of either hormone may cause a moderate elevation in serum Mg^{2+}. Adrenal insufficiency may cause a mild elevation as a result of decreased renal excretion of Mg^{2+}.[12]

CASE STUDY 11.1, PART 3

Alyssa had some additional labs tests before IV fluids were started.

Analyte	Patient Value	Reference Range
Na^+	129 mmol/L	135–145 mmol/L
K^+	3.0 mmol/L	3.5–5.1 mmol/L
Cl^-	77 mmol/L	98–107 mmol/L
HCO_3^-	9 mmol/L	22–28 mmol/L
Mg^+	0.45 mmol/L	0.66–1.07 mmol/L
Ionized Ca^{2+}	1.0 mmol/L	1.15–1.33 mmol/L

2. Explain how the sodium, potassium, chloride, and urine sodium results relate to Alyssa's condition.
3. Explain why her ionized calcium is low, and how it's affected by magnesium.

$MgSO_4$ may be used therapeutically with preeclampsia, cardiac arrhythmia, or myocardial infarction; Mg^{2+} is a vasodilator and can decrease uterine hyperactivity in eclamptic states and increase uterine blood flow. This therapy can lead to maternal hypermagnesemia, as well as neonatal hypermagnesemia due to the immature kidney of the newborn. Premature infants are at greater risk to develop actual symptoms.[12] Studies have shown that IV Mg^{2+} therapy in myocardial infarction patients may reduce early mortality.[12] Dehydration can cause a pseudohypermagnesemia, which can be corrected with rehydration. Because of increased bone loss, mild serum Mg^{2+} elevations can occur in individuals with multiple myeloma or bone metastases.

Symptoms of hypermagnesemia. Symptoms of hypermagnesemia typically do not occur until the serum level exceeds 1.5 mmol/L.[12] The most frequent symptoms involve cardiovascular, dermatologic, GI, neurologic, neuromuscular, metabolic, and hemostatic abnormalities (**Table 11.17**). Mild to moderate symptoms, such as hypotension, bradycardia, skin flushing, increased skin temperature, nausea, vomiting, and lethargy, may occur when serum levels are 1.5 to 2.5 mmol/L.[12] Life-threatening symptoms, such as ECG changes, heart block, asystole, sedation, coma, respiratory depression or arrest, and paralysis, can occur when serum levels reach 5.0 mmol/L.[12]

Elevated Mg^{2+} levels may inhibit PTH release and target tissue response. This may lead to hypocalcemia and hypercalcuria.[12] Normal hemostasis is a Ca^{2+}-dependent process that may be inhibited as a result of competition between increased levels of Mg^{2+} and Ca^{2+} ions. Thrombin generation and platelet adhesion are two processes in which interference may occur.[12]

Treatment of hypermagnesemia. Treatment of Mg^{2+} excess associated with increased intake is to discontinue the source of Mg^{2+}. Severe symptomatic hypermagnesemia requires immediate supportive therapy for cardiac, neuromuscular, respiratory, and/or neurologic abnormalities. Patients with renal failure require hemodialysis. Patients with normal renal function may be treated with a diuretic and IV fluid.

Table 11.17 Symptoms of Hypermagnesemia

Cardiovascular	Neuromuscular
Hypotension	*Decreased Reflexes*
Bradycardia	Dysarthria
Heart block	Respiratory depression Paralysis
Dermatologic	*Metabolic*
Flushing	Hypocalcemia Warm skin
Gastrointestinal	*Hemostatic*
Nausea	Decreased thrombin generation
Vomiting	Decreased platelet adhesion
Neurologic	
Lethargy Coma	

Data from Polancic JE. Magnesium: metabolism, clinical importance, and analysis. *Clin Lab Sci.* 1991;4(2):105–109.

Determination of Magnesium

Specimen. Nonhemolyzed serum or lithium heparin plasma may be analyzed. Because the Mg^{2+} concentration inside erythrocytes is 10 times greater than that in the ECF, hemolysis must be avoided, and the serum should be separated from the cells as soon as possible. Oxalate, citrate, and EDTA anticoagulants are unacceptable because they will bind with Mg^{2+}. A 24-hour urine sample is preferred for urine analysis because of a diurnal variation in excretion. The urine must be acidified with HCl to avoid precipitation.

Methods. The three most common colorimetric methods for measuring total serum Mg^{2+} are calmagite, formazan dye, and methylthymol blue. In the calmagite method, Mg^{2+} binds with calmagite to form a reddish-violet complex that may be read at 532 nm. In the formazan dye method, Mg^{2+} binds with the dye to form a colored complex that may be read at 660 nm. In the methylthymol blue method, Mg^{2+} binds with the chromogen to form a colored complex. Most methods use a Ca^{2+} shelter to prohibit interference from this divalent cation. The reference method for measuring Mg^{2+} is AAS.

Although the measurement of total Mg^{2+} concentrations in serum remains the usual diagnostic test for detection of Mg^{2+} abnormalities, it has limitations. First, because approximately 25% of Mg^{2+} is protein bound, total Mg^{2+} may not reflect the physiologically active, free ionized Mg^{2+}. Second, because Mg^{2+} is primarily an intracellular ion, serum concentrations will not necessarily reflect the status of intracellular Mg^{2+}; even when tissue and cellular Mg^{2+} is depleted by as much as 20%, serum Mg^{2+} concentrations may remain normal.

Reference Ranges

See **Table 11.18**.

Table 11.18 Reference Range for Magnesium

Serum, colorimetric	0.66–1.07 mmol/L (1.7–2.4 mg/dL)

Phosphate

Phosphate Physiology

Found everywhere in living cells, phosphate (PO_4^-) compounds participate in many of the most important biochemical processes. The genetic materials deoxyribonucleic acid and ribonucleic acid are complex phosphodiesters. Most coenzymes are esters of phosphoric or pyrophosphoric acid. The most important reservoirs of biochemical energy are ATP, creatine phosphate, and PEP. Phosphate deficiency can lead to ATP depletion, which is ultimately responsible for many of the clinical symptoms observed in **hypophosphatemia** (decreased concentration of phosphate).

Alterations in the concentration of 2,3-bisphosphoglycerate (2,3-BPG) in RBCs affect the affinity of hemoglobin for oxygen, with an increase facilitating the release of oxygen in tissue and a decrease making oxygen bound to hemoglobin less available. By affecting the formation of 2,3-BPG, the concentration of inorganic phosphate indirectly affects the release of oxygen from hemoglobin.

Understanding the cause of an altered phosphate concentration in the blood is often difficult because transcellular shifts of phosphate are a major cause of hypophosphatemia in blood. That is, an increased shift of phosphate into cells can deplete phosphate in the blood. Once phosphate is taken up by the cell, it remains there to be used in the synthesis of phosphorylated compounds (e.g., ATP, creatine phosphate). As these phosphate compounds are metabolized, inorganic phosphate slowly leaks out of the cell into the blood, where it is regulated principally by the kidney.

Regulation

Phosphate in blood may be absorbed in the intestine from dietary sources, released from cells into blood, and lost from bone. In healthy individuals, all these processes are relatively constant and easily regulated by renal excretion or reabsorption of phosphate.

Disturbances to any of these processes can alter phosphate concentrations in the blood; however, the loss of regulation by the kidneys will have the most profound effect. Although other factors, such as vitamin D, calcitonin, growth hormone, and acid–base status, can affect renal regulation of phosphate, the most important factor is PTH, which overall lowers blood concentrations by increasing renal excretion.

Vitamin D acts to increase phosphate in the blood. Vitamin D increases both phosphate absorption in the intestine and phosphate reabsorption in the kidney.

Growth hormone, which helps regulate skeletal growth, can affect circulating concentrations of phosphate. In cases of excessive secretion or administration of growth hormone, phosphate concentrations in the blood may increase because of decreased renal excretion of phosphate.

Distribution

Although the concentration of all phosphate compounds in blood is about 12 mg/dL (3.9 mmol/L), most of that is organic phosphate, and only about 3 to 4 mg/dL is inorganic phosphate. Phosphate is the predominant intracellular anion, with intracellular concentrations varying, depending on the type of cell. About 80% of the total body pool of phosphate is contained in bone, 20% in soft tissues, and less than 1% is active in the serum/plasma.[10]

Clinical Applications

Hypophosphatemia. Hypophosphatemia occurs in about 1% to 5% of hospitalized patients. The incidence of hypophosphatemia increases to 20% to 40% in patients with the following disorders: diabetic ketoacidosis, chronic obstructive pulmonary disease, asthma, malignancy, long-term treatment with total parenteral nutrition, inflammatory bowel disease, anorexia nervosa, and alcoholism. The incidence increases to 60% to 80% in ICU patients with sepsis. In addition, hypophosphatemia can also be caused by increased renal excretion, as with hyperparathyroidism, and decreased intestinal absorption, as with vitamin D deficiency or antacid use. Although most cases are moderate and seldom cause problems, severe hypophosphatemia (<1.0 g/dL or 0.3 mmol/L) requires monitoring and possible replacement therapy. There is a 30% mortality rate in those who have severe hypophosphatemia versus a 15% rate in those with mild to moderate hypophosphatemia.

Hyperphosphatemia. Patients at greatest risk for **hyperphosphatemia** (increased concentration of phosphate) are those with renal disease.[10] An

increased intake of phosphate or increased release of cellular phosphate may also cause hyperphosphatemia. Because they may not yet have developed mature PTH and vitamin D metabolism, neonates are especially susceptible to hyperphosphatemia caused by increased phosphate intake, such as from cow's milk or laxatives. Increased breakdown of cells can sometimes lead to hyperphosphatemia, as with severe infections, intensive exercise, neoplastic disorders, or intravascular hemolysis. Because immature lymphoblasts have about four times the phosphate content of mature lymphocytes, patients with lymphoblastic leukemia are especially susceptible to hyperphosphatemia. Hypoparathyroidism may also cause hyperphosphatemia.

Determination of Inorganic Phosphorus

Specimen. Serum or lithium heparin plasma is acceptable for analysis. Oxalate, citrate, or EDTA anticoagulants should not be used because they interfere with the analytic method. Hemolysis should be avoided because of the higher concentrations inside RBCs. Circulating phosphate levels are subject to circadian rhythm, with highest levels in late morning and lowest in the evening. Urine analysis for phosphate requires a 24-hour sample collection because of significant diurnal variations.

Methods. Most of the current methods for phosphorus determination involve the formation of an ammonium phosphomolybdate complex. This colorless complex can be measured by ultraviolet absorption at 340 nm or can be reduced to form molybdenum blue, a stable blue chromophore, which is read between 600 and 700 nm.

Reference Ranges

Phosphate values vary with age. Divided into age groups, the ranges are shown in **Table 11.19**.

Table 11.19 Reference Ranges for Inorganic Phosphorus

Serum	
Neonate	1.45–2.91 mmol/L (4.5–9.0 mg/dL)
Child ≤15 y	1.29–2.26 mmol/L (4.0–7.0 mg/dL)
Adult	0.81–1.45 mmol/L (2.5–4.5 mg/dL)
Urine (24 h)	13–42 mmol/d (0.4–1.3 g/d)

Lactate

Lactate Biochemistry and Physiology

Lactate is a byproduct of an emergency mechanism that produces a small amount of ATP when oxygen delivery is severely diminished.[13] Pyruvate is the normal end product of glucose metabolism (glycolysis). The conversion of pyruvate to lactate is activated when a deficiency of oxygen leads to an accumulation of excess NADH (**Figure 11.8**). Normally, sufficient oxygen maintains a favorably high ratio of NAD to NADH. Under these conditions, pyruvate is converted to acetyl-coenzyme A (acetyl-CoA), which enters the citric acid cycle and produces 38 mol of ATP for each mole of glucose oxidized. However, under hypoxic conditions, acetyl-CoA formation does not occur, and NADH accumulates, favoring the conversion of pyruvate to lactate through anaerobic metabolism. As a result, only 2 mol of ATP are produced for each mole of glucose metabolized to lactate, with the excess lactate released into the blood. This release of lactate into blood has clinical importance because the accumulation of excess lactate in blood is an early, sensitive, and quantitative indicator of the severity of oxygen deprivation (Figure 11.8).[13]

Regulation

Because lactate is a by-product of anaerobic metabolism, it is not specifically regulated, as with K^+ or Ca^{2+}, for example. As oxygen delivery decreases below a critical level, blood lactate concentrations rise rapidly and indicate tissue hypoxia earlier than does pH (**Figure 11.9**). The liver is the major organ for removing lactate by converting lactate back to glucose by a process called *gluconeogenesis*.

Clinical Applications

Measurements of blood lactate are useful for metabolic monitoring in critically ill patients, for indicating the severity of the illness, and for objectively determining patient prognosis. There are two types of lactic acidosis. Type A is associated with hypoxic conditions, such as shock, myocardial infarction, severe CHF, pulmonary edema, or severe blood loss. Type B is of metabolic origin, such as with diabetes mellitus, severe infection, leukemia, liver or renal disease, and toxins (ethanol, methanol, or salicylate poisoning).

Determination of Lactate

Specimen Handling. Special care should be practiced when collecting and handling specimens for lactate analysis. Ideally, a tourniquet should

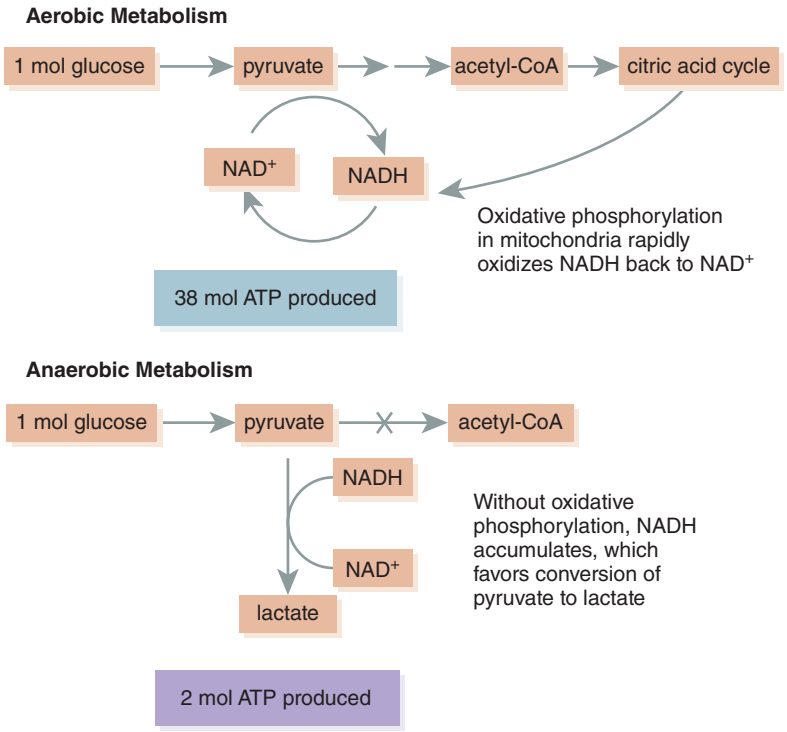


Figure 11.8 Aerobic versus anaerobic metabolism of glucose.

not be used because venous stasis will increase lactate levels. If a tourniquet is used, blood should be collected immediately and the patient should not exercise the hand before or during collection. After sample collection, glucose is converted to lactose by way of anaerobic glycolysis and should be prevented. Heparinized blood may be used but must be delivered on ice, and the plasma must be quickly separated. Iodoacetate and fluoride, which inhibit glycolysis without affecting coagulation, are usually satisfactory additives.

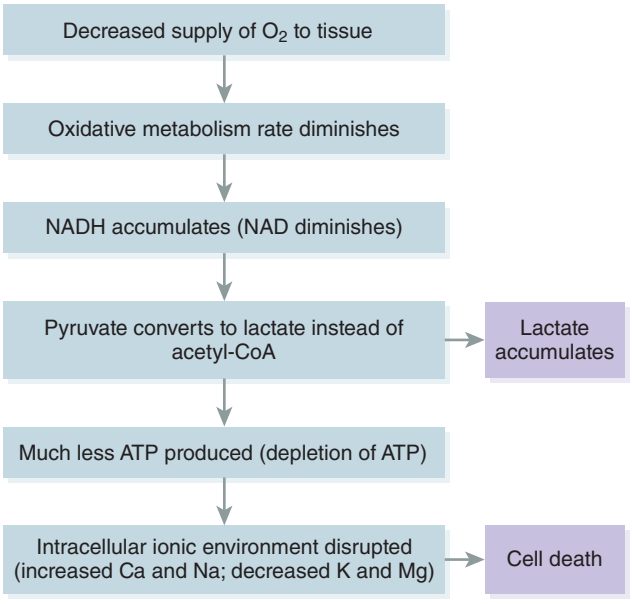


Figure 11.9 Metabolic effects of hypoxia, leading to cell death.

Methods. Although lactate is a sensitive indicator of inadequate tissue oxygenation, the use of blood lactate measurements has been hindered because older methods were slow and laborious. Other means of following perfusion or oxygenation have been used, such as indwelling catheters that measure blood flow, pulse oximeters, base-excess determinations, and measurements of oxygen consumption (V_{O2}). Current enzymatic methods make lactate determination readily available.

The most commonly used enzymatic method uses lactate oxidase to produce pyruvate and H_2O_2:

$$\text{Lactate} + O_2 \xrightarrow{\text{Lactate oxidase}} \text{pyruvate} + H_2O_2 \quad \textbf{(Eq. 11.5)}$$

One of two couple reactions may then be used. Peroxidase may be used to produce a colored chromogen from H_2O_2:

$$H_2O_2 + \text{H donar} + \text{chromogen} \xrightarrow{\text{Peroxidase}} \text{colored dye} + 2H_2O \quad \textbf{(Eq. 11.6)}$$

CASE STUDY 11.2, PART 2

The lab results for Chris, the teenager with diabetes, are back, and the attending physician is not surprised by them.

© simona pilolla 2/Shutterstock.

Analyte	Patient Value	Reference Range
Sodium	145 mmol/L	135–145 mmol/L
Potassium	6.2 mmol/L	3.5–5.1 mmol/L
Chloride	87 mmol/L	98–107 mmol/L
Total CO_2	9 mmol/L	22–28 mmol/L
BUN	35 mg/dL	6–20 mg/dL
Creatinine	1.4 mg/dL	0.9–1.3 mg/dL
Glucose	985 mg/dL	70–100 mg/dL
Lactate	5.1 mg/dL	0.3–2.0 mg/dL
Osmolality	387 mOsm/kg	275 –295 mOsm/kg
pH	7.10	7.35–7.45
Urine glucose	4+	Negative

1. Based on these results, what is the diagnosis?
2. Why are the chloride and bicarbonate decreased?

Reference Ranges

See **Table 11.20**.

Table 11.20 Reference Ranges for Lactate

	Enzymatic Method, Plasma
Venous	0.3–2.0 mmol/L (2.7–18 mg/dL)
Arterial	0.3–1.6 mmol/L (2.7–13.5 mg/dL)

© Jones & Bartlett Learning.

Anion Gap

Routine measurement of electrolytes usually involves only Na^+, K^+, Cl^-, and HCO_3^- (as total CO_2). These values may be used to approximate the **anion gap (AG)**, which is the difference between unmeasured anions and unmeasured cations. It is important to note that there is never an actual "gap" between total cationic charges and anionic charges due to constant electroneutrality. The AG is calculated by the concentration difference between commonly measured cations (Na^+, K^+) and commonly measured anions (Cl^-, HCO_3^-), as shown in **Figure 11.10**. AG is useful in indicating an increase in one or more of the unmeasured anions in the serum and also as a form of quality control for the analyzer used to measure these electrolytes. Consistently abnormal AGs in serum from healthy persons may indicate an instrument calibration problem rather than physiologic issues.

There are two commonly used methods for calculating the AG. The first equation is:

$$AG = Na^+ - (Cl^- + HCO_3^-) \qquad \textbf{(Eq. 11.7)}$$

It is equivalent to unmeasured anions minus the unmeasured cations in this way:

$$(PO_4^- + 2SO_4^{2-}) - (K^+ + 2Ca^{2+} + Mg^{2+}) \qquad \textbf{(Eq. 11.8)}$$

The reference range for the AG using this calculation is 7 to 16 mmol/L. The second calculation method is:

$$AG = (Na^+ + K^+) - (Cl^- + HCO_3^-) \qquad \textbf{(Eq. 11.9)}$$

It has a reference range of 10 to 20 mmol/L.[2]

An *elevated AG* may be caused by uremia/renal failure, which leads to PO_4^- and SO_4^- retention; ketoacidosis, as seen in cases of starvation or diabetes; methanol, ethanol, ethylene glycol, or salicylate poisoning; lactic acidosis; hypernatremia; and

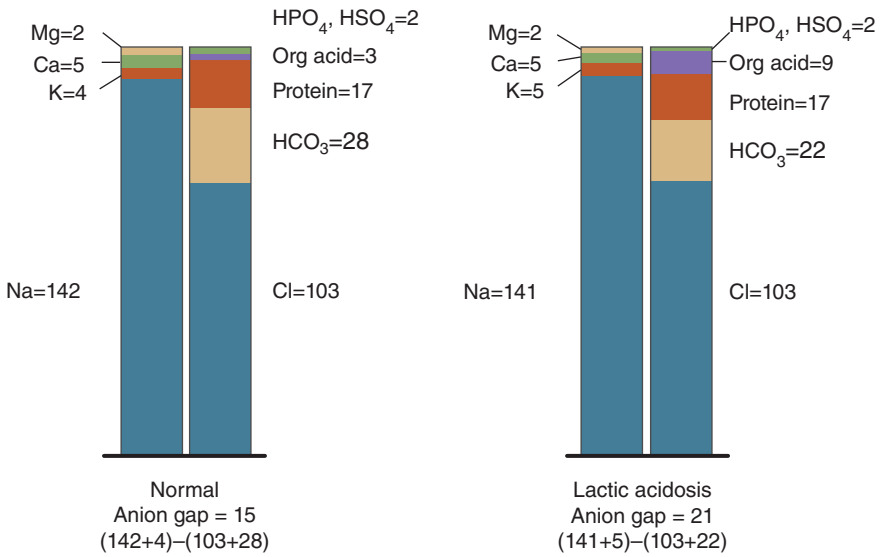


Figure 11.10 Demonstration of anion gap from concentrations of anions and cations in normal state and in lactic acidosis.

instrument error. *Low AG* values are rare but may be seen with hypoalbuminemia (decrease in unmeasured anions) or severe hypercalcemia (increase in unmeasured cations).

Electrolytes and Renal Function

The kidney is central to the regulation and conservation of electrolytes in the body (**Figure 11.11**). The following is a summary of electrolyte excretion and conservation in a healthy individual:

1. Glomerulus: This portion of the nephron acts as a filter, retaining large proteins and protein-bound constituents while most other plasma constituents pass into the filtrate. The concentrations in the filtered plasma should be approximately equal to ECF without protein.
2. Renal tubules:
   a. Phosphate reabsorption is inhibited by PTH and increased by 1,25-$(OH)_2$-D_3 (vitamin D). Excretion of PO_4^- is stimulated by calcitonin.
   b. Ca^{2+} is reabsorbed under the influence of PTH and 1,25-$(OH)_2$-D_3. Calcitonin stimulates excretion of Ca^{2+}.
   c. Mg^{2+} reabsorption occurs largely in the thick ascending limb of Henle's loop.
   d. Sodium reabsorption can occur through three mechanisms:
      - Approximately 70% of the Na^+ in the filtrate is reabsorbed in the proximal

CASE STUDY 11.2, PART 3

Let's take a look at the repeated results for Chris.

Analyte	Patient Value	Reference Range
Sodium	145 mmol/L	135–145 mmol/L
Potassium	5.8 mmol/L	3.5–5.1 mmol/L
Chloride	87 mmol/L	98–107 mmol/L
Total CO_2	8 mmol/L	22–28 mmol/L

3. Calculate the anion gap. What is the cause for the AG results in this patient?

tubules by isosmotic reabsorption. It is limited, however, by the availability of Cl^- to maintain electrical neutrality.

- Na^+ is reabsorbed in exchange for H^+. This reaction is linked with and depends on carbonic anhydrase.
- Stimulated by aldosterone, Na^+ is reabsorbed in exchange for K^+ in the distal tubules. (H^+ competes with K^+ for this exchange.)

e. Cl^- is reabsorbed, in part, by passive transport in the proximal tubule along the concentration gradient created by Na^+.

f. K^+ is reabsorbed by two mechanisms:

- Active reabsorption in the proximal tubule almost completely conserves K^+.
- Exchange with Na^+ is stimulated by aldosterone. H^+ competes with K^+ for this exchange.

g. Bicarbonate is recovered from the glomerular filtrate and converted to CO_2 when H^+ is excreted in the urine.

- Henle's loop: With normal AVP function, it creates an osmotic gradient that enables water reabsorption to be increased or decreased in response to body fluid changes in osmolality.
- Collecting ducts: Also, under AVP influence, this is where final adjustment of water excretion is made.

For a review of kidney structure, refer to Chapter 21, *Renal Function*.

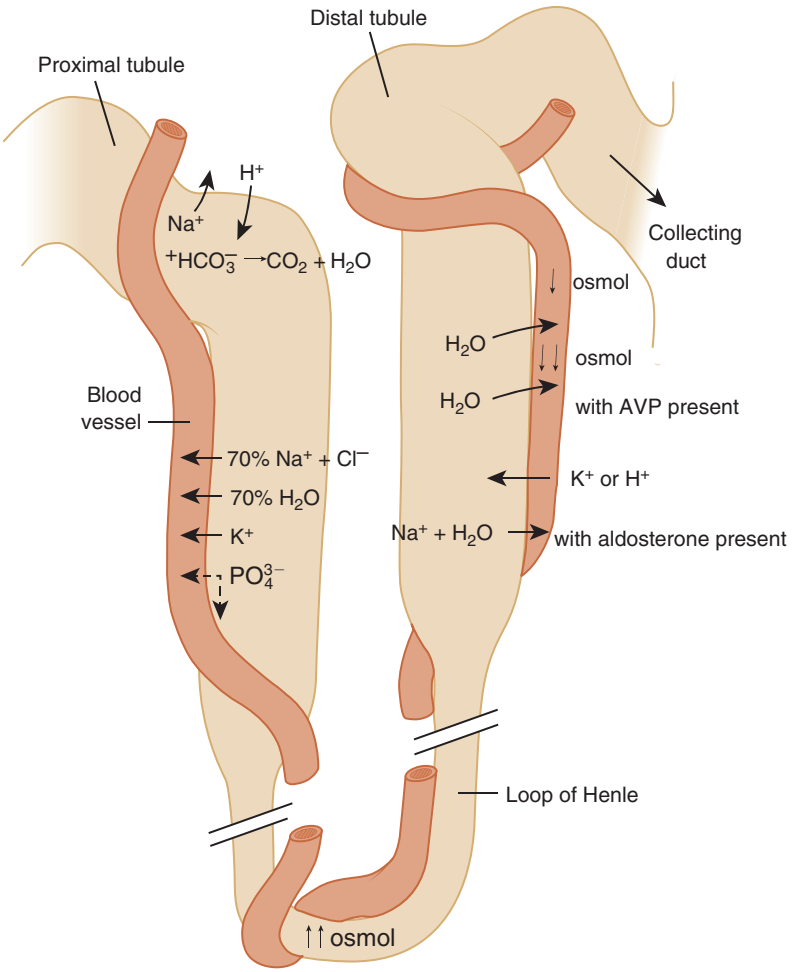
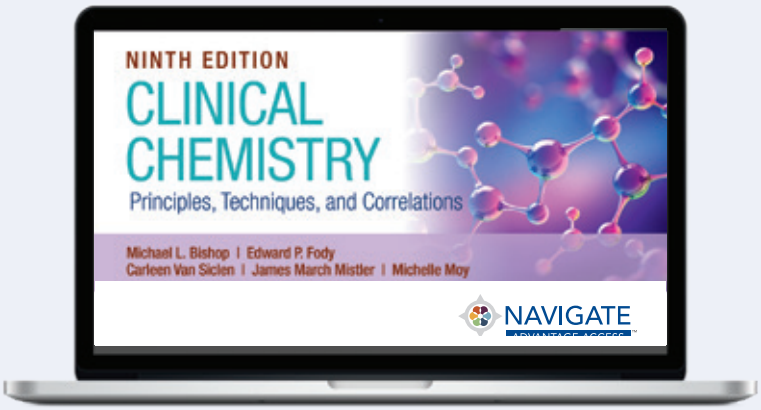


Figure 11.11 Summary of electrolyte movements in the renal tubules. AVP, arginine vasopressin hormone.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Muhsin SA, Mount DB. Diagnosis and treatment of hypernatremia. *Best Pract Res Clin Endocrinol Metab*. 2016;30(2):189–203.
2. Rifai N, Rita Horvath A, Wittwer C, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 8th ed. St. Louis, MO: Elsevier Saunders; 2019.
3. Rondon-Berrios H, Agaba EI, Tzamaloukas AH. Hyponatremia: pathophysiology, classification, manifestations and management. *Int Urol Nephrol* 2014;46:2153–2165.
4. Verbalis JG, Goldsmith SR, Greenberg A, et al. Diagnosis, evaluation, and treatment of hyponatremia: expert panel recommendations. *Am J Med*. 2013;126(10, suppl 1): S1–S42.
5. Kumar S, Tomas B. Sodium. *Lancet*. 1998;352:220–228.
6. Ortho Clinical Diagnostics. *Vitros Na+ Package Insert, Version 5.0*. Rochester, NY: Ortho Clinical Diagnostics; 2015.
7. Gumz M, Rabinowitz L, Wingo C. An integrated view of potassium homeostasis. *N Engl J Med*. 2015;373(1):60–72.
8. Palmer B, Clegg D. Physiology and pathophysiology of potassium homeostasis: core curriculum 2019. *Am J Kidney Dis*. 2019;74(5):682–695.
9. Yamamoto M, Yamaguchi T, Yamauchi M, et al. Acute-onset hypomagnesemia-induced hypocalcemia caused by the refractoriness of bones and renal tubules to parathyroid hormone. *J Bone Miner Metab*. 2011;29:752–755.
10. Blaine J, Chonchol M, Levi M. Renal control of calcium, phosphate, and magnesium homeostasis. *Clin J Am Soc Nephrol*. 2014;10(7):1257–1272.
11. Dong J, Xun P, He K, Qin L. Magnesium intake and risk of type 2 diabetes. *Diabetes Care*. 2011;34(9):2116–2122.
12. Polancic JE. Magnesium: metabolism, clinical importance, and analysis. *Clin Lab Sci*. 1991;4:105–109.
13. Rimachi R, Bruzzi de Carvahlo F, Orellano-Jimenez C, Cotton F, Vincent JL, De Backer D . Lactate/pyruvate ratio as a marker of tissue hypoxia in circulatory and septic shock. *Anaesth Intens Care* 2012;40(3):427–432.

CHAPTER 12

Blood Gases, pH, and Buffer Systems

Kyle B. Riding

CHAPTER OUTLINE

KEY TERMS

Acidemia
Acidosis
Alkalemia
Alkalosis
Amperometry
Base excess
Compensation
Deoxyhemoglobin
Hypercarbia
Hypercapnia
Hypoxia
Hypoxemia
Oxyhemoglobin
Oxygen saturation
Potentiometry

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Describe the chemical and physiological systems involved in acid–base balance, including the bicarbonate, phosphate, protein, and hemoglobin buffer systems along with the role of the kidneys and lungs.
- Use the Henderson-Hasselbalch equation to predict the pH or bicarbonate of a patient sample when given appropriate data.
- Define and describe the four major states of acid–base imbalance.

- Identify common causes of metabolic acidosis, metabolic alkalosis, respiratory acidosis, and respiratory alkalosis.
- Explain compensation mechanisms for acid–base imbalances including renal Na-H exchange, HCO_3^- reabsorption, NH_4^+ formation, and H_2PO_4 formation.
- Given laboratory data, use the Henderson-Hasselbalch equation to determine the acid–base disorder including uncompensated or compensated conditions.
- Describe oxygen transport and identify mechanisms by which oxygen status is assessed in patients.
- Draw the hemoglobin–oxygen dissociation curve and the impact that pH, pCO_2, 2,3-diphosphoglycerate (2,3-DPG) (also referred to as 2,3-biphosphoglycerate [2,3-BPG]), and temperature have on its shape and release of oxygen.
- Describe the principles involved in the measurement of pH, pCO_2, pO_2, and the various hemoglobin species.
- Explain the clinical significance of the following pH and blood gas parameters: pH, pCO_2, pO_2, actual bicarbonate, carbonic acid, base excess, oxygen saturation, fractional oxyhemoglobin, hemoglobin oxygen (binding) capacity, oxygen content, and total CO_2.
- Discuss problems and precautions in collecting and handling samples for pH and blood gas analysis. Include syringes, anticoagulants, mixing, icing, and capillary and venous samples as well as arterial samples in the discussion.
- Describe instrumental approaches to measuring various hemoglobin species and pH and blood gas parameters.
- Describe approaches to quality assurance, including quality control, proficiency testing, and delta checks to assess analytic quality.
- Discuss the reasons for possible discrepancies, given oxygen saturation data calculated by the blood gas analyzer and measured by the CO-oximeter.

An important aspect of clinical biochemistry is information on a patient's acid–base balance and blood gas homeostasis. These data often are used to assess patients in life-threatening situations. These test parameters are interrelated, and focusing on only one test result can be misleading.

This chapter discusses exchange of gases, carbon dioxide and oxygen, together with the body's mechanisms to maintain an acid–base balance. The interpretation of data from measurement of pH and other blood gas parameters, and the techniques and instrumentation used in these measurements, are also described. In addition, preanalytical considerations and quality assurance of blood gas analysis are reviewed.

Acid–Base Balance

Maintenance of H^+

An *acid* is a substance that can donate hydrogen ions (H^+) when dissolved in water. A *base* is a substance that can accept hydrogen ions. pH of a solution is defined as the negative log of the hydrogen ion concentration. Accordingly, a decrease in one pH unit represents a 10-fold increase in H^+ concentration.

The logarithmic pH scale expresses H^+ concentration (c is concentration):

$$\text{pH} = \log \frac{1}{c\,\text{H}^+} = -\log c\text{H}^+ \qquad \textbf{(Eq. 12.1)}$$

CASE STUDY 12.1, PART 1

Meet Doug, a 21-year-old male with a history of type 1 diabetes mellitus. Doug is taken to a local emergency department. Upon his arrival, it is noted that Doug seems easily confused. He complains of a headache, nausea, dry mouth, and excessive thirst. His breathing is rapid, and one of the nurses caring for him notes a fruity odor to his breath. Doug self-reports adherence with his insulin therapy.

© AJR_photo/Shutterstock.

CASE STUDY 12.2, PART 1

Kelsey, a 2-year old female, fell into her family's pool after the protective fencing was left open. Kelsey's mother found her daughter unresponsive in the pool two minutes later and immediately rescued her. She immediately called 911 and began performing CPR on her daughter. Her daughter's airway was restored during the CPR but Kelsey was still only able to breathe via slow, shallow breaths.

© Julián Rovagnati/ShutterStock.

Through normal metabolic processes, the body produces approximately 150 g of H^+ each day. Yet the concentration of H^+ in the extracellular body fluids is maintained within a narrow range of 36 to 44 nmol/L. This equates to a pH range of 7.35 to 7.45. Through mechanisms that involve the lungs and kidneys, the body controls and excretes H^+ in order to maintain pH homeostasis. Imbalances between the rate of acid formation and excretion can occur and can lead to alterations in consciousness, neuromuscular irritability, tetany, coma, and death.

The relative strengths of acids and bases are defined by their ability to dissociate in water and are described by their dissociation constant (K_a). Tables of dissociation constants can be found online and in most biochemistry texts. The pK_a is defined as the negative log of the dissociation constant. The dissociation constant is also known as the ionization constant. If the pH equals pK_a, then the solution is in equilibrium, and protonated and unprotonated species are present in equal concentrations. However, if the pH of the solution is less than pK_a then majority of the components will largely be protonated due to the excess in H^+ ions present. Many species have more than one pK_a, meaning they can accept or donate more than one H^+ ion.

Buffer Systems: Regulation of H⁺ and the Henderson-Hasselbalch Equation

A *buffer* consists of a weak acid and a salt of its conjugate base, and it allows a solution to resist changes in pH upon adding acid or base. The effectiveness of a buffer depends on the pK_a of the buffering system and the pH of the environment in which it is placed. In plasma, the bicarbonate–carbonic acid system is one of the principal buffers:

$$\underset{\text{Carbonic acid}}{H_2CO_3} \leftrightarrow \underset{\text{Bicarbonate}}{HCO_3^-} + H^+ \qquad \textbf{(Eq. 12.2)}$$

This buffer system uses carbonic acid (H_2CO_3) and a salt of its conjugate base, bicarbonate (HCO_3^-). H_2CO_3 is a weak acid because it does not completely dissociate into H^+ and HCO_3^-, whereas a strong acid, such as HCl, completely dissociates into H^+ and Cl^- in solution.

When hydrogen ions are added to the bicarbonate–carbonic acid system, the HCO_3^- will combine with the H^+ to form H_2CO_3 (**Eq. 12.2**). Therefore, the reaction is driven toward the formation of the weak acid, increasing the amount of carbonic acid and consuming bicarbonate ion. Conversely, when a strong base is added, H_2CO_3 will combine with the added OH^- ions to form H_2O and the weaker conjugate base HCO_3^-. Although the bicarbonate–carbonic acid system has low buffering capacity, it is still the most important buffer system in extracellular fluids for three reasons: (1) H_2CO_3 dissociates into CO_2 and H_2O, allowing CO_2 to be eliminated by the lungs and H^+ as water; (2) changes in CO_2 modify the ventilation (respiratory) rate; and (3) HCO_3^- concentration can be changed by the kidneys.

Proteins and phosphates are also buffer systems. They are primarily in the intracellular fluids and to a minor extent in the extracellular spaces. Most circulating proteins have a net negative charge and are capable of binding H^+. The phosphate buffer system ($HPO_4^{2-} - H_2PO_4^-$) is the primary buffer in urine and is involved in the exchange of sodium ion in the urine filtrate. Hemoglobin plays a role in buffering the CO_2 as it is transported to the lungs for excretion.

The interrelationship of the weak acid, conjugate base, and pH is depicted by the *Henderson-Hasselbalch equation* (**Eq. 12.3**). The Henderson-Hasselbalch equation expresses acid–base relationships in a mathematical formula:

$$pH = pK_a' + \log\frac{cA^-}{cHA} \qquad \textbf{(Eq. 12.3)}$$

where A^- is the proton acceptor or base (e.g., HCO_3^-), HA is the proton donor or weak acid (e.g., H_2CO_3), and pK_a' is the pH at which there is an equal concentration of protonated and unprotonated species for the given buffer system. Knowing any of the three variables allows for the calculation of the fourth.

Regulation of Acid–Base Balance: Lungs and Kidneys (Transport of Carbon Dioxide)

Carbon dioxide, the end product of most aerobic metabolic processes, easily diffuses out of the tissue where it is produced and into the plasma and red blood cells of the surrounding capillaries. In plasma, a small amount of CO_2 is physically dissolved or combined with amino acid residues of proteins to form carbamino compounds. However, most of the CO_2 freely diffuses into the red blood cells and reacts with water to form H_2CO_3 in the presence of the catalyst carbonic anhydrase. The newly formed H_2CO_3 dissociates into H^+ and HCO_3^- (**Figure 12.1**).

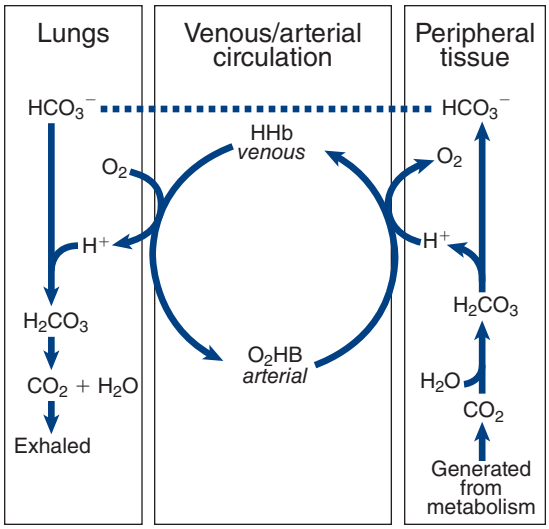


Figure 12.1 Interrelationship of the bicarbonate and hemoglobin buffering systems.

The dissociation of H_2CO_3 causes development of a concentration gradient due to the increase in HCO_3^- concentration within the red blood cell. Accordingly, bicarbonate ions diffuse from the red blood cells into the plasma. To maintain electroneutrality (the same number of positively and negatively charged ions on each side of the red cell membrane), chloride diffuses into the cell. This is known as the *chloride shift* (also known as the Hamburger shift). The H^+ generated is buffered by binding with the recently deoxygenated hemoglobin molecules. Therefore, the abbreviation HHb is used to designate **deoxyhemoglobin** since the molecule has hydrogen bound instead of oxygen.

In the lungs, this process is reversed. Inspired O_2 diffuses from the alveoli into the blood and is bound to hemoglobin, forming **oxyhemoglobin** (O_2Hb). The H^+ that was carried on the (reduced) hemoglobin in the venous blood is released to recombine with HCO_3^- to form H_2CO_3, which dissociates into H_2O and CO_2. The CO_2 diffuses into the alveoli and is eliminated through ventilation. The net effect of these systems is a minimal change in H^+ concentration between the venous and arterial circulation.

When the lungs do not remove CO_2 at the rate of its production, it accumulates in the blood, causing an increase in H^+ concentration. However, if CO_2 removal is faster than its production (hyperventilation), the H^+ concentration will be decreased. Therefore, the pH of the blood will increase. Therefore, an alteration in CO_2 is often considered to be the equivalent of the addition or subtraction of acid to the system even though CO_2 as a chemical species is not itself an acid or a base. A change in the H^+ concentration of blood that results from nonrespiratory disturbances causes the respiratory center to respond by altering the rate of ventilation in an effort to restore the blood pH to normal. The lungs together with the bicarbonate buffer system provide the first line of defense to change the acid–base status.

The kidneys regulate the excretion of both acid and base, making them an important buffer in the regulation of acid–base balance. Specifically, the kidney's main role is to reabsorb bicarbonate ion, HCO_3^-, from the glomerular filtrate in the proximal tubules (**Figure 12.2A**). If bicarbonate ions were not reabsorbed and excreted in the urine, this would impact the bicarbonate–carbonic acid buffer system, resulting in a markedly increased levels of H^+ ions.

Because the glomerular filtrate contains the same HCO_3^- concentration as plasma, passive reabsorption of HCO_3^- across the tubule membrane into the blood cannot occur. Instead, sodium (Na^+) in the glomerular filtrate is exchanged for H^+ in the tubular cell. This is known as the Na^+-H^+ exchange and is a major mediator of the kidneys' buffering capacity. The excreted H^+ combines with HCO_3^- in the filtrate to form H_2CO_3, which is subsequently converted into H_2O and CO_2 by carbonic anhydrase. The CO_2 diffuses into the tubular cells and reacts with H_2O to reform H_2CO_3 within the cell, which then largely dissociates into hydrogen and bicarbonate ions. The bicarbonate ion is reabsorbed into the blood along with sodium. Through this cascade of events, the hydrogen ion provides a vehicle to reabsorb bicarbonate ion. In a healthy individual, virtually all the filtered bicarbonate ion is reabsorbed by the kidney through this process.

During periods of increased plasma pH, the peritubular capillary supply has diminished H^+ ion concentrations. This limits the activity of the Na^+-H^+ exchange, and the kidney has limited reabsorption of HCO_3^-, which compensates for the elevated blood pH. The opposite is also true, involving decreased plasma pH, increased Na^+-H^+ exchange, and subsequent increased HCO_3^- reabsorption. This exchange between H^+ and Na^+ suggests, in part, why clinicians order pH and blood gases together, along with electrolytes (Na^+, K^+, and Cl^-), to assess the acid–base status of a patient.

Beyond reabsorption of bicarbonate, the other major function of the kidneys' buffering capacity is excreting hydrogen ions. Under normal conditions,

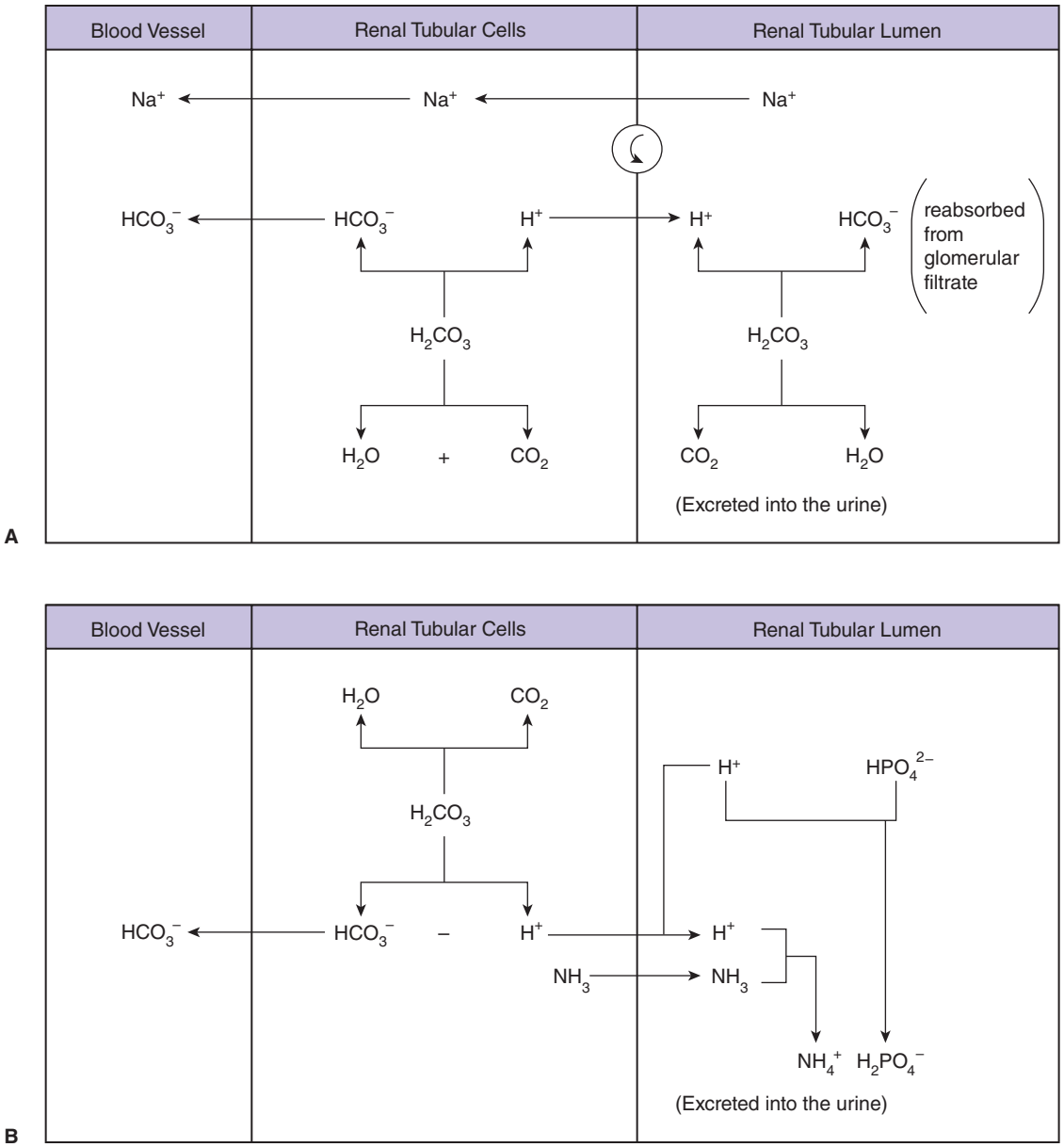


Figure 12.2 **(A)** Bicarbonate reabsorption by the proximal tubule cell. **(B)** Phosphate Buffering System.

the body produces a net excess (50 to 100 mmol/L) of acid (H^+) each day that must be excreted by the kidney. The minimum urine pH that can be generated is 4.6, which limits the amount of free hydrogen ion excretion. The phosphate buffering system (**Figure 12.2B**) allows for greater hydrogen ion excretion. Most of the H^+ ions in the glomerular filtrate combine with monohydrogen phosphate (HPO_4^{2-}) and ammonia (NH_3) and are excreted in the urine as dihydrogen phosphate ($H_2PO_4^-$) and ammonium ions (NH_4^+). Ammonia is produced by the deamination of glutamine and other amino acids. During acidosis, ammonia production is increased to help buffer the excess hydrogen ions.

Assessment of Acid–Base Homeostasis

The Bicarbonate Buffering System

In assessing acid–base homeostasis, components of the bicarbonate buffering system are measured and calculated. For the bicarbonate–carbonic acid buffer system, the dissolved CO_2 (dCO_2) is in equilibrium with CO_2 gas in the lungs, which can be expelled. The dCO_2, expelled by the lungs, is the *respiratory component*. The lungs participate rapidly in the regulation of blood pH through hypoventilation or

hyperventilation. The kidneys control the bicarbonate concentration by reabsorption or excretion, which is the *metabolic component*.

Reflecting on the how the Henderson-Hasselbalch equation relates to this buffer system is essential to understanding its utility in clinical assessment. As discussed, the equation relates the pH, pK_a, weak acid, and conjugate base together via a mathematical equation. The pH is directly measurable, and the pK_a of this buffer system is constant at 6.1 under normal physiological conditions. Carbonic acid cannot be directly measured since it rapidly degrades into dCO_2. But this relationship proves beneficial since the H_2CO_3 concentration is directly proportional to the partial pressure exerted by the dCO_2, which is measured and reported as the pCO_2 in units of mmHg. To convert pCO_2 in mmHg to an equivalent of carbonic acid concentration in mmol/L in plasma at 37°C requires the use of the conversion factor 0.0307 mmol/L/mm Hg. Both pH and pCO_2 are directly measured in blood gas analysis, and the pK_a' is a constant; therefore, can be calculated:

$$\text{pH} = pK_a' + \log\frac{c\text{HCO}_3^-}{0.0307 \times pCO_2} \qquad \textbf{(Eq. 12.4)}$$

When the kidneys and lungs are functioning properly, the ratio of HCO_3^- to H_2CO_3 is 20:1, corresponding to a pH of 7.40. This can be calculated by substituting normal values (**Table 12.1**) for HCO_3^- and pCO_2 from the preceding equation:

$$\frac{24 \text{ mmol/L}}{(0.0307 \text{ mmol/L} - \text{mm Hg}) \times 40 \text{ mm Hg}} = \frac{24}{1.2} = \frac{20}{1} \qquad \textbf{(Eq. 12.5)}$$

The normal pH of 7.4 can be derived by adding the log of 20 (1.3) to the pK_a' of the bicarbonate system (7.40 = 6.1 + 1.3).

Table 12.1 Arterial Blood Gas Reference Range at 37°C

pH	7.35–7.45
pCO_2 (mm Hg)	35–45
HCO_3^- (mmol/L)	22–29
Total CO_2 content (mmol/L)	23–27
pO_2 (mmol/L)	85–105
SO_2 (%)	>95
O_2Hb (%)	>95

The numerator (HCO_3^-) denotes the kidney function, whereas the denominator (pCO_2, which represents H_2CO_3) denotes the lung function. The lungs will impact pH through retention or elimination of CO_2 by changing the rate and volume of ventilation. The kidneys regulate pH by excreting acid, primarily in the ammonium ion, and by reclaiming HCO_3^- from the glomerular filtrate. It is critical to remember that pH is the dependent variable that responds to changes based upon the ratio of HCO_3^- and H_2CO_3, not the reverse, when assessing patient status.

Acid–Base Disorders: Acidosis and Alkalosis

Acid–base disorders result from imbalances in the acid base equilibrium. When blood pH is less than the reference range (7.35 to 7.45), it is termed **acidemia/acidosis**. Similarly, a pH greater than the reference range is termed **alkalemia/alkalosis**. Technically, the suffix *-osis* refers to a process in the body and the suffix *-emia* refers to the corresponding state in blood. Therefore, *-osis* is the cause of the *-emia*. Acidemia will result if the hydrogen ion concentration increases through increased pCO_2 concentrations or decreases in the bicarbonate concentration. Alkalemia will result if the hydrogen ion concentration decreases, either from decreased pCO_2 or increased concentrations of bicarbonate.

A disorder caused by respiratory dysfunction (a change in the pCO_2, the respiratory component) is termed *primary respiratory acidosis* or *alkalosis*. A disorder resulting from a change in the bicarbonate level (a consequence of renal or metabolic dysfunction) is termed a *nonrespiratory or metabolic disorder.*

The response to maintain acid–base homeostasis is termed **compensation** and is accomplished by the organ not associated with the primary process. For example, if the imbalance is of metabolic (nonrespiratory) origin, the body compensates by altering ventilation. Conversely, for disturbances of the respiratory component, the kidneys will compensate by selectively excreting or reabsorbing specific ions. Mixed acid–base disorders arise from the presence of more than one process or compensatory mechanism in response to the primary disorder.

The lungs can immediately compensate by retaining or expelling carbon dioxide; however, this response is short term and often incomplete, whereas the kidneys are slower to respond (2 to 4 days), but the response is long term and sustained. *Fully*

compensated status implies that the pH has returned to the normal range (the ratio of HCO_3^- to H_2CO_3 of 20:1 has been restored); *partially compensated* implies that the pH is approaching normal. *Uncompensated* implies the pH is abnormal and the body has not started compensating for the acid–base imbalance.

Acidosis may be caused by a primary metabolic or respiratory imbalance. In primary metabolic acidosis, the amount of acid exceeds the capacity of the buffer systems, and there is a decrease in bicarbonate. The resulting decrease in pH is due to the decreased ratio of the metabolic to respiratory component in the Henderson-Hasselbalch equation:

$$\text{pH} \propto \frac{c\text{HCO}_3^-}{N(0.0307 \times p\text{CO}_2)} < \frac{20}{1} \qquad \textbf{(Eq. 12.6)}$$

where N is a normal value and $\propto$ indicates proportional.

The body compensates for metabolic acidosis through *hyperventilation*, which is an increase in the rate or depth of breathing. By expelling (i.e., "blowing off") CO_2, the base-to-acid ratio will normalize by reducing the pCO_2 and thereby elevating the pH. Secondary compensation occurs when the kidneys begin to correct the ratio by excreting hydrogen ions and reabsorbing bicarbonate ions.

Metabolic acidosis may be caused by the following:

- Overdose of acid-producing substances (e.g., aspirin, ethanol, methanol, ethylene glycol).
- Excess production of acidic ketone bodies, as seen in diabetic ketoacidosis.
- Reduced excretion of hydrogen ions, as seen in renal tubular acidosis.
- Excessive loss of bicarbonate from diarrhea, as seen in hyperchloremic acidosis.

Primary *respiratory acidosis* results from a decrease in alveolar ventilation (*hypoventilation*), causing a decreased elimination of CO_2 by the lungs. This causes the patient to develop **hypercarbia** (i.e., **hypercapnia**), a decrease in pCO_2.

$$\text{pH} \propto \frac{N c\text{HCO}_3^-}{\uparrow (0.0307 \times p\text{CO}_2)} < \frac{20}{1} \qquad \textbf{(Eq. 12.7)}$$

In primary respiratory acidosis, the compensation occurs through metabolic/renal processes. However, as previously mentioned, the renal buffering processes, unlike respiratory system buffering, require hours to days to affect the pH. Therefore, it takes days to weeks for full compensation to occur. The kidneys respond to respiratory acidosis by increasing the excretion of H^+ and reclamation of HCO_3^-. As a result of this, HCO_3^- in the blood increases, causing the pH to return to normal.

Respiration is regulated by the medulla oblongata of the brain. Chemoreceptors present in the aortic arch and the carotid sinus also influence respiratory rate by responding to the levels of H^+ (pH), O_2, and CO_2 in the blood and cerebrospinal fluid. *Respiratory acidosis* may be caused by conditions that affect those regulatory systems including the following:

- Ineffective removal of CO_2 from the blood, as seen in lung disorders such as asthma.
- Airway obstruction, as seen in chronic obstructive pulmonary disease (COPD).
- Overdose of drugs (e.g., barbiturates, morphine, alcohol) that slow the respiratory center causing hypoventilation, which subsequently increases blood pCO_2 levels.
- Decreased cardiac output, as observed with congestive heart failure, will result in less blood being presented to the lungs for gas exchange and, therefore, an elevated pCO_2.

As with acidosis, alkalosis can result from metabolic and respiratory causes. Primary *metabolic alkalosis* results from a gain in HCO_3^-, causing an increase in the pH:

$$\text{pH} \propto \frac{N c\text{HCO}_3^-}{\uparrow (0.0307 \times p\text{CO}_2)} > \frac{20}{1} \qquad \textbf{(Eq. 12.8)}$$

The hydrogen ion sensors in the brain signal the respiratory center to depress respiration, which results in hypoventilation and increases the retention of CO_2, thereby increasing pCO_2 and lowering the pH. However, this response is erratic, and an increase in pCO_2 is a stimulus itself to increase respiration. Ultimately, the primary cause of metabolic alkalosis needs to be corrected.

Metabolic alkalosis maybe caused by the following reasons:

- Excessive loss of stomach acid through vomiting or nasogastric suctioning.
- Prolonged use of diuretics causes increased renal excretion of H^+ resulting in acidic urine and alkaline blood.
- Excess administration of sodium bicarbonate or excess ingestion of antacids.
- Hypokalemia causes H^+ ions to shift intracellularly. This intracellular shift enhances HCO_3^- reabsorption in the kidneys.

CASE STUDY 12.1, PART 2

© AJR_photo/Shutterstock.

In the ED, Doug had a positive urine ketone test, and the point of care glucometer reading was 543 mg/dL. An arterial blood gas, electrolytes, glucose, and lactate were ordered and sent to the clinical laboratory for analysis.

The arterial blood sample provides the following results:

Analyte	Patient Value	Reference Range
pH	7.255	7.35–7.45
pCO_2	35 mm Hg	35–45 mm Hg
HCO_3^-	13 mmol/L	22–26 mmol/L
pO_2	85 mmol/L	85–105 mmol/L
Sodium	126 mmol/L	Adults: 135–145 mmol/L
		24 h: 120–240 mmol/d, varies with diet
		CSF: 136–150 mmol/L
Potassium	5.3 mmol/L	Adults: 3.5–5.1 mmol/L
		24 h: 15-105 mmol/24 hours
Chloride	110 mmol/L	Adults: 98–107 mmol/L
		24 h: 35–285 mmol/24 hours
Bicarbonate	12 mmol/L	22–29 mmol/L
Glucose	675 mg/dL	Fasting: 70–100 mg/dL
Lactate	pending	0.3–2.0 mmol/L

1. What acid–base disorder is Doug most likely experiencing based upon these manifestations?
2. Predict his lactate result: increased, decreased, or normal?

CASE STUDY 12.2, PART 2

© Julián Rovagnati/ShutterStock.

Remember Kelsey. Emergency Medical Services arrived promptly and rushed her to the emergency department (ED). The hospitalist ordered and collected an arterial blood gas (ABG).

1. Explain what happened to Kelsey's gas exchange during this trauma incident.

The emergency department's stat lab obtained ABG results two minutes after collection:

Analyte	Patient Value	Reference Range
pH	7.12	7.35–7.45
pCO_2	92 mm Hg	35–45 mm Hg
HCO_3	29 mmol/L	22–26 mmol/L

2. What acid-base disorder is Kelsey classified as having?
3. What is Kelsey's compensation status?
4. What compensation mechanisms are happening in the kidneys? In the lungs?

CASE STUDY 12.1, PART 3

Doug was admitted, and treatment of diabetic ketoacidosis was initiated. His blood glucose began decreasing with therapy. Three hours after admission, another blood gas was ordered. The results are as follows:

Analyte	Patient Value	Reference Range
pH	7.550	7.35–7.45
pCO_2	55 mm Hg	35–45 mm Hg
HCO_3^-	44 mmol/L	22–26 mmol/L

© AJR_photo/Shutterstock.

3. What acid–base disorder does Doug have?
4. What is Doug's compensation status?
5. What are three compensation mechanisms occurring in the kidneys and one compensation mechanism happening in the lungs?

Primary *respiratory alkalosis* from an increased rate of alveolar ventilation causes excessive elimination of CO_2 by the lungs:

$$\text{pH} \propto \frac{Nc\text{HCO}_3^-}{\downarrow(0.0307 \times p\text{CO}_2)} > \frac{20}{1} \qquad \text{(Eq. 12.9)}$$

The kidneys compensate by excreting HCO_3^- in the urine and reclaiming H^+ to the blood through decreased activity of the Na^+-H^+ exchange.

The causes of respiratory alkalosis include:

- High altitudes decrease pCO_2 resulting in an increased respiratory rate (hyperventilation)
- Anxiety-induced hyperventilation
- Aspirin overdose stimulates the respiratory center of the brain leading to hyperpnea
- Pulmonary embolism or pulmonary fibrosis in which pulmonary oxygen exchange is impaired

Determining Patient Acid–Base Status

As with all chemistry panels with interrelated analytes, assessment of an arterial blood gas panel can be challenging. It is best to begin by assessing what analytes are outside of the reference ranges. Next, determine if the pH demonstrates alkalosis or acidosis. Third, determine if it is a metabolic or respiratory disorder. This task becomes more challenging when the patient results show partial or full compensation. There are several schemes used to help simplify this complex process and are beneficial to learners. However, in practice, the user should always ensure the clinical symptoms and laboratory results support the arrived at conclusion, understanding that these schemes are of limited use in mixed acid–base disturbances.

Oxygen and Gas Exchange

Oxygen and Carbon Dioxide

The role of oxygen in metabolism is crucial to life. In the mitochondria, the electron transport chain reduces molecular oxygen to water. For adequate tissue oxygenation, the following seven conditions are necessary: (1) available atmospheric oxygen, (2) adequate ventilation, (3) gas exchange between the lungs and arterial blood, (4) binding of O_2 onto hemoglobin, (5) adequate hemoglobin, (6) adequate blood flow to tissues, and (7) release of O_2 to the tissue. Any disturbances in these conditions can result in **hypoxia**, poor tissue oxygenation. pO_2 is often measured to evaluate a patient's oxygen status and is typically measured along with pH and pCO_2 in routine blood gas analysis.

The amount of O_2 available in atmospheric air depends on the barometric pressure and altitude. At sea level, the barometric pressure is 760 mm Hg. (In the International System of Units, 1 mm Hg = 0.133 kPa, where 1 Pa = 1 N/m^2.) Dalton's law states that total atmospheric pressure is the sum of the partial pressures of each of the gases in the atmosphere. One atmosphere exerts 760 mm Hg of pressure and is made up of O_2 (20.93%), CO_2 (0.03%), nitrogen (78.1%), and inert gases (~1%). The percentage for each gas is the same at all altitudes; the *partial pressure* for each gas in the atmosphere is equal to the barometric

pressure at a particular altitude times the appropriate percentage for each gas. The vapor pressure of water (47 mm Hg) at 37°C must be accounted for in calculating the partial pressure for the individual gases (**Figure 12.3**). In the body, these gases are always fully saturated with water. For example:

Partial pressure of O_2 at sea level (in the body)
= (760 mm Hg – 47 mm Hg) × 20.93%
= 149 mm Hg (at 37°C)

Partial pressure of CO_2 at sea level (in the body)
= (760 mm Hg – 47 mm Hg) × 0.03%
= 2 mm Hg (at 37°C)

Air is moved into the lungs through expansion of the thoracic cavity, which creates a temporary negative pressure gradient, causing air to move into the numerous tracheal branches that terminate at the alveoli. Gas exchange between alveolar air and pulmonary blood occurs in the alveoli. At the beginning of inspiration, these airways are still filled with air (gas) retained from the previously expired breath. This air, termed *dead space air*, dilutes the air being inspired. The inspired air is warmed to 37°C and becomes saturated with water vapor and mixes with air in the alveoli. The resulting pO_2 in the alveoli averages about 110 mm Hg instead of the anticipated 149 mm Hg. There are many factors that can influence the amount of O_2 that moves to the alveoli, and some of the factors are as follows:

- *Destruction of the alveoli.* The normal surface area of the alveoli is as large as a tennis court. Diminished surface area in diseases such as emphysema will cause an inadequate amount of O_2 to move into the blood.
- *Pulmonary edema.* Gas diffuses from the alveoli into the capillary through a small space. With pulmonary edema, fluid "leaks" into this space and inhibits diffusion of gases.
- *Airway blockage.* In patients with asthma or bronchitis, airways are blocked, which prevents air from the atmosphere from reaching the alveoli.

There are factors that can influence the amount of O_2 delivered to the tissue, and some common causes of inadequate perfusion are as follows:

- *Inadequate blood supply.* When the blood supply to the lung is inadequate, O_2 enters the blood in the lungs, but not enough blood is being carried away to the tissue where it is needed. This may be the consequence of a blockage in a pulmonary blood vessel (pulmonary embolism), pulmonary hypertension, or heart failure.
- *Diffusion of CO_2 and O_2.* O_2 diffuses 20 times slower than CO_2, and it is more sensitive to problems with diffusion. In instances such as pulmonary edema, **hypoxemia** will develop but with minimal alteration to pCO_2. This type of hypoxemia is generally treated with supplemental O_2. The percentage of O_2 can be increased temporarily when needed; however, concentrations of O_2

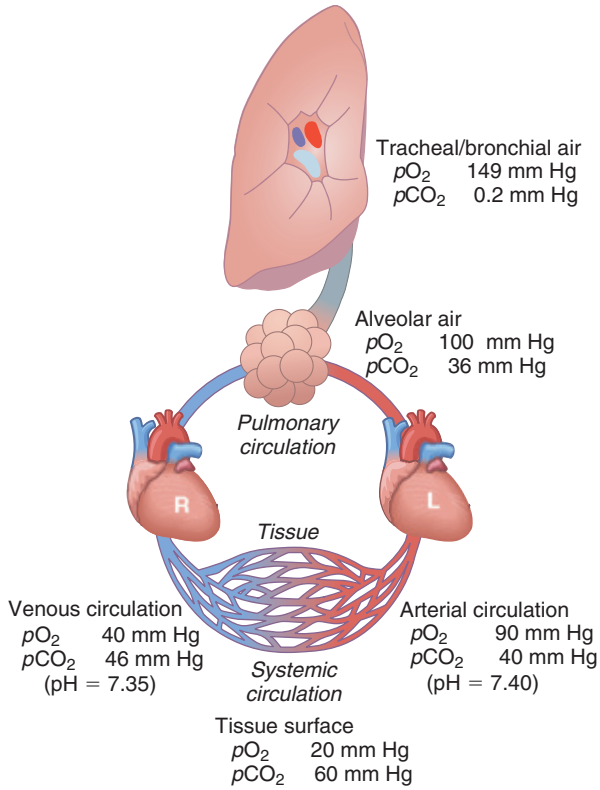


Figure 12.3 Gas content in lungs and pulmonary and systemic circulation.

CASE STUDY 12.1, PART 4

6. Predict Doug's oxygen saturation as increased, decreased, or normal at the time of admission.

≥60% must be used with caution because they can be toxic to the lungs.

- *Intrapulmonary shunting*. Nonfunctional alveoli prohibit venous blood from being oxygenated that results in decreased oxygen content.
- *Anemia*. Blood with less hemoglobin will carry proportionally less oxygen.

Oxygen Transport

Most O_2 in arterial blood is transported to the tissue by hemoglobin. In the adult, a single hemoglobin (A_1) molecule can combine reversibly with four molecules of O_2. The actual amount of O_2 loaded onto hemoglobin depends on the availability of O_2; concentration and type(s) of hemoglobin present; presence of competing molecules, such as carbon monoxide (CO); pH; temperature of the blood; and levels of pCO_2 and 2,3-diphosphoglycerate (2,3-DPG, also referred to as 2,3-biphosphoglycerate [2,3-BPG]). When adequate atmospheric and alveolar O_2 is available along with normal diffusion of O_2 to the arterial blood, more than 95% of "functional" hemoglobin (hemoglobin capable of reversibly binding O_2) will bind O_2.

The *fraction of inspired oxygen* (FiO_2) is the concentration of oxygen in the gas mixture. Patients having difficulty breathing are placed on oxygen therapy. The FiO_2 can be increased by breathing gas mixtures up to 100% O_2. Increasing FiO_2 further saturates hemoglobin. However, once the hemoglobin is 100% saturated, an increase in O_2 to the alveoli serves only to increase the concentration of dO_2 in the arterial blood. Prolonged administration of high concentrations of O_2 may cause oxygen toxicity and, in some cases, decreased ventilation, which leads to hypercarbia. Blood gas analyzers may require manual entry of FiO_2 and the patient's body temperature.

Hemoglobin exists in one of four forms:

1. Oxyhemoglobin (O_2Hb) is hemoglobin containing ferrous iron (Fe^{2+}) in the heme group that is reversibly bound to O_2.
2. Deoxyhemoglobin (HHb), also known as reduced hemoglobin, is hemoglobin without O_2. It is formed when oxyhemoglobin releases its O_2 and is capable of forming a bond when O_2 is available. Notably, hemoglobin increases its affinity to hydrogen ions as it loses O_2.
3. Carboxyhemoglobin (COHb) is hemoglobin bound to carbon monoxide (CO), which makes it unavailable for O_2 transport. This is because the bond between CO and Hb is 200 times stronger than the bond between O_2 and Hb.
4. Methemoglobin (MetHb) is hemoglobin unable to bind O_2 because iron is in an oxidized (Fe^{3+}) rather than reduced state. The Fe^{3+} can be reduced by the enzyme methemoglobin reductase, which is found in red blood cells.

Dedicated spectrophotometers (CO-oximeters), discussed later in this chapter, are used to determine the relative concentrations (relative to total hemoglobin) of these hemoglobin molecules.

Assessment of a Patient's Oxygen Status

Four parameters are commonly used to assess a patient's oxygen status:

1. Oxygen saturation (SO_2)
2. Fractional (percent) oxyhemoglobin (FO_2Hb)
3. Trends in oxygen saturation assessed by transcutaneous (TC) and pulse oximetry (SpO_2)
4. Amount of O_2 dissolved in plasma (pO_2).

Oxygen saturation, SO_2, represents the percentage of the functional hemoglobin that is saturated with O_2 compared with the total amount of hemoglobin capable of binding O_2:

$$SO_2 = \frac{cO_2Hb}{(cO_2Hb + cHHb)} \times 100 \qquad \textbf{(Eq. 12.10)}$$

SO_2 is frequently measured by pulse oximetry (discussed later). These calculated results, however, can differ significantly from those determined by direct CO-oximeter measurement due to the assumption that normal, healthy subjects have fully functional hemoglobin and the oxyhemoglobin dissociation curve has a specific shape and location. These algorithms do not account for the presence of other hemoglobin species, such as COHb and MetHb, that are incapable of reversibly binding O_2. These shortcomings can generate erroneous results; therefore, SO_2 should not be used to assess oxygenation status.[1,2]

Fractional (or percent) oxyhemoglobin (FO_2Hb) is calculated from the ratio of the concentration of oxyhemoglobin to the concentration of total hemoglobin ($ctHb$):

$$FO_2Hb = \frac{cO_2Hb}{ctHb} = \frac{cO_2Hb}{cO_2Hb + cHHb + cdysHb} \qquad \textbf{(Eq. 12.11)}$$

where the $cdysHb$ represents *dyshemoglobins* which are dysfunctional hemoglobin derivatives, such as COHb, MetHb, and sulfhemoglobin that cannot reversibly

bind with O_2 but still contribute to the "total" hemoglobin measurement.

These two terms, SO_2 and FO_2Hb, can be confused easily since in most healthy individuals (and even those individuals with some disease states) the numeric values for SO_2 are equal or close to the FO_2Hb value. However, in the presence of dysfunctional hemoglobins, the values for FO_2Hb and SO_2 will deviate. They will also deviate when the patient is a smoker since CO preferentially binds to hemoglobin as opposed to O_2.

Partial pressure of oxygen dissolved in plasma (pO_2) accounts for little of the body's O_2 stores. A healthy adult breathing room air will have a pO_2 of 90 to 95 mm Hg. In an adult with a blood volume of 5 L, only 13.5 mL of O_2 will be available as dissolved oxygen (pO_2) in plasma, compared with more than 1000 mL of O_2 carried as O_2Hb.

Noninvasive measurements for following trends in oxygenation status can be attained with *pulse oximetry* (SpO_2). These devices pass light of two or more wavelengths through the tissue in the capillary bed of the toe, finger, or earlobes. Pulse oximeters that calculate oxyhemoglobin saturation based only on oxyhemoglobin and deoxyhemoglobin will overestimate the oxyhemoglobin saturation in the presence of dyshemoglobins. In addition, the accuracy of pulse oximetry can be compromised by many factors, including poor perfusion and severe anemia. Recent technological advances in pulse oximetry have allowed for measurement of COHb and MetHb.

The maximum amount of O_2 that can be carried by hemoglobin in a given quantity of blood is the *hemoglobin oxygen (binding) capacity*. The molecular weight of tetramer hemoglobin is 64,458 g/mol. One mole of a perfect gas occupies 22,414 mL. Therefore, each gram of hemoglobin carries 1.39 mL of O_2:

$$\frac{22{,}414 \text{ mL/mol}}{64{,}458 \text{ g/mol}} = 1.39 \text{ mL/g} \qquad \textbf{(Eq. 12.12)}$$

When the total hemoglobin (tHb) is 15 g/dL, and the hemoglobin is 100% saturated with O_2, the O_2 capacity is

$$15 \text{ g/100 mL} \times 1.39 \text{ mL/g} = 20.8 \text{ mL } O_2\text{/mL of blood} \qquad \textbf{(Eq. 12.13)}$$

Oxygen content is the total O_2 in blood and determined from the sum of the O_2 bound to hemoglobin (O_2Hb) and the amount dissolved in the plasma (pO_2). Normally, 97% to 99% of the available hemoglobin is saturated with O_2. Assuming a tHb of 15 g/dL, the O_2 content for every 100 mL of blood plasma becomes 20.5 mL.

$$0.3 \text{ mL} + (20.8 \text{ mL} \times 0.97) = 20.5 \text{ mL} \qquad \textbf{(Eq. 12.14)}$$

pO_2 and pCO_2 are only indices of gas exchange efficiency in the lungs; they do not reveal the *content* of either gas in the blood. For every 1 mm Hg pO_2, 0.00314 mL of O_2 will be dissolved in 100 mL of plasma at 37°C. For example, if the pO_2 is 100 mm Hg, 0.3 mL of O_2 will be dissolved in every 100 mL of blood plasma. The amount of dissolved O_2 is usually not clinically significant. However, it is of clinical significance when a patient is put in a hyperbaric chamber to assess if hemoglobin delivery is impaired. This is accomplished by increasing the pressure of inspired air to force more O_2 in solution, and changes can be subsequently evaluated.

Hemoglobin–Oxygen Dissociation

While adequate ventilation and gas exchange with the pulmonary circulation are important for tissue oxygenation, hemoglobin must also be able to release oxygen to prevent hypoxia. Hemoglobin has four heme molecules that transport and release oxygen to the tissues and subsequently transport H^+ and CO_2 from the tissue to the lungs. Increases in pO_2 will saturate hemoglobin with oxygen, whereas increased H^+ concentration and pCO_2 levels in the tissue change the molecular configuration of O_2Hb, facilitating O_2 release.

Oxygen dissociates from adult hemoglobin (A_1) in a characteristic fashion. This sigmoidal relationship is graphed (**Figure 12.4**) with pO_2 on the x-axis and percent SO_2 on the y-axis as a measure of hemoglobin's relative affinity for oxygen. Once one oxygen molecule is bound, it becomes easier for additional oxygen molecules to bind. A significant decrease in pO_2 can occur before changing the affinity of hemoglobin to oxygen, and this equates to approximately 60 mm Hg. However, below this threshold, oxygen will readily be released. The position of the different S-shaped oxygen dissociation curves reflects the differing *affinity* of hemoglobin to O_2 that reflects rate of this dissociation.

The affinity of hemoglobin for O_2 depends on hydrogen ion concentration (pH), pCO_2 and CO levels, body temperature, and 2,3-diphosphoglycerate (2,3-DPG)(2,3-biphosphoglycerate [2,3-BPG]) concentration. Changes in these variables will affect the position and shape of the oxygen dissociation curve. In actively metabolizing tissue, the conditions in the microenvironment promote the release of oxygen.

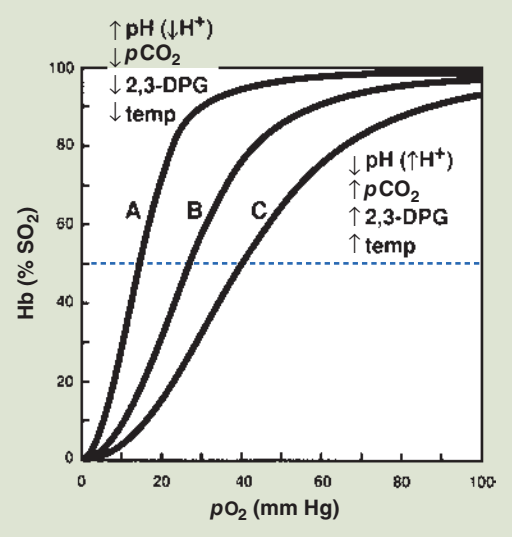


Figure 12.4 Oxygen dissociation curves. *Curve B* is the normal human curve. *Curves A* and *C* are from blood with increased affinity and decreased affinity, respectively. 2,3-DPG(2,3-BPG), 2,3-diphosphoglycerate (2,3-biphosphoglycerate).

Oxidative metabolism increases the temperature, H^+, CO_2, and 2,3-DPG(2,3-BPG) concentrations, which results in a right shift of the dissociation curve. This decreased affinity of hemoglobin to O_2 facilitates oxygen uptake for the tissue. In the lungs, temperature, H^+, pCO_2, and 2,3-DPG(2,3-BPG) decrease relative to tissue levels, causing a shift in the oxygen dissociation curve slightly to the left. This enhances O_2 binding to hemoglobin and improves O_2 uptake.

The metabolic byproduct 2,3-DPG(2,3-BPG) is also involved in two seemingly unrelated adaptations to potentially hypoxic conditions. When the β-chains of the hemoglobin molecule bind 2,3-DPG(2,3-BPG), oxyhemoglobin dissociation curve shifts to the right, which results in release of oxygen for oxygen uptake by the tissue. Anemic patients have elevated levels of 2,3-DPG(2,3-BPG) as a compensatory mechanism that helps minimize the degree of hypoxia seen in these patients. In addition, as an adaptation mechanism, 2,3-DPG levels increase at high altitudes, which results in the release of oxygen from hemoglobin. Other causes of changes in oxygen disassociation include the presence of dysfunctional hemoglobins and an elevation in CO from cigarette smoking.

The preceding discussion refers to normal adult (A_1) hemoglobin. In newborns or patients with hemoglobinopathies, the pattern of dissociation differs. For example, fetal hemoglobin, a form of hemoglobin present before birth and in the first few months of life, has a stronger affinity to oxygen compared to adult hemoglobin. Therefore, the fetal dissociation curve will be shifted to the left.

In summary, the unique structure of hemoglobin allows it to act as both an acid–base buffer and O_2 buffer. As hemoglobin moves through the body, exposure to the various microenvironments promotes association and dissociation of O_2, CO_2, and H^+. In tissue, exposure to elevated CO_2 and H^+ results in enhanced O_2 release (oxygen buffering). This release of oxygen from hemoglobin accelerates the uptake of CO_2 and H^+ by hemoglobin (acid–base buffering). In the lungs, the microenvironment promotes uptake of O_2 and release of CO_2.

Measurement

Blood Gas Analyzers: pH, pCO_2, and pO_2

Blood gas analyzers use *electrodes* (macroelectrochemical or microelectrochemical sensors) as sensing devices to measure pO_2, pCO_2, and pH. The pO_2 measurement is amperometric. **Amperometry** is defined as measurement of amperes. Ampere is the unit of measure for electric current. The reduction of oxygen produces a current that is proportional to the amount of oxygen present in the sample. Conversely, the pCO_2 and pH measurements are potentiometric. **Potentiometry** measures the electric potential between two electrodes, in which a change in voltage indicates the concentration of each analyte. Advances in microsensor technology have greatly expanded the analytic menu of whole blood analyzers. In addition to pH and blood gas measurements, many manufacturers include hemoglobin and/or hematocrit, sodium, potassium, chloride, glucose, lactate, creatinine, blood urea nitrogen, and CO-oximetry with their instrumentation. The blood gas analyzer can also calculate several additional parameters: bicarbonate, total CO_2, base excess, and SO_2.

Measurement of pO_2

The pO_2 electrode is an amperometric electrode, often referred to as a Clarke electrode. An oxygen-permeable membrane covers the tip of the electrode to selectively allow O_2 to diffuse into the electrolyte solution. A small, constant polarizing potential (approximately –0.65 V) is applied between the anode and cathode. As a result, the electrons are drawn from the anode to the cathode. The oxygen diffuses from the sample to the cathode, where it is

subsequently reduced. The current produced in this circuit is proportional to the amount of O_2 reduced at the cathode. A *microammeter* placed in the circuit between the anode and cathode measures the movement of electrons (current). The semipermeable membrane will also allow other gases to pass, such as CO_2 and N_2, but these gases will not be reduced at the cathode if the polarizing voltage is tightly controlled.

The primary source of error for pO_2 measurement is associated with the buildup of protein material on the surface of the membrane. This buildup attenuates oxygen diffusion and slows the electrode response.

It is of utmost importance to not to expose the sample to room air when collecting, transporting, and making O_2 measurements. Contamination of the sample with room air can result in significant error and can cause spurious increases in the pO_2 result. This is due to the pO_2 of room air being greater than 150 mm Hg. Another source of error in pO_2 determinations arises when a patient's sample is not analyzed immediately. This is due to the leukocytes in the sample continuing to use O_2 in metabolic processes. In addition, if a patient has a markedly increased leukocyte count, this reduction in pO_2 can happen more rapidly.

Measurement of pH and pCO_2

To measure pH, a glass membrane sensitive to H^+ is placed around an internal Ag–AgCl electrode to form a measuring electrode. Potential develops at the glass membrane as a result of H^+ diffusion from the unknown solution into the membrane's surface. Hydrogen ions are proportional to the difference in cH^+ between the unknown sample and the buffer solution inside the electrode. The developed potential at the glass membrane is measured by placing a reference electrode in the solution, and both electrodes must be connected to a pH (volt) meter. The indicator voltage is compared to the reference electrode (commonly either a calomel [Hg–HgCl] or an Ag–AgCl half-cell) that provides a steady reference voltage. The pH meter reflects the potential difference between the two electrodes. For the cell described, the Nernst equation predicts that a change of +59.16 mV, at 25°C, results in a 10-fold increase in H^+ activity, equivalent to a decrease of one pH unit (e.g., pH 7.0 to 6.0). However, this response is temperature specific. For example, at 37°C, a change of one pH unit elicits a 61.5 mV change.

pCO_2 method is based on the pH electrode. It has an outer semipermeable membrane that allows CO_2 to diffuse and dissolve into an internal layer of electrolyte, usually a bicarbonate buffer. The CO_2 that diffuses across the membrane reacts with the buffer, forming *carbonic acid*. Carbonic acid then dissociates into bicarbonate and hydrogen ions. The change in H^+ activity is measured by the pH electrode and is related to the amount of CO_2 present in the sample.

As with any electrode, the buildup of protein material on the membrane will affect diffusion and cause errors. pCO_2 electrodes are slower to respond because of the chemical reaction that must be completed. As previously mentioned, electrodes are also sensitive to temperatures and barometric pressure. Other sources of error include erroneous calibration caused by incorrect or contaminated calibration materials.

Types of Sensors

Macroelectrode sensors have been used in blood gas instruments since the beginning of the clinical measurement of blood gases. These have been modified over time in an effort to simplify their use and minimize the required sample volume and maintenance. *Microelectrodes* basically are miniaturized macroelectrodes. Miniaturization became possible with better manufacturing capabilities and with the development of the sophisticated electronics required to handle minute changes in signal.

Thick and thin film technology is a further modification of electrochemical sensors. Although the measurement principle is identical, the sensors are reduced to tiny wires embedded in a printed circuit card. The special card has etched grooves to separate components. A special paste material containing the required components (similar in function to the electrolytes of macroelectrodes) is spread over the sensors. To reduce the required sample volume, several sensors can be placed on a single small card. These sensors are disposable and less expensive to manufacture, which reduces maintenance.

Optical sensors are another form of technology that is used for blood gas measurements. Optical sensors (*optodes*) measure fluorescence or phosphorescence of organic dyes with O_2, CO_2, and H^+. The dye is separated from the sample by a membrane, as with electrodes, and the analyte diffuses into the dye, causing either an increase in or a quenching of fluorescence proportional to the amount of analyte. Calibration is used to establish the relationship between concentration and fluorescence. Normally, a single calibration will suffice for long periods because this

technology is not subject to the drifts seen in electrochemical technology.

Optical technology has been applied to indwelling blood gas systems.[3] Fiberoptic bundles carry light to sensors positioned at the tip of catheters, and other bundles carry light back, allowing changes in fluorescence to be measured in a catheter within the patient's arterial system. The commercial development of indwelling systems has been limited by the increased probability of thrombogenesis and protein buildup on the membrane, separating the sample from the fluorescing dyes. This buildup impedes free sample diffusion into the measuring chamber.

Calibration

Calibration of a blood gas analyzer will vary among manufacturers. Normally, two different gas mixtures with known pCO_2 and pO_2 levels are used. One of the mixtures is used to calibrate the lower end, and the composition of this mixture is 0% O_2 and 5% CO_2, whereas the second mixture is used to calibrate the gases on the upper end, and the composition of the mixture is 20% O_2 and 10% CO_2.

Most blood gas instruments will calibrate automatically at specified time intervals since electrodes are drifted over time. Most electrodes will require calibration every 30 to 60 minutes. Instruments are programmed to prompt the operator if the values deviate from the expected value. For example, if the value(s) obtained during calibration exceed(s) a programmed tolerance limit, flagging of a drift error will occur at the time of calibration, and corrective action will need to be taken before patient samples can be analyzed.

The pH electrode is calibrated against two primary traceable buffer solutions traceable to standards that meet specifications set by the National Institute of Standards and Technology (NIST). Typically, there are two calibrators, a low and high calibrator, with pH values of 6.8 and 7.38, respectively. The calibrators must be stored at the stated temperature and not exposed to room air, since pH changes with the absorption of CO_2.

pH and blood gas measurements are extremely sensitive to temperature. It is critical that the electrode sample chamber be maintained at constant temperature for all measurements. All blood gas analyzers have electrode chambers thermostatically controlled to 37°C ± 0.1°C, as small temperature variations can drastically change the pH and blood gas values. Accordingly, the Nernst equation is temperature specific; if the temperature of the measurement system changes, the output (voltage) will change. The solubility of gases in a liquid medium also depends on the temperature; as the temperature goes down, the solubility of the gas increases.

Correction for Temperature

Values for pH, pCO_2, and pO_2 are temperature dependent. By convention, these measurements are made at 37°C. But what happens when a patient's actual body temperature differs from 37°C? After manual entry of the patient's body temperature, the blood gas analyzer software performs these corrections. However, the data can be difficult to interpret because appropriate reference ranges for the patient's temperature must be used, but these are not readily available. When temperature-corrected results are reported, it is critical that results also be reported with the 37°C (noncorrected) results.

Calculated Parameters

Several acid–base parameters can be calculated from measured pH and pCO_2 values. Manufacturers of blood gas instruments include algorithms to perform these calculations. No calculated parameter is universally used; many physicians have preferred parameters for identifying various pathologies.

HCO_3^- measurement is based on the Henderson-Hasselbalch equation and can be calculated when pH and pCO_2 are known. The Henderson-Hasselbalch equation assumes that the pK_a of the bicarbonate buffer system in plasma at 37°C is 6.1.

Carbonic acid (H_2CO_3) concentration can be calculated using the solubility coefficient of CO_2 in plasma at 37°C. The solubility constant to convert pCO_2 to mmol/L of H_2CO_3 is 0.0307. If the temperature or the composition of plasma changes (e.g., an increase in lipids, in which gases are more soluble), the constant will also change.

Total carbon dioxide content ($ctCO_2$) is the bicarbonate plus the dCO_2 (carbonic acid) plus the associated CO_2 with proteins (carbamates). A blood gas analyzer approximates $ctCO_2$ by adding the bicarbonate and carbonic acid values ($ctCO_2 = cHCO_3^- + [0.0307\ pCO_2]$).

Some clinicians use **base excess** to assess the metabolic component. Base excess is calculated from an algorithm that uses the patient's pH, pCO_2, and hemoglobin. A positive base excess value indicates an excess of bicarbonate and suggests *metabolic alkalosis*. Conversely, a negative value (base deficit) indicates a deficit of bicarbonate and

suggests *metabolic acidosis*. Because the nonrespiratory alkalosis or acidosis may be a result of primary disturbances or compensatory mechanisms, base excess values should not be used alone in assessing a patient's acid–base status.

Spectrophotometric Determination of Oxygen Saturation (CO-Oximetry)

The *actual percent oxyhemoglobin* (O_2Hb) can be determined spectrophotometrically using a CO-oximeter designed to directly measure the various hemoglobin derivatives. Each hemoglobin derivative has a characteristic absorbance curve (**Figure 12.5**). The number of hemoglobin derivatives measured will depend on the number and specific wavelengths incorporated in the instrumentation. For example, two-wavelength instrument systems can measure only two hemoglobin species (i.e., O_2Hb and HHb), which are expressed as a fraction or percentage of the total hemoglobin.

Ideally, CO-oximeters should have four wavelengths for measurements of HHb, O_2Hb, COHb, and MetHb. Instruments with five wavelengths can also measure sulfhemoglobin as well as recognize dyes and pigments, turbidity, and abnormal proteins. Some CO-oximeters employ hundreds of wavelengths, which has reduced measurement interferences. Microprocessors control the sequencing of multiple wavelengths of light through the sample and apply the necessary matrix equations after absorbance readings are made to calculate the percentage of the individual hemoglobin derivative:

$$\begin{aligned} O_2HB &= a_1A_1 + a_2A_2 + \ldots + a_nA_n \\ HHb &= b_1A_1 + \ldots + b_nA_n \\ COHB &= c_1A_1 + c_2A_2 + \ldots + c_nA_n \\ MetHb &= d_1A_1 + d_2A_2 + d_nA_n \end{aligned} \qquad \textbf{(Eq. 12.15)}$$

where a_1, a_n, b_n, etc., are coefficients that are analogues of the absorption constant (a) that are derived from established methods and A_1, A_2 . . . An are the absorbances of the sample. The matrix equations will change depending on the number of wavelengths of light (which is manufacturer specific) passed through the sample.

The results from these instruments should not be confused with a calculated SO_2 from a blood gas analyzer, which *estimates* the SO_2 value from a measured pO_2 and an empirical equation based on the oxygen–hemoglobin dissociation curve. Only measured O_2Hb values reflect the patient's true status because calculated SO_2 and O_2Hb values will be different in the presence of dyshemoglobins. For example, in carbon monoxide poisoning, the SO_2 will likely be normal with a significantly decreased O_2Hb level.

As with any spectrophotometric measurement, potential sources of error exist, including faulty calibration of the instrument and spectral-interfering substances. The presence of any substances absorbing light at the same wavelengths used in the

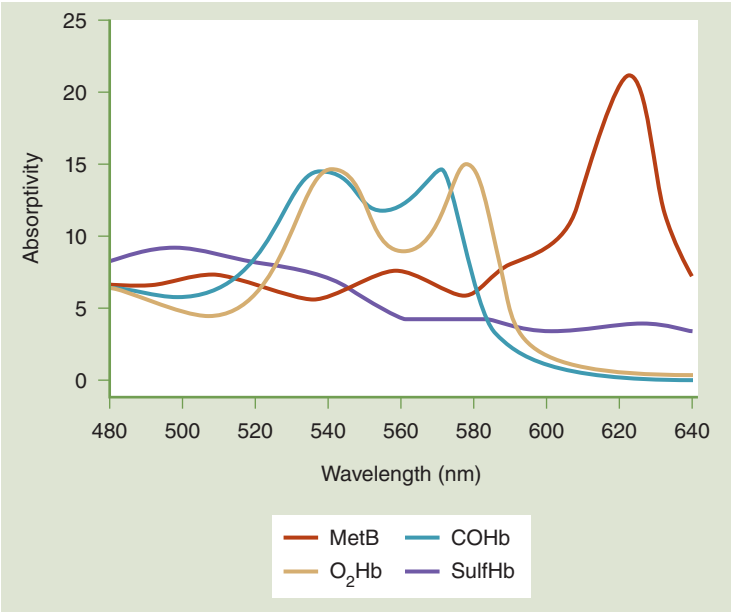


Figure 12.5 Optical absorption of hemoglobin fractions.

Reproduced from *Clin Chem News* 1990 [January], with permission.

measurement is a potential source of error. Product claims for specific instruments must be consulted for interferences.

The primary purpose of determining O_2Hb is to assess oxygen transport from the lungs. Before blood sample collection, it is best to stabilize the patient's ventilation status. Following changes in supplemental O_2 or mechanical ventilation, wait an appropriate period of time before the sample is drawn. Heparinized blood samples should be collected under anaerobic conditions and mixed immediately. All samples should be analyzed promptly to avoid changes in saturation resulting from the consumption of oxygen by metabolizing cells.

Quality Assurance

Preanalytic Considerations

Blood gas measurements, like all laboratory measurements, are subject to preanalytic, analytic, and postanalytic errors. Blood gas measurements are particularly affected by *preanalytic errors*—those introduced during sample collection and transportation of samples before analysis.[1,4]

The quality assurance cycle is shown in **Figure 12.6**. The steps included in the analytic phase are under the direct control of the laboratory. Notably, most of the quality assurance cycle lies outside the laboratory; therefore, the laboratorian must take an active role in educating *all* individuals involved to ensure quality. The preanalytic considerations start with proper patient identification, which is essential before any blood specimen is collected. Once collected, the specimen must be correctly labeled with accurate information. Only trained personnel who have the knowledge about possible sources of error and the experience with the blood drawing equipment and technique should draw samples for blood gas analyses.

Because arterial blood collection may be painful and results in patient hyperventilation, which lowers the pCO_2 and increases the pH, the ability to reassure the patient is essential. The choice of site (radial, brachial, femoral, or temporal artery) depends on the predominant patient population (e.g., pediatric patients, burn patients, and outpatients) within an institution. The Clinical and Laboratory Standards Institute (CLSI) publication *Procedures for the Collection of Arterial Blood Specimens* is an excellent reference.[4]

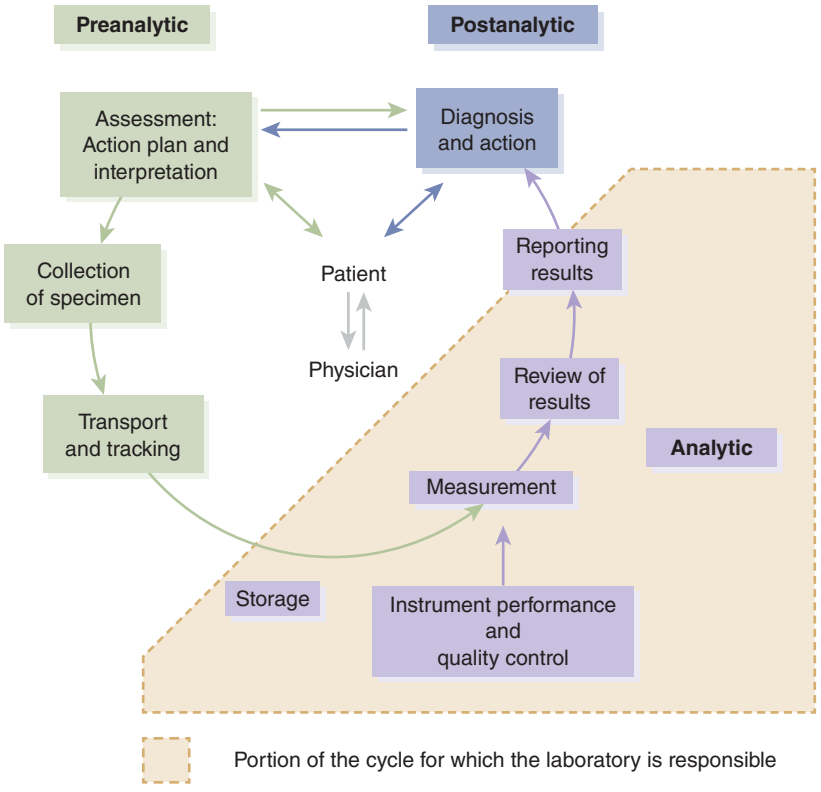


Figure 12.6 Blood gas analysis quality assurance cycle.

Although arterial samples for pH and blood gas studies are recommended, peripheral venous samples can be used if pulmonary function or O_2 transport is not being assessed. For venous samples, the specimen source must be clearly specified and the appropriate (venous) reference ranges included with the results for data interpretation. Depending on the patient, capillary blood may need to be collected to assess pH and pCO_2. The correlation with arterial blood is good for pH and pCO_2. However, capillary pO_2 values do not correlate well with arterial pO_2 values because of exposure to room air. Central venous (pulmonary artery) blood samples are obtained to assess O_2 consumption, which is calculated from the O_2 content of arterial and pulmonary artery blood and the cardiac output. For sample collection from an indwelling arterial line, an appropriate blood volume must initially be withdrawn and discarded from the line to assure that the actual collected sample contains only arterial blood. Proper flushing procedure minimizes the chance of specimen contamination with solutions (i.e., liquid heparin, medication, or electrolyte fluids) that may be in the line.[1]

Sources of error in the collection and handling of blood gas specimens include the following:

- Collection device
- Form and concentration of heparin used for anticoagulation
- Speed of syringe filling
- Maintenance of the anaerobic environment
- Mixing of the sample to ensure dissolution and distribution of the heparin and transport and storage time before analysis

For proper interpretation of blood gas results, the patient's status—ventilation (on room air or supplemental O_2) and body temperature—at the time of sample collection must also be documented.

In most instances, the ideal collection device for arterial blood sampling is a 1- to 3-mL self-filling, plastic, disposable syringe, containing the appropriate type and amount of anticoagulant. Evacuated collection tubes are not appropriate for blood gases. While both dry (lyophilized) and liquid lithium heparin are acceptable anticoagulants, the liquid form is not recommended because liquid heparin can dilute the sample and possibly alter the sample due to equilibration with room air.[1,4] Once drawn, the blood in the syringe must be mixed thoroughly with the heparin to prevent microclots from forming. Adequate mixing is also important immediately before analysis to resuspend the settled cells. Although sodium and lithium salts of heparin are recommended for pH and blood gas analysis, other forms are available: ammonium, zinc, electrolyte balanced, and calcium titrated. Selection of the proper type of heparin is particularly important with instruments combining blood gas, electrolyte, and metabolite measurements. It is important to consult the manufacturer's product insert.

Slow filling of the syringe may be caused by a mismatch of syringe and needle sizes. While a needle that is too small reduces patient's pain and therefore the likelihood of arteriospasm and hematoma, it may produce bubbles that affect pCO_2 and pO_2 measurement. It may also cause hemolysis of the blood, which significantly alters potassium levels, an electrolyte frequently ordered along with pH and blood gases. Maintenance of an anaerobic environment is critical to correct results. Any air trapped in the syringe during the draw should be immediately expelled at the completion of the draw.

Transport time and analysis should be minimal to reduce cellular metabolism. Anaerobic condition results in oxygen and glucose consumption and carbon dioxide and lactate production. Placing the capped blood gas syringe in ice water promptly after the draw minimizes cell metabolism. In uncapped syringes, there is potential for pO_2 to increase due to oxygen diffusing from the atmosphere into the anaerobically collected specimen. In addition, lower temperatures cause increased oxygen solubility in blood and a left shift in the oxyhemoglobin dissociation curve, resulting in more oxygen combining with hemoglobin. Consequently, when the sample is heated by the blood gas analyzer, the measured pO_2 is falsely elevated.[5] The best practice in avoiding many of these preanalytic errors is to analyze the sample as quickly as possible. Oxygen and carbon dioxide levels in blood kept at cool room temperatures for up to 20 to 30 minutes are minimally affected except in the presence of an elevated leukocyte or platelet count. In the presence of elevated leukocytes, glycolysis will also be elevated, which has a pronounced effect on pH, pO_2, and pCO_2. The CLSI guidelines advocate samples be kept at room temperature and analyzed in less than 30 minutes.[1] Preanalytical errors of samples that will be utilized on multianalyte instruments should also be considered. For example, prolonged ice water slurry storage can result in falsely elevated potassium in whole blood samples. Seeing as sample procurement and handling are the source of most errors in blood gas analysis, it is necessary that procedures and policies are carefully constructed and that adherence is monitored to ensure quality. At this time, no quality

CASE STUDY 12.1, PART 5

Doug's blood gas sample was collected and delayed in transport by 45 minutes.

7. What results, if any, were altered by this delay?

© AJR_photo/Shutterstock.

assessment approach can detect all preanalytic problems associated with blood gas analysis.

Analytic Assessments: Quality Control and Proficiency Testing

Quality control (QC) assesses the process in the three components of analytic phase: preanalytic, analytic, and postanalytic testing process. Ideally, to evaluate the performance of a test method, a control should reflect an actual patient sample matrix; however, this is not possible for blood gases. Nonetheless, there are several approaches to assess quality control for blood gases.[1]

Surrogate *liquid control materials* are the basis of most of the traditional QC practices.[6] Usually, these are sold in sealed glass containers with solutions equilibrated with gases of known concentration. The containers can be snapped open and then analyzed in a manner similar to a patient's sample. Alternatively, the instrument can automatically assess the liquid at programmed intervals for analysis. Surrogate liquid controls are typically available in at least three levels, corresponding to values observed with low, expected (i.e., "normal"), and elevated values for each of the measured analytes. The QC materials vary in stability and are susceptible to temperature variation in storage and handling. Furthermore, liquid control materials have significantly different matrices than fresh whole blood; therefore, the laboratorian must be aware that they may not detect problems that can affect patient samples, or they may detect errors induced by improper handling of the commercial controls. Aqueous-based controls consist of buffered medium sealed in vials with gas mixtures. These have low O_2 solubility, making them sensitive to factors that affect pO_2. Aqueous controls must be at room temperature for analysis, and manufacturer recommendations must be followed closely or pO_2 results may be unreliable. Hemoglobin-containing and emulsion-based controls have increased O_2 solubility to better resist O_2 changes.

Tonometry is the equilibration of a fluid with gases of known concentration and under controlled conditions, such as constant temperature, barometric pressure, and humidification.[1] When whole blood is used, it is considered the reference procedure to establish the accuracy for pCO_2 and pO_2. However, tonometry is rarely used today because the technique is considered to be cumbersome and time consuming.

Split sample analysis using two or more instruments for simultaneous analysis of a patient sample is another quality control approach. The delta checks, or the difference in values obtained on the two instruments, often identify problems that might be missed by routine QC. The allowable difference in duplicates should be tighter than those observed with surrogate liquid controls. Further investigation is usually needed, as discrepancies between results provide no clues regarding which data point is wrong or which instrument is malfunctioning. Consequently, the duplicate assay approach cannot be used as the sole method of QC. It can be a useful approach for detection of errors and for troubleshooting instruments.

Internal controls have become particularly popular for testing devices used in point of care. This category includes a variety of quality assurance mechanisms that are integrated into the design of the device, such as electronic QC (which simulates sensor signals to test electronic components), automated procedural controls (which ensures that certain steps of the method occur appropriately), and automated internal checks (which may, for example, ensure the quality of a raw electronic signal). Such controls may check all, but usually just a portion, of the test system's analytic components each time the test is performed.

Whatever the QC approach may be, the QC needs of the blood gas laboratory contrast sharply with those of the general laboratory, which can analyze many patient samples as a group and include multiple control specimens with each run. In blood gas instruments, the critical nature of the measurements and the limited patient sample volume do not always allow for repeat analyses, if problems exist. Consequently, the blood gas laboratory must perform prospective QC because instruments must be

CASE STUDY 12.2, PART 3

Remember Kelsey. The little girl has been admitted to the pediatric wing and continues to improve.

© Julián Rovagnati/ShutterStock.

5. Predict Kelsey's pO_2 value based on the ABG results from the ED. Increased, decreased, or normal?

Analyte	Patient Value	Reference Range
pH	7.12	7.35–7.45
pCO_2	92 mm Hg	35–45 mm Hg
HCO_3^-	29 mmol/L	22–26 mmol/L

6. Based upon her arterial pH of 7.12, was her hemoglobin's affinity for oxygen increased, decreased, or normal?

As her ABG values began to normalize, her samples were run on a point-of-care ABG device.

7. What type of sensors could this device use?

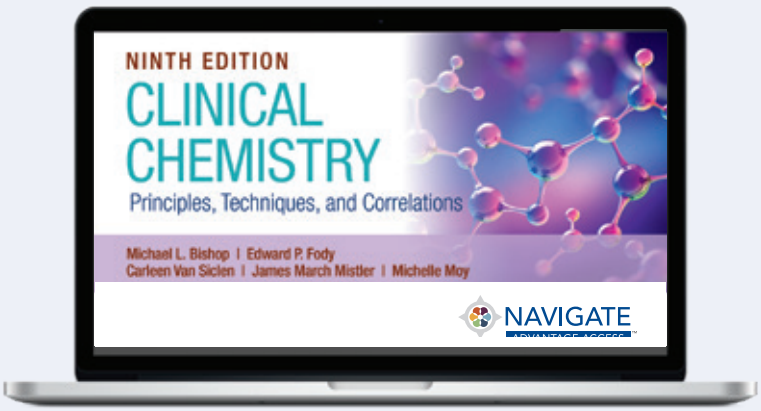
prequalified to ensure proper performance before the patient sample arrives for analysis.

Several manufacturers have devised onboard QC systems that greatly reduce operator handling errors. Advances in computer technology and software monitoring algorithms now can automatically analyze QC products, continuously monitor instrument performance, and detect some preanalytical problems (e.g., microclots and hemolysis) for improved reliability and error detection in blood gas and multianalyte instruments. Participating in external, interlaboratory surveys or proficiency testing programs is another essential component of ensuring the quality of blood gas measurements.[7]

Ongoing comparisons of results through proficiency testing help ensure that systematic (accuracy) errors do not slowly increase and go undetected by internal QC procedures. A rigorous internal QC program ensures internal consistency. Good performance in a proficiency testing program ensures the absence of significant bias relative to other laboratories and further confirms the validity of a laboratory's patient results. If an individual analyzer does not produce proficiency testing results consistent with its peer laboratories (those using the same method/instrument) or if the differences between values change over time, suspicion of the instrument's performance is warranted.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. D'Orazio P, Ehrmeyer SS, Jacobs E, Toffaletti JG, Wandrup JH. *Blood Gas and pH Analysis and Related Measurements; Approved Guideline*. Wayne, PA: CLSI; 2009.
2. Ehrmeyer S, Ancy J, Laessig R. Oxygenation: measure the right thing. *J Respir Care Pract.* 1998;11:25–30.
3. Roupie EE. Continuous assessment of arterial blood gases. *Crit Care.* 1997;1(1):11–14.
4. Blonshine S, Fallon KD, Lehman CM, Sitting S. *Procedures for the Collection of Arterial Blood Specimens; Approved Standard*. 4th ed. Wayne, PA: CLSI; 2012.
5. Ancy J. Blood gases and preanalytic error prevention. *RT Magazine*. 2012:7–10.
6. Parvin CA, Person NB, Baumann N, et al. *Statistical Quality Control for Quantitative Measurement Procedures: Principles and Definitions*. 4th ed. Wayne, PA: CLSI; 2016.
7. Ehrmeyer S, Laessig R. Benefits for voluntary proficiency testing. *Adv Lab Admin.* 2006;15(8):3.

PART 3

Endocrinology and Clinical Correlation

CHAPTER 13

Basic Endocrinology

Scott Moore

CHAPTER OUTLINE

KEY TERMS

Adrenocorticotropic hormone (ACTH)
Amines
Diurnal variation
Effector/trophic hormone
Endocrine
Exocrine
Glucocorticoids
Growth hormone (GH)
Growth hormone-releasing hormone (GHRH)
Hormone
Hypothalamic-pituitary-thyroid axis
Mineralocorticoids
Peptides
Steroids
Thyroxine-binding globulin (TBG)
Trophic hormone
Tropic hormone

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define basic terms in endocrinology.
- Outline hormone classifications and basic structures.
- Discuss hormone biosynthesis, metabolism, catabolism, transport, and elimination.
- Sketch positive and negative feedback mechanisms and provide examples.
- Correlate laboratory results with disease states and endocrine disorders.
- Explain the most common screening and diagnostic tests for endocrine disorders.
- Interpret patient results with awareness of common factors that affect hormone levels.
- Describe classification, source, effect, and mechanism of action of each hormone.
- Apply endocrinology knowledge to case studies throughout the chapter.
- Investigate discrepant endocrinology laboratory results.

The endocrine system is composed of glands that produce hormones, which are secreted directly into the bloodstream for use throughout the body (**Figure 13.1**). A **hormone** is a chemical substance that sends a message to another cell in the body. This message can either regulate or control the activity of other bodily functions. These cellular messages are sent to other cells via the bloodstream (**endocrine**), the gastrointestinal tract (**exocrine**), neurologically (*neurocrine*), or in interstitial fluid (*paracrine*). This chapter will discuss hormones secreted by the following endocrine glands including the hypothalamus,

CASE STUDY 13.1, PART 1

Meet Sarah, a 26-year-old female who presents 5 weeks postpartum with seizure, inability to breastfeed, and constantly feeling cold after a challenging delivery that required 8 units of packed red blood cells (PRBCs). Her TSH, ACTH, and prolactin levels are low.

© Pablo Calvog/Shutterstock.

CASE STUDY 13.2, PART 1

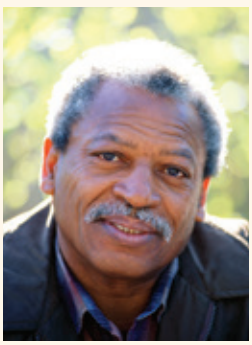
Meet Paulina, a 47-year-old female, who reports symptoms consistent with polyuria, polydipsia, and polyphagia for the last 3 months. The physician orders a HbA1c. Her result is 4.6%. She is seen by an endocrinologist, who determines that a peptide hormone secreted in the posterior pituitary gland is the cause of her symptoms.

© Krakenimages.com/Shutterstock.

CASE STUDY 13.3, PART 1

Meet Mike, a 59-year-old male, who presents to the physician and reports that he has been noticing purple "stretch marks" on his abdomen and has been urinating 10 to 15 times per day and feels like he can never quench his thirst. He has smoked 1 pack per day for 39 years. His serum electrolytes are shown below. A biopsy reveals a small-cell lung carcinoma. Based on signs and symptoms, he is referred to endocrinology.

Electrolyte Panel	Result	Reference Range
Sodium	136 mmol/L	136–145 mmol/L
Potassium	2.1 mmol/L	3.4–5.0 mmol/L
Chloride	78 mmol/L	98–107 mmol/L
Carbon dioxide	24 mmol/L	23-29 mmol/L



© Thinkstock.

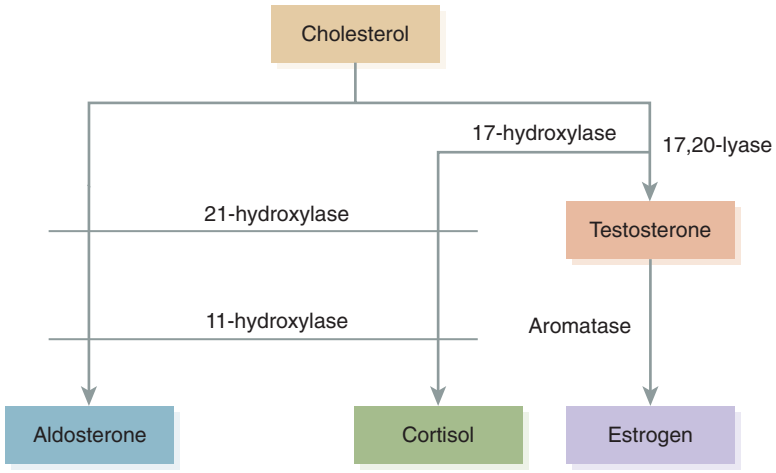


Figure 13.1 Simplified adrenal steroid production. All adrenal steroids are made from the cholesterol backbone. Patients must have functioning 17-hydroxylase and 17,20-lyase enzymes in order to produce sex steroids. Likewise, patients will require a functioning 21-hydroxylase and 11-hydroxylase to produce aldosterone with the addition of 17-hydroxylase to produce cortisol. Deficiencies in specific enzymes can be detected by presenting symptoms of patients. Patients with a deficient 17-hydroxylase will have hypertension from the aldosterone produced, likely be infertile or lack secondary sexual characteristics, and may have episodes of hypoglycemia.

pituitary, thyroid, parathyroid, adrenal, pancreas, ovaries, and testes.

Hormone Classification

Hormones can be classified by their chemical structure, mechanism of action, stimulation of endocrine glands, and/or effect of hormone (see **Table 13.1**). Hormones are created from different substrates, and this provides a useful classification system. There are several structural categories of hormones: **amine**, **peptide**, protein, glycoprotein, **steroids**, and fatty acid (eicosanoid) hormones. Amine, peptide, and protein hormones are hydrophilic, are largely water soluble, and have variable half-lives (i.e., they exert their effects for a short time in the body). Lipid hormones, on the other hand, are hydrophobic and so have longer half-lives—that is, they will stay in the body and exert their effects longer than protein hormones. The classification of hormones and their characteristics are discussed below.

Amine

Amino acids like tryptophan or tyrosine are modified to create *amine hormones*. Examples include epinephrine from tyrosine and melatonin from tryptophan. Other amine hormones include norepinephrine, triiodothyronine, thyroxine, and serotonin, as well as urinary 5-hydroxyindoleacetic acid. Amine hormones have very short half-lives compared to other protein hormones.

Peptide

Peptide hormones are made of chains of amino acids. A peptide hormone is generally considered to be an unbroken chain of amino acids of 50 or less, whereas a protein hormone has an amino acid length greater than 50. Most peptide hormones found in humans have around 20 amino acids. These peptide hormones are hydrophilic (water soluble), which means they cannot cross cellular membranes easily. Instead, peptide hormones must first bind to membrane-bound receptors. When the receptors are activated, they characteristically initiate a cascade of intracellular signals to elicit a response. Examples of peptide hormones include vasopressin and oxytocin.

Protein

Like peptide hormones, *protein hormones* are made of chains of amino acids. Just like other proteins in mammalian bodies, these hormone proteins will have *primary, secondary*, and *tertiary* structures to them. These hormones are formed from a large number of amino acid residues in any length greater than 50 amino acids. There are many protein hormones including **adrenocorticotropic hormone (ACTH)**, calcitonin, insulin, glucagon, and oxytocin, among others.

Table 13.1 Classification of Major Hormones

	Organ Sources	Classification	Effects	Regulation
GH[28]	Anterior Pituitary	Peptide	Body growth, basal metabolic functions, and acute phase stress reactant	Production is tightly regulated through several complex feedback mechanisms and lifestyle factors
ACTH[29]	Anterior Pituitary	Protein	Regulates cortisol and androgen production	Negative feedback regulation by serum glucocorticoid concentration and neural input
TSH[30]	Anterior Pituitary	Protein	Tropic hormone to regulate T4 production	Thyrotropin releasing hormone (TRH) release from the hypothalamus and negative feedback by circulating serum T4 levels
FSH[8]	Anterior Pituitary	Glycoprotein	Women: Initiate follicular maturation and estrogen production Men: Spermatogenesis	Pulsatile GnRH secretion from hypothalamus Women: Negative feedback from estrogen levels Men: Negative feedback from inhibin B
LH[9]	Anterior Pituitary	Glycoprotein	Women: Initiation of ovulation Men: Activate testicular testosterone production	Pulsatile GnRH secretion from hypothalamus Women: Negative feedback from estrogen levels Men: Negative feedback from inhibin B
PRL[10]	Anterior Pituitary	Peptide	Lactation	Dopamine-mediated hypothalamic regulation, nipple stimulation, light, olfaction, and stress
ADH[20]	Posterior Pituitary	Peptide	Activation of ADH-sensitive water pores in the collecting duct to regulate osmolarity	Osmotic receptors in the hypothalamus
Calcitonin[31]	C-cells of Thyroid	Protein	Inhibits osteoclastic breakdown	Negative feedback from elevated Ca^{2+}
PCT[32]	Parafollicular C-cells of the Thyroid	Peptide	Prohormone to calcitonin	CALC-1 gene expression, linked with calcitonin regulation
PTH[33]	Parathyroid cells	Peptide	Increases circulating calcium ions, by promoting Ca^{2+} absorption in the intestines, promotes Ca^{2+} reabsorption, and blocks PO_4^- reabsorption	Negative feedback from elevated serum Ca^{2+} and PO_2^-

(continues)

Table 13.1 Classification of Major Hormones *(continued)*

	Organ Sources	Classification	Effects	Regulation
Insulin[34]	β cells of the pancreas	Protein	Drops plasma glucose levels by binding directly to insulin receptors on cell surfaces to allow glucose entry into cells through GLUT4 channels	Secreted primarily in response to glucose, but fats and amino acids can create more robust insulin secretion
Glucagon[35]	α cells of the pancreas	Protein	Increases glycogenolysis and gluconeogenesis by binding to G-coupled protein receptors throughout the body	α cells do not express a glucagon receptor and are therefore considered to regulate their own secretion through insulin
Gastrin[36]	G cells the pyloric antrum of the stomach	Peptide	Induces HCl secretion from the parietal cells in the stomach	Directly stimulated by vagal neurons and presence of amino acids in the stomach
Secretin[37]	S cells in the duodenum	Peptide	Induces HCl secretion from parietal cells and bicarbonate release from the pancreas	Stimulated by gastric acid
β-hCG[38]	Syncytiotrophoblast cells of the placenta	Peptide	Stimulates the corpus luteum in the ovary to produce progesterone to maintain pregnancy	Regulated by several hormones, growth factors, cytokines, PPARγ ligands, and the homeobox gene (DLX3)
Cortisol[39]	Zona fasciculata in the adrenal cortex	Steroid	Immune response, stress response, and glucose homeostasis	Hypothalamic-pituitary-adrenal (HPA) axis
Aldosterone[40]	Zona glomerulosa in the adrenal cortex	Steroid	Increase sodium absorption in the distal tubules and proximal collecting ducts of the nephrons	Renin–angiotensin–aldosterone system
Testosterone[41]	Zona reticularis in the adrenal cortex	Steroid	Men: Regulate sex determination, male sex characteristics, spermatogenesis, and fertility	HPA axis
DHEA[40]	Zona reticularis in the adrenal cortex	Steroid	Relatively weak androgen and serves as a precursor for testosterone and estrogen	HPA axis
Progesterone[42]	Zona reticularis in the adrenal cortex, gonads, and corpus luteum of pregnancy in the ovary	Steroid	Promote development of the endometrial lining and allow for an increase in FSH during the menstrual cycle	Regulated by several factors that affect the duration of pregnancy

(continues)

Table 13.1 Classification of Major Hormones *(continued)*

	Organ Sources	Classification	Effects	Regulation
Estrogens[43]	Zona reticularis in the adrenal cortex	Steroid	Breast development, endometrial proliferation, proliferation of vaginal mucosal cells, bone development, increase HDL-C and decrease LDL-C	Regulation by sex hormone binding globulin (SHBG), albumin binding, and negative feedback in the hypothalamus and anterior pituitary gland
Catecholamines[44]	Adrenal Medulla	Amine neurotransmitter	Dopamine, epinephrine, and norepinephrine create the "fight or flight" response including vasoconstriction, tachycardia, sweating, anxiety	Multiple factors including availability of tyrosine, glucocorticoid upregulation, and degradation via catechol-o-methyl transferase (COMT) or monoamine oxidase (MAO) activity
Thyroxine[45]	Thyroid follicular cells	Steroid	Activate metabolism and increase the basal metabolic rate	Regulated by albumin binding, thyroxine binding globulin (TBG), and negative feedback through TSH by circulating serum T4 levels
Serotonin[46]	Neural cells	Amine neurotransmitter	CNS: Perception, memory, mood Intestines: Intestinal motility Platelets: Blood vessel tone and functionality	Mainly in the central nervous system nuclei, it has also been found in the enterochromaffin cells located throughout the digestive tract and platelets

Glycoprotein

Glycoprotein hormones are conjugated proteins bound to carbohydrates, which include galactose, mannose, or fructose. Examples of glycoproteins include follicle-stimulating hormone (FSH), luteinizing hormone (LH), and thyroid-stimulating hormone (TSH). These hormones have similar characteristics of other protein hormones in solubility and half-life.

Steroid

Steroid hormones are all derived from a lipid, typically cholesterol. Because these hormones are made from lipids, they are hydrophobic and must be transported in the blood bound to carrier proteins. However, unlike protein hormones, steroid hormones can cross cellular membranes easily because of their lipid content. Steroid hormones characteristically act in the nucleus of cells, binding to nuclear receptors in order to modify transcription and translation of genes. As a rule, these hormones have a delayed effect because genes are affected, which need to be produced to exert their effects. Examples of steroid hormones include aldosterone, cortisol, estrogens, progesterone, testosterone, and other androgens like dehydroepiandrosterone (DHEA).

Fatty Acid

This category of hormones is made up of small fatty acid derivatives of arachidonic acid. Examples include eicosanoids, leukotrienes, prostaglandins, and thromboxanes. Unlike steroid hormones, these hormones are rapidly degraded and are only effective in the body for seconds.

Introduction to Hormones

The remainder of the chapter will focus on introducing the classifications, mechanisms of action, organ sources, and regulatory mechanisms for each selected hormone, starting with hormones secreted by the anterior pituitary gland.

Growth Hormone (GH)

The pituitary gland is vital for normal growth, and if the pituitary gland is removed, growth ceases. Even if hormones from other endocrine glands that are acted on by the pituitary are replaced (thyroxine, adrenal steroids, and gonadal steroids), growth is not restored until **growth hormone (GH)** is administered. When GH is given in isolation without the other hormones, growth is not promoted. Therefore, it takes complete functioning of the pituitary for growth in an individual. It also takes adequate nutrition, normal levels of insulin, and overall good health to achieve a person's genetic growth potential.

GH, also called *somatotropin*, is structurally related to prolactin and human placental lactogen. A single peptide with two intramolecular disulfide bridges, it belongs to the direct **effector/trophic** class of anterior pituitary hormones. A **trophic hormone** stimulates the target tissue to grow or increase in size or number. The somatotrophs, pituitary cells that produce GH, constitute over one-third of normal pituitary weight. Release of somatotropin from the pituitary is stimulated by the hypothalamic peptide **growth hormone–releasing hormone (GHRH)**; GH's secretion is inhibited by somatostatin (SS).[1] GH is secreted in pulses, with an average interval of 2 to 3 hours, and with the most reproducible peak occurring at the onset of sleep.[2] Between these pulses, the level of GH may fall below the detectable limit, which may result in an incorrect clinical evaluation of GH deficiency if based on a single measurement. Ghrelin, an enteric hormone, plays an important role in nutrient sensing, appetite, glucose, and regulation. It is also a potent stimulator of GH secretion.[3]

No other hypothalamic–hypophyseal system more vividly illustrates the concept of an open-loop paradigm than that seen with GH. The on-and-off functions of GHRH/SS and the basic pattern of secretory pulses of GH are heavily modulated by other factors (**Table 13.2**).

GH is considered an amphibolic hormone because it directly influences both anabolic and catabolic processes. It has many diverse effects on carbohydrate, lipid, and protein metabolism.

Table 13.2 Modifiers of Growth Hormone Secretion

Stimulates Growth Hormone Secretion	Inhibits Growth Hormone Secretion
Sleep	Glucose loading
Exercise	β-Agonists (e.g., epinephrine)
Physiologic stress	α-Blockers (e.g., phentolamine)
Amino acids (e.g., arginine)	Emotional/psychogenic stress
Hypoglycemia	Nutritional deficiencies
Sex steroids (e.g., estradiol)	Insulin deficiency/ hyperglycemia
α-Agonists (e.g., norepinephrine)	Thyroxine deficiency
β-Blockers (e.g., propranolol)	

During carbohydrate metabolism, GH maintains glucose levels within the normal range. It allows an individual to effectively transition from a nonfasting state to a fasting state without experiencing a shortage of substrates required for normal intracellular oxidation. GH directly antagonizes the effect of insulin on glucose metabolism in peripheral tissues, promotes gluconeogenesis in the liver, and stimulates lipolysis.[4,5] Lipolysis is the degradation of lipids by hydrolysis.

During lipolysis, triglycerides are broken down to fatty acids and glycerol. Growth hormone enhances the utilization of fat for energy by stimulating triglyceride breakdown and oxidation in adipose cells. Isolated GH deficiency in children may be accompanied by hypoglycemia. However, hypoglycemia is more likely to occur if both GH and another hormone, ACTH (discussed next), are deficient.[6]

During protein metabolism, growth hormone stimulates protein anabolism in many tissues. Anabolism requires energy to create molecules. On the other hand, catabolism breaks down molecules and releases energy. The anabolic effects of GH are reflected by increased amino acid uptake, increased protein synthesis, and decreased oxidation of proteins. This is translated into a positive nitrogen balance and phosphate retention.

Adrenocorticotropic Hormone (ACTH) and the Adrenal Cortical Hormones

The major adrenal cortical hormones—aldosterone, cortisol, and dehydroepiandrosterone sulfate (DHEA-S)—are uniquely synthesized from the common precursor cholesterol by cells located in one of three functionally distinct zonal layers of the adrenal cortex. These zonal layers, from outer to inner, are zona glomerulosa, zona fasciculata, and zona reticularis, respectively (**Figure 13.2**). Zona glomerulosa (G-zone) cells (outer 10%) synthesize aldosterone, a **mineralocorticoid** critical for sodium retention, potassium excretion, acid–base homeostasis, and regulation of blood pressure. They have low cytoplasmic-to-nuclear ratios and small nuclei with dense chromatin with intermediate lipid inclusions.

Zona fasciculata (F-zone) cells in the middle layer synthesize **glucocorticoids**, such as cortisol, corticosterone, and 11-deoxycorticosterone. Glucocorticoids are hormones produced by the adrenal cortex that regulate glucose metabolism and have anti-inflammatory activity. Fasciculata cells are cords of clear cells, with a high cytoplasmic-to-nuclear ratio and lipids laden with "foamy" cytoplasm.

Zona reticularis cells located in the inner zone of the adrenal cortex secrete sex steroids including weak androgens and small amounts of estrogens. The primary adrenal androgens are dehydroepiandrosterone (DHEA), dehydroepiandrosterone sulfate (DHEA-S), and androstenedione, as well as small amounts of testosterone and dihydrotestosterone. Adrenal androgens are the primary source of androgens in females. Adrenal cell types are presumed to arise from stem cells. A proposed tissue layer between the zona glomerulosa and fasciculata may serve as a site for progenitor cells to regenerate zonal cells.[7]

Cortex Steroidogenesis

Control of steroid hormone biosynthesis is complex, including adrenocorticotropic hormone (ACTH) and angiotensin II (AT II). It occurs via substrate availability, enzyme activity, and inhibitory feedback loops that are layer specific.

All adrenal steroids are derived by sequential enzymatic conversion of a common substrate, cholesterol. Adrenal parenchymal cells accumulate and store circulating low-density lipoproteins (LDLs). The adrenal gland can also synthesize additional cholesterol using the enzyme acetyl-CoA, ensuring that adrenal steroidogenesis remains normal in patients with variable lipid disorders and in patients on lipid-lowering agents.

Only free cholesterol can enter steroidogenic pathways in response to ACTH. The availability of free intracellular cholesterol is metabolically regulated by ACTH (stimulatory) and LDL (inhibitory) through multiple mechanisms. Corticotropin-releasing hormone (CRH) is secreted from the hypothalamus in response to circadian signals, low serum cortisol levels, and stress. CRH stimulates release of stored ACTH from the anterior pituitary gland, which stimulates transport of free cholesterol into adrenal mitochondria, initiating steroid production.

Conversion of cholesterol to pregnenolone is a rate-limiting step in steroid biosynthesis: six carbon atoms are removed from cholesterol by enzyme cytochrome P450 (CYP450) present in the mitochondrial membrane (**Figure 13.3**). Newly synthesized pregnenolone is then returned to the cytosol for subsequent zonal conversion by microsomal enzymes in each layer by F-zone enzymes and/or androgens by enzymes in the R-zone (**Figure 13.4**).

High serum glucocorticoids suppress release of CRH and ACTH via a negative feedback mechanism. Cortisol is the primary feedback regulator

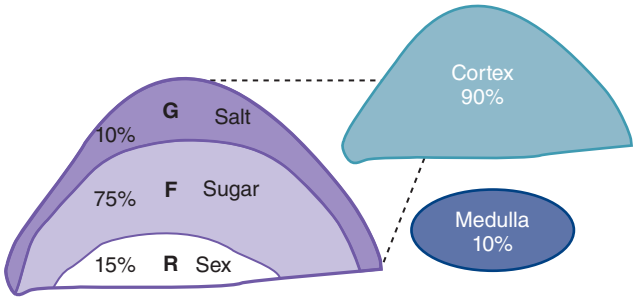


Figure 13.2 Adrenal gland by layer. Zona glomerulosa (G zone); Zona fasciculate (F zone); Zona reticularis (R zone).

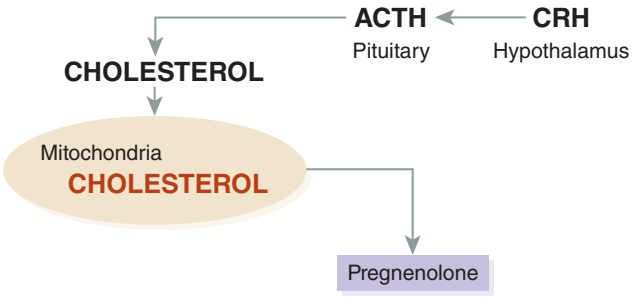


Figure 13.3 Conversion of cholesterol to pregnenolone. ACTH, adrenocorticotropic hormone; CRH, corticotropin-releasing hormone.

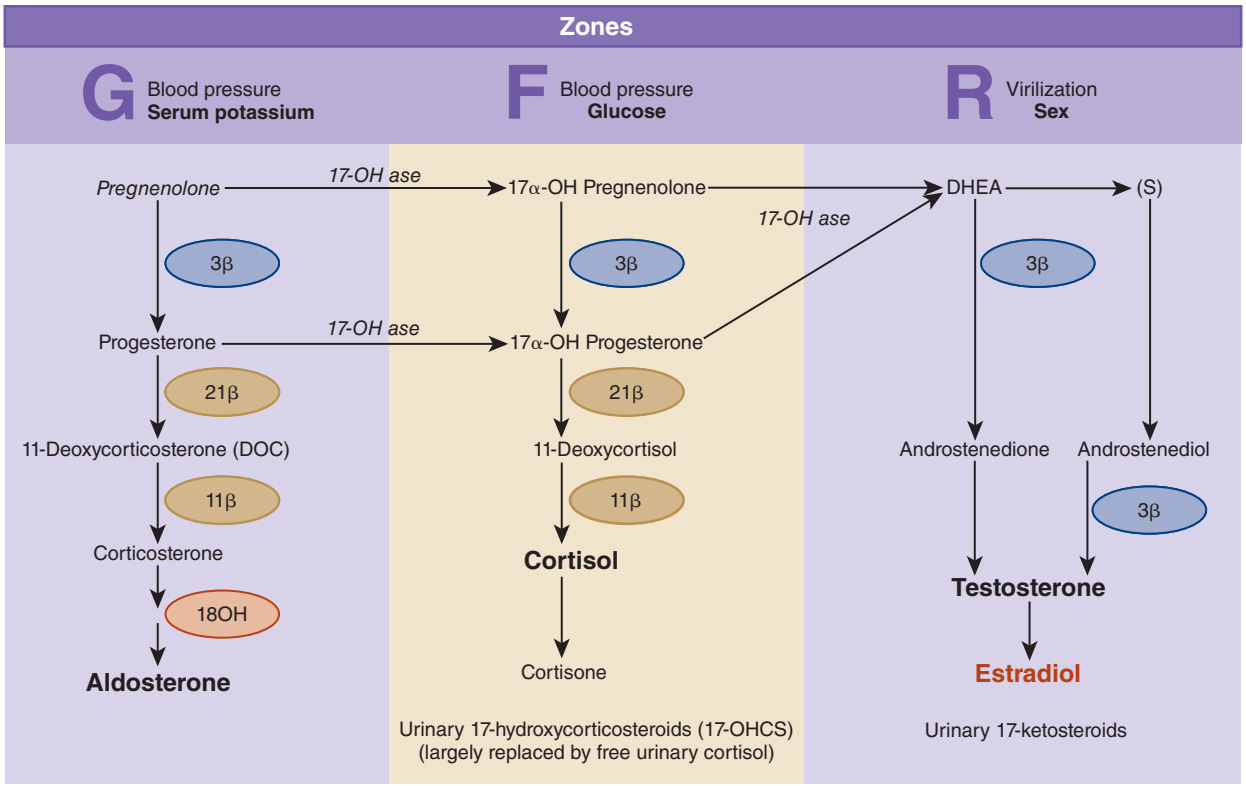


Figure 13.4 Adrenocortical hormone synthesis by zone. Zona glomerulosa (G zone); Zona fasciculate (F zone); Zona reticularis (R zone).

of ACTH-stimulated hormone production in the adrenal cortex. ACTH generally does not impact G-zone aldosterone synthesis, although cortisol has mineralocorticoid action.

Decreased activity of any enzyme required for biosynthesis can occur as an acquired or inherited (autosomal recessive) trait. Defects that decrease the production of cortisol cause increases in ACTH and CRH secretion, which stimulates the production of increased cortisol levels and leads to adrenal hyperplasia or overproduction of androgens, depending on the affected enzyme.

Evaluation of adrenal function requires measuring relevant adrenal hormones, metabolites, and regulatory secretagogues. Diagnosis is based on the correlation of clinical and laboratory findings.[8]

Follicle Stimulating Hormone and Luteinizing Hormone

Follicle stimulating hormone (FSH)[9] and luteinizing hormone (LH)[10] are secreted in the anterior pituitary gland and are both integral hormones in supporting reproduction, which have different roles in men and women. FSH, as its name implies, assists the ovarian follicle to develop in females and promotes spermatocyte development in males. LH and estrogen are connected in a positive feedback loop, which allows the estrogen levels to gradually rise until a set point is reached and the LH concentration spikes, signaling ovulation (feedback loops are discussed later in the chapter). Once the oocyte has escaped and the corpus luteum is present, the luteal phase is initiated, and the positive feedback cycle is broken. FSH release is stimulated by gonadotropin-releasing hormone (GnRH). The unique thing about this relationship is that GnRH is released in a pulsatile manner with low frequencies leading to higher FSH release.

Prolactin (PRL)

The hormone prolactin is a peptide hormone secreted from the anterior pituitary and primarily responsible for lactation and mammary gland development in nursing mothers.[11] However, prolactin also takes part in numerous other homeostatic reactions. Prolactin is most similar in structure to GH and human placental lactogen.

CASE STUDY 13.1, PART 2

Remember Sarah, the 26-year-old female who was 5 weeks postpartum. Address the following factors related to Sarah's condition.

1. What endocrine gland secretes TSH, ACTH, and prolactin?
2. What other hormones does this gland secrete?
3. In hypopituitarism, would tropic or trophic hormones be affected?

Androgens

Androgens are produced as by-products of cortisol synthesis that are regulated by ACTH. Although prolactin, proopiomelanocortin peptides, and T lymphocytes are known stimulators of androgens, regulatory mechanisms of biosynthesis in the zona reticularis remain not fully understood (**Figure 13.5**). R-cells in the zona reticularis primarily produce DHEA and multiple 19-carbon steroids (androgens and estrogens) from 17α-hydroxylated pregnenolone and progesterone. DHEA is sulfated to DHEA-S by sulfotransferase, an adrenal enzyme, and secreted daily.

Both DHEA and DHEA-S are precursors to more active androgens (e.g., androstenedione, testosterone, and 5-dihydrotestosterone) and estrogens (e.g., estradiol and estrone). Although DHEA and DHEA-S have minimal androgenic activity, adverse effects are caused by conversion to active androgens in the adrenal and peripheral tissue (e.g., hair follicles, sebaceous glands, genitalia, adipose, and prostate tissue). Although men derive less than 5% of their testosterone from adrenal or peripheral sources, women rely on the adrenals for 40% to 65% of their daily testosterone production.

Although observational data demonstrate adrenal androgen production increases in both genders in late childhood and correlates with the onset of pubic hair (adrenarche), it peaks in young adults and progressively declines with age.

Progesterone

Progesterone is a 21-carbon compound within the steroid family and is produced by the corpus luteum, a hormone-secreting body formed from the ovary's follicle after egg release. Progesterone induces the secretory activity of those endometrial glands that have been primed by estrogen, readying the endometrium for embryo implantation. Other effects include thickening of the cervical mucus, reduction of uterine contractions, and thermogenic

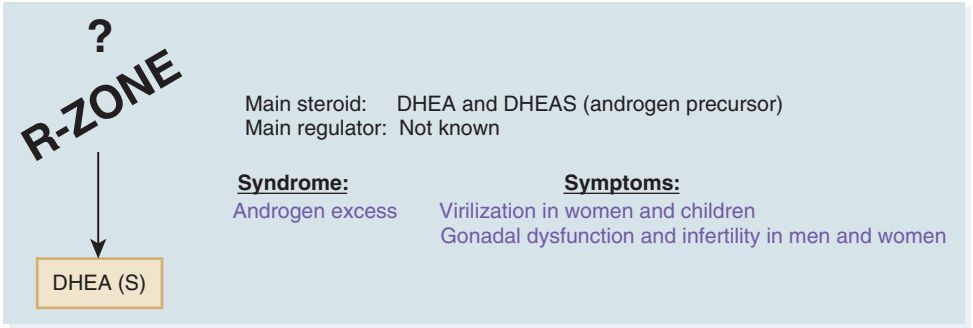


Figure 13.5 R-zone (zona reticularis) function and pathology. Manifestations of adrenal hyperandrogenism vary with age, onset, and gender. DHEA, dehydroepiandrosterone; DHEA-S, dehydroepiandrosterone sulfate.

effect, in which basal body temperature rises after ovulation. This effect is of clinical use in marking the occurrence of ovulation. Progesterone is the dominant hormone responsible for the luteal phase, and deficiency results in failure of implantation of the embryo.[12]

Estrogens/Estradiol/Estriol

Naturally synthesized estrogens are 18-carbon compounds. The principal estrogen produced in the ovary is estradiol. Estrone and estriol are primarily metabolites of intraovarian and extraglandular conversion of androstenedione and testosterone. Estrogens promote the development of secondary sex characteristics in females including breast, uterine, and vaginal development and also affect the skin, vascular smooth muscles, bone cells, and the central nervous system.[13] The lack of estrogen that naturally occurs with the onset of menopause leads to atrophic changes in these organs.[14] During the reproductive period, it is estrogen that is responsible for follicular phase changes in the uterus, with deficiency resulting in irregular and incomplete development of the endometrium.

Catecholamines

Catecholamines have a benzene ring with a catechol nucleus and two hydroxyl side groups linked to a sidechain amine, hence the name "catecholamine," and include dopamine (DA), epinephrine (EPI), and norepinephrine (NE). Dopamine is a neurotransmitter and helps with both physical movement and emotional well-being. Epinephrine targets a vast majority of cells in the body and is responsible for increasing heart rate, blood pressure, respiration rate, and metabolism. Norepinephrine increases blood pressure and pupil dilation, improves breathing, and promotes blood clotting. Epinephrine and norepinephrine are responsible for the "flight or fight" response.

Biosynthesis begins with the sequential conversion of phenylalanine substrates in a tightly regulated, compartmentalized manner. All reactions take place in the cytoplasm, except for the production of NE, which occurs within lipid vesicles or outer mitochondrial membranes, as illustrated in **Figure 13.6**.

In the cytosol, NE is converted into EPI by a cortisol-dependent enzyme called phenylethanolamine *N*-methyltransferase (PNMT). Any form of stress that increases cortisol levels stimulates EPI production.

In the adrenal medulla, EPI and NE are stored in neurosecretory vesicles but are in dynamic equilibrium with the surrounding cytoplasm. They diffuse into the cytoplasm and actively re-enter the vesicles via vesicular monoamine transporters (VMATs). They enter the circulation via exocytosis of the vesicles. The cytoplasmic catecholamines may also be metabolized by enzyme catecholamine-*O*-methyltransferase (COMT) into metanephrines and normetanephrines, which enter circulation via diffusion.

EPI is the main catecholamine secreted by the adrenal medulla, and NE is the principal

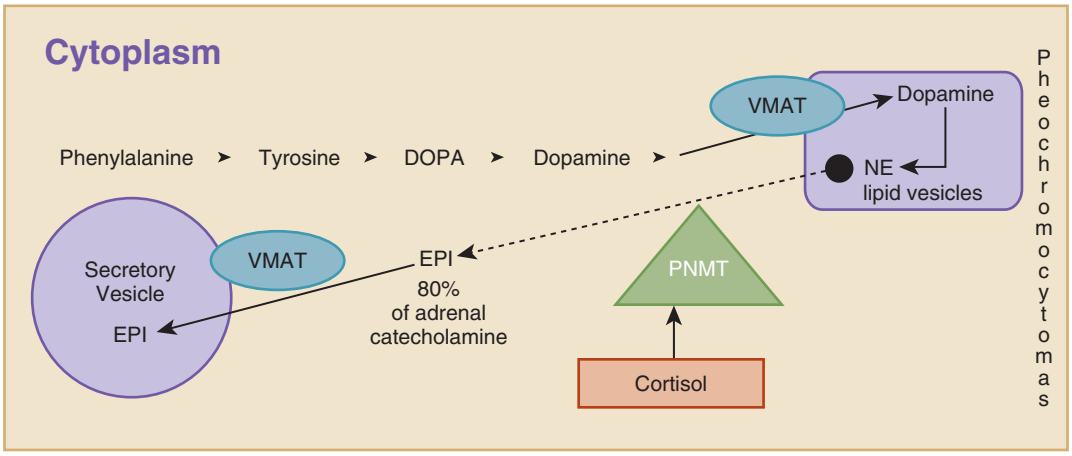


Figure 13.6 Biosynthesis and storage of catecholamines. VMAT, vesicular monoamine transporter; NE, norepinephrine; EPI, epinephrine; PNMT, phenylethanolamine *N*-methyltransferase.

catecholamine synthesized in the central nervous system. The ratio of NE to EPI in the serum is approximately 9:1. In adrenal insufficiency (low cortisol), that ratio increases to 45:1 in females and 24:1 in males.

In sympathetic neurons, cytoplasmic dopamine is sequestered into vesicles, converted into NE, and stored until nerve stimulation causes its release into a synapse by exocytosis.

Metabolism and Excretion of Catecholamines

All catecholamines are rapidly eliminated from target cells and the circulation by three mechanisms:

1. Reuptake into secretory vesicles
2. Uptake in non-neuronal cells (mostly liver)
3. Degradation

Degradation relies on two enzymes—COMT (in non-neuronal tissues) and monoamine oxidase (MAO) (within neurons)—to produce metabolites (metanephrines and vanillylmandelic acid [VMA]) from free catecholamines. Metabolites and free catecholamines are eliminated by direct filtration into the urine and excreted as free NE (5%), conjugated NE (8%), metanephrines (20%), and VMA (30%) (**Figure 13.7**).

Thyroid Hormone (T_3, T_4) and Thyroid Stimulating Hormone (TSH)

Activity of thyroid hormone depends on the location and number of iodine atoms. Approximately 80% of T_4 is metabolized into either T_3 (35%) or reverse T_3 (rT_3; 45%). Outer-ring deiodination of T_4(5′-deiodination) leads to production of 3,5,3′-triiodothyronine (T_3). T_3 is three to eight times more metabolically active than T_4 and is often considered to be the active form of thyroid hormone, with T_4 considered the "pre"-hormone (with thyroglobulin being the "prohormone"). In addition to its "pre"-hormone activity, however, inner-ring deiodination of T_4 results in the production of metabolically inactive rT_3(**Figure 13.8**).

There are three forms of iodothyronine 5′-deiodinase. Type 1 iodothyronine 5′-deiodinase, the most abundant form, found mostly in the liver and kidney, is the largest contributor to the circulating T_3 pool. Certain drugs (e.g., propylthiouracil, glucocorticoids, and propranolol) slow the activity of this deiodinase and are used in the treatment of severe thyroid hormone excess, or hyperthyroidism. Type 2 iodothyronine 5′-deiodinase, found in the brain and pituitary gland, functions to maintain constant levels of T_3 in the central nervous system. Its activity

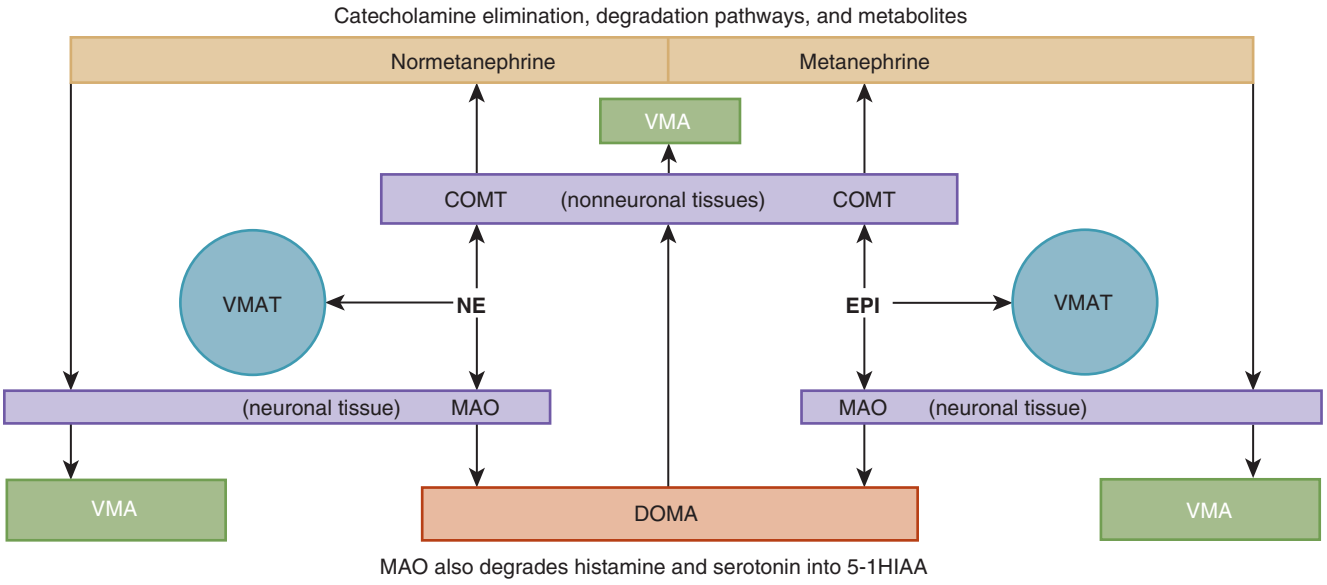


Figure 13.7 Catecholamine degradation. Free catecholamines (EPI and NE) are either sequestered into VMAT-containing vesicles or converted into metabolites, DOMA by neuronal MAOs, and metanephrines by nonneuronal COMTs. These metabolites are ultimately degraded to VMA (DOMA by COMTs and metanephrines by MAOs) and excreted. EPI, epinephrine; NE, norepinephrine; VMAT, vesicular monoamine transporter; DOMA, 3,4-dihydroxymandelic acid; MAO, monoamine oxidase; COMT, catechol methyltransferases; VMA, vanillylmandelic acid.

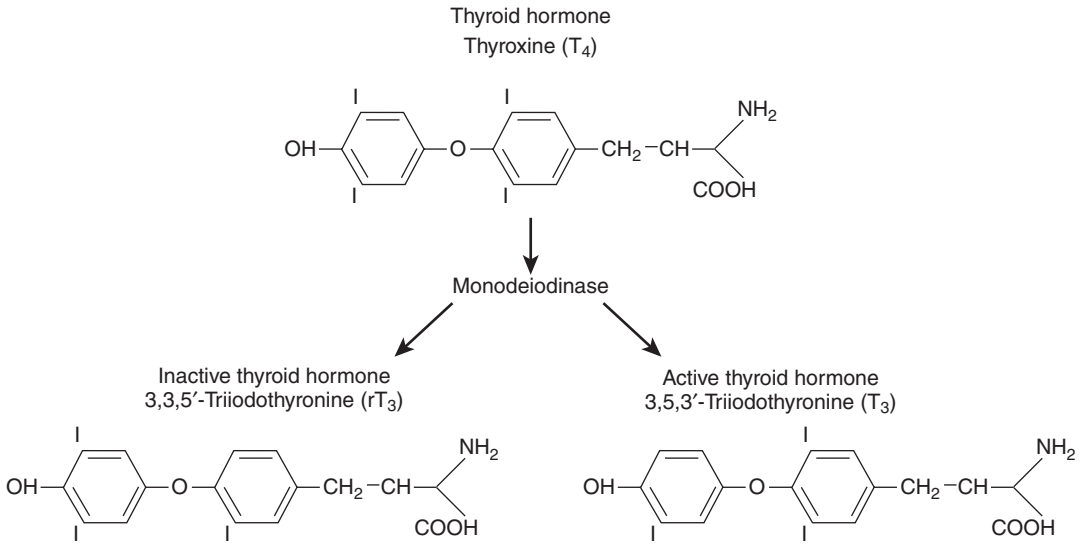


Figure 13.8 Metabolism of thyroxine.

is decreased when levels of circulating T_4 are high and increased when levels are low. Activity of the deiodination enzymes gives another level of control of thyroid hormone activity beyond the thyrotropin-releasing hormone (TRH) and thyrotropin (TSH) control of the **hypothalamic-pituitary-thyroid axis** (**Figure 13.9**).[15]

Protein Binding of Thyroid Hormones

When released into the circulation, only 0.04% of T_4 and 0.4% of T_3 are unbound by proteins and available for hormonal activity. The three major binding proteins, in order of significance, are **thyroxine-binding globulin (TBG)**, thyroxine-binding prealbumin, and albumin. Alterations to the concentrations of binding proteins significantly affect circulating quantities of T_4 and T_3. For example, high estrogen levels during pregnancy lead to increased TBG production by the liver, which results in higher levels of bound thyroid hormones, leading to high levels of total T_3 and total T_4. Typically, however, levels of the unbound active, or free, thyroid hormones, remain in the normal range and the individual remains euthyroid. In some instances, however, measurement of free T_4 and free T_3 may be necessary to eliminate any confusion caused by abnormal binding protein levels.

Control of Thyroid Function

Understanding the hypothalamic-pituitary-thyroid axis is essential for correctly interpreting thyroid function testing. This axis is central to the regulation of thyroid hormone production. TRH is synthesized by neurons in the supraoptic and supraventricular nuclei of the hypothalamus and stored in the median eminence of the hypothalamus. When secreted, this hormone stimulates cells in the anterior pituitary gland to manufacture and release TSH. TSH, in turn, circulates to the thyroid gland and leads to increased production and release of thyroid hormones (TH). When the hypothalamus and pituitary sense that there is an inadequate amount of thyroid hormone in circulation, TRH and TSH secretion increases and stimulates increased thyroid hormone production. If TH levels are high, TRH and TSH release will be inhibited, leading to lower levels of thyroid hormone production. This feedback loop requires a normally functioning hypothalamus, pituitary, and thyroid gland, as well as an absence of any interfering agents or agents that mimic TSH action (Figure 13.9).

Actions of Thyroid Hormone

Once released from the thyroid gland, TH circulates in the bloodstream where free T_4 and T_3 are available to travel across the cell membrane. In the cytoplasm, T_4 is deiodinated into T_3, the active form of thyroid hormone. T_3 combines with its nuclear receptor on thyroid hormone-responsive genes, leading to production of messenger RNA that, in turn, leads to production of proteins that influence metabolism and development. Effects of TH include tissue growth, brain maturation, increased heat production, increased oxygen consumption, and increased expression of β-adrenergic receptors. Clinically,

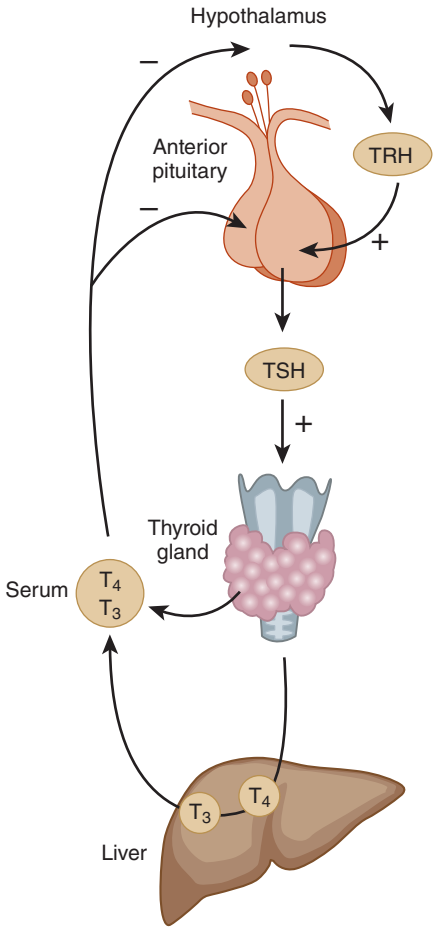


Figure 13.9 The hypothalamic-pituitary-thyroid axis. Thyrotropin releasing hormone (TRH) stimulates the production and release of thyrotropin (TSH). TSH stimulates the thyroid gland to synthesize and secrete thyroid hormone. T_4 that is released by the thyroid gland is mostly converted to T_3 by the liver and kidney. T_3 and T_4 feedback inhibit TSH release directly through action at the pituitary and indirectly by decreasing TRH release from the hypothalamus.

individuals who have excess thyroid hormone (hyperthyroidism) will have symptoms of increased metabolic activity such as tachycardia and tremor, while individuals with hypothyroidism (decreased levels of TH) note symptoms of lowered metabolic activity like edema and constipation.

Posterior Pituitary Hormones

The posterior pituitary is an extension of the forebrain and represents the storage region for arginine vasopressin (AVP, previously called antidiuretic hormone [ADH]) and oxytocin. Both small peptide hormones are synthesized in the supraoptic and paraventricular nuclei of the hypothalamus and transported to the neurohypophysis via their axons in the hypothalamoneurohypophyseal tract. This tract transits the median eminence of the hypothalamus and continues into the posterior pituitary through the pituitary stalk. The synthesis of each of these hormones is tightly linked to the production of neurophysin,[16] a larger protein whose function is poorly understood. Both hormones are synthesized outside of the hypothalamus in various tissues, and it is plausible they have an autocrine or a paracrine function.

Oxytocin

Oxytocin is a cyclic nonapeptide, with a disulfide bridge connecting amino acid residues 1 and 6. As a posttranslational modification, the C-terminus is amidated. Oxytocin has a critical role in lactation[17] and plays a major role in labor and parturition.[18] Oxytocin is also unique because its secretion responds to a positive feedback loop, meaning that circulating levels of oxytocin perpetuate further hormone secretion, instead of suppressing further hormone secretion as is the case with most anterior pituitary hormones. In this way, uterine contractions propagate oxytocin release, which causes further uterine contractions, which causes further oxytocin release until parturition occurs. Synthetic oxytocin, *pitocin*, is used in obstetrics to induce labor. Recent studies have linked oxytocin to a variety of biosocial behaviors to include maternal nurturing and mother–infant bonding.[19] In addition to its reproductive and prosocial effects, oxytocin has been shown to have effects on pituitary, renal, cardiac, metabolic, and immune function.

Arginine Vasopressin

Similar to oxytocin structurally, arginine vasopressin (AVP) is a cyclic nonapeptide with an identical disulfide bridge. It differs from oxytocin by only two amino acids. AVP's major action is to regulate renal free water excretion and, therefore, has a central role in water balance. The vasopressin receptors in the kidney (V_2) are concentrated in the renal collecting tubules and the ascending limb of the loop of Henle. They are coupled to adenylate cyclase, and once activated, they induce insertion of aquaporin-2, a water channel protein, into the tubular luminal membrane.[20] Vasopressin is also a potent pressor agent and effects blood clotting[21] by promoting factor VII release from hepatocytes and von Willebrand factor release from the endothelium. These vasopressin receptors (V_{1a} and V_{1b}) are coupled to phospholipase C.

Hypothalamic osmoreceptors and vascular baroreceptors regulate the release of vasopressin from the

posterior pituitary. The osmoreceptors are extremely sensitive to even small changes in plasma osmolality, with an average osmotic threshold for vasopressin release in humans of 284 mOsm/kg. As plasma osmolality increases, vasopressin secretion increases. The consequence is a reduction in renal free water clearance, a lowering of plasma osmolality, and a return to homeostasis. The vascular baroreceptors (located in the left atrium of the heart, aortic arch, and carotid arteries) initiate vasopressin release in response to a fall in blood volume or blood pressure. A 5% to 10% fall in arterial blood pressure in normal humans will trigger vasopressin release; however, in contrast to an osmotic stimulus, the vasopressin response to a baroreceptor-induced stimulus is exponential. In fact, baroreceptor-induced vasopressin secretion will override the normal osmotic suppression of vasopressin secretion.

Diabetes insipidus (DI), characterized by increased production of urine (polyuria) and increased thirst (polydipsia), is a consequence of vasopressin deficiency. However, total vasopressin deficiency is unusual, and the typical patient presents with a partial deficiency. The causes of hypothalamic DI include apparent autoimmunity to vasopressin-secreting neurons, trauma, diseases affecting pituitary stalk function, and various pituitary tumors. A sizable percentage of patients (up to 30%) will have idiopathic DI.[22]

Depending on the degree of vasopressin deficiency, diagnosis of DI can be readily apparent or may require extensive investigation. Documenting an inappropriately low vasopressin level with an elevated plasma osmolality would yield a reasonably secure diagnosis of DI. In less obvious cases, the patient may require a water deprivation test. During this procedure, fluids are withheld from the patient and serial determinations of serum and urine osmolality are performed to document the patient's ability to conserve water. Under selected circumstances, a health care provider may simply offer a therapeutic trial of vasopressin or a synthetic analog such as desmopressin (dDAVP) and assess the patient's response. In this circumstance, amelioration of both polyuria and polydipsia would be considered a positive response, and a presumptive diagnosis of DI is made. However, if the patient has primary polydipsia (also known as compulsive water drinking), a profound hypo-osmolar state (water intoxication) can ensue due to the continued ingestion of copious amounts of fluids and a reduced renal excretion of free water. This scenario illustrates the importance of carefully evaluating each patient prior to therapy.

Vasopressin excess may also occur and is much more difficult to treat. Since excess vasopressin leads to the pathologic retention of free water, restricting free water intake to small amounts each day has been a historical cornerstone of treatment. Recently, conivaptan and tolvaptan,[23] vasopressin V_2 receptor antagonists, have been approved for the management of euvolemic hyponatremia due to vasopressin excess.

Calcitonin and Procalcitonin

Procalcitonin is a precursor of the hormone calcitonin and has been shown to be a useful marker of infection and sepsis.[24] It is a biomarker that aids in the risk assessment of critically ill patients on admission and continually for progression to severe sepsis and/or septic shock. The percent change in procalcitonin level over time aids in the prediction of mortality in these critically ill patients.

Calcitonin, on the other hand, is not used for critically ill patients. This hormone decreases the plasma calcium load by increasing the storage of calcium in the bones. Calcitonin works in the opposite manner as parathyroid hormone (discussed below).

Parathyroid Hormone (PTH)

Because our bodies need calcium for so many basic functions, including muscle contraction, it is tightly regulated. If the plasma calcium concentration becomes too low, then the parathyroid gland will release the parathyroid hormone (PTH). PTH increases the plasma calcium concentration by three primary mechanisms:

1. Increases bone resorption of calcium
2. Increases intestinal absorption of calcium
3. Increases removal of phosphate in the proximal and distal convoluted tubules in the kidneys

Insulin

Insulin allows glucose to enter our peripheral cells, and it is the only hormone that decreases blood glucose concentrations. It is produced by the β-cells of the islets of Langerhans in the pancreas by translating a long protein called pre-proinsulin. Pre-proinsulin is modified to create proinsulin, and then an accessory chain called c-peptide is cleaved, resulting in insulin formation. Insulin deficiency is the hallmark of type 1 diabetes mellitus, while a reduction in the cellular response to insulin results in type 2 diabetes mellitus.

Glucagon

Opposing the action of insulin is glucagon, which increases the plasma glucose concentration. This is a useful survival advantage because the brain operates almost entirely on glycolysis and needs a constant

CASE STUDY 13.2, PART 2

Remember Paulina, a 47-year-old woman with polyuria, polydipsia, and polyphagia. Address the following questions related to her symptoms and diagnosis.

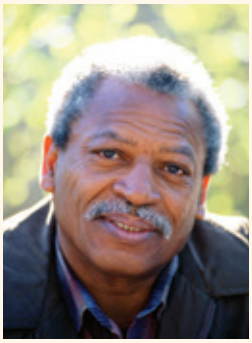
1. Why did the physician order a HbA_1c test for Paulina?
2. Other than diabetes mellitus, what other disorder should be included in the differential diagnosis?
3. What laboratory tests are performed to make a definitive diagnosis?

© Krakenimages.com/Shutterstock.

CASE STUDY 13.3, PART 2

Remember Mike, a 59-year-old male, who has a small-cell lung carcinoma. Address the following related to his symptoms and diagnosis.

1. What are most likely the two hormones that are responsible for Mike's symptoms?
2. What laboratory tests may be helpful in the diagnosis of syndrome of inappropriate ADH secretion (SIADH)?
3. What suppression test would be ordered to confirm Cushing syndrome?



© Thinkstock.

supply of glucose to function properly. Glucagon helps the body to properly regulate the blood glucose concentration, increasing the glucose concentration by promoting glycogenolysis and inhibiting glycolysis.

Gastrin and Secretin

Gastrin is released from the G cells located in the pyloric antrum in the stomach, while secretin is formed in the duodenal mucosa. These hormones are grouped together as they both help to regulate the acid–base balance in the lumen of the intestine. Gastrin secretion is associated with the release of hydrochloric acid as well as proliferation of the gastric epithelium. The acid helps to digest the food, and the gastric epithelium helps to protect the stomach. Stomach acid content when making contact with the duodenal mucosa will stimulate secretin and subsequent bicarbonate release into the intestinal lumen, thereby neutralizing the acid.

Human Chorionic Gonadotropin (hCG)

Human chorionic gonadotropin is a hormone produced during pregnancy by the syncytiotrophoblast cells[25] of the placenta that may perpetuate umbilical cord growth and development as well as provide multiple immunomodulatory effects.[26] There is also new data to suggest that hCG may also play a role in implantation of the embryo into the endometrium. The medical profession uses hCG as a marker for pregnancy and several types of malignancies (e.g., prostate, colorectal, lung, gynecological, and breast cancers).

Serotonin/5-hydroxyindolacetic Acid (5-HIAA)

Serotonin is a neurotransmitter and not a hormone. Although it is not a hormone, it has a similar function and sends chemical messages throughout the body. It is integral to regulating many of our behaviors, like memory and mood, and is also present in the cardiovascular system, bone resorption, and coagulation.[27] Serotonin is synthesized from the essential amino acid tryptophan. Serotonin is metabolized to 5-hydroxyindoleacetic acid (5-HIAA) which is excreted by the kidneys and measured in the urine. It is useful in the differential diagnosis of carcinoid syndrome.

Hormone Metabolism

Metabolism is defined as the sum of chemical processes that occur within a living organism to maintain life and includes both catabolism and anabolism. This can take many forms and is vital to understanding the hormonal homeostasis in the human body. Hormones and all other cellular signals are metabolized, and the speed at which they are created (anabolism) or broken down (catabolism) determines the extent to which they are capable of binding to receptors and eliciting their intended effects. There are many conditions that modify the metabolism of hormones. Hormone levels in the blood will be modified if either the speed of production or breakdown are changed. For instance, the concentrations of protein-bound hormones will be greatly affected by the amount of proteins that are available for binding. Enzymatic breakdown of hormones can be accentuated by disease states such as cirrhosis of the liver.

Alcohol Consumption

In the short term, alcohol appears to increase the degradation of testosterone.[47] If alcohol is consumed for an extended duration or in large quantities that may cause liver damage, then cirrhosis may ensue. In cirrhosis of the liver, less albumin and other binding proteins will be produced. When fewer proteins are available to bind the lipophilic hormones, then there are reduced hormone concentrations in the blood and, consequently, a reduction in the effects of these hormones on tissues.

Adrenal Steroid Hormone Synthesis

Many steroid hormones are produced in the adrenal gland by a cascade of enzymatic reactions demonstrated in Figure 13.1. A deficiency in any of these enzymes will cause classical and stereotypical effects on hormone production. While a deficiency in 11-β-hydroxylase will cause a decrease in aldosterone and glucocorticoids, it will also shunt all cholesterol into the sex steroid pathway as those enzymes are functional and in adequate concentration. This usually results in masculinization, amenorrhea, and hypoglycemia. These patients also typically have asymptomatic hypertension because the 21-hydroxylase will produce deoxycorticosterone, which is not the intended end product of aldosterone but can still acts as a mineralocorticoid, which can stimulate aldosterone receptors.

Mechanisms of Elimination

While many mechanisms do exist, there are two predominant organs for eliminating hormones: the kidneys and liver. Consider the blood to be a conveyor belt throughout the body. A hormone is generally removed at the end of the conveyor belt, and the end of this conveyor belt is the kidneys or the liver. The bloodstream passes by an inspector (the liver), which places tags on certain hormones for destruction or creation of different molecules. Steroid hormones are eliminated by inactivating metabolic pathways and excretion in urine or bile. Thyroid hormones are inactivated by intracellular deiodinases. Catecholamines are rapidly degraded within the blood circulation. The fatty acid derivatives are rapidly inactivated by metabolism and are typically only active for a short period of time (seconds).

Hormone Transport

After hormones are released from their endocrine gland, they circulate in the bloodstream in one of two forms: free (unbound) or protein bound. Free-floating hormones are susceptible to degradation and metabolism in the blood and can be exposed to many different enzymes or products, which can cause a permanent change in structure of the hormone. This change in structure impacts the functionality of the hormone. There are a few different ways that hormones can travel to distant sites in the body.

Solubility

Hormones can be characterized as either water-soluble (hydrophilic) or fat-soluble (lipophilic). The simpler to understand mechanism of transport in the blood is how many water-soluble molecules tend to travel. If the hormone is water soluble, then it will easily flow to distant sites in the bloodstream. Hormones that are free-floating in the plasma are vulnerable and open to be metabolized by an enzyme or other substance that would confer a permanent change in its structure. For instance, examples include the peptide and protein hormones, which are water soluble.

Carrier Proteins

Unlike water-soluble hormones, fat-soluble hormones are transported bound to a carrier protein. Consider these hormones needing a guardian. This guardian is a protein that will chaperone the hormone during its travel throughout the bloodstream. The carrier protein will have different effects on the hormones. First, it will protect the hormone from being acted upon by an enzyme or undergoing a chemical reaction. Second, it will increase the water solubility of lipophilic hormones. This allows fatty hormones to travel in mostly aqueous environments, like the blood. Examples of carrier proteins include albumin, corticosteroid-binding globulin, sex hormone-binding globulin (SHBG) and thyroid-binding globulin (TBG). A significant change in the concentration of these carrier proteins affects the concentration of free hormones in circulation.

Micelles

Amphipathic and lipid particles will coalesce and create small lipid droplets in the blood. Consider a vinaigrette salad dressing that is left to settle: the

aqueous substances and lipid substances will separate from one another. This is a natural phenomenon that occurs in our bodies as well. Lipids do not want to travel with water; if we try to force them to travel together, the lipids will self-assemble in aggregates called *micelles*. Micelles are enclosed packages containing lipid-rich substances with a hydrophilic outer layer and hydrophobic inner layer. Hormones and other lipid-soluble substances prefer to travel in these environments when in the blood or other aqueous environments.

Feedback Mechanisms

Common processes in the body that regulate hormone secretion are called feedback mechanisms, also known as feedback loops (**Figure 13.10**). These are ways to turn secretion on and off. This occurs because certain stimuli inform our bodies that an action needs to be performed. For example, if it is cold in your house in the winter, then you may go check to see if the front door has been accidentally left open. If it is, you will close it. In this case, low

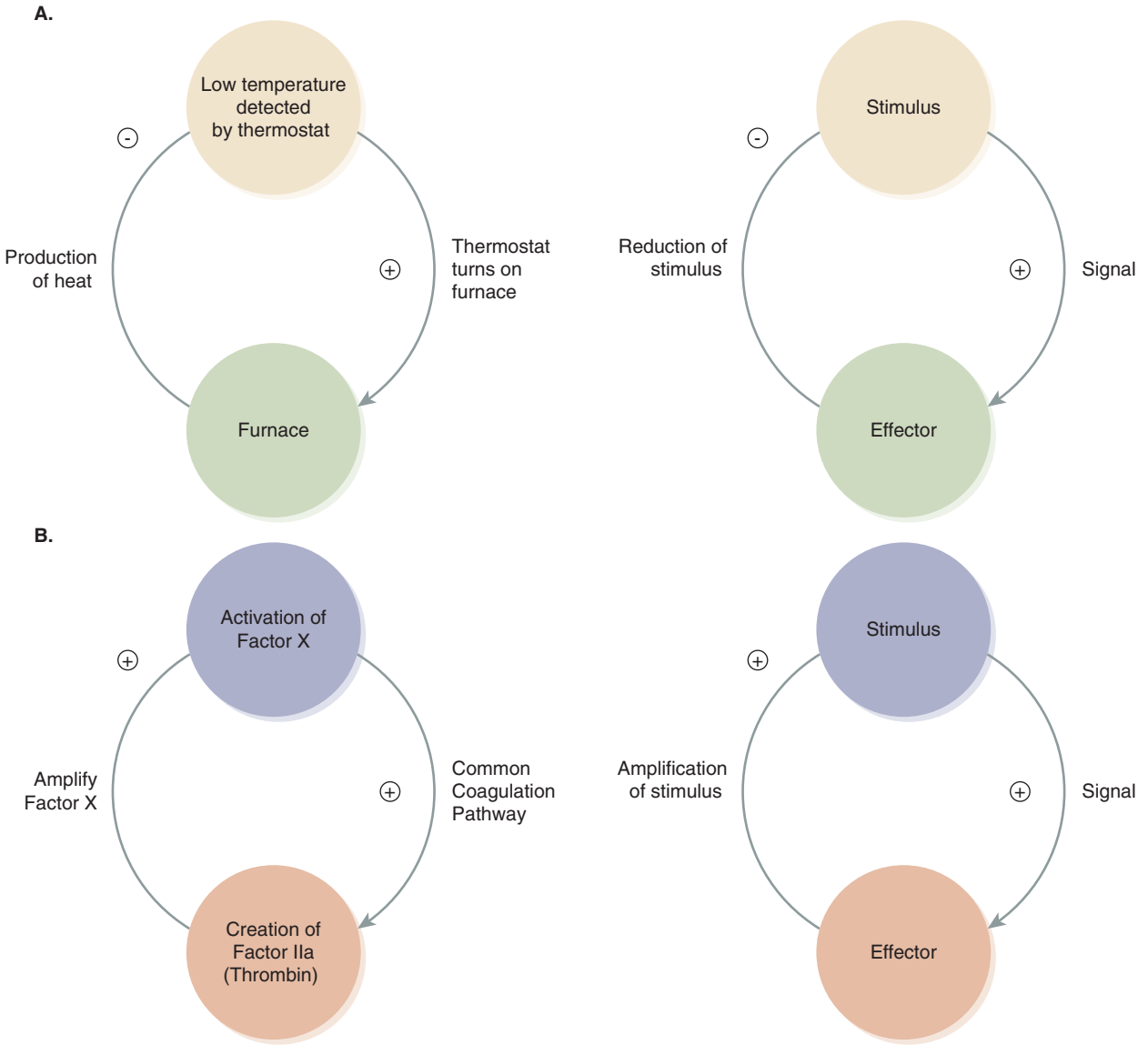


Figure 13.10 The primary endocrine regulatory systems are demonstrated here with examples. **(A)** Demonstrates the negative feedback pattern in the case of a thermostat and furnace working together to keep a comfortable temperature in your home. The temperature drops, which signals the thermostat to turn the furnace on. This raises the temperature and decreases the initial stimulus, the low temperature. **(B)** Demonstrates the positive feedback pattern in the case of the extrinsic coagulation cascade. The initial signal is tissue factor from damaged tissue, which then amplifies the original signal until a definite endpoint is reached. These systems require good control mechanisms or else they can get out of control, as demonstrated with disseminated intravascular coagulopathy.

ambient temperature is the stimulus, closing the door is the action, and maintaining heat in your home environment is the effect. Feedback mechanisms provide monitoring and control of the cellular environment.

Negative Feedback Mechanisms

The most common feedback mechanism in the body is the negative feedback loop. The most important aspect of these mechanisms is to understand that a stimulus will feedback upstream to decrease the production of itself. The thyroid is the classic example of a negative feedback loop (**Figure 13.11**).

Think of a negative feedback loop as a laboratory involving a manager, a supervisor, and a laboratorian. The manager monitors the blood, and when it senses that there's not enough thyroid hormone, it will tell the supervisor that more thyroid hormone needs to be produced. The supervisor tells the laboratorians that they need to make more thyroid hormone. Then, after more thyroid hormone is released and the supervisor notices that there is enough thyroid hormone in the system, the increased production will no longer be needed, which is also relayed to the manager. In the body, the hypothalamus (the manager) secretes thyrotropin releasing hormone (TRH), which stimulates the anterior pituitary (the supervisor) to secrete thyroid stimulating hormone (TSH), which stimulates the thyroid (the laboratorian) to secrete triiodothyronine (T_3) and thyroxine (T_4). Increased T_3 and T_4 will inhibit (negative feedback) the pituitary from secreting TSH and the hypothalamus from releasing TRH, decreasing both of their secretions, ultimately decreasing the levels of T_3 and T_4 secretion; this is the epitome of a negative feedback loop.

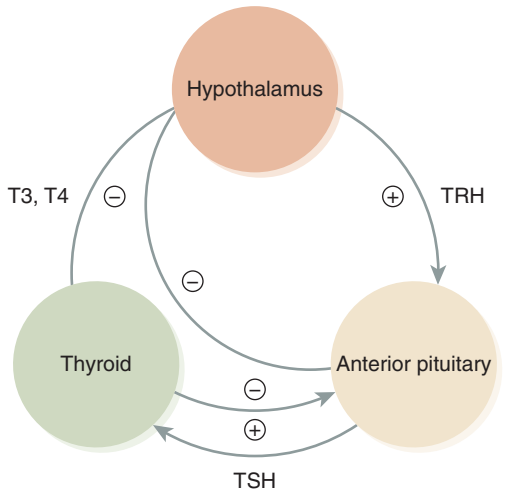


Figure 13.11 Stereotypical negative feedback system for the human body. The example of thyroid hormone is shown. The effector hormones are T_3 or T_4, which regulate their own production through negative feedback at the hypothalamus. The hypothalamus initiates the secretion of a tropic hormone, in this case thyrotropin releasing hormone (TRH), which stimulates the production of another hormone in the anterior pituitary gland, thyroid stimulating hormone. Thyroid stimulating hormone (TSH) stimulates the release of T_3 and T_4 in the thyroid gland, which will negatively feedback at the level of both the hypothalamus and anterior pituitary to decrease the stimulatory signals being released, TRH and TSH.

Positive Feedback Mechanisms

Positive feedback mechanisms are control mechanisms that increase the stimulus received until a distinct endpoint is achieved. These mechanisms can easily go out of control, as the mechanism exists to increase the stimulus, so there must be another counter-regulatory mechanism at the ready in case it gets out of control. We will discuss two examples of a positive feedback loop.

The first example that most people think of with positive feedback is that of giving birth, or parturition. The feedback cycle begins when the stretch receptors on the cervix begin to feel pressure, which likely means that the baby has developed enough to be born. When this happens, the nerves stimulate the hypothalamus and subsequently the posterior pituitary gland to release the hormone oxytocin. Oxytocin circulates around the body until it reaches the myometrium of the uterus and causes the rate and strength of contractions to increase. This increase in contraction pressure creates greater pressure on the cervix, causing more stimulation of the stretch receptors and more neurologic signals to be sent to the hypothalamus. This, in turn, results in more oxytocin being released, further increasing the strength and rate of contractions. This cycle continues until delivery. When the baby is born, its head is no longer exerting pressure on the cervix, ending the positive feedback loop. Another relevant example of positive feedback in the medical laboratory is that of blood coagulation.[48]

Primary, Secondary, and Tertiary Disorders

Disorders of the endocrine glands are characterized by the organ of disfunction—in other words, whether the disorder affects a regulating endocrine gland (secondary or tertiary) or the final endocrine gland

Table 13.3 Laboratory Analysis of Thyroid Disorders

	T_4	TSH	TRH
Primary Hypothyroidism	↓	↑	↑
Secondary Hypothyroidism	↓	↓	↑
Tertiary Hypothyroidism	↓	↓	↓
Primary Hyperthyroidism	↑	↓	↓
Secondary Hyperthyroidism	↑	↑	↓
Tertiary Hyperthyroidism	↑	↑	↑

that is regulated (primary). We will use the example of the thyroid hormone above and discuss the hypothalamic-pituitary-thyroid (HPT) axis to explain endocrine disorders. This feedback loop requires normal functioning endocrine glands, as well as an absence of any interfering agents or agents that mimic TSH action (refer back to Figure 13.9).

TRH is synthesized in and secreted from the hypothalamus when there is an inadequate concentration of thyroid hormones in circulation. When secreted, this hormone stimulates cells in the anterior pituitary gland to manufacture and release TSH. TSH, in turn, circulates to the thyroid gland and leads to increased production and release of thyroid hormones (TH). If there is an issue with the thyroid gland itself, this is called *primary* disease. This means the thyroid gland is not responsive to stimuli from the pituitary or hypothalamus. A *secondary* disorder of thyroid function is one in which there is an issue with the anterior pituitary, which is not responsive from either hypothalamus or thyroid glands. A *tertiary* disorder is when the hypothalamus is affected and therefore is unable to control the anterior pituitary or thyroid glands. As shown in **Table 13.3**, the control of the hypothalamus-pituitary-thyroid (HPT) axis is off, allowing for increases and decreases in circulating hormones depending on the endocrine organ affected. Primary, secondary, and tertiary endocrine disorders are characterized for other endocrine glands in this same manner.

Other Factors That Affect Hormone Levels

There are many factors that can affect the concentration or activity of a hormone. There are many considerations that need to be contemplated when interpreting endocrinology test results.

Emotional Stress

Stress can lead to a well-adapted or a maladaptive endocrinologic response. The purpose of our endocrine system is to help our bodies to survive the different stressors that are placed upon it. Just like physical stresses cause our body to change, psychological stresses also take their toll. Individuals who are perpetually stressed, or have long-term stressors in their lives, do have detectable clinical and laboratory changes in their hormone levels.[49] Some of the more common hormones that are affected by psychological or emotional stress are glucocorticoids, catecholamines, growth hormone, and prolactin. Some of the hormones released during stressful situations are useful in the fight-or-flight response, which evolutionarily has enabled the possessor of those genes to produce a robust fight-or-flight response.

Stress can also perpetuate and potentiate other endocrine disorders and lead to an imbalance of adrenal hormones. One of the hormones that is typically measured is cortisol, which correlates closely with the level of emotional stress that an individual is experiencing. Cortisol and its upstream tropic hormones (CRH and ACTH) may increase between two- and five-fold when a patient is experiencing stress. It has been proposed that the chronic stress response can become part of a pathologic positive feedback loop, whereas individuals who seem resilient and able to cope with stress have a homeostatic stress response.[50]

Time of Day

Another useful feature of the hypothalamic–pituitary unit is the cyclic nature of hormone secretion. The nervous system usually regulates this function through external signals, such as light–dark changes or the ratio of daylight to darkness. The term *zeitgeber* ("time giver") refers to the process of entraining or synchronizing these external cues into the function of internal biologic clocks. As a result, many pituitary hormones are secreted in different amounts, depending on the time of day. These circadian rhythms, or the **diurnal variations**, are typified by ACTH and TSH secretion. Diurnal variation is a fluctuation in the daily blood or urine concentration of an analyte. An analogy would be the ocean with highest water levels seen at high tide and the lowest water levels seen at low tide. For example, cortisol is assessed as a first morning specimen at 8:00 am; early morning is when cortisol spikes, and the concentration tapers off throughout the day. With ACTH, the nadir (lowest) of secretion is between 11:00 pm and 3:00 am, and the peak occurs on awakening or

around 6:00 to 9:00 am.[51] The circadian rhythm of ACTH is a result of variations in pulse amplitude and not alterations in pulse frequency.[52] The nocturnal levels of TSH are approximately twice the daytime levels; the nocturnal rise in TSH is a result of increased pulse amplitude.[53]

Menstrual Cycle

By convention, the menstrual cycle is considered to start on the first day of menses (day 1). The menstrual cycle consists of two phases of parallel events occurring at the ovaries and endometrium. Within the ovaries, these events are known as the follicular and luteal phases, while the concurrent endometrial events are known as the proliferative and secretory phases.[54]

The Follicular Phase

The follicular phase begins with the onset of menses and ends on the day of LH surge. Early in the follicular phase, the ovary secretes very little estrogen or progesterone. A rise in FSH, however, stimulates estrogen production. The estrogen secreted by the developing follicle within the ovary stimulates uterine epithelial cells proliferation, blood vessel growth, and endometrial gland development to increase the thickness of the endometrium.

The Luteal Phase

Estrogen levels peak the day before ovulation, at which point a positive feedback system results in an LH surge. The start of the luteal phase is marked by the extrusion of the ovum approximately 36 hours after the LH surge, with subsequent luteinization of the graafian follicle to form the corpus luteum. The corpus luteum secretes progesterone to aid in the implantation of the embryo. The intense secretory capacity of the uterine glands aids the implantation of the embryo. In the absence of fertilization, with a gradual decline in the production of progesterone and estrogen by the corpus luteum, there is a loss of endometrial blood supply, which results in shedding of the endometrium approximately 14 days after ovulation has occurred. The typical duration of menstrual bleeding is 3 to 5 days, with blood loss averaging 50 mL. Onset of menses marks the end of the luteal phase.

Ovulation

The central control of FSH and LH secretion resides in the gonadotropin-releasing hormone (GnRH) pulse generator of the hypothalamus. Positive and negative feedback responses exist among estrogen, progesterone, LH, and FSH production. It is because of the lack of estrogen after menopause that both FSH and LH levels rise.[55] During reproductive years, FSH levels are elevated early in the follicular phase. A midcycle surge in LH production stimulates a series of events that culminate in ovulation, with FSH levels falling after this event. Any injury to the hypothalamus or the presence of either psychosocial or physical stressors may lead to changes in these hormonal processes and may result in anovulation and/or amenorrhea.[56]

Menopause

Females who experience amenorrhea for 12 months without another underlying cause are considered to have entered menopause, with a typical onset at age 51. Typically, the ovaries cease producing adequate levels of estrogen to continue fertility, thus signaling an end of the reproductive years. This normal physiologic state is permanent and is characterized by excessive FSH levels. FSH is secreted in an unregulated fashion because there is a lack of estrogen to provide the normal negative feedback to the pituitary for FSH.

Hormone Replacement Therapy

Hormone replacement therapy (HRT), also referred to as estrogen replacement, remains a contentious issue. The Women's Health Initiative study enrolled 16,608 postmenopausal women who were placed on conventional hormone replacement combinations consisting of estrogen and progestin. The study showed an overall increased incidence of invasive breast cancer (hazard ratio, 1.26), stroke, venous clot formation, and coronary heart disease (CHD) events, and no benefit in cognitive decline, with combined estrogen plus progestin therapy (however, *unopposed* estrogen therapy consisting of estrogen alone was not associated with increase in either breast cancer or CHD, suggesting that these risks may be conferred by the progestin component). With HRT, reductions in bone loss, colon polyp formation, and menopausal symptoms (hot flashes and vaginal dryness) were noted. The recently published ELITE trial (Early versus Late Intervention Trial with Estradiol) results showed decreased progression of subclinical atherosclerosis (as seen by slower increase in carotid intima–media thickening) in menopausal women who were started on HRT within 6 years of menopause, lending further credence to the "timing hypothesis": that is, younger women may indeed

benefit from HRT in terms of decreased CHD risk.[57-60] In large meta-analyses of randomized controlled trials, including the Women's Health Initiative, a significantly lower risk of CHD (hazard ratio, 0.68; 95% CI, 0.48 to 0.96) and of death from any cause (hazard ratio, 0.61; 95% CI, 0.39 to 0.95) was found with hormone therapy than with placebo among women who were younger than 60 years of age, less than 10 years past menopause, or both when they underwent randomization.[61]

Testosterone replacement is also a hotly debated topic, as there are males who wish to increase their testosterone levels for many different reasons. The primary indication for testosterone therapy is for hypogonadism, though there are still locales where one may be able to acquire a prescription for testosterone supplementation without a proper diagnosis of hypogonadism. There is current worry about the cardiovascular disease and prostate cancer risks with testosterone supplementation in individuals who are not deficient. For those with accurately diagnosed hypogonadism, testosterone supplementation confers many benefits without the concomitant cardiovascular disease or prostatecancer risks.[62]

Food Intake/Diet

Food and water intake will impact hormone secretion. For example, a lipid-based hormone, leptin,[63] is one of the body's satiety (satisfaction) signals. A leptin deficiency is associated with hyperphagia and obesity. Another example is unregulated growth hormone (GH) production. Under normal circumstances, growth hormone is suppressed by a glucose load. Commonly, a GH suppression test is performed by assessing the GH concentrations both before and after giving a patient a glucose load to detect if the body will properly regulate the GH concentration. Last, AVP is a good example of how fluid levels in our body will determine how much hormone is present. It helps the body conserve water in times of dehydration. When baroreceptors in the right atrium of the heart, aortic arch, and carotid artery sense a decreased pressure, they will signal the posterior pituitary through the vagus nerve to produce AVP. In addition, there is a second mechanism to increase AVP secretion, which is via osmoreceptors in the hypothalamus.

Drugs

Many drugs change hormone secretion or activity. Drugs may also have adverse side effects, like infertility and hypogonadism with testicular shrinkage in patients who use testosterone-based performance-enhancing drugs. Below are some examples of deleterious effects from drugs given for treatment of pathologic conditions.

ACE Inhibitors

ACE inhibitors are typically used in patients with hypertension to decrease the activity of angiotensin converting enzyme and subsequently decrease the amount of angiotensin, which is formed from the conversion of angiotensinogen. Less angiotensin means that there will be less aldosterone created, and with less aldosterone, the kidneys will reabsorb less sodium and retain less water, causing the blood pressure to drop. ACE inhibitors also have a kidney protective effect, and these are typically given to patients with diabetes to protect their kidneys.

Methimazole

Another drug with an intentional dampening effect on hormones is methimazole, a medication used to treat hyperthyroidism. This drug will essentially cause the cessation of thyroid hormone synthesis.[64] Methimazole works by inhibiting the enzyme thyroid peroxidase, which interferes with iodination of the tyrosine residues and subsequently, the formation of active T_3 and T_4.

Corticosteroids

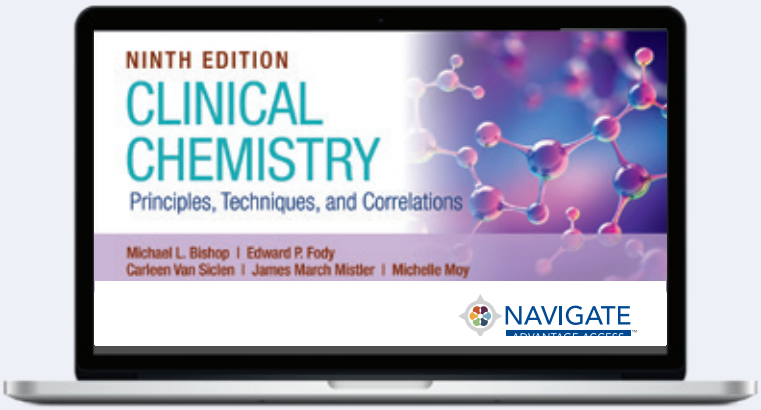
Exogenous corticosteroids will act just like the endogenous cortisol on the receptors in the body and usually increase the glucose concentrations in the plasma. When this occurs, the glucose will negatively loop back at the level of the hypothalamus, turning off CRH secretion. Decreased CRH secretion will decrease secretion of ACTH in the anterior pituitary gland and lastly decrease the endogenous adrenal glucocorticoid production in order for the body to attempt to normalize theblood glucose.

Psychotropic Drugs

Psychotropic drugs are medications effect one's mood and state of mind. The drugs in question are typically selective serotonin reuptake inhibitors (SSRIs) and other antidepressants.[65] An undesired effect may be the syndrome of inappropriate ADH secretion (SIADH). This pathology will typically result in a great amount of water retention and hyponatremia, with serum sodium levels approaching 107 mg/dL or lower.[66] The mechanism of SIADH due to SSRIs is not yet fully elucidated.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Córdoba-Chacón J, Gahete MD, Castaño JP, et al. Somatostatin and its receptors contribute in a tissue specific manner to the sex-dependent metabolic (fed/fasting) control of growth hormone axis in mice. *Am J Physiol Endocrinol Metab*. 2011;300:E46–E54.
2. Ribeiro-Oliveira A, Barkan AL. Growth hormone pulsatility and its impact on growth and metabolism in humans. *Growth Hormone Relat Dis Ther Contemp Endocrinol*. 2011;(pt 1): 33–56.
3. Pradhan G, Samson SL, Sun X. Ghrelin: much more than a hunger hormone. *Curr Opin Clin Nutr Metab Care*. 2013;16:619–624.
4. Chia DJ. Minireview: mechanisms of growth hormone mediated gene regulation. *Mol Endocrinol*. 2014;28: 1012–1025.
5. Møller N, Jørgensen JO. Effects of growth hormone on glucose, lipid, and protein metabolism in human subjects. *Endocr Rev*. 2009;30(2):152–177.
6. Cambiaso P, Schiaffini R, Pontrelli G, et al. Nocturnal hypoglycemia in ACTH and GH deficient children: role of continuous glucose monitoring. *Clin Endocrinol (Oxf)*. 2013;79:232–237.
7. Miller W, Chrousos GP. The adrenal cortex. In: Felig P, Frohman LA, eds. *Endocrinology and Metabolism*. New York, NY: McGraw-Hill; 2001:387–493.
8. Lacroix A. In: Pinchera A, ed. *Endocrinology and Metabolism*. London, UK: McGraw-Hill; 2001:285–297.
9. Orlowski M, Sarao MS. Physiology, follicle stimulating hormone. [Updated 2020 Jul 10]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK535442/
10. Nedresky D, Singh G. Physiology, luteinizing hormone. [Updated 2020 Sep 30]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539692/
11. Al-Chalabi M, Bass AN, Alsalman I. Physiology, prolactin. [Updated 2020 Jul 10]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507829/
12. Carmina E, Campagna AM, Lobo RA. A 20-year follow-up of young women with polycystic ovary syndrome. *Obstet Gynecol*. 2012;119(2):263–269.
13. Poutanen M. Understanding the diversity of sex steroid action. *J Endocrinol*. 2012;212:1–2.
14. te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update*. 2002;8(2):141–154.
15. Greenspan FS. The thyroid gland. In: Greenspan FS, Strewler GJ, eds. *Basic & Clinical Endocrinology*. 5th ed. New York, NY: Appleton & Lange; 1997.
16. Elphick MR. NG peptides: a novel family of neurophysin associated neuropeptides. *Gene*. 2010;458(1–2):20–26.
17. Lee H-J, Macbeth AH, Pagani J, et al. Oxytocin: the great facilitator of life. *Prog Neurobiol*. 2009;88:127–151.
18. Vrachnis N, Iliodromiti Z. The oxytocin-oxytocin receptor system and its antagonists as tocolytic agents. *Int J Endocrinol*. 2011;2011:350546.
19. Ross HE, Young LJ. Oxytocin and neural mechanisms regulating social cognition and affiliative behavior. *Front Neuroendocrinol*. 2009;30:534–547.
20. Wilson JLL, Miranda CA, Knepper MA. Vasopressin and the regulation of aquaporin-2. *Clin Exp Nephrol*. 2013;17: 751–764.
21. Ozier Y, Bellamy L. Pharmacological agents: antifibrinolytics and desmopressin. *Best Pract Res Clin Anaesthesiol*. 2010;24(1):107–119.
22. Jane JA Jr, Vance ML, Laws ER. Neurogenic diabetes insipidus. *Pituitary*. 2006;9(4):327–329.
23. Cassagnol M, Shogbon AO, Saad M. The therapeutic use of vaptans for the treatment of dilutional hyponatremia. *J Pharm Pract*. 2011;24(4):391–399.
24. Taylor R, Jones A, Kelly S, Simpson M, Mabey J. A review of the value of procalcitonin as a marker of infection. *Cureus*. 2017;9(4):e1148. Published 2017, Apr. 10. doi:10.7759/cureus.1148
25. Betz D, Fane K. Human Chorionic Gonadotropin. [Updated 2020 Aug 30]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK532950/

26. Heidegger H, Jeschke U. Human chorionic gonadotropin (hCG)—an endocrine regulator of gestation and cancer. *Int J Mol Sci*. 2018;19(5):1502. https://doi.org/10.3390/ijms19051502
27. Bamalan OA, Al Khalili Y. Physiology, serotonin. [Updated 2020 Oct 3]. In: *StatPearls* [Internet]. Treasure Island (FL): *StatPearls* Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545168/
28. Brinkman JE, Tariq MA, Leavitt L, et al. Physiology, growth hormone. [Updated 2021 Feb 9]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482141/
29. Allen MJ, Sharma S. Physiology, adrenocorticotropic hormone (ACTH) [Updated 2020 Sep 2]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK500031/
30. Pirahanchi Y, Toro F, Jialal I. Physiology, thyroid stimulating hormone. [Updated 2020 Jun 28]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK499850/
31. Stevenson JC. Regulation of calcitonin and parathyroid hormone secretion by oestrogens. *Maturitas*. 1982 Apr;4(1):1–7. doi: 10.1016/0378-5122(82)90013-5. PMID: 6896553.
32. Rule JA, Hynan LS, Attar N, et al. Procalcitonin identifies cell injury, not bacterial infection, in acute liver failure. *PLoS One*. 2015;10(9):e0138566. https://doi.org/10.1371/journal.pone.0138566
33. Khan M, Jose A, Sharma S. Physiology, parathyroid hormone. [Updated 2020 Jul 26]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK499940/
34. Fu Z, Gilbert ER, Liu D. Regulation of insulin synthesis and secretion and pancreatic beta-cell dysfunction in diabetes. *Curr Diabetes Rev*. 2013;9(1):25–53.
35. Morris CH, Baker J. Glucagon. [Updated 2020 Jun 14]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK559195/
36. Prosapio JG, Sankar P, Jialal I. Physiology, gastrin. [Updated 2020 Oct 27]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK534822/
37. DiGregorio N, Sharma S. Physiology, secretin. [Updated 2020 Jul 10]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK537116/
38. Nwabuobi C, Arlier S, Schatz F, Guzeloglu-Kayisli O, Lockwood CJ, Kayisli UA. hCG: biological functions and clinical applications. *Int J Mol Sci*. 2017;18(10):2037. Published 2017, Sept. 22. doi:10.3390/ijms18102037
39. Thau L, Gandhi J, Sharma S. Physiology, cortisol. [Updated 2021 Feb 9]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538239/
40. Scott JH, Menouar MA, Dunn RJ. Physiology, aldosterone. [Updated 2021 Feb 15]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK470339/
41. Nassar GN, Leslie SW. Physiology, testosterone. [Updated 2021 Jan 9]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK526128/
42. Cable JK, Grider MH. Physiology, progesterone. [Updated 2020 Jun 4]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK558960/
43. Delgado BJ, Lopez-Ojeda W. Estrogen. [Updated 2020 Aug 24]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK538260/
44. Paravati S, Rosani A, Warrington SJ. Physiology, catecholamines. [Updated 2020 Jul 26]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK507716/
45. Eghtedari B, Correa R. Levothyroxine. [Updated 2020 Oct 12]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK539808/
46. Bamalan OA, Al Khalili Y. Physiology, serotonin. [Updated 2020 Oct 3]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545168/
47. Gordon G, Altman K, Southern, L, et al. Effect of alcohol (ethanol) administration on sex-hormone metabolism in normal men. *N Engl J Med*. 1976;295:793–797.
48. Smith SA, Travers RJ, Morrissey JH. How it all starts: initiation of the clotting cascade. *Crit Rev Biochem Mol Biol*. 2015;50(4):326–336. doi:10.3109/10409238.2015.1050550
49. Ranabir S, Reetu K. Stress and hormones. *Indian J Endocrinol Metab*. 2011;15(1):18–22. doi:10.4103/2230-8210.77573
50. Goldstein D. Catecholamines and stress. *Endocr Reg*. 2003;37:69–80.
51. Dijk DJ, Duffy JF, Silva EJ, Shanahan TL, Boivin DB, Czeisler CA. Amplitude reduction and phase shifts of melatonin, cortisol and other circadian rhythms after a gradual advance of sleep and light exposure in humans. *PLoS One*. 2012;7(2):e30037. doi: 10.1371/journal.pone.0030037. Epub 2012, Feb. 17. PMID: 22363414; PMCID: PMC3281823.
52. Spiga F, Waite EJ, Liu Y, et al. ACTH-dependent ultradian rhythm of corticosterone secretion. *Endocrinology*. 2011;152:1448–1457.
53. Tonsfeldt KJ, Chappell PE. Clocks on top: the role of the circadian clock in the hypothalamic and pituitary regulation of endocrine physiology. *Mol Cell Endocrinol*. 2012;349:3–12.
54. Jabbour HN, Kelly RW, Fraser HM, et al. Endocrine regulation of menstruation. *Endocr Rev*. 2006;27(1):17–46.
55. Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: mechanisms and clinical consequences. *Endocr Rev*. 2009;30(5):465–493.
56. Santoro N. Update in hyper and hypogonadotropic amenorrhea. *J Clin Endocrinol Metab*. 2011;96(11):3281–3288.
57. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, et al. Women's Health Initiative Investigators. Estrogen plus progestin and colorectal cancer in postmenopausal women. *N Engl J Med*. 2004;350(10):991–1004.
58. Manson JE, Hsia J, Johnson KC, et al. Women's Health Initiative Investigators. Estrogen plus progestin and the risk of coronary heart disease. *N Engl J Med*. 2003;349(6):523–534.
59. Hays J, Ockene JK, Brunner RL, et al. Women's Health Initiative Investigators. Effects of estrogen plus

progestin on health-related quality of life. *N Engl J Med*. 2003;348(19):1839–1854.
60. Hodis HN, Mack W, Henderson VW, et al. Vascular effects of early versus late postmenopausal treatment with estradiol. *N Engl J Med*. 2016;374(13):1221–1231. doi: 10.1056/NEJMoa1505241.
61. Boardman HM, Hartley L, Eisinga A, et al. Hormone therapy for preventing cardiovascular disease in post-menopausal women. *Cochrane Database Syst Rev*. 2015;(3):CD002229.
62. Corona G, Sforza A, Maggi M. Testosterone replacement therapy: long-term safety and efficacy. *World J Mens Health*. 2017;35(2):65–76. doi:10.5534/wjmh.2017.35.2.65
63. Stanley S, Wynne K, McGowan B, Bloom S. Hormonal regulation of food intake. *Physiol Rev*. 2005;85(4):1131–1158
64. Singh G, Correa R. Methimazole. [Updated 2020 Oct 12]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021. Available from: https://www.ncbi.nlm.nih.gov/books/NBK545223/
65. Spigset O, Hedenmalm K. Hyponatraemia and the syndrome of inappropriate antidiuretic hormone secretion (SIADH) induced by psychotropic drugs. *Drug-Safety*. 1995;12:209–225. https://doi.org/10.2165/00002018-199512030-00006
66. Kirpekar VC, Joshi PP. Syndrome of inappropriate ADH secretion (SIADH) associated with citalopram use. *Indian J Psychiatr*. 2005;47(2):119–120. doi:10.4103/0019-5545.55960

Thyroid Function

Yukari Nishizawa-Brennen

CHAPTER OUTLINE

KEY TERMS

Follicular cells
Free T_3
Free T_4
Graves' disease
Hashimoto's thyroiditis
Hyperthyroidism
Hypothalamic–pituitary–thyroid axis
Hypothyroidism
Parathyroid glands
Subacute thyroiditis
Subclinical hyperthyroidism
Subclinical hypothyroidism
Thyroglobulin
Thyroid peroxidase (TPO)
Thyroid stimulating hormone (TSH)
Thyrotoxicosis
Thyrotropin
Thyrotropin-releasing hormone (TRH)
Thyroxine (T_4)
Thyroxine-binding globulin (TBG)
Thyroxine-binding prealbumin (TBPA)
Triiodothyronine (T_3)
TSH receptor antibodies

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Discuss the biosynthesis, secretion, transport, and action of the thyroid hormones.
- Define terms associated with thyroid disorders.
- Sketch and label the hypothalamic–pituitary–thyroid axis.
- Summarize the hypothalamic–pituitary–thyroid axis feedback loops regulating thyroid hormone production.
- Specify the recommended diagnostic test to screen for thyroid dysfunction.
- Explain the principle of each thyroid function test discussed in this chapter.
- Correlate findings with suspected thyroid disorders given thyroid test results.

- Differentiate between primary, secondary, and tertiary thyroid disorders.
- Describe the appropriate laboratory thyroid function testing protocol to use to effectively evaluate or monitor patients with suspected thyroid disease.
- Apply theoretical knowledge of thyroid function to resolve chapter case studies.

The Thyroid

The thyroid gland is responsible for the production of thyroid hormone and calcitonin. Calcitonin is secreted by the parafollicular cells (C cells) and is involved in calcium homeostasis. Thyroid hormone is critical in regulating body metabolism, neurologic development, and numerous other body functions. Clinically, conditions affecting thyroid hormone levels are much more common than those affecting calcitonin and are the major focus of this chapter.[1]

Thyroid Anatomy and Development

The thyroid gland is positioned in the lower anterior neck and is shaped like a butterfly. It is made up of two lobes resting on each side of the trachea, bridged by the isthmus, with a band of thyroid tissue running anterior to the trachea. Posterior to the thyroid gland lie the **parathyroid glands**, which regulate serum calcium levels, and the recurrent laryngeal nerves innervating the vocal cords. The locations of these structures become important during thyroid surgery, when injury could lead to hypocalcemia or permanent hoarse voice.

The fetal thyroid develops from an outpouching of the foregut at the base of the tongue that migrates to its final location over the thyroid cartilage in the first 4 to 8 weeks of gestation. By week 11 of gestation, the thyroid gland begins to produce measurable amounts of thyroid hormone.[1] Iodine is an essential component of thyroid hormone. In parts of the world where severe iodine deficiency exists, neither the mother nor the fetus can produce adequate amounts of thyroid hormone, and both develop **hypothyroidism**. As thyroid hormone is critical to fetal neurologic development, neonatal hypothyroidism can lead to mental impairment and cretinism. In areas where iodine deficiency is not an issue, other problems with thyroid development may occur, including congenital hypothyroidism, which occurs in 1 of 1000–4000 live births; the incidence rate varies with demographics. If the mother has normal thyroid function, small amounts of maternal thyroid hormone crossing the placenta protect the fetus during development. Immediately postpartum, however, these newborns require initiation of thyroid hormone treatment or their neurologic development will be significantly impaired. In much of the developed world, screening tests are performed on newborns to detect congenital hypothyroidism.[2]

Microscopically, the thyroid gland is composed of thyroid follicles. Each follicle contains an outer layer of follicular endothelial cells that surround the thyroid lumen filled with *colloid*. Colloid is predominantly thyroglobulin, a protein used to produce thyroid hormones. Parafollicular cells, also called C cells, are located between the follicular endothelial cells and secrete calcitonin.

Thyroid Hormone Synthesis

Thyroid hormone is made primarily of the trace element iodine, making iodine metabolism a key determinant in thyroid function.[1] Iodine is found in

CASE STUDY 14.1, PART 1

Ailani, a 29-year-old woman, recently underwent radioactive iodine ablation treatment for Graves' disease.

CASE STUDY 14.2, PART 1

Ebony, a 67-year-old woman, is referred for treatment of hyperlipidemia as a follow up to a health screening at her assisted living facility.

seafood, dairy products, and vitamins. Interestingly, iodine is used in high concentrations in the contrast medium used in many radiologic procedures, including computed tomography (CT) scans and heart catheterization. It is also present in amiodarone, a medication used to treat certain heart conditions. The recommended daily intake of iodine for adults is 150 µg, although most individuals in developed countries ingest far more than this amount. In the United States, the average adult daily intake of iodine is sufficient at an estimated 190 to 300 µg, well above the U.S. National Academy of Medicine's recommended adult dietary allowance of 150 µg/day.[3,4] If iodine intake drops below 50 µg daily, the thyroid gland is unable to manufacture adequate amounts of thyroid hormone, and thyroid hormone deficiency, commonly known as **hypothyroidism**, results.[5]

Thyrocytes (thyroid cells) are organized into spheres surrounding a central core of fluid called *colloid*. These structures are called follicles. The major component of colloid, **thyroglobulin**, is a glycoprotein manufactured exclusively by thyroid **follicular cells** and rich in the amino acid tyrosine. Some of these tyrosyl residues will be iodinated, producing the building blocks of thyroid hormone. On the outer side of the follicle, iodine is actively transported into the follicular cells by the sodium/iodide (Na^+/I^-) symporter located on the basement membrane. Inside the thyroid cell, iodide diffuses across the cell to the apical side of the follicle, which abuts the core of colloid. Here, catalyzed by a membrane-bound enzyme called **thyroid peroxidase (TPO)**, iodide is oxidized and bound with tyrosyl residues on thyroglobulin. This results in production of monoiodothyronine (MIT) and diiodothyronine (DIT). This same enzyme also aids in the coupling of two tyrosyl residues to form **triiodothyronine (T_3)** (one MIT residue + one DIT residue) or **thyroxine (T_4)** (two DIT residues). These are the two predominant thyroid hormones. This thyroglobulin matrix, bound to T_4 and T_3, is stored in the thyroid lumen filled with colloid. **Thyroid-stimulating hormone (TSH)**, released from the anterior pituitary, signals the follicular cell to ingest a droplet of colloid by endocytosis. Inside the follicular cell, these droplets are digested by intracellular lysosomes into T_4, T_3, and other products. T_4 and T_3 are then secreted by the follicular cells into the circulation (**Figure 14.1**).

Activity of thyroid hormone depends on the location and number of iodide molecules. Approximately 80% of T_4 is metabolized into either T_3 (35%) or reverse T_3 (rT_3) (45%). Outer-ring deiodination of T_4 (5′-deiodination) leads to production of 3,3′,5-triiodothyronine (T_3). T_3 is three to eight times more metabolically active than is T_4 and often considered to be the biologically active form of thyroid hormone, with T_4 considered the prehormone

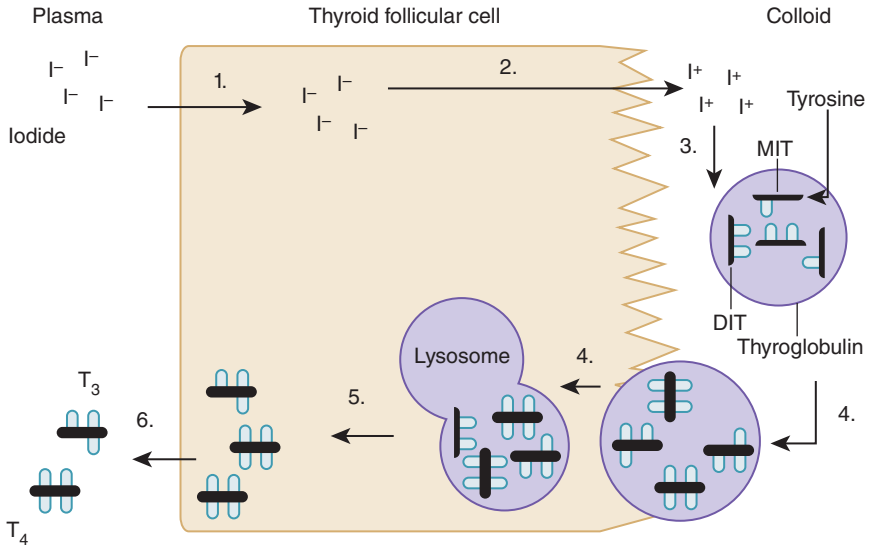


Figure 14.1 Biosynthesis of thyroid hormones includes the following steps: (1) iodide (I^-) trapping by thyroid follicular cells; (2) diffusion of iodide to the apex of the cell and transport into the colloid; (3) oxidation of inorganic iodide to iodine and incorporation of iodine into tyrosine residues within thyroglobulin molecules in the colloid; (4) combination of two diiodotyrosine (DIT) molecules to form tetraiodothyronine (thyroxine, T_4) or of monoiodotyrosine (MIT) with DIT to form triiodothyronine (T_3); (5) uptake of thyroglobulin from the colloid into the follicular cell by endocytosis, fusion of the thyroglobulin with a lysosome, and proteolysis and release of T_4 and T_3; and (6) release of T_4 and T_3 into the circulation.

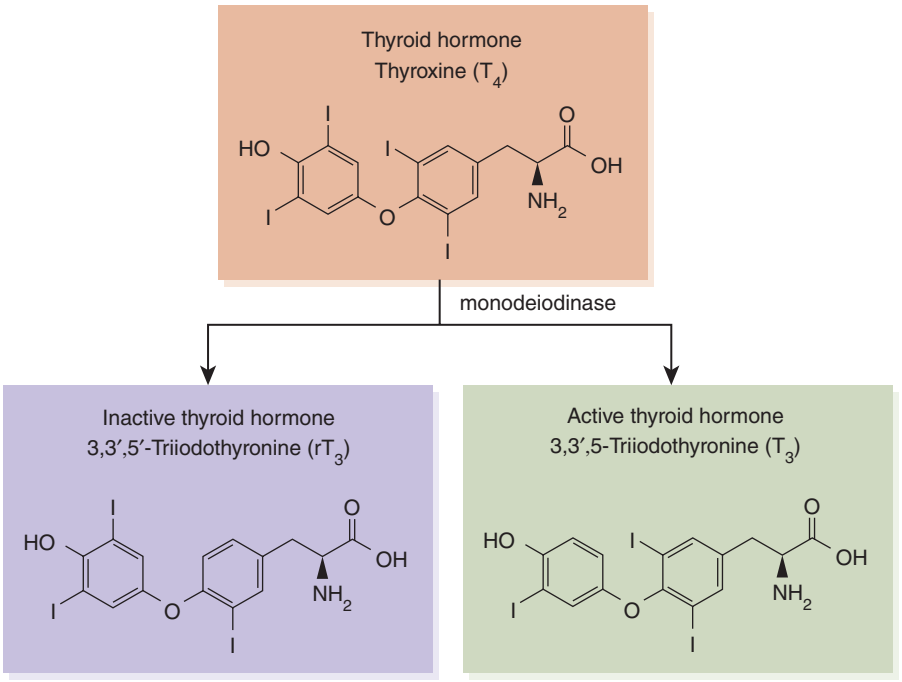


Figure 14.2 Metabolism of thyroxine.

and thyroglobulin being the prohormone. A prohormone is a precursor with minimal hormonal effect. In addition to its "pre"-hormone activity, however, inner-ring deiodination of T_4 results in the production of metabolically inactive rT_3 (**Figure 14.2**).

There are two main forms of iodothyronine 5′-deiodinase. Type 1 iodothyronine 5′-deiodinase is the most abundant form and found mostly in the liver and kidney; this is the largest contributor to the circulating T_3 pool. Certain drugs (e.g., propylthiouracil, glucocorticoids, and propranolol) used to manage symptoms of severe **hyperthyroidism** are known to slow the activity of type 1 deiodinase. Type 2 iodothyronine 5′-deiodinase, found in the brain and pituitary gland, functions to maintain constant levels of T_3 in the central nervous system. Its activity is decreased when levels of circulating T_4 are high and increased when levels are low. Activity of the deiodination enzymes gives another level of control of thyroid hormone activity beyond the **thyrotropin-releasing hormone (TRH)** and TSH (also known as **thyrotropin**) control of the hypothalamic–pituitary–thyroid axis (**Figure 14.3**).[1]

Protein Binding of Thyroid Hormone

When released into the circulation, only 0.04% of T_4 and 0.4% of T_3 are unbound by proteins and available for hormonal activity. The three major binding proteins, in order of significance, are **thyroxine-binding globulin (TBG)**, **thyroxine-binding prealbumin (TBPA)**, and albumin. Alterations to the concentrations of binding proteins significantly affect circulating quantities of T_4 and T_3. For example, high estrogen levels during pregnancy lead to increased TBG production by the liver, which results in higher levels of bound thyroid hormones, leading to higher levels of total T_3 and total T_4. Typically, however, levels of the unbound active, or free, thyroid hormones remain in the normal range and the individual remains euthyroid. In some instances, however, measurement of **free T_4** and **free T_3** may be necessary to eliminate any confusion caused by abnormal binding protein levels.[1]

Control of Thyroid Function

Understanding the **hypothalamic–pituitary–thyroid axis** is essential for correctly interpreting thyroid function test results. This axis is central to the regulation of thyroid hormone production. TRH is synthesized by neurons in the supraoptic and supraventricular nuclei of the hypothalamus and stored in the median eminence of the hypothalamus. When secreted, this hormone stimulates cells in the anterior pituitary gland to manufacture and release TSH. TSH, in turn, circulates to the thyroid gland and leads to increased production and release of thyroid hormones. When the hypothalamus and pituitary sense that an inadequate amount of thyroid hormone is in circulation, TRH and TSH secretion increases, and both stimulate

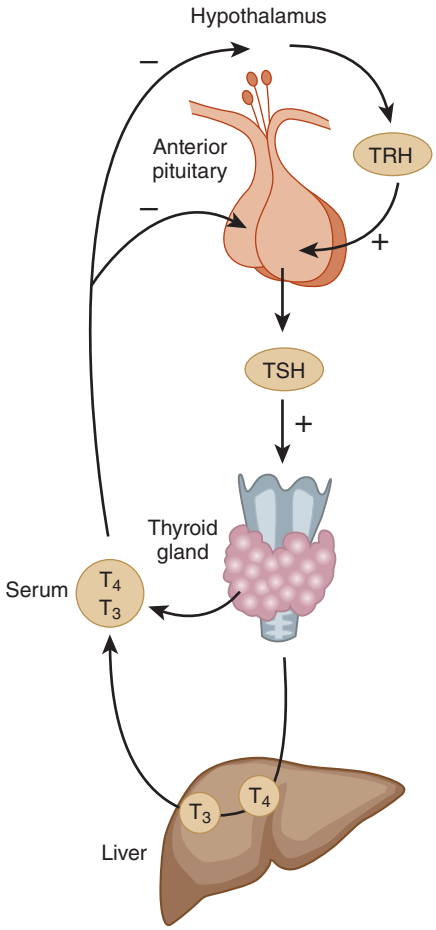


Figure 14.3 Hypothalamic–pituitary–thyroid axis. Thyrotropin-releasing hormone (TRH) stimulates the production and release of thyroid stimulating hormone (TSH, thyrotropin). TSH stimulates the thyroid gland to synthesize and secrete thyroid hormone. T_4 that is released by the thyroid gland is mostly converted to T_3 by the liver and kidney. T_3 and T_4 feedback inhibit TSH release directly through action at the pituitary and indirectly by decreasing TRH release from the hypothalamus.

increased thyroid hormone production. If thyroid hormone levels are high, TRH and TSH release will be inhibited, leading to lower levels of thyroid hormone production. This negative feedback loop requires a normally functioning hypothalamus, pituitary, and thyroid gland, as well as an absence of any interfering agents or agents that mimic TSH action (Figure 14.3).[1]

Actions of Thyroid Hormones

Once released from the thyroid gland, thyroid hormones circulate in the bloodstream either bound to protein or as free, active hormone. Free T_4 and T_3, that circulate unbound, are able to bind to cellular membrane transporters and are then moved across the cell membrane. In the cytoplasm, T_4 is deiodinated into T_3, the more biologically active form. T_3 combines with its nuclear receptor on thyroid hormone-responsive genes, leading to production of messenger RNA that, in turn, leads to production of proteins that influence metabolism and development. Effects of thyroid hormone include tissue growth, brain maturation, increased energy production, increased oxygen consumption, increased expression of β-adrenergic receptors, and specific effects to each organ (e.g., heart, liver, kidneys, skeletal muscle, and skin). Clinically, individuals who have excess thyroid hormone production (hyperthyroidism) or increased circulation of thyroid hormone from any cause (**thyrotoxicosis**), will have symptoms of increased metabolic activity such as tachycardia and tremor. In contrast, individuals with hypothyroidism (decreased levels of thyroid hormone) note symptoms of lowered metabolic activity such as edema and constipation.[1]

Tests for Thyroid Function

Thyroid-Stimulating Hormone

The most useful test for assessing thyroid function is the TSH, currently in its third generation. All TSH immunoassays are capable of screening for thyroid dysfunction, including primary hypothyroidism with elevated levels of TSH. Second-generation TSH immunometric assays, with detection limits of 0.1 mU/L, effectively screen for hyperthyroidism, but the third-generation TSH chemiluminometric assays, with increased sensitivity to detection limits of 0.01 mU/L, give fewer false-negative results and more accurately distinguish between euthyroidism and hyperthyroidism. Although a fourth-generation assay exists providing a 10-fold increase in sensitivity compared to third-generation assays, it is used primarily for research purposes. The third-generation ultrasensitive (s-TSH) assays are the preferred method for monitoring and adjusting thyroid hormone replacement therapy and screening for abnormal thyroid hormone production in the clinical setting.[6] The reference range for s-TSH is 0.3–4.2 mIU/L.

The sensitivity of the third-generation TSH assays led to the ability to detect what is termed *subclinical disease*, which is a mild degree of thyroid dysfunction, due to the large, reciprocal change in TSH levels seen for even small changes in free T_4.[7] In **subclinical hypothyroidism**, the TSH is minimally increased while the free T_4 stays within the normal range. Likewise, in **subclinical hyperthyroidism**, the TSH is low while the free T_4 is normal (**Table 14.1**).

Table 14.1 Interpretation of Thyroid Tests

	Low Free T_4	Normal Free T_4	High Free T_4
Low TSH	Secondary hypothyroidism	Subclinical hyperthyroidism	Primary hyperthyroidism
	Severe nonthyroidal illness	Nonthyroidal illness	
Normal TSH	Secondary hypothyroidism	Normal	Test artifact
	Severe nonthyroidal illness		Secondary hyperthyroidism
			Preanalytic error caused by blood drawn within 6–9 h of thyroxine dose
High TSH	Primary hypothyroidism	Subclinical hypothyroidism	Test artifact
			Secondary hyperthyroidism
			Thyroid hormone resistance[1]

Thyroxine (T_4)

Serum total T_4 levels are usually measured by chemiluminometric assay or similar immunometric technique. As previously mentioned, because more than 99.9% of thyroid hormone is protein bound, alteration in thyroid hormone–binding proteins frequently leads to total T_4 levels outside of the normal range without representing true clinical thyroid dysfunction. Because of this, assays to measure free T_4 and T_3, the biologically active hormone forms, were developed. At least partially because of lower processing costs and ease of interpretation, free T_4 kits now replace total T_4 assessment at the clinical level.[8] Currently available assay kits for measuring free T_4 levels are not error proof, though, and can still be affected by some binding protein abnormalities.[9] When this is suspected, measurement of free T_4 levels is performed by dialysis.[9] The reference range for T_4 is 0.4–4.0 mIU/L.

Triiodothyronine (T_3)

The thyroid gland produces only approximately 20% of triiodothyronine (T_3), which contains three atoms of iodine. The other 80% of the total T_3 is derived from the deiodination or conversion of T_4 to T_3. Quantitatively, there is more circulating T_4 than T_3 levels in the blood, but T_3 is the most biologically active thyroid hormone. It is more potent than T_4 but also has a shorter half-life than T_4. Most of the triiodothyronine circulates in the bloodstream bound to carrier proteins (TBG, albumin, and prealbumin), and only a small fraction circulates unbound (free). Only the free T_3 (FT_3) is biologically active. Total T_3 is composed of both bound and free fractions. In hyperthyroidism, both T_4 and T_3 levels are elevated. However, in T_3 thyrotoxicosis, only the T_3 is elevated. In hypothyroidism, both T_4 and T_3 levels are decreased. The reference range for total T_3 is 80–200 ng/dL. Abnormal levels of TBG and/or albumin may cause abnormal T_3 concentrations. In addition, autoantibodies to thyroid hormones can interfere with the immunoassay.

Free Thyroxine (FT_4)

Free thyroxine (FT_4) comprises a small fraction of total thyroxine because most T_4 is bound to thyroid carrier proteins. FT_4 is available to enter the tissues and therefore is the biologically active fraction of T_4. Clinically, it is used as a second-line test after TSH in the evaluation of suspected thyroid disorders and, when used in conjunction with TSH, allows for a more accurate diagnosis of thyroid function. Increased FT_4 values suggest hyperthyroidism but may also be caused by excess thyroid hormone replacement (exogenous thyroxine). Conversely, decreased values suggest hypothyroidism. Abnormal levels of TBG and/or albumin may cause abnormal total T_4 concentrations; therefore FT_4 measurements provide more reliable patient results. The FT_4 reference range is 0.9–1.7 ng/dL. Known interfering substances include thyroid autoantibodies, biotin, and human anti-mouse antibodies (HAMA).

Free Triiodothyronine (FT_3)

Most T_3 circulates in the blood bound to TBG and albumin. Approximately 99.7% of total T_3 (TT_3) is bound, and only 0.3% is unbound (FT_3). The FT_3 is

the biologically active form. In hyperthyroidism, FT_3 levels are elevated. Due to the low fraction of FT_3, it is measured less frequently. It is a third-level test of thyroid function, providing further confirmation of hyperthyroidism, supplementing T_4, sensitive thyrotropin (sTSH), and T_3 assays. However, FT_3 levels may be required to evaluate clinically euthyroid patients who have an altered distribution of thyroid binding proteins as seen during pregnancy or in patients with dysalbuminemia. The reference range for FT_3 is 2.8–4.4 pg/mL. Increased FT_3 values are commonly associated with hyperthyroidism or with excess thyroid hormone replacement.

Thyroglobulin

Thyroglobulin is a protein synthesized and secreted exclusively by thyroid follicular cells. This prohormone in the circulation is proof of the presence of thyroid tissue. This fact makes thyroglobulin an ideal tumor marker and post-treatment thyroid cancer surveillance test. Patients with well-differentiated thyroid cancer successfully treated with surgery and radioactive iodine ablation should have undetectable thyroglobulin levels.

Thyroglobulin is currently measured by an immunoenzymatic (sandwich) assay as well as enzyme-linked immunoassay (ELISA), immunoradiometric assay (IRMA), and immunochemiluminescent assay (ICMA) methods. The accuracy of the thyroglobulin assay is primarily dependent on the specificity of the antibody used and the absence of antithyroglobulin autoantibodies. Even with modern assays, antithyroglobulin autoantibodies interfere with measurements and lead to unreliable thyroglobulin results. For this reason, it is critically important to screen for autoantibodies whenever thyroglobulin is measured. If antibodies are present, the value of the thyroglobulin assay is marginal. Approximately 25% of patients with well-differentiated thyroid cancer have antithyroglobulin autoantibodies. If a patient with well-differentiated thyroid cancer and antithyroglobulin autoantibodies has been successfully treated with surgery and radioactive iodine ablation, autoantibodies should disappear over time.[10]

Thyroid Autoimmunity

Many diseases of the thyroid gland are related to autoimmune processes. In autoimmune thyroid disease, antibodies are directed at thyroid tissue with variable responses. The most common cause of hyperthyroidism is an autoimmune disorder called **Graves' disease**. The antibodies in this condition are directed at the TSH receptor (TSHR), stimulating the receptor and leading to growth of the thyroid gland (known as a goiter) and production of excessive amounts of thyroid hormone. This condition can be diagnosed with tests that detect TSHR antibodies. Two types of assays exist to detect an autoimmune etiology to hyperthyroidism: competition-based assays that detect TSH receptor antibodies based upon their ability to compete for TSHR with a known ligand, and thyroid-stimulating immunoglobulin (TSI) assays that detect cyclic adenosine monophosphate (cAMP) production in patients' sera.[11] **TSH receptor antibodies** are positive in 70% to 100% of patients with Graves' disease. **Hashimoto's thyroiditis**, also known as *chronic lymphocytic thyroiditis*, is at the other end of the autoimmune continuum. In this condition, antibodies lead to decreased thyroid hormone production by destruction of the thyroid gland, which is the most common cause of hypothyroidism. The most appropriate test for this condition is the TPO antibody, which is present in 10% to 15% of the general population and 80% to 99% of patients with autoimmune hypothyroidism (**Table 14.2**).[12]

Thyroglobulin Antibody

Thyroglobulin autoantibodies are antibodies against thyroglobulin (Tg), which plays a crucial role in thyroid hormone synthesis, storage, and release.

Table 14.2 Prevalence of Thyroid Autoantibodies

Antibody	General Population	Graves' Disease	Hashimoto's thyroiditis
Thyroglobulin antibody (anti-Tg)	3%	12–30%	35–60%
Thyroperoxidase antibody (anti-TPO)	10–15%	45–80%	80–99%
Thyrotropin receptor antibody (anti-TSHR/TRAb)	1–2%	70–100%	6–60%

Tg is not secreted into the bloodstream under normal circumstances. However, destruction through autoimmune hypothyroidism or inflammation of the thyroid (*thyroiditis*) may result in leakage of Tg into the bloodstream. This results in the formation of autoantibodies to Tg (anti-Tg) in a percentage of individuals. The same destructive process may result in the formation of autoantibodies to other thyroid antigens, in particular TPO (anti-TPO). Anti-Tg autoantibodies are most commonly associated with Hashimoto's thyroiditis. In individuals with hypothyroidism due to autoimmune disease, 35% to 60% will have anti-Tg autoantibodies. In Graves' disease, anti-Tg autoantibodies are observed in 12% to 30% of patients. The reference range for thyroglobulin antibodies is <4.0 IU/mL.

Thyroperoxidase Antibody

Thyroperoxidase (TPO) is an enzyme found in the thyroid gland that is involved in thyroid hormone synthesis. It is also known as *iodide peroxidase* and is the enzyme responsible for the synthesis of T_4 and T_3. TPO antibodies activate, complement, and play an important role in pathogenesis of hypothyroidism. The determination of TPO antibodies is the most sensitive test for detecting autoimmune thyroid diseases including Hashimoto's thyroiditis and Graves' disease. These thyroid diseases have detectable concentrations of TPO antibodies. The highest TPO antibody levels are observed in patients with Hashimoto's thyroiditis, and TPO antibodies are present in approximately 80–99% of these patients. These autoantibodies also frequently occur (45–80%) in Graves' disease. Clinically, the TPO antibody test is used as a diagnostic tool in deciding whether to treat a patient with subclinical hypothyroidism. The reference range for TPO autoantibodies is <9.0 IU/mL, and values above 9.0 IU/mL are associated with autoimmune thyroiditis. Elevated levels are also seen in other autoimmune diseases.

Thyrotropin Receptor Antibody

Autoimmune thyroid disease is characterized by the presence of autoantibodies against Tg, TPO, and the thyrotropin/TSH receptor (TSHR). Among these autoantibodies, thyrotropin receptor autoantibodies (TRAb) are most closely associated with disease pathogenesis. All forms of autoimmune thyrotoxicosis, including Graves' disease, are caused by the production of stimulating TRAb. These autoantibodies, also known as long-acting thyroid stimulator (LATS) or thyroid-stimulating immunoglobulins (TSI), bind to the thyrotropin receptor and stimulate the thyroid gland, independent of the normal feedback mechanism regulated by TSH. Approximately 70% to 100% of patients with Graves' disease have TRAb. TRAb may be detected prior to clinical manifestations of autoimmune thyrotoxicosis. TSI are IgG antibodies, and they can therefore cross the placenta, causing neonatal thyrotoxicosis. The reference range for TRAb is ≤ 1.75 IU/L.

Other Tools for Thyroid Evaluation

Nuclear Medicine Evaluation

Radioactive iodine is useful in assessing the metabolic activity of thyroid tissue and assisting in the evaluation and treatment of thyroid cancer. When radioactive iodine is given orally, a percentage of the dose is taken up by the thyroid gland. This percentage is called the radioactive iodine uptake (RAIU). High uptake suggests that the gland is metabolically active and producing significant amounts of thyroid hormone. Low uptake suggests that the gland is metabolically inactive. Because TSH stimulates iodine uptake by the thyroid gland, it is important to interpret the scan in conjunction with an assessment of TSH levels. An undetectable or low TSH should turn off the thyroid gland's uptake of iodine. If the uptake is normal or high despite an undetectable or low TSH, the thyroid must be either acting autonomously without regard to the hypothalamic–pituitary–thyroid access feedback system or through a TSH surrogate. Such is the case with Graves' disease, where an immunoglobulin activates the TSHR on the thyroid gland, leading to high rates of thyroid hormone production and a high RAIU. The high level of thyroid hormone in the circulation feeds back to the pituitary and hypothalamus, turning off TSH and TRH, but this has no effect on the levels of TSI. Conversely, if the RAIU is low in the presences of an undetectable or low TSH, the differential diagnosis includes excess exogenous thyroid hormone ingestion, high iodine intake, or a condition in which stored thyroid hormone is leaking from the thyroid gland (typically in a setting known as **subacute thyroiditis**).[13]

Radioactive iodine is also useful to detect and treat metastases and recurrent thyroid cancer. Thyroid cancer tissue takes up radioactive iodine that can be detected by whole-body imaging. Higher dosages

of radioactive iodine would subsequently be used to destroy the cancer tissue.[14]

Thyroid Ultrasound

The significance of thyroid ultrasound in the assessment of thyroid anatomy and characterization of palpable thyroid abnormalities has progressively increased in the last few decades. Ultrasound imaging uses high-frequency sound waves to produce images of the thyroid gland. Gel is placed on the neck. and a probe is used to collect the sounds that bounce back. This noninvasive procedure is able to detect non-palpable thyroid nodules. Small (<1.5 cm) thyroid nodules can be seen in approximately 50% of clinically normal thyroid glands.[15]

Fine-Needle Aspiration

Thyroid fine-needle aspiration (FNA) biopsy is often the first step and most accurate tool in the evaluation of thyroid nodules in the absence of hyperthyroidism. The routine use of FNA allows prompt identification and treatment of thyroid malignancies and avoids unnecessary surgery in most individuals with benign thyroid lesions. In this procedure, a small-gauge needle is inserted into the thyroid nodule, and cells are aspirated for cytologic evaluation in the laboratory. It is becoming more commonly used with ultrasound-guided imaging. FNA biopsy results are reported by a pathologist according to the Bethesda System for Reporting Thyroid Cytopathology as falling into one of six categories: nondiagnostic/unsatisfactory, benign, atypia/follicular lesion of undetermined significance, follicular neoplasm/suspicious for follicular neoplasm, suspicious for malignancy, and malignant. These categories dictate subsequent treatment, ranging from routine ultrasound monitoring to surgical excision.[16,17]

Disorders of the Thyroid

Hypothyroidism

Hypothyroidism, defined as a low FT_4 level with a normal or high TSH, is one of the most common disorders of the thyroid gland, occurring in up to 20% of women over the age of 65. Symptoms of hypothyroidism vary, depending on the degree of hypothyroidism and the rapidity of its onset (**Box 14.1**). When thyroid hormone is significantly

Box 14.1 Signs and Symptoms of Hypothyroidism

Signs

Delayed relaxation phase of deep tendon reflex testing
Bradycardia
Diastolic hypertension
Coarsened skin, yellowing of skin (carotenemia)
Edema in the face and hands
Thinning of eyebrows/loss of lateral aspect of brows
Pleural/pericardial effusion
Ascites

Symptoms

Cold intolerance
Depression
Intellectual disability
Menorrhagia
Growth failure (children)
Pubertal delay
Dry skin
Edema
Constipation
Hoarseness
Dyspnea on exertion

CASE STUDY 14.1, PART 2

Remember Ailani? She is seen by her endocrinologist for her 6-week follow-up appointment. Her laboratory test results were as follows:

Analyte	Patient Value	Reference Range
TSH	8.01	(0.4–4.0 mIU/mL)
FT_4	0.79	(0.78–2.19 ng/dL)

1. Did the treatment remove too little, too much, or just the right amount of thyroid tissue?
2. What treatment does Ailani need now?

decreased, symptoms of cold intolerance, fatigue, dry skin, constipation, hoarseness, dyspnea on exertion, cognitive dysfunction, hair loss, and weight gain have been reported.[1] On physical examination, those with severe hypothyroidism may have low body temperature, bradycardia, delay in the relaxation phase of deep tendon reflexes, yellow discoloration of the skin (from hypercarotenemia), hair loss, diastolic hypertension, pleural and pericardial effusions, menstrual irregularities, and edema.[1,2]

Because of the diffuse distribution of thyroid hormone receptors and the many metabolic effects of thyroid hormone, hypothyroidism can lead to a variety of other abnormalities. Hyponatremia can occur from the combination of increased urinary sodium excretion and an inability to maximally dilute urine due to inappropriate release of antidiuretic hormone. Severe and prolonged hypothyroidism can lead to myopathy and elevated levels of creatine kinase (CK). Anemia can also be seen, either as a result of a decreased demand for oxygen carrying capacity or through an associated autoimmune pernicious anemia, menorrhagia, or malabsorption of iron and folic acid.[1,18,19] Fifty percent or more of those with uncorrected hypothyroidism will have hyperlipidemia that improves with thyroid hormone replacement.[20,21] In the presence of these clinical abnormalities (hyponatremia, unexplained elevation of CK, anemia, or hyperlipidemia), evaluation for hypothyroidism as a potential secondary cause should be considered.

Box 14.2 Types of Hypothyroidism

Primary
Thyroid gland dysfunction

Secondary
Pituitary dysfunction

Tertiary
Hypothalamic dysfunction

Hypothyroidism can be divided into primary, secondary, or tertiary disease, dependent on the location of the defect (**Box 14.2**). The most common cause of hypothyroidism is Hashimoto's thyroiditis. This disorder is an autoimmune disease targeting the thyroid gland, often associated with an enlarged gland, or goiter.[1] TPO antibody testing is positive in 80% to 99% of patients with chronic lymphocytic thyroiditis[12] (Hashimoto's thyroiditis). Other common causes of hypothyroidism include thyroid surgery, radiation therapy, radioactive iodine treatment for hyperthyroidism, and medications. (**Table 14.3**). Occasionally, individuals will experience transient hypothyroidism associated with inflammation of the thyroid gland. Examples of transient hypothyroidism include recovery from nonthyroidal illness and the hypothyroid phase of any of the forms of subacute thyroiditis (painful thyroiditis, postpartum thyroiditis, and painless thyroiditis).[1]

Table 14.3 Causes of Hypothyroidism

	Condition	Comments
Primary	Hashimoto's thyroiditis (Chronic lymphocytic thyroiditis)	Anti-TPO is positive in 80–99% of cases Anti-Tg is positive in 35–60% of cases
	Treatment for toxic goiter (i.e., subtotal thyroidectomy or radioactive iodine)	History and physical exam (neck scar) are key to diagnosis
	Excessive iodine intake	History and urinary iodine measurement useful
	Subacute thyroiditis	Usually transient
Secondary	Hypopituitarism	Caused by adenoma, radiation therapy, or destruction of pituitary
Tertiary	Hypothalamic dysfunction	Rare

Anti-Tg, thyroglobulin antibodies; Anti-TPO, thyroid peroxidase antibodies.

Screening

The American Thyroid Association and American Association of Clinical Endocrinologists guidelines for hypothyroidism screening call for initial measurement of TSH in asymptomatic patients at age 35, then routine screening every 5 years after the age of 35. Individuals with symptoms or risk factors such as goiter, presence of another autoimmune disease, a first-degree relative with autoimmune thyroid disease, lithium use, or amiodarone use should be monitored more closely via TSH testing and physical examination.[22]

Treatment

Hypothyroidism is treated with thyroid hormone replacement therapy.[22,23] Levothyroxine (T_4) is the treatment of choice. In primary hypothyroidism, the goal of therapy is to achieve a normal TSH level. If hypothyroidism is of secondary or tertiary origin, TSH levels will not be useful in managing the condition, and a mid-normal FT_4 level becomes the treatment target. Levothyroxine has a half-life of approximately 7 days. When doses of exogenous thyroid hormone are changed, it is important to wait at least five half-lives (35 days) before rechecking thyroid function tests in order to achieve a new steady state.[1]

Box 14.3 Signs and Symptoms of Thyrotoxicosis

Signs

Tachycardia
Tremor
Warm, moist, flushed, smooth skin
Lid lag, widened palpebral fissures
Ophthalmopathy (Graves' disease)
Goiter
Brisk deep tendon reflexes
Muscle wasting and weakness
Dermopathy/pretibial myxedema (Graves' disease)
Osteopenia, osteoporosis

Symptoms

Nervousness, irritability, anxiety
Tremor
Palpitations
Fatigue, weakness, decreased exercise tolerance
Weight loss
Heat intolerance
Hyperdefecation
Menstrual changes (oligomenorrhea)
Prominence of eyes

Thyrotoxicosis

Thyrotoxicosis is a constellation of findings that result when peripheral tissues are presented with, and respond to, an excess of thyroid hormone. Thyrotoxicosis can be the result of excessive thyroid hormone ingestion, leakage of stored thyroid hormone from storage in the thyroid follicles, or excessive thyroid gland production of thyroid hormone. The manifestations of thyrotoxicosis vary, depending on the degree of thyroid hormone elevation and the status of the affected individual. Symptoms often include anxiety, emotional lability, weakness, tremor, palpitations, heat intolerance, increased perspiration, and weight loss despite a normal or increased appetite (**Box 14.3**).

Hyperthyroidism

Graves' disease is the most common cause of hyperthyroidism. It is an autoimmune disease in which antibodies are produced that activate the TSHR. Features of Graves' disease include thyrotoxicosis, goiter, ophthalmopathy (eye changes associated with inflammation and infiltration of periorbital tissue), and dermopathy (skin changes in the lower extremities that have an orange-peel texture). There is a strong familial disposition to Graves' disease. Approximately 15% of patients will have a close relative with this condition, and women are five times more likely to develop it than men. Laboratory testing will usually document a high free T_4 and/or T_3 level with a low or undetectable TSH. TSIs and TSH receptor antibodies are usually positive in this condition. Radiologic findings include elevated RAIU, and the thyroid scan will show diffuse uptake (**Table 14.4**).[1]

Approximately 50% of patients with Graves' hyperthyroidism develop clinically obvious Graves' ophthalmopathy, a particularly concerning manifestation. With more sensitive testing, such as orbital CT scanning or magnetic resonance imaging (MRI), most patients with Graves' hyperthyroidism will be shown to have ophthalmopathy.[24] Findings in Graves' ophthalmopathy include orbital soft tissue swelling, injection of the conjunctivae, proptosis (forward protrusion of the eye secondary to infiltration of retro-orbital muscles and fat), double vision (secondary to orbital muscle involvement and fibrosis), and corneal disease (often related to trauma because of difficulty closing the eyelids). Treatment of Graves' ophthalmopathy may include noninvasive modalities such as moisturizing and protecting the cornea with drops and ointment or more invasive therapies such

Table 14.4 Disorders Associated with Thyrotoxicosis

	Condition	Pathogenic Mechanism	TSH Level	RAIU	Other Tests
Hyperthyroidism	Graves' disease	Anti-TSHR	↓	↑	Anti-TSHR positive; TSI positive
	Toxic adenoma	Benign nodule	↓	↑	Seen on thyroid scan
	Toxic multinodular goiter	Foci of functional autonomy	↓	↑	Seen on thyroid scan
	TSH-secreting tumor	Benign pituitary tumor	Normal/↑	↑	Pituitary MRI
Non-hyperthyroidism	Painful thyroiditis	Leakage of thyroid hormone	↓	↓	↑ Tg
	Postpartum thyroiditis	Leakage of thyroid hormone	↓	↓	↑ TPO Ab
	Exogenous hormone	Excess medication	↓	↓	
	Ectopic thyroid tissue	Metastatic thyroid cancer; struma ovarii	↓	↓	Distant metastasis seen on radioactive iodine scan

RAIU, radioactive iodine uptake by thyroid; Tg, thyroglobulin; TSHR Ab, TSH receptor antibodies; TSI, thyroid-stimulating immunoglobulin; Anti-TPO, thyroperoxidase antibodies.

as injection of glucocorticoids retro-orbitally and, less frequently, surgical decompression of the orbits to prevent optic nerve injury and blindness.[25]

Thyroid disease associated with Graves' disease is treated with medication, radioactive iodine, surgery, or a combination of these. Initially, many thyrotoxic patients require β-blockers to control symptoms of adrenergic excess, such as tremor and tachycardia. Thyroperoxidase inhibitors such as propylthiouracil (PTU) or methimazole (MMI) can be added to inhibit thyroid hormone biosynthesis and secretion.[1,26] The antithyroidal medications, as they are known, carry a risk profile that includes rash and, rarely, hepatotoxicity, agranulocytosis, and aplastic anemia.[1] Long-term remission rate is approximately 50%. Patients with small goiters and mild hyperthyroidism are more likely to achieve remission. Low dietary iodine increases the chance of staying in long-term remission.[26]

When radioactive iodine or surgery is used, the goal is to destroy or remove enough thyroid tissue so that the patient becomes hypothyroid. Subsequent lifelong treatment with thyroid hormone replacement therapy is usually required. Radioactive iodine therapy has been used for treatment of Graves' disease for more than 50 years and is considered both safe and effective. Surgery is associated with risk of recurrent laryngeal nerve injury, leading

CASE STUDY 14.2, PART 2

Recall Ebony. Ebony scheduled an appointment with her primary care physician. The physician noted her cholesterol and triglycerides were high from the screening. Additional blood work confirmed the high cholesterol and triglycerides. Ebony's history review indicated hair loss (she wears a wig), hoarseness to her voice, cold intolerance, and fatigue. She indicated adherence with her lipid-lowering medication.

1. What laboratory test should be ordered to screen Ebony for thyroid disease?
2. If diagnosed with thyroid dysfunction, what treatment might Ebony require?
3. What other abnormal laboratory results other than hyperlipidemia are commonly seen in Ebony's thyroid disorder?

to permanent hoarseness, and injury to the parathyroid glands, causing hypoparathyroidism leading to hypocalcemia.[1]

Because of their comparative risks, radioactive iodine is generally the preferred treatment modality in the United States. Antithyroidal medications are typically used either because of patient preference or during pregnancy and breastfeeding. There are two situations in Graves' disease in which surgery is preferred to other forms of therapy: (1) if there is concern that the patient may have thyroid cancer in addition to Graves' disease, surgery is the best way to ensure removal of the potential cancer, and (2) in patients with severe ophthalmopathy, some experts in Graves' disease management prefer surgery because of concern that radioactive iodine treatment may cause an acute flaring of associated eye problems. Surgery may also be used during the second trimester of pregnancy in women for whom antithyroidal medication is not an option.[1]

Toxic Adenoma and Multinodular Goiter

Toxic adenomas and multinodular goiter are two relatively common causes of hyperthyroidism. These conditions are caused by autonomously functioning thyroid tissue. In these instances, neither TSH nor TSHR-stimulating immunoglobulin is required to stimulate thyroid hormone production. In some toxic nodules, receptor mutations have been identified. These mutations have the same effect as chronic stimulation of the TSHR on thyroid hormone production. Clinically, toxic adenomas present with signs and symptoms of hyperthyroidism, possibly with palpable nodule(s). On a thyroid scan, the nodules are "hot"—that is, they avidly take up radioactive iodine. The RAIU within the nodule is also inappropriately high for the suppressed level of TSH. In toxic multinodular goiter, there are multiple areas within the thyroid gland that are autonomously producing thyroid hormone (Table 14.4). Treatment for these two conditions involves surgery, radioactive iodine, or thyroperoxidase inhibitor medications. Although the medications can block thyroid hormone production, they are not expected to lead to remission in these two conditions. Often, the toxic nodules produce so much thyroid hormone that the rest of the thyroid gland is suppressed and metabolically inactive. When radioactive iodine is given, it tends to destroy only the hyperactive (autonomous) nodules, leaving normal (suppressed) thyroid tissue undamaged. Because the normal thyroid tissue is hypofunctioning and takes up little of the radioactive iodine, when treatment is given, the patient may be left with normal thyroid function without the need for thyroid hormone replacement therapy.[1]

Drug-Induced Thyroid Dysfunction

Amiodarone-Induced Thyroid Disease

Several drugs other than thyroperoxidase inhibitors can also affect thyroid function. Amiodarone, a drug used to treat cardiac arrhythmias, is a fat-soluble drug with a long half-life (50 days) that interferes with normal thyroid function.[27] The fact that 37% of the molecular weight of amiodarone is iodine accounts for a significant part of the thyroid dysfunction observed. Iodine, when given in large doses, acutely leads to inhibition of thyroid hormone production. This is called the Wolff-Chaikoff effect.[1] Amiodarone also blocks T_4-to-T_3 conversion. The combination of these two actions leads to hypothyroidism in 8% of patients on chronic amiodarone therapy. Amiodarone can also lead to hyperthyroidism in 3% of patients treated chronically with this medication.[27] Certain patients may develop hyperthyroidism as they escape the Wolff-Chaikoff effect and use the excess iodine for thyroid hormone production. In others, the medication may induce inflammation of the gland (subacute thyroiditis) with subsequent leakage of stored thyroid hormone into the circulation.[1]

Subacute Thyroiditis

Several conditions occur that lead to transient changes in thyroid hormone levels.[28] These conditions are associated with inflammation of the thyroid gland, leakage of stored thyroid hormone, followed by repair of the gland. Although nomenclature varies between authors, grouping together postpartum thyroiditis, painless thyroiditis, and painful thyroiditis as forms of subacute thyroiditis is one of the simplest classification schemes. These conditions are often associated with a thyrotoxic phase when thyroid hormone is leaking into the circulation, a hypothyroid phase when the thyroid gland is repairing itself, and a euthyroid phase when the gland is repaired, with the duration of these phases lasting from weeks to months.[1]

Postpartum thyroiditis is the most common form of subacute thyroiditis. It occurs in 5% to 9% of women in the postpartum period.[12] It is strongly associated with the presence of TPO antibodies and chronic lymphocytic thyroiditis. Patients may experience a period of thyrotoxicosis followed by hypothyroidism or simply hypothyroidism or hyperthyroidism. Thyroid hormone levels usually return to normal after several months; however, even 7 to 9 years postpartum, 25% of patients have persistent hypothyroidism.[1,12] During the thyrotoxic phase, β-blockers can be used if treatment is necessary. During the hypothyroid phase, thyroid hormone replacement therapy can be given if symptoms require, usually for 3 to 6 months, with continuation if permanent hypothyroidism develops. The thyrotoxic phase of this condition, as well as other forms of subacute thyroiditis, can be distinguished from Graves' disease by a low RAIU and an absence of TSI or TSH receptor antibodies (Table 14.4). Painless thyroiditis or subacute lymphocytic thyroiditis shares many characteristics of postpartum thyroiditis, except there is no associated pregnancy.

Painful thyroiditis, also called subacute granulomatous, subacute nonsuppurative thyroiditis, or De Quervain's thyroiditis, is characterized by neck pain, low-grade fever, myalgia, a tender diffuse goiter, and swings in thyroid function tests (as discussed earlier). Viral infections are felt to trigger this condition. TPO antibodies are usually absent; erythrocyte sedimentation rate and thyroglobulin levels are often elevated.[1]

Nonthyroidal Illness

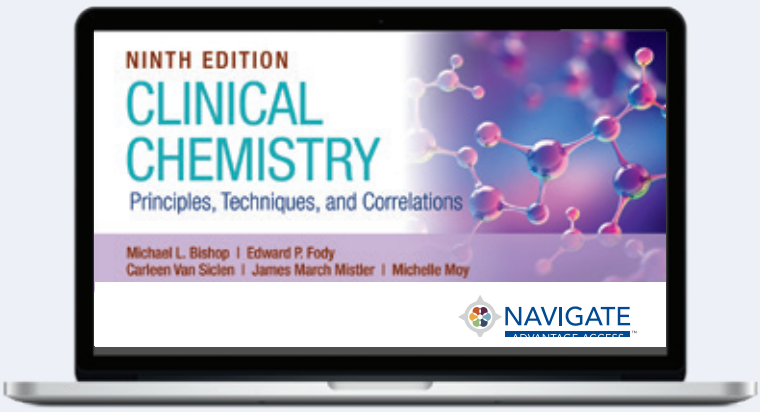
Hospitalized patients, especially critically ill patients, often have abnormalities in their thyroid function tests without thyroid dysfunction, a condition known as nonthyroidal illness or euthyroid sick syndrome. Typically, the laboratory pattern is one of normal or low TSH, low T_3, and low free T_4. Because illness decreases 5′-monodeiodinase activity, less T_4 is converted to active T_3. This leads to decreased levels of T_3 and higher levels of reverse T_3. There also seems to be an element of central hypothyroidism and thyroid hormone–binding changes associated with severe illness. It is believed that many of these changes are an appropriate adaptation to illness, and thyroid hormone replacement therapy is not indicated.[1]

Thyroid Nodules

Thyroid nodules are common and clinically apparent thyroid nodules are present in 6.4% of adult women and 1.5% of adult men, according to Framingham data.[29] Autopsy studies, however, find nodules in 50% of normal thyroid glands.[30] Despite the frequency of thyroid nodules, only 6% to 9% of nonpalpable nodules prove to be thyroid cancer.[31] FNA of these nodules, with cytologic examination of the aspirate, has become a routine practice to help distinguish the nodules that require surgical removal from those that do not.[12]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Cooper DS, Ladenson PW. The thyroid gland. In: Gardner DG, Shoback D, eds. *Greenspan's Basic & Clinical Endocrinology*. 10th ed. New York, NY: McGraw-Hill Education; 2021.
2. Lavin L. *Manual of Endocrinology and Metabolism*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2019:489.
3. Caldwell KL, Makhmudov A, Ely E, et al. Iodine status of the U.S. populations, national health and nutrition examination survey, 2005–2006 and 2007–2008. *Thyroid.* 2011;21: 419–427.
4. Institute of Medicine, Food and Nutrition Board. Iodine. In: *Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc*. Washington, DC: National Academies Press; 2001.
5. National Research Council. In: *Distribution and Administration of Potassium Iodide in the Event of a Nuclear Incident*. Washington, DC: National Academies Press; 2004.
6. Ross DS, Ardisson LJ, Meskell MJ. Measurement of thyrotropin in clinical and subclinical hyperthyroidism using a new chemiluminescent assay. *J Clin Endocrinol Metab.* 1989;69:684–688.
7. Spencer CA, LoPresti JS, Patel A, et al. Applications of a new chemiluminometric thyrotropin assay to subnormal measurement. *J Clin Endocrinol Metab.* 1990;70:453–460.
8. Ekins R. The free hormone hypothesis and measurement of free hormones [editorial]. *Clin Chem.* 1992;38:1289–1293.
9. Wong TK, Pekary AE, Hoo GS, et al. Comparison of methods for measuring free thyroxine in nonthyroidal illness. *Clin Chem.* 1992;38:720–724.
10. Gianoukakis, AG. Thyroglobulin antibody status and differentiated thyroid cancer: what does it mean for prognosis and surveillance? *Curr Opin Oncol.* 2015;27:26–32.
11. Barbesino G, Tomer Y. Clinical utility of TSH receptor antibodies. *J Clin Endocrinol Metab.* 2013;98:2247–2255.
12. Saravanan P, Dayan CM. Thyroid autoantibodies. *Endocrinol Metab Clin North Am.* 2001;30:315–337.
13. Salvatore D, Cohen R, Kopp PA, Larsen PR. Thyroid pathophysiology and diagnostic evaluation. In: Melmed S, Auchus RJ, Goldfine AB, Koenig RJ, Rosen CJ, eds. *Williams Textbook of Endocrinology.* 14th ed. Philadelphia, PA: Elsevier; 2020: 332–363.
14. Cavalieri RR. Nuclear imaging in the management of thyroid carcinoma. *Thyroid.* 1996;6:485–492.
15. Tan GH, Gharib H. Thyroid incidentalomas: management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. *Ann Intern Med.* 1997;126:226–231.
16. Haugen BR, Alexander EK, Bible KC, et al. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association Guidelines Task Force on thyroid nodules and differentiated thyroid cancer. *Thyroid.* 2016;26:1–133.
17. Crippa S, Mazzucchelli L, Cibas ES, et al. The Bethesda System for reporting thyroid fine-needle aspiration specimens. *Am J Clin Pathol.* 2010;134:343–351.
18. Mariani L, Berns J. The renal manifestations of thyroid disease. *J Am Soc Nephrol.* 2012;23:22–26.
19. Khaleeli AA, Gohil K, McPhail G, et al. Muscle morphology and metabolism in hypothyroid myopathy: effects of treatment. *J Clin Pathol.* 1983;36:519–526.
20. Diekman T, Lansberg PJ, Kastelein JJ, et al. Prevalence and correction of hypothyroidism in a large cohort of patients referred for dyslipidemia. *Arch Intern Med.* 1995;155: 1490–1495.
21. Duntas LH. Thyroid disease and lipids. *Thyroid*. 2002;12: 287–293.
22. Ladenson PW, Singer PA, Ain KB, et al. American Thyroid Association guidelines for detection of thyroid dysfunction. *Arch Intern Med.* 2000;160:1573–1575.
23. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Thyroid.* 2012;22:1200–1235.
24. Wiersinga WM, Bartalena L. Epidemiology and prevention of Graves' ophthalmopathy. *Thyroid.* 2002;12:855–860.
25. Burch HB, Wartofsky L. Graves' ophthalmopathy: current concepts regarding pathogenesis and management. *Endocr Rev.* 1993;14:747–793.
26. Solomon BL, Evaul JE, Burman KD, et al. Remission rates with antithyroid drug therapy: continuing influence of iodine intake? *Ann Intern Med.* 1987;107:510–512.
27. Nademanee K, Singh BN, Callahan B, et al. Amiodarone, thyroid hormone indexes, and altered thyroid function: long-term serial effects in patients with cardiac arrhythmias. *Am J Cardiol.* 1986;58:981–986.
28. Pearce EN, Farwell AP, Braverman LE. Thyroiditis. *N Engl J Med.* 2003;348:2646–2655.
29. Vander JB, Gaston EA, Dawber TR. The significance of nontoxic thyroid nodules. Final Report of a 15-year study of the incidence of thyroid malignancy. *Ann Intern Med.* 1968;69:537–540.
30. Brander A, Viikinkoski P, Nickels J, et al. Thyroid gland: US screening in a random adult population. *Radiology.* 1991;181:683–687.
31. Papini E, Guglielmi R, Bianchini A, et al. Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-Doppler features. *J Clin Endocrinol Metab.* 2002;87:1941–1946.

CHAPTER 15

Hypothalamic and Pituitary Function

Stephanie L. Jacobson

CHAPTER OUTLINE

KEY TERMS

Adrenocorticotropin hormone (ACTH)
Anterior pituitary
Arginine vasopressin (AVP)
Corticotropin-releasing hormone (CRH)
Diurnal variation
Dopamine
Follicle-stimulating hormone (FSH)
Gonadotropin-releasing hormone (GnRH)
Growth hormone (GH)
Growth hormone–releasing hormone (GHRH)
Hypothalamic–hypophyseal portal system
Hypothalamus
Insulin-like growth factor (IGF)
Luteinizing hormone (LH)
Neurohypophysis
Oxytocin
Posterior pituitary
Prolactin
Prolactin inhibitory factor (PIF)
Pulsatile secretion
Somatostatin (SS)
Thyroid-stimulating hormone (TSH)
Thyrotropin-releasing hormone (TRH)
Tropic hormone

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Describe the functions of and list the hormones secreted by hypothalamus and the anterior and posterior pituitary glands.
- Diagram the anatomic relationship between the pituitary and hypothalamus.
- Explain the feedback loops of the hypothalamic and pituitary hormones.

- Differentiate the effects of pulsatility and diurnal variation on the results of hormone measurements discussed in this chapter.
- Examine the regulation of prolactin secretion.
- State the causes of prolactin elevation.
- Correlate the clinical features, diagnostic testing, and laboratory results for the hormones discussed in this chapter.

The term *pituitary* (derived from both Latin and Greek) literally means to "spit mucus," reflecting the primitive notion of pituitary function. In one way, the ancient physiologists were correct—they believed that the brain was responsible for signaling the pituitary to secrete mucus; however, instead of mucus, it was later discovered that the brain directs the pituitary to secrete hormones that regulate other endocrine glands. When this was recognized, the pituitary was designated the "master gland." Without the pituitary, there is a cessation of growth, together with profound alterations in intermediary metabolism and failure of gonadal, thyroidal, and adrenal function. The pituitary is also referred to as the *hypophysis*, from Greek meaning "undergrowth," attesting to its unique position under the hypothalamus.

Our concept of pituitary function and its role in regulating other endocrine glands has changed. Rather than being viewed as the master gland, it is more appropriately recognized as a transponder that translates neural input into hormonal products. Features that distinguish the function of the pituitary include feedback loops, **pulsatile secretions**, **diurnal variations**, and environmental or external modification of the gland's performance. These characteristics of pituitary operation can distort the clinical evaluation of suspected endocrine disease or, alternatively, lend incredible insight into subtle defects in endocrinologic function.

Embryology and Anatomy

The three distinct parts of the pituitary are the **anterior pituitary**, or *adenohypophysis*; the intermediate lobe, or pars intermedialis; and the **posterior pituitary**, or **neurohypophysis**. The intermediate lobe is poorly developed in humans and has little functional capacity other than to confuse health care providers by forming nonfunctional, benign, cystic enlargements of the pituitary. The posterior pituitary, which arises from the diencephalon, is responsible for the storage and release of **oxytocin and arginine vasopressin (AVP)**, previously called *antidiuretic hormone* (ADH). The anterior pituitary, the largest portion of the gland, originates from the Rathke pouch, an evagination of buccal ectoderm that progressively extends upward and is eventually enveloped by the sphenoid bone. The creation of the median eminence, the inferior portion of the hypothalamus, and the pituitary stalk is the other critical event in the formation of the hypothalamic–hypophyseal unit. Pituitary function can be detected between the seventh and ninth weeks of gestation. The ultimate determination of anterior pituitary cell types is dependent the formation of lactotrophs (prolactin-secreting cells), somatotrophs (growth hormone [GH]-secreting cells), thyrotrophs (thyroid-stimulating hormone [TSH]-secreting cells), corticotrophs (adrenocorticotropin hormone

CASE STUDY 15.1, PART 1

Gustav, a 48-year-old man, is experiencing muscle weakness, frequent headaches, and excessive sweating. He admits to poorly controlled hypertension and has noticed a gradual increase in both his glove and his shoe size, as well as a reduction in libido.

CASE STUDY 15.2, PART 1

Sarah is a 26-year-old female who presents 5 weeks postpartum with seizure, inability to breastfeed, and constantly feeling cold after a challenging delivery that required 8 units of packed red blood cells (PRBCs).

[ACTH]-secreting cells), and gonadotrophs (luteinizing hormone [LH], and follicle-stimulating hormone [FSH]-secreting cells).[1]

The pituitary resides in a pocket of the sphenoid (the sella turcica, meaning "Turkish saddle") and is surrounded by dura mater. The reflection of dura that separates the superior portion of the pituitary from the hypothalamus, the diaphragma sella, is penetrated by the infundibulum, or pituitary stalk, that connects the adenohypophysis to the median eminence and **hypothalamus**. The pituitary stalk contains both neural and vascular structures that terminate in the hypophysis. The posterior pituitary is connected to the supraoptic and paraventricular hypothalamic nuclei (where AVP and oxytocin are produced) by way of two distinct neurosecretory tracts, the supraopticohypophyseal and tuberohypophyseal tracts, which pass through the stalk. The anterior pituitary receives 80% to 90% of its blood supply and many hypothalamic factors via the **hypothalamic–hypophyseal portal system**, also contained in the stalk. The primary plexus of this portal system is located in the median eminence and is composed of capillaries lacking a blood–brain barrier (fenestrated capillaries), where the hypothalamic nuclei that modulate pituitary function terminate their axons. In turn, these long portal vessels connect the primary plexus to the anterior pituitary and serve as a conduit for the hypothalamic–hypophysiotropic hormones. These anatomic relationships are illustrated in **Figure 15.1**.

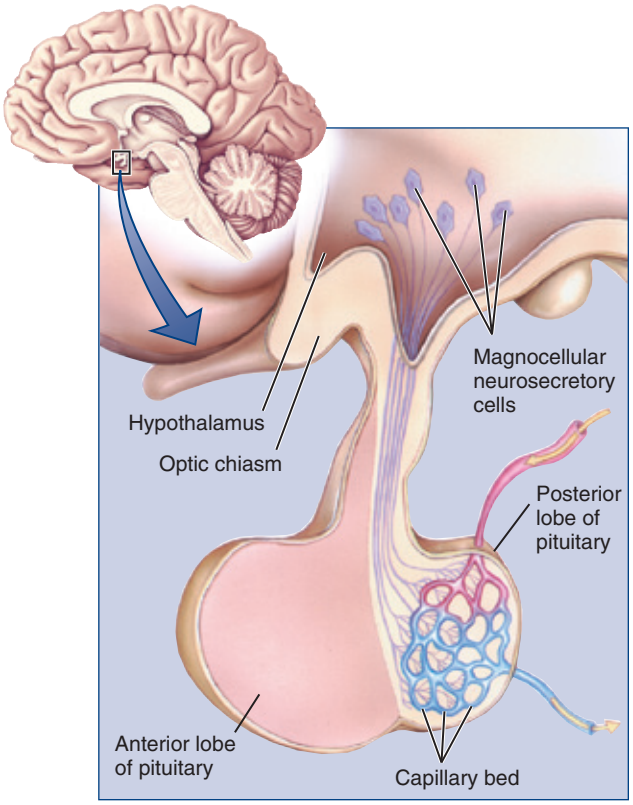


Figure 15.1 Relational anatomy of the pituitary and hypothalamus.

Reproduced with permission from Bear MF, Connors BW, Paradiso MA. Neuroscience: *Exploring the Brain.* 2nd ed. Baltimore, MD: Lippincott Williams & Wilkins; 2001:501

Functional Aspects of the Hypothalamic–Hypophyseal Unit

Afferent pathways (inputs) to the hypothalamus are integrated in various specialized nuclei, processed, and then resolved into specific responses. Because the hypothalamus has many efferent neural connections (outputs) to higher brain centers (such as the limbic system, the autonomic nervous system, and the pituitary), these responses are similar for each specific pituitary hormone and characterized by negative feedback mechanisms, pulsatility, and diurnal variation (see Chapter 13, *Basic Endocrinology*).

An example of an endocrine feedback loop is the hypothalamic–pituitary–thyroidal axis. The hypothalamus produces the hypophysiotropic hormone, **thyrotropin releasing hormone (TRH)**, and releases it into the portal system where it directs the thyrotrophs (or TSH-producing cells) in the anterior pituitary to secrete **thyroid-stimulating hormone (TSH)**. TSH circulates to the thyroid and stimulates several steps in the thyroid that are critical in the production and release of thyroid hormones (T_3 and T_4). Thyroxine is released in the blood and circulates to the hypothalamus and pituitary to suppress further TRH and TSH production. This axis can also be partially inhibited by adrenal steroids (glucocorticoids) and by cytokines; as a result, thyroid hormone production may decline during periods of severe physiologic stress.[2] The feedback of thyroxine at the level of the pituitary is called a *short feedback loop*, and feedback at the level of the hypothalamus is called a *long feedback loop*. Feedback between the pituitary and hypothalamus (when present) is called an *ultrashort feedback loop*. **Figure 15.2** illustrates this simple feedback loop.

All anterior pituitary hormones are secreted in a pulsatile fashion. The pulse frequency of secretion is generally regulated by neural modulation and is specific for each hypothalamic–pituitary–end-organ unit. Perhaps the best example of pituitary pulsatility is the secretion of the hormones that regulate

gonadal function (**luteinizing hormone [LH]** and **follicle-stimulating hormone [FSH]**). In normal male subjects, the median interpulse interval for LH is 55 minutes, and the average LH peak duration is 40 minutes.[3] The pulse frequency of the regulatory hypothalamic hormone, **gonadotropin-releasing hormone (GnRH)**, has profound effects on LH secretion profiles: increasing the frequency of GnRH pulses reduces the gonadotrope secretory response, while decreasing the GnRH pulse frequency increases the amplitude of the subsequent LH pulse.[4]

Another feature of the hypothalamic–pituitary unit is the cyclic nature of hormone secretion. The nervous system usually regulates this function through external signals, such as light–dark changes or the ratio of daylight to darkness. The term *zeitgeber* ("time giver") refers to the process of entraining or synchronizing these external cues into the function of internal biologic clocks. As a result, many pituitary hormones are secreted in different amounts, depending on the time of day. These circadian, or **diurnal variations**, are typified by **adrenocorticotropic hormone (ACTH)**, or TSH secretion. With ACTH, the nadir (trough) of secretion is between 11:00 pm and 3:00 am, and the peak occurs on awakening or around 6:00 to 9:00 am.[5] This circadian rhythm of ACTH is a result of variations in pulse amplitude and not alterations in pulse frequency.[6] The nocturnal levels of TSH are approximately twice the daytime levels; the nocturnal rise in TSH is a result of increased pulse amplitude.[7]

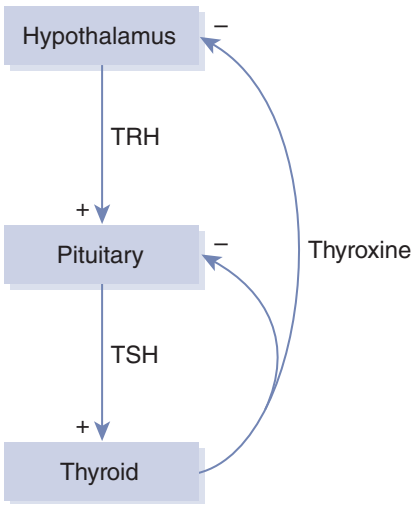


Figure 15.2 Simple feedback loop. TRH, thyrotropin-releasing hormone; TSH, thyroid-stimulating hormone.

Pituitary Tumors

According to autopsy studies, up to 20% of people harbor clinically silent pituitary adenomas, and findings consistent with pituitary tumors are observed in 10% to 30% of normal individuals undergoing MRI examinations. In addition, pituitary tumors account for 91% of the lesions removed from carefully selected patients who have undergone transsphenoidal surgery.[8,9] Close medical follow-up is recommended if an incidentally discovered lesion is hormonally silent and is less than 1 cm in diameter. In terms of frequency, prolactin-secreting pituitary tumors are the most common, followed by nonfunctioning or null cell tumors, while tumors that secrete GH, gonadotropins, ACTH, or TSH account for the remainder. The WHO defines "atypical pituitary tumors" as tumors that have an MIB-1 proliferative index greater than 3%, excessive *p53* immunoreactivity, and increased mitotic activity.[10] MIB-1 is a monoclonal antibody used to detect the Ki-67 antigen, a marker of cell proliferation, and a high "proliferation index" suggests a higher degree of atypia.[11] Most of these "atypical" tumors are macroadenomas (i.e., >1 cm in diameter) and show invasion into surrounding structures like the cavernous sinuses. They may or may not be hormonally active (i.e., produce hormones) and commonly stain for GH or ACTH, although they may not produce clinically evident syndromes.[10]

Physiologic enlargement of the pituitary can be seen during puberty and pregnancy. The enlargement seen during pregnancy is due to lactotroph hyperplasia. Thyrotroph and lactotroph or gonadotroph hyperplasia can also be seen in longstanding primary thyroidal or gonadal failure, respectively.

Hypophysiotropic or Hypothalamic Hormones

The hypothalamus produces many different products; however, only those that have a direct effect on classic pituitary function will be discussed in this chapter. Most products are peptides; however, bioactive amines are also synthesized and transported from the hypothalamus. Hypothalamic hormones may have multiple actions. For example, TRH stimulates the secretion of both TSH and **prolactin**; GnRH stimulates both LH and FSH production; and **somatostatin (SS)** inhibits GH and TSH release from the pituitary. In addition to its effects on water metabolism, AVP can also stimulate ACTH secretion. The

Table 15.1 Hypophysiotropic Hormones

Hormone	Structure	Action
TRH	3 amino acids	Releases TSH and prolactin
GnRH	10 amino acids	Releases LH and FSH
CRH	41 amino acids	Releases ACTH
GHRH	44 amino acids	Releases GH
Somatostatin	14 and 28 amino acids	Inhibits GH and TSH release
Dopamine (prolactin inhibitory factor)	1 amino acid	Inhibits prolactin release

ACTH, adrenocorticotropin hormone; CRH, corticotropin-releasing hormone; FSH, follicle-stimulating hormone; GHRH, growth hormone–releasing hormone; GnRH, gonadotropin-releasing hormone; LH, luteinizing hormone; TRH, thyrotropin-releasing hormone; TSH, thyroid-stimulating hormone.

main stimulus for ACTH secretion is **corticotropin-releasing hormone (CRH)**. These hypophysiotropic hormones are found throughout the central nervous system and various other tissues, including the gut, pancreas, and other endocrine glands. Their function outside the hypothalamus and pituitary is poorly understood. The action of hypophysiotropic hormones on anterior pituitary function is summarized in **Table 15.1**.

Anterior Pituitary Hormones

The hormones secreted from the anterior pituitary are larger and more complex than those synthesized in the hypothalamus. These pituitary hormones are either tropic, meaning their actions are specific for another endocrine gland, or they are direct effectors, because they act directly on peripheral tissue. TSH and its unique role in regulating thyroid function provides an example of a **tropic hormone**, and an example of a direct effector is GH. GH has direct effects on substrate metabolism in numerous tissues and stimulates the liver to produce growth factors that are critical in enhancing linear growth. The tropic hormones are LH, which directs ovulation in women and testosterone production from Leydig cells in men; FSH, which is responsible for ovarian recruitment and early folliculogenesis in women and spermatogenesis in men; TSH, which directs thyroid hormone production from the thyroid; and ACTH, which regulates adrenal steroidogenesis. Both GH and prolactin are direct effectors. A general summary of relationships among anterior pituitary hormones and their target organs and feedback effectors is provided in **Table 15.2**.

The actions of the tropic hormones are discussed in more detail in chapters devoted to their specific

Table 15.2 Anterior Pituitary Hormones

Pituitary Hormone	Target Gland	Structure	Feedback Hormone
LH	Gonad (tropic)	Dimeric glycoprotein	Sex steroids (E_2/T)
FSH	Gonad (tropic)	Dimeric glycoprotein	Inhibin
TSH	Thyroid (tropic)	Dimeric glycoprotein	Thyroid hormones (T_3/T_4)
ACTH	Adrenal (tropic)	Single peptide derived from POMC	Cortisol
Growth hormone	Multiple (direct effector)	Single peptide	IGF-1
Prolactin	Breast (direct effector)	Single peptide	Unknown

ACTH, adrenocorticotropin hormone; E_2, estradiol; FSH, follicle-stimulating hormone; IGF-1, insulin-like growth factor; LH, luteinizing hormone; POMC, pro-opiomelanocortin; T_3, triiodothyronine; T_4, thyroxine; TSH, thyroid-stimulating hormone; T, testosterone.

target glands (see Chapter 14, *Thyroid Function*, Chapter 16, *Adrenal Function*, and Chapter 17, *Gonadal Function*).

Growth Hormone

The pituitary is vital for normal growth. Growth ceases if the pituitary is removed, and if the hormonal products from other endocrine glands that are acted on by the anterior pituitary are replaced (thyroxine, adrenal steroids, and gonadal steroids), growth is not restored until **growth hormone (GH)** is administered. However, if GH is given in isolation without the other hormones, growth is not promoted. In other words, it takes complete functioning of the anterior pituitary and its hormones to establish conditions ripe for growth of the individual. It also takes adequate nutrition, normal levels of insulin, and overall good health to achieve a person's genetic growth potential.

GH, also called *somatotropin*, is structurally related to prolactin and human placental lactogen, a placental hormone responsible for metabolism and growth effects. A single peptide with two intramolecular disulfide bridges, it belongs to the direct effector class of anterior pituitary hormones. The somatotrophs, pituitary cells that produce GH, constitute over one-third of normal pituitary weight. Release of somatotropin from the pituitary is stimulated by the hypothalamic peptide **growth hormone–releasing hormone (GHRH)**; somatotropin's secretion is inhibited by SS.[12] GH is secreted in pulses, with an average interpulse interval of 2 to 3 hours, with the most reproducible peak occurring at the onset of sleep.[13] Between these pulses, the level of GH may fall below the detectable limit, resulting in the clinical evaluation of GH deficiency if based on a single measurement. Ghrelin, an enteric hormone that plays important roles in nutrient sensing, appetite and glucose regulation, is also a potent stimulator of GH secretion.[14] The on-and-off functions of GHRH/SS and the basic pattern of secretory pulses of GH are heavily modulated by many other factors (**Table 15.3**).

Table 15.3 Other Modifiers of Growth Hormone Secretion

Stimulate Growth Hormone Secretion	Inhibit Growth Hormone Secretion
Sleep	Glucose loading
Exercise	β-Agonists (e.g., epinephrine)
Physiologic stress	α-Blockers (e.g., phentolamine)
Amino acids (e.g., arginine)	Emotional/psychogenic stress
Hypoglycemia	Nutritional deficiencies
Sex steroids (e.g., estradiol)	Insulin deficiency/ hyperglycemia
α-Agonists (e.g., norepinephrine)	Thyroxine deficiency
β-Blockers (e.g., propranolol)	

Actions of GH

GH has many diverse effects on metabolism; it is considered an amphibolic hormone because it directly influences both anabolic and catabolic processes. One major effect of GH is that it allows an individual to effectively transition from a fed state to a fasting state without experiencing a shortage of substrates required for normal intracellular oxidation. GH directly antagonizes the effect of insulin on glucose metabolism, promotes hepatic gluconeogenesis, and stimulates lipolysis.[15,16] Enhanced lipolysis provides oxidative substrates for peripheral tissue, such as skeletal muscle, and yet conserves glucose for the central nervous system by stimulating the hepatic delivery of glucose and opposing insulin-mediated glucose catabolism. Isolated GH deficiency in children may be accompanied by hypoglycemia; however, hypoglycemia is more likely to occur if both GH and ACTH are deficient.[17]

The anabolic effects of GH are reflected by enhanced protein synthesis in skeletal muscle and other tissues. This is translated into a positive nitrogen balance and phosphate retention.

Although GH has direct effects on many tissues, it also has indirect effects that are mediated by factors that were initially called *somatomedins*. In early experiments, it became apparent that GH supplementation in hypophysectomized (pituitary removed) animals induced the production of an additional protein that stimulated the incorporation of sulfate into cartilage.[18,19] As this protein was purified, it was evident that more than one somatomedin existed, and because of its structural homology to proinsulin, the nomenclature shifted to **insulin-like growth factor (IGF)**.[20,21] For example, somatomedin C, the major growth factor induced by GH, is now IGF-1.[22] IGFs also have cell surface receptors that are distinct from insulin; however,

increased levels of IGF-2 can cross-react with the insulin receptor and cause hypoglycemia,[23] and hyperinsulinemia can partially activate IGF-1 receptors.[24] GH stimulates the production of IGF-1 from the liver, and as a result, IGF-1 becomes a biologic amplifier of GH levels. IGFs are complexed to specific serum binding proteins that have been shown to affect the actions of IGFs in multifaceted ways.[25] IGF-binding protein 3 (IGFBP-3) is perhaps the best-studied member of the IGFBP family. Recently, IGFBPs and specifically IGFBP-3, have been shown to directly play a role in the pathophysiology of several human cancers; this may be independent of IGFs and their receptor–mediated pathways.[26] The tumor suppressor gene, *p53*, has been shown to upregulate active IGFBP-3 secretion, which in turn inhibits IGF-1–signaled mutagenesis and inhibits neoplastic cell proliferation.[27] Low levels of IGFBP-3 with increased IGF-1 were associated with increased risk of cancers, including prostate, colorectal, lung, and premenopausal breast cancers.[26,28]

Testing

Growth Hormone

As noted above, a single, random measurement of GH is rarely diagnostic. Due to the fact that growth hormone levels in normal and disease populations overlap, growth hormone suppression and stimulation tests are more accurate in assessing conditions of excess and deficiency. Baseline measurements are helpful, for instance, when interpreted in conjunction with results from glucose suppression tests for the diagnosis of acromegaly or in conjunction with results from the GH stimulation test for the diagnosis of GH deficiency. Both the suppression test and the stimulation test measure GH levels at timed intervals. The GH reference range is 0.01–1.0 ng/mL for males and 0.01–3.5 ng/mL for females. GH testing has limited value in assessing GH secretion in normal children; therefore, IGF-1 is recommended as the first test for assessing deficient or excess growth during childhood.

Insulin-like Growth Factor 1 (IGF-1)

The IGF-1 test is preferred for assessing GH deficiency or excess during childhood and/or adolescent development. It is also useful in the evaluation of excess or deficient growth disorders in both adults and children. IGF-1 is also used for monitoring recombinant human growth hormone treatment and for treatment follow-up of individuals with acromegaly and gigantism. IGFBP has a lower diagnostic sensitivity and specificity, so this test adds little value unless ordered in combination with the IGF-1. Persons with malnutrition have low IGF-1 levels, so IGF-1 is a more sensitive indicator than prealbumin for monitoring nutritional status.

The current testing paradigms for GH are soundly based on the dynamic physiology of the GH axis. For example, circulating levels of IGF-1, and perhaps IGFBP-3, reasonably integrate the peaks of GH secretion. Elevated levels of both are consistent with, but may not be diagnostic of, a sustained excess of GH. Other conditions, such as a hepatoma, can be associated with high levels of IGF-1, and levels of IGFBP-3 may be inappropriately normal in some patients with active acromegaly. Conversely, low IGF-1 levels may reflect inadequate production of GH, and low IGF levels are also seen in patients with poorly controlled diabetes, malnutrition, or other chronic illnesses.[29] The inherent high biologic variability and assay performance issues further confound the use of IGF-1 measurements in the clinical setting.[30] Recently, new recommendations to improve the assay performance for both IGF-1 and GH have been published.[31] Definitive testing for determining the autonomous production of GH relies upon the normal suppressibility of GH by oral glucose loading.[22,32,33] This test is performed after an overnight fast, and the patient is given a 75-g oral glucose load. GH is measured at time zero (the time when the glucose is consumed), 30, 60, 90 and 120 minutes after glucose ingestion. Following oral glucose loading, GH levels are undetectable in normal individuals; however, in patients with acromegaly (discussed below), GH levels fail to suppress and paradoxically may even rise.

Testing patients for suspected GH deficiency is more complicated. There are several strategies to stimulate GH, and new protocols are currently evolving. Once considered the gold standard, insulin-induced hypoglycemia is being replaced by less uncomfortable testing schemes.[34] Combination infusions of GHRH and the amino acid L-arginine, or an infusion of L-arginine coupled with oral levodopa (L-DOPA) are the most widely used and are better tolerated by the patient.[34] If GH levels rise above 3–5 ng/mL, it is unlikely that the patient is GH deficient;[34] however, a lower threshold may be adopted because of improved sensitivity of the newer, two-site GH assays.[35] On the other hand, several studies have shown that provocative GH testing may not be necessary in patients with low IGF-1 levels and otherwise documented panhypopituitarism.[29]

GH Excess

Acromegaly results from pathologic or autonomous GH excess and, in the majority of patients, is a result of a pituitary tumor. There have been isolated case reports of tumors causing acromegaly as a result of the ectopic production of GHRH; however, the ectopic production of GHRH or GH remains rare.[36-38] Recent reports have documented mutations in the aryl hydrocarbon–interacting protein gene (AIP)[39] in cases of familial acromegaly and polymorphisms in the SS receptor type 5 gene in rare cases.[40] If a GH-producing tumor occurs before epiphyseal closure of the long bones, the patient develops gigantism,[41] and the child may grow to an impressive height. If the GH-producing tumor happens after puberty, the patient develops classical, but insidious, features of acromegaly such as bony and soft-tissue overgrowth.[42] These features include progressive enlargement of the hands and feet as well as growth of facial bones, including the mandible and bones of the skull. In advanced cases, the patient may develop significant gaps between their teeth. Diffuse (not longitudinal if the condition occurred following puberty) overgrowth of the ends of long bones or the spine can produce a debilitating form of arthritis.[43] Because GH is an insulin antagonist, glucose intolerance or overt diabetes can occur. Hypertension, accelerated atherosclerosis, and proximal muscle weakness, resulting from acquired myopathy,[44] may be seen late in the illness. Sleep apnea is common. Organomegaly, especially thyromegaly, is common, but hyperthyroidism is exceedingly rare unless the tumor cosecretes TSH. GH excess is also a hypermetabolic condition, and as a result, acromegalic patients may complain of excessive sweating or heat intolerance. The features of acromegaly develop slowly over time, and changes in facial features may go unrecognized by the patient and/or their family. In these cases, the patient's complaints may center on the local effects of the tumor (headache or visual complaints) or symptoms related to the loss of other anterior pituitary hormones (hypopituitarism). A careful, retrospective review of older photographs may be crucial in differentiating coarse features due to inheritance from the classical consequences of acromegaly. If left untreated, acromegaly shortens life expectancy because of increased risk of heart disease, resulting from the combination of hypertension, coronary artery disease, and diabetes/insulin resistance. Because patients with acromegaly also have a greater lifetime risk of developing cancer, cancer surveillance programs (especially regular colonoscopy) are recommended.[22]

Cosecretion of prolactin can be seen in up to 40% of patients with acromegaly.[45] Only a few TSH/GH-secreting tumors have also been reported.[46]

Confirming the diagnosis of acromegaly is relatively easy; however, some patients with acromegaly have normal random levels of GH. An elevated level of GH that does not suppress normally with glucose loading equates to a straightforward diagnosis. In those patients with normal, but inappropriately sustained, random levels of GH, elevated levels of IGF-1 are helpful; however, nonsuppressibility of GH with glucose loading is the definitive test.[32,33]

While testing for acromegaly is usually straightforward, treatment can be challenging. The goal of treatment is tumor ablation, with continued function of the remainder of the pituitary. Transsphenoidal adenomectomy is the procedure of choice.[22] If normal GH levels and kinetics (normal suppressibility to glucose) are restored following surgery, the patient is likely cured. Unfortunately, GH-producing

CASE STUDY 15.1, PART 2

Remember Gustav. He decides to see his healthcare provider for his symptoms. A review of older photographs of the man documents coarsening of facial features, progressive prognathism, and broadening of the nose. His healthcare provider suspects that he has acromegaly.

1. What screening tests are available?
2. What is the definitive test for autonomous growth hormone production?

tumors may be too large or may have invaded into local structures, which precludes complete surgical removal, and the patient is left with a smaller, but hormonally active, tumor. External beam or focused irradiation is frequently used at this point, but it may take several years before GH levels decline.[47,48] In the interim, efforts are made to suppress GH. Three different classes of agents, SS analogs (octreotide, pasireotide, and lanreotide), dopaminergic agonists (cabergoline and bromocriptine), and GH receptor antagonists (pegvisomant) may be employed for GH suppression.[49,50]

GH Deficiency

GH deficiency occurs in both children and adults. In children, it may be genetic or due to tumors, such as craniopharyngiomas. In adults, it is a result of structural or functional abnormalities of the pituitary (see the section on *hypopituitarism* in this chapter); however, a decline in GH production is an inevitable consequence of aging, and the significance of this phenomenon is poorly understood.[51,52]

Although GH deficiency in children is manifested by growth failure, not all patients with short stature have GH deficiency.[53] There have been several genetic defects identified in the GH axis, and the most common type is a recessive mutation in the GHRH gene that causes a failure of GH secretion. A rarer mutation, loss of the GH gene itself, has also been observed. Mutations that result in GH insensitivity have also been reported. These mutations may involve the GH receptor, IGF-1 biosynthesis, IGF-1 receptors, or defects in GH signal transduction. As a result, patients with GH insensitivity do not respond normally to exogenously administered GH. Finally, structural lesions of the pituitary or hypothalamus may also cause GH deficiency and may be associated with other anterior pituitary hormone deficiencies.

An adult GH deficiency syndrome has been described in patients who have complete or even partial failure of the anterior pituitary. The symptoms of this syndrome are extremely vague and include social withdrawal, fatigue, loss of motivation, and a diminished feeling of well-being,[54] but several studies have documented increased mortality in children who are GH deficient, although this relationship is less clear in adults.[55] Osteoporosis and alterations in body composition (i.e., reduced lean body mass) are frequent concomitants of adult GH deficiency.[56]

GH replacement therapy has become relatively simple with the advent of recombinant human GH.[57] Currently, the cost of GH is the major limiting factor for replacement. GH has been employed by athletes as a performance-enhancing substance, as an aid in injury recovery, and to combat the effects of aging; however, the effectiveness of GH for these purposes is controversial.[58]

Prolactin

Prolactin is structurally related to GH and human placental lactogen. Considered a stress hormone, it has vital functions in relationship to reproduction. Prolactin is classified as a direct effector hormone (as opposed to a tropic hormone) because it effects diffuse target tissues and lacks a single endocrine end organ.

Prolactin is unique among the anterior pituitary hormones because its major mode of hypothalamic regulation is tonic inhibition rather than intermittent stimulation. **Prolactin inhibitory factor (PIF)** was once considered a polypeptide hormone capable of inhibiting prolactin secretion; **dopamine**, however, is the only neuroendocrine signal that inhibits prolactin and is now considered to be the elusive PIF. Any compound that affects dopaminergic activity in the hypothalamus will also alter prolactin secretion. Examples of medications that are known to cause hyperprolactinemia include phenothiazines, butyrophenones, metoclopramide, reserpine, tricyclic antidepressants, α-methyldopa, and antipsychotics that antagonize the dopamine D2 receptor. Any disruption of the pituitary stalk (e.g., tumors, trauma, or inflammation) causes an elevation in prolactin as a result of interruption of the flow of dopamine from the hypothalamus to the lactotrophs, the pituitary prolactin-secreting cells. TRH directly stimulates prolactin secretion, and increases in TRH (as seen in primary hypothyroidism) elevate prolactin levels.[59] Estrogens also directly stimulate lactotrophs to synthesize prolactin. Pathologic stimulation of the neural suckling reflex is the likely explanation of hyperprolactinemia associated with chest wall injuries. Hyperprolactinemia may also be seen in renal failure and polycystic ovary syndrome. Physiologic stressors, such as exercise and seizures, also elevate prolactin. The feedback effector for prolactin is unknown. Although the primary regulation of prolactin secretions is tonic inhibition (e.g., dopamine), it is also regulated by several hormones, including GnRH, TRH, and vasoactive intestinal polypeptide. Stimulation of breasts, as in nursing, causes the release of prolactin-secreting hormones from the hypothalamus through a spinal reflex arc.

As mentioned, the physiologic effect of prolactin is lactation. The usual consequence of prolactin excess is hypogonadism, either by suppression of gonadotropin secretion from the pituitary or by inhibition of gonadotropin action at the gonad.[60] The suppression of ovulation seen in lactating postpartum mothers is related to this phenomenon.

Prolactinoma

A prolactinoma is a pituitary tumor that directly secretes prolactin, and it represents the most common type of functional pituitary tumor. The clinical presentation of a patient with a prolactinoma depends on the age and gender of the patient and the size of the tumor. Premenopausal women most frequently complain of menstrual irregularity/amenorrhea, infertility, or galactorrhea, while men and postmenopausal women generally present with symptoms of a pituitary mass, such as headaches or visual complaints. Occasionally, males may present with reduced libido or complaints of erectile dysfunction. The reason(s) for the varied presentations of a prolactinoma are somewhat obscure but likely relate to the dramatic, noticeable alteration in menses or the abrupt onset of a breast discharge in younger women. By contrast, the decline in reproductive function in older patients may be overlooked as an inexorable consequence of "aging." One recently recognized complication of prolactin-induced hypogonadism is osteoporosis.[61]

Other Causes of Hyperprolactinemia

There are many physiologic, pharmacologic, and pathologic causes of hyperprolactinemia, and a common error by clinicians is to ascribe any elevation in prolactin to a "prolactinoma." Generally, substantial elevations in prolactin (>150 ng/mL) indicate prolactinoma, and the degree of elevation in prolactin is correlated with tumor size.[62] Modest elevations in prolactin (25 to 100 ng/mL) may be seen with pituitary stalk interruption, use of dopaminergic antagonist medications, or other medical conditions such as primary thyroidal failure, renal failure, or polycystic ovary syndrome. Breast or genital stimulation may also modestly elevate prolactin. Significant hyperprolactinemia is also encountered during pregnancy. Under most circumstances, the principal form of prolactin is a 23-kD peptide; however, a 150-kD form may also be secreted. This larger prolactin molecule has a markedly reduced biologic potency and does not share the reproductive consequences of the 23-kD variety. If the 150-kD form of prolactin predominates, this is called macroprolactinemia; the clinical consequences are unclear, but most patients are relatively asymptomatic.[62] The prevalence of macroprolactinemia has been estimated at 10% to 22% of hyperprolactinemic samples[63] and can be excluded by precipitating serum samples with polyethylene glycol prior to measuring prolactin.

Clinical Evaluation of Hyperprolactinemia

A careful history and physical examination are usually sufficient to exclude most common, nonendocrine causes of hyperprolactinemia. It is essential to obtain TSH and free T_4 to eliminate primary hypothyroidism as a cause for the elevated prolactin. If a pituitary tumor is suspected, a careful assessment of other anterior pituitary function (i.e., ACTH/cortisol, LH, FSH, and gender-specific gonadal steroids) and an evaluation of pituitary anatomy with a high-resolution MRI should be obtained. The reference range for prolactin in males is 4.0–15.2 ng/mL and in females is 4.8–23.3 ng/mL.

Management of Prolactinoma

The therapeutic goals are correction of symptoms that result from local invasion or extension of the tumor by reducing tumor mass, restoration of normal gonadal function and fertility, prevention of osteoporosis, and preservation of normal anterior and posterior pituitary function. The different therapeutic options include simple observation, surgery, radiotherapy, or medical management with dopamine agonists.[64] However, the management of prolactinoma also depends on the size of the tumor; macroadenomas with tumor size >10 mm are less likely to be "cured" than are microadenomas with tumor size < 10 mm.[65]

Oral medication using drugs known as dopamine agonists are commonly used to treat prolactinomas. Tumor shrinkage is noted in more than 90% of patients treated with either bromocriptine mesylate (Parlodel) or cabergoline (Dostinex) dopamine receptor agonists. Both drugs also shrink prolactin-secreting macroadenomas.[64] A resumption of menses and restoration of fertility is also frequently seen during medical therapy. The adverse effects of bromocriptine include orthostatic hypotension, dizziness, and nausea. The gastrointestinal adverse effects of bromocriptine can be ameliorated through intravaginal administration,

and its efficacy is otherwise uncompromised. Cabergoline has fewer adverse effects and interacts with a serotonergic receptor, although has been linked to the development of valvular heart disease.[66] Either agent should be discontinued during pregnancy unless tumor regrowth has been documented.

Surgery to excise the tumor is not a primary mode of prolactinoma management. The indications for neurosurgical intervention include pituitary tumor apoplexy (hemorrhage), acute visual loss due to macroadenoma, cystic prolactinoma, intolerance to medication, or tumor resistance to dopaminergic agonists. Surgical cure rates are inversely proportional to tumor size and the degree of prolactin elevation. Radiation therapy is generally reserved for high-surgical-risk patients with locally aggressive macroadenomas who are unable to tolerate dopamine agonists.

Idiopathic Galactorrhea

Lactation occurring in women with normal prolactin levels is defined as *idiopathic galactorrhea*. This condition is usually seen in women who have been pregnant several times and has no pathologic implication; it may be a manifestation of a localized increased sensitivity to prolactin in breast tissue and is a diagnosis of exclusion.

Hypopituitarism

The failure of either the pituitary or the hypothalamus results in the loss of anterior pituitary function. Complete loss of function is termed *panhypopituitarism*. There also may be a loss of only a single pituitary hormone, which is referred to as a *monotropic hormone deficiency*. The loss of a tropic hormone (ACTH, TSH, LH, and FSH) is reflected in functional cessation of the affected endocrine gland. Loss of the direct effector hormones (GH and prolactin) may not be readily apparent. This section concentrates on the causes of hypopituitarism and certain subtleties involved in the therapy of panhypopituitarism.

The laboratory diagnosis of hypopituitarism is relatively straightforward. In contrast to the primary failure of an endocrine gland that is accompanied by dramatic increases in circulating levels of the corresponding pituitary tropic hormone, secondary failure (hypopituitarism) is associated with low or normal levels of tropic hormone. In primary hypothyroidism, for example, the circulating levels of thyroxine are low and TSH levels are extremely high. As a result of pituitary failure in hypothyroidism, TSH levels are inappropriately low and associated with low free thyroxine levels. See Chapter 13, *Basic Endocrinology*, for more information on primary and secondary hormone deficiency.

Etiology of Hypopituitarism

The many causes of hypopituitarism are listed in **Table 15.4**. Direct effects of pituitary tumors, or the sequelae of treatment of tumors, are the most common causes of pituitary failure. Pituitary tumors may cause panhypopituitarism by compressing or replacing normal tissue or interrupting the flow of hypothalamic hormones by destroying the pituitary stalk. Large, nonsecretory pituitary tumors

CASE STUDY 15.1, PART 3

Remember Gustav. On a follow-up visit with his healthcare provider, Gustav expresses concern regarding his decreased libido. The provider decides to order prolactin levels.

© Juanmonino/E+/Getty.

3. What would you expect the prolactin levels to be?
4. Are prolactin abnormalities common in individuals with acromegaly?

CASE STUDY 15.2, PART 2

Remember Sarah. She sees her physician for her 5-week follow up. The physician orders hormone testing to help understand her symptoms. The results are as follows:

TSH: 0.1 mIU/L (Reference Range: 0.3–4.2 mIU/)
ACTH: 7.3 pg/mL (Reference Range: 7.2–63 pg/mL at 8:00 am)
Prolactin: 3 ng/mL (Reference Range: 4.8–23.3 ng/mL)

© mimagephotography/Shutterstock.

1. What do these results indicate about Sarah's overall pituitary function?

Table 15.4 Causes of Hypopituitarism

1. Pituitary tumors
2. Parapituitary/hypothalamic tumors
3. Trauma
4. Radiation therapy/surgery
5. Infarction
6. Infection
7. Infiltrative disease
8. Immunologic
9. Familial
10. Idiopathic

(chromophobe adenomas or null cell tumors) or macroprolactinomas are most commonly associated with this phenomenon. Parasellar tumors (meningiomas and gliomas), metastatic tumors (breast and lung), and hypothalamic tumors (craniopharyngiomas or dysgerminomas) can also cause hypopituitarism through similar mechanisms. Hemorrhage into a pituitary tumor is rare; however, when it occurs, it frequently causes complete pituitary failure.[67] Postpartum ischemic necrosis of the pituitary following a complicated delivery (Sheehan syndrome) typically presents as profound, unresponsive shock or as failure to lactate after childbirth. Infiltrative diseases, such as hemochromatosis, sarcoidosis, or histiocytosis, can also affect pituitary function. Fungal infections, tuberculosis, and syphilis can involve the pituitary or hypothalamus and may cause impairment of function. Lymphocytic hypophysitis,[68] an autoimmune disease of the pituitary, may only affect a single cell type in the pituitary, resulting in a monotropic hormone deficiency, or can involve all cell types, yielding total loss of function. Ipilimumab, a monoclonal antibody used in melanoma patients, has been associated with lymphocytic hypophysitis in up to 5% of treated patients.[69] Severe head trauma may shear the pituitary stalk or may interrupt the portal circulation. Similarly, surgery involving the pituitary may compromise the stalk and/or blood supply to the pituitary or may iatrogenically diminish the mass of functioning pituitary tissue. Panhypopituitarism can result from radiotherapy used to treat a primary pituitary tumor or a pituitary that was inadvertently included in the radiation port; loss of function, however, may be gradual and may occur over several years. There have been rare instances of familial panhypopituitarism[70] or monotropic hormone deficiencies. In Kallmann's syndrome, for example, GnRH is deficient, and the patient presents with secondary hypogonadism. Lastly, there may not be an apparent identified cause for the loss of pituitary function, and the patient is classified as having idiopathic hypopituitarism, although it is always prudent to continue the search for an underlying cause.

Treatment of Panhypopituitarism

In the average patient, replacement therapy for panhypopituitarism is the same as for primary target organ failure. Patients are treated with thyroxine, glucocorticoids, and gender-specific sex steroids. It is less clear about GH replacement in adults, and additional studies are needed to clarify this issue. Replacement becomes more complicated in panhypopituitary patients who desire fertility. Pulsatile GnRH infusions have induced puberty and restored fertility in patients with Kallmann's syndrome,[71] and gonadotropin preparations have restored ovulation/spermatogenesis in people with gonadotropin deficiency.[72]

Posterior Pituitary Hormones

The posterior pituitary is an extension of the forebrain and represents the storage region for arginine vasopressin (AVP) and oxytocin. Both of these small peptide hormones are synthesized in the supraoptic and paraventricular nuclei of the hypothalamus and transported to the neurohypophysis via their axons in the hypothalamoneurohypophyseal tract. This tract transits the median eminence of the hypothalamus and continues into the posterior pituitary through the pituitary stalk. The synthesis of each of these hormones is tightly linked to the production

CASE STUDY 15.2, PART 3

Recall Sarah.

2. What is the difference between hypopituitarism and panhypopituitarism?

3. What is the probable cause of Sarah's condition?

of neurophysin,[73] a larger protein with a poorly understood function. Both hormones are synthesized outside of the hypothalamus in various tissues, and it is plausible they have an autocrine or a paracrine function.

Oxytocin

Oxytocin is a cyclic nonapeptide, with a disulfide bridge connecting amino acid residues 1 and 6. As a posttranslational modification, the C-terminus is amidated. Oxytocin has a critical role in lactation[74] and plays a major role in labor and parturition (childbirth).[75] Oxytocin is also unique because its secretion responds to a positive feedback loop, meaning that circulating levels of oxytocin actually perpetuate further hormone secretion, instead of suppressing further hormone secretion as is the case with most anterior pituitary hormones. In this way, uterine contractions propagate oxytocin release, which causes further uterine contractions, which cause further oxytocin release until parturition occurs. Synthetic oxytocin, Pitocin, is used in obstetrics to induce labor. Recent studies have linked oxytocin to a variety of biosocial behaviors to include maternal nurturing and mother–infant bonding.[76] In addition to its reproductive and prosocial effects, oxytocin has been shown to have effects on pituitary, renal, cardiac, metabolic, and immune function.

Arginine Vasopressin

Structurally similar to oxytocin, AVP is a cyclic nonapeptide with an identical disulfide bridge; it differs from oxytocin by only two amino acids. The major action of AVP (formerly called antidiuretic hormone) is to regulate renal free water excretion and, therefore, has a central role in water balance. The vasopressin receptors in the kidney (V_2) are concentrated in the renal collecting tubules and the ascending limb of the loop of Henle. They are coupled to adenylate cyclase, and once activated, they induce insertion of aquaporin-2, a water channel protein, into the tubular luminal membrane.[77] AVP is also a potent pressor agent and effects blood clotting[78] by promoting factor VII release from hepatocytes and von Willebrand factor release from the endothelium. These vasopressin receptors (V_{1a} and V_{1b}) are coupled to phospholipase C.

Hypothalamic osmoreceptors and vascular baroreceptors regulate the release of vasopressin from the posterior pituitary. The osmoreceptors are extremely sensitive to even small changes in plasma osmolality, with an average osmotic threshold for vasopressin release in humans of 284 mOsm/kg (reference range is 275–295 mOsm/kg). As plasma osmolality increases, vasopressin secretion increases. The consequence is a reduction in renal free water clearance, a lowering of plasma osmolality, and a return to homeostasis. The vascular baroreceptors (located in the left atrium, aortic arch, and carotid arteries) initiate vasopressin release in response to a fall in blood volume or blood pressure. A 5% to 10% fall in arterial blood pressure in normal humans will trigger vasopressin release; however, in contrast to an osmotic stimulus, the vasopressin response to a baroreceptor-induced stimulus is exponential. In fact, baroreceptor-induced vasopressin secretion will override the normal osmotic suppression of vasopressin secretion.

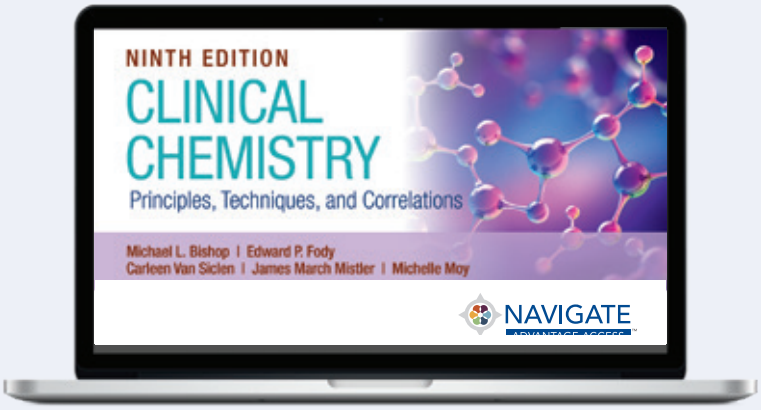
Diabetes insipidus (DI), characterized by copious production of urine (polyuria) and intense thirst (polydipsia), is a consequence of vasopressin deficiency or resistance. However, total vasopressin deficiency is unusual, and the typical patient presents with a partial deficiency. The causes of hypothalamic DI include autoimmunity to vasopressin-secreting neurons, trauma, diseases affecting pituitary stalk function, and various central nervous system or pituitary tumors. A sizable percentage of patients (up to 30%) will have idiopathic DI.[79]

Depending on the degree of vasopressin deficiency, diagnosis of DI can be readily apparent or may require extensive investigation. Documenting an inappropriately low vasopressin level with an elevated plasma osmolality would yield a reasonably secure diagnosis of DI. In less obvious cases, the patient may require a water deprivation test in which fluids are withheld from the patient and serial determinations of serum and urine osmolality are performed in an attempt to document the patient's ability to conserve water. Under selected circumstances, a healthcare provider may simply offer a therapeutic trial of vasopressin or a synthetic analog such as desmopressin (dDAVP) and assess the patient's response. In this circumstance, amelioration of both polyuria and polydipsia would be considered a positive response, and a presumptive diagnosis of DI is made. However, if the patient has primary polydipsia (also known as compulsive water drinking), a profound hypo-osmolar state (water intoxication) can ensue due to the continued ingestion of copious amounts of fluids and a reduced renal excretion of free water. This scenario illustrates the importance of carefully evaluating each patient prior to therapy.

AVP excess may also occur and is much more difficult to treat. Since excess AVP leads to the pathologic retention of free water, restricting free water intake to small amounts each day has been a historical cornerstone of treatment. Recently, conivaptan and tolvaptan,[80] vasopressin V_2 receptor antagonists, have been approved for the management of euvolemic hyponatremia due to vasopressin excess.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Kelberman D, Rizzoti K, Lovell-Badge R, et al. Genetic regulation of pituitary gland development in human and mouse. *Endocr Rev.* 2009;30(7):790–829.
2. Adler SM, Wartofsky L. The nonthyroidal illness syndrome. *Endocrinol Metab Clin North Am.* 2007;36(3):657–672.
3. Veldhuis JD, Keenan DM, Pincus SM. Regulation of complex pulsatile and rhythmic neuroendocrine systems: the male gonadal axis as a prototype. *Prog Brain Res.* 2010;181:79–110.
4. Nunemaker CS, Satin LS. Episodic hormone secretion: a comparison of the basis of pulsatile secretion of insulin and GnRH. *Endocrine.* 2014;47:49–63.
5. Dijk DJ, Duffy JF, Silva EJ, et al. Amplitude reduction and phase shifts of melatonin, cortisol and other circadian rhythms after a gradual advance of sleep and light exposure in humans. *PLoS One.* 2012;7(2):e30037.
6. Spiga F, Waite EJ, Liu Y, et al. ACTH-dependent ultradian rhythm of corticosterone secretion. *Endocrinology.* 2011;152:1448–1457.
7. Tonsfeldt KJ, Chappell PE. Clocks on top: the role of the circadian clock in the hypothalamic and pituitary regulation of endocrine physiology. *Mol Cell Endocrinol.* 2012;349:3–12.
8. Valassi E, Biller BM, Klibanski A, et al. Clinical features of non-pituitary sellar lesions in a large surgical series. *Clin Endocrinol (Oxf).* 2010;73:798–807.
9. Couldwell WT, Altay T, Krisht K, et al. Sellar and parasellar metastatic tumors. *Int J Surg Oncol.* 2012;2012:647256.
10. Zada G, Woodmansee WW, Ramkissoon S, et al. Atypical pituitary adenomas: incidence, clinical characteristics, and implications. *J Neurosurg.* 2011;114:336–344.
11. Marrelli D, Pinto E, Nari A, et al. Mib-1 proliferation index is an independent predictor of lymph node metastasis in invasive breast cancer: a prospective study on 675 patients. *Oncol Rep.* 2006;15(2):425–429.
12. Córdoba-Chacón J, Gahete MD, Castaño JP, et al. Somatostatin and its receptors contribute in a tissue specific manner to the sex-dependent metabolic (fed/fasting) control of growth hormone axis in mice. *Am J Physiol Endocrinol Metab.* 2011;300:E46–E54.
13. Ribeiro-Oliveira A, Barkan AL. Growth hormone pulsatility and its impact on growth and metabolism in humans. *Growth Hormone Relat Dis Ther Contemp Endocrinol.* 2011;(pt 1):33–56.
14. Pradhan G, Samson SL, Sun X. Ghrelin: much more than a hunger hormone. *Curr Opin Clin Nutr Metab Care.* 2013;16:619–624.
15. Chia DJ. Minireview: mechanisms of growth hormone mediated gene regulation. *Mol Endocrinol.* 2014;28:1012–1025.
16. Møller N, Jørgensen JO. Effects of growth hormone on glucose, lipid, and protein metabolism in human subjects. *Endocr Rev.* 2009;30(2):152–177.
17. Cambiaso P, Schiaffini R, Pontrelli G, et al. Nocturnal hypoglycemia in ACTH and GH deficient children: role of continuous glucose monitoring. *Clin Endocrinol (Oxf).* 2013;79:232–237.
18. Salmon WD, Daughaday WH. A hormonally controlled serum factor which stimulates sulfate incorporation by cartilage in vitro. *J Lab Clin Med.* 1957;49:825–826.
19. Daughaday WH, Hall K, Raben MS, et al. Somatomedin: proposed designation for sulphation factor. *Nature.* 1972;235:107.
20. Klapper DG, Svoboda ME, Van Wyk JJ. Sequence analysis of somatomedin-C: confirmation of identity with insulin-like growth factor I. *Endocrinology.* 1983;112:2215–2217.

21. Rinderknecht E, Humbel RE. The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin. *J Biol Chem.* 1978;253:2769–2776.
22. Melmed S, Colao A, Barkan A, et al. Guidelines for acromegaly management: an update. *J Clin Endocrinol Metab.* 2009;94(5):1509–1517.
23. Khowaja A, Johnson-Rabbett B, Bantle J, et al. Hypoglycemia mediated by paraneoplastic production of insulin-like growth factor-2 from a malignant renal solitary fibrous tumor—clinical case and literature review. *BMC Endocr Disord.* 2014;14:49.
24. Li G, Barrett EJ, Ko S-H, et al. Insulin and insulin-like growth factor-I receptors differentially mediate insulin-stimulated adhesion molecule production by endothelial cells. *Endocrinology.* 2009;150:3475–3482.
25. Forbes BE, McCarthy P, Norton RS. Insulin-like growth factor binding proteins: a structural perspective. *Front Endocrinol (Lausanne).* 2012;3:38.
26. Jogie-Brahim S, Feldman D, Oh Y. Unraveling insulin-like growth factor binding protein-3 actions in human disease. *Endocr Rev.* 2009;30(5):417–437.
27. Torng P-L, Lin C-W, Chan MWY, et al. Promoter methylation of IGFBP-3 and p53 expression in ovarian endometrioid carcinoma. *Mol Cancer.* 2009;8:120.
28. Jing M, Pollack MN, Giovannucci E, et al. Prospective study of colorectal cancer risk in men and plasma levels of insulinlike growth factor (IGF)-I and IGF-binding protein-3. *J Natl Cancer Inst.* 1999;91(7):620–625.
29. Kwan AYM, Hartman ML. IGF-I measurements in the diagnosis of adult growth hormone deficiency. *Pituitary.* 2007;10(2):151–157.
30. Bancos I, Algeciras-Schimnich A, Grebe S, et al. Evaluation of variables influencing the measurement of insulin-like growth factor-I. *Endocr Pract.* 2014;20:421–426.
31. Clemmons DR. Consensus statement on the standardization and evaluation of growth hormone and insulin-like growth factor assays. *Clin Chem.* 2011;57:555–559.
32. Ben-Shlomo A, Melmed S. Acromegaly. *Endocrinol Metab Clin North Am.* 2008;37(1):101–122.
33. Melmed S. Medical progress: acromegaly. *N Engl J Med.* 2006;355(24):2558–2573.
34. Molitch ME, Clemmons DR, Malozowski S, et al. Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2011;96(6):1587–1609.
35. Bidlingmaier M, Freda PU. Measurement of human growth hormone by immunoassays: current status, unsolved problems and clinical consequences. *Growth Horm IGF Res.* 2010;20(1):19–25.
36. Faglia G, Arosio M, Bazzoni N. Ectopic acromegaly. *Endocrinol Metab Clin North Am.* 1992;21:575–596.
37. Ezzat S, Ezrin C, Yamashita S, et al. Recurrent acromegaly resulting from ectopic growth hormone gene expression by a metastatic pancreatic tumor. *Cancer.* 1993;71:66–70.
38. Gola M, Doga M, Bonadonna S, et al. Neuroendocrine tumors secreting growth hormone-releasing hormone: pathophysiological and clinical aspects. *Pituitary.* 2006;9(3):221–229.
39. Chahal HS, Stals K, Unterländer M, et al. AIP mutation in pituitary adenomas in the 18th century and today. *N Engl J Med.* 2011;364:43–50.
40. Ciganoka D, Balcere I, Kapa I, et al. Identification of somatostatin receptor type 5 gene polymorphisms associated with acromegaly. *Eur J Endocrinol.* 2011;165:517–525.
41. de Herder WW. Acromegaly and gigantism in the medical literature. Case descriptions in the era before and the early years after the initial publication of Pierre Marie (1886). *Pituitary.* 2009;12(3):236–244.
42. Reddy R, Hope S, Wass J. Acromegaly. *BMJ.* 2010;341:c4189.
43. Killinger Z, Rovenský J. Arthropathy in acromegaly. *Rheum Dis Clin North Am.* 2010;36(4):713–720.
44. McNab TL, Khandwala HM. Acromegaly as an endocrine form of myopathy: case report and review of literature. *Endocr Pract.* 2005;11(1):18–22.
45. Lopes MB. Growth hormone-secreting adenomas: pathology and cell biology. *Neurosurg Focus.* 2010;29(4):E2.
46. Johnston PC, Hamrahian AH, Prayson RA, et al. Thyrotoxicosis with the absence of clinical features of acromegaly in a TSH and GH-secreting, invasive pituitary macroadenoma. *Endocrinol Diabetes Metab Case Rep.* 2015;2015:140070.
47. Del Porto LA, Liubinas SV, Kaye AH. Treatment of persistent and recurrent acromegaly. *J Clin Neurosci.* 2011;18(2):181–190.
48. Rowland NC, Aghi MK. Radiation treatment strategies for acromegaly. *Neurosurg Focus.* 2010;29(4):E12.
49. Sherlock M, Woods C, Sheppard MC. Medical therapy in acromegaly. *Nat Rev Endocrinol.* 2011;7(5):291–300.
50. Brue T, Castinetti F, Lundgren F, et al. Which patients with acromegaly are treated with pegvisomant? An overview of methodology and baseline data in ACROSTUDY. *Eur J Endocrinol.* 2009;161(suppl 1):S11–S17.
51. Mullis PE. Genetics of isolated growth hormone deficiency. *J Clin Res Pediatr Endocrinol.* 2010;2:52–62.
52. Bartke A. Growth hormone and aging: a challenging controversy. *Clin Interv Aging.* 2008;3(4):659–665.
53. Allen DB, Cuttler L. Clinical practice. Short stature in childhood—challenges and choices. *N Engl J Med.* 2013;368:1220–1228.
54. Reed ML, Merriam GR, Kargi AY. Adult growth hormone deficiency—benefits, side effects, and risks of growth hormone replacement. *Front Endocrinol (Lausanne)* 2013;4:64.
55. Friedrich N, Schneider H, Dörr M, et al. All-cause mortality and serum insulin-like growth factor I in primary care patients. *Growth Horm IGF Res.* 2011;21(2):102–106.
56. Woodhouse LJ, Mukerjee A, Shalet SM, et al. The influence of growth hormone status on physical impairments, functional limitations, and health-related quality of life in adults. *Endocr Rev.* 2006;27(3):287–317.
57. Svensson J, Bengtsson BA. Safety aspects of GH replacement. *Eur J Endocrinol.* 2009;161(suppl 1):S65–S74.
58. Erotokritou-Mulligan I, Holt RIG, Sönksen PH. Growth hormone doping: a review. *J Sports Med.* 2011;2:99–111.
59. Hekimsoy Z, Kafesçiler S, Güçlü F, et al. The prevalence of hyperprolactinemia in overt and subclinical hypothyroidism. *Endocr J.* 2010;57(12):1011–1015.
60. Shibli-Rahhal A, Schlechte J. Hyperprolactinemia and infertility. *Endocrinol Metab Clin North Am.* 2011;40(4):837–846.
61. Shibli-Rahhal A, Schlechte J. The effects of hyperprolactinemia on bone and fat. *Pituitary.* 2009;12(2):96–104.
62. Chahal J, Schlechte J. Hyperprolactinemia. *Pituitary.* 2008;11(2):141–146.
63. Gibney J, Smith TP, McKenna TJ. The impact on clinical practice of routine screening for macroprolactin. *J Clin Endocrinol Metab.* 2005;90:3927–3932.

64. Melmed S, Montori VM, Schlechte JA, et al.; Endocrine Society. Diagnosis and treatment of hyperprolactinemia: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2011;96(2):273–288.
65. Dekkers OM, Lagro J, Burman P, et al. Recurrence of hyperprolactinemia after withdrawal of dopamine agonists: systematic review and meta-analysis. *J Clin Endocrinol Metab.* 2010;95(1):43–51.
66. Delgado V, Biermasz NR, van Thiel SW, et al. Changes in heart valve structure and function in patients treated with dopamine agonists for prolactinomas, a 2-year follow-up study. *Clin Endocrinol (Oxf).* 2012;77(1):99–105.
67. Nawar RN, Abdelmannan D, Selman WR, et al. Pituitary tumor apoplexy: a review. *J Intens Care Med.* 2008;23:75–90.
68. Caturegli P, Lupi I, Landek-Salgado M, et al. Pituitary autoimmunity: 30 years later. *Autoimmun Rev.* 2008;7: 631–637.
69. Hodi FS, O'Day SJ, McDermott DF, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med.* 2010;363:711–723.
70. Romero CJ, Nesi-Franca S, Radovick S. Molecular basis of hypopituitarism. *Trends Endocrinol Metab.* 2009;20:506–516.
71. Fechner A, Fong S, McGovern P. A review of Kallmann syndrome: genetics, pathophysiology, and clinical management. *Obstet Gynecol Surv.* 2008;63(3):189–194.
72. Du X, Yuan Q, Yao Y, et al. Hypopituitarism and successful pregnancy. *Int J Clin Exp Med.* 2014;7:4660–4665.
73. Elphick MR. NG peptides: a novel family of neurophysin associated neuropeptides. *Gene.* 2010;458(1–2):20–26.
74. Lee H-J, Macbeth AH, Pagani J, Young WS 3rd. Oxytocin: the great facilitator of life. *Prog Neurobiol.* 2009;88:127–151.
75. Vrachnis N, Iliodromiti Z. The oxytocin-oxytocin receptor system and its antagonists as tocolytic agents. *Int J Endocrinol.* 2011;2011:350546.
76. Ross HE, Young LJ. Oxytocin and neural mechanisms regulating social cognition and affiliative behavior. *Front Neuroendocrinol.* 2009;30:534–547.
77. Wilson JLL, Miranda CA, Knepper MA. Vasopressin and the regulation of aquaporin-2. *Clin Exp Nephrol.* 2013;17: 751–764.
78. Ozier Y, Bellamy L. Pharmacological agents: antifibrinolytics and desmopressin. *Best Pract Res Clin Anaesthesiol.* 2010;24(1):107–119.
79. Jane JA Jr, Vance ML, Laws ER. Neurogenic diabetes insipidus. *Pituitary.* 2006;9(4):327–329.
80. Cassagnol M, Shogbon AO, Saad M. The therapeutic use of vaptans for the treatment of dilutional hyponatremia. *J Pharm Pract.* 2011;24(4):391–399.

CHAPTER 16

Adrenal Function

Stephanie L. Jacobson

CHAPTER OUTLINE

KEY TERMS

Adrenocorticotropic hormone (ACTH)
Aldosterone
Angiotensin II (AT II)
Atrial natriuretic peptide (ANP)
Cardiovascular disease
Corticotropin-releasing hormone (CRH)
Cortisol
Dehydroepiandrosterone (DHEA)
Dehydroepiandrosterone sulfate (DHEA-S)
Dopamine (DA)
5-Dihydrotestosterone
Epinephrine (EPI)
Homovanillic acid (HVA)
Hypertension
Monoamine oxidase (MAO)
Norepinephrine (NE)
Phenylethanolamine *N*-methyltransferase (PNMT)
Pheochromocytoma
Pheochromocytoma–Paraganglioma (PPGL)
Primary aldosteronism (PA)
Vasoactive inhibitory peptide (VIP)
Vesicular monoamine transporters (VMATs)
Zona fasciculata (F-zone)
Zona glomerulosa (G-zone)
Zona reticularis (R-zone)

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define catecholamines and specify the recommended specimen(s) needed to quantify each.
- Explain how the adrenal gland functions to maintain blood pressure, potassium, and glucose homeostasis.
- Describe steroid biosynthesis, regulation, and actions according to anatomic location within the adrenal gland.
- Discuss the pathophysiology of adrenal cortex disorders, namely, Cushing's syndrome, and Addison's disease.
- Interpret laboratory tests to differentially diagnose primary and secondary Cushing's syndrome and Addison's disease.
- Differentiate the adrenal enzyme deficiencies and their blocking pathways in establishing a diagnosis.
- Sketch the biosynthesis, storage, and metabolism of catecholamines.
- State the most useful measurements in supporting the diagnosis of pheochromocytoma.
- List the clinical findings associated with hypertension that suggest an underlying adrenal etiology is causing high blood pressure.
- Apply knowledge of adrenal function to address case studies.

The adrenal gland is a multifunctional organ that produces steroid hormones and neuropeptides essential for life. Despite the complex actions of adrenal hormones, most pathological conditions of the adrenal gland are manifested by their impact on blood pressure, electrolyte balance, and androgen excess.[1] An adrenal etiology should be considered in the differential diagnosis when patients presents with (1) **hypertension**, hypokalemia, and metabolic alkalosis (suspect hyperaldosteronism); (2) hypertension, spells of anxiety, palpitations, dizziness, and diaphoresis (suspect **pheochromocytoma**); (3) hypertension, rapid unexplained weight gain, red/purple stretch marks, and proximal muscle weakness (suspect Cushing's syndrome); (4) inappropriate hirsutism/virilization and inability to conceive (suspect congenital adrenal hyperplasia); and (5) loss of appetite, unintentional weight loss, and pigmented skin (suspect primary adrenal insufficiency).

In clinical practice, patients often present with underproduction or overproduction of one or more adrenal hormones. Adrenal nodules are also discovered incidentally on abdominal imaging performed for other reasons. Hypofunction is treated with hormone replacement, and hyperfunction is treated with pharmacologic suppression and/or surgery.

Embryology and Anatomy

The adrenal gland is composed of two embryologically distinct tissues: the outer adrenal cortex and inner adrenal medulla. The cortex is derived from mesenchymal cells located near the urogenital ridge that differentiate into three structurally and functionally distinct zones (**Figure 16.1**). The medulla arises from neural crest cells that invade the cortex during the second month of fetal development. By adulthood, the medulla contributes 10% of total adrenal weight.

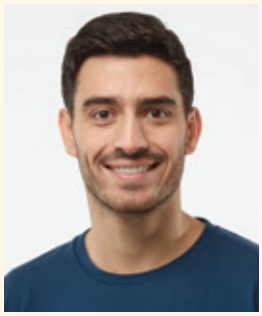
CASE STUDY 16.1, PART 1

Meet Ciera, a 32-year-old woman who goes to her physician for checkup with issues of decreased appetite, fatigue, and nausea.

© Dean Drobot/Shutterstock.

CASE STUDY 16.2, PART 1

Meet Jacob, a 24-year-old male. He goes to his physician with complaints of weakness and sudden weight gain; he tells his physician that his shirts don't fit anymore.



© Damir Khabirov/Shutterstock.

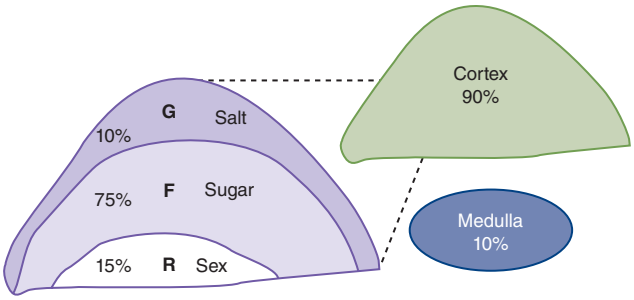


Figure 16.1 Adrenal gland by layer.

Adult adrenal glands are shaped like pyramids, located superior and medial to the upper pole of the kidneys in the retroperitoneal space (also known as suprarenal glands). On cross section, both regions remain distinct; the cortex appears yellow, while the medulla is dark mahogany.[2] Adrenal arterial supply is symmetric, which ensures rapid dissemination of hormones throughout the body in response to stress. Small arterioles branch to form a dense subcapsular plexus that drains into the sinusoidal plexus of the cortex. There is no direct arterial blood supply to the middle and inner zones. In contrast, venous drainage from the central vein displays laterality. After crossing the medulla, the right adrenal vein empties into the inferior vena cava, and the left adrenal vein drains into the left renal vein. There is a separate capillary sinusoidal network from the medullary arterioles that also drains into the central vein and limits the exposure of cortical cells to medullary venous blood. Glucocorticoids from the cortex are carried directly to the adrenal medulla via the portal system, where they stimulate enzyme production of **epinephrine (EPI)** (Figure 16.1).

Sympathetic and parasympathetic axons reach the medulla through the cortex. En route, these axons release neurotransmitters (e.g., catecholamines and neuropeptide Y) that modulate cortex blood flow, cell growth, and function. Medullary projections into the cortex potentially influence cortex function, and they have been found to contain cells that also synthesize and release neuropeptides, such as **vasoactive inhibitory peptide (VIP)**, adrenomedullin, and **atrial natriuretic peptide (ANP)**.

The Adrenal Cortex

The major adrenal cortical hormones, **aldosterone**, **cortisol**, and **dehydroepiandrosterone sulfate (DHEA-S)**, are uniquely synthesized by cells located in one of three functionally distinct zonal layers of the adrenal cortex from the common precursor cholesterol. These zonal layers are zona glomerulosa, zona fasciculata, and zona reticularis, respectively (Figure 16.1).

The **zona glomerulosa (G-zone)** cells constitute the outer 15% of the cortex and synthesize aldosterone, a mineralocorticoid critical for sodium retention, potassium excretion, acid–base homeostasis, and regulation of blood pressure. They have low cytoplasmic-to-nuclear ratios and small nuclei with dense chromatin with intermediate lipid inclusions. Aldosterone release is regulated by the renin-angiotensin-aldosterone system (RAAS). See Chapter 11, *Electrolytes*.

The **zona fasciculata (F-zone)** cells are found in the middle 75% of the cortex and are responsible for synthesizing the glucocorticoids, such as cortisol, cortisone and corticosterone, which are crucial for glucose homeostasis and blood pressure. Fasciculata cells are cords of clear cells, with a high cytoplasmic-to-nuclear ratio and lipids laden with "foamy" cytoplasm.

Zona reticularis (R-zone) cells constitute the inner 10% of the cortex and secrete mineralocorticoids (aldosterone), adrenal androgens, and estrogens. The adrenal androgens include **dehydroepiandrosterone (DHEA)**, androstenedione and testosterone. DHEA and DHEA-sulfate (DHEA-S) are the main adrenal androgens. This zone is sharply demarcated with lipid-deficient cords of irregular, dense cells with lipofuscin deposits.

Adrenal cell types are presumed to arise from stem cells. A proposed tissue layer between the zona glomerulosa and fasciculata may serve as a site for progenitor cells to regenerate zonal cells.[3]

Glucocorticoid Physiology

Cortisol synthesis (8–15 mg/d) is critical to hemodynamics and glucose homeostasis. F-zone disorders manifest with blood pressure and glucose abnormalities (**Figure 16.2**). Glucocorticoids, such as cortisol, maintain blood glucose by inducing lipolysis and causing amino acid release from muscle for conversion into glucose (gluconeogenesis) and storage as liver glycogen.

Cortex Steroidogenesis

Control of steroid hormone biosynthesis is complex, including **adrenocorticotropic hormone (ACTH)** and **angiotensin II (AT II)**. It occurs via substrate availability, enzyme activity, and inhibitory feedback loops that are layer specific.

All adrenal steroids are derived by sequential enzymatic conversion of a common substrate, cholesterol.

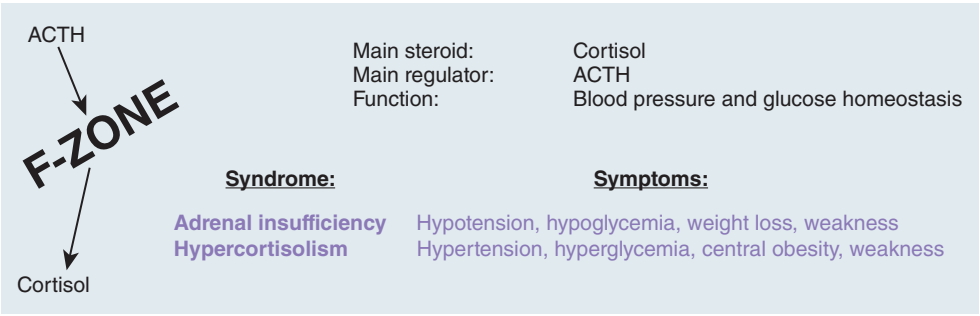


Figure 16.2 F-zone function and physiology. ACTH, adrenocorticotropic hormone.

Adrenal parenchymal cells accumulate and store circulating low density lipoproteins (LDLs). The adrenal gland can also synthesize additional cholesterol using the enzyme acetyl CoA, which ensures that adrenal steroidogenesis remains normal in patients with variable lipid disorders, and even in patients on lipid-lowering agents.

Only free cholesterol can enter steroidogenic pathways in response to ACTH. The availability of free intracellular cholesterol is metabolically regulated by ACTH (stimulatory) and LDL (inhibitory) through multiple mechanisms. **Corticotropin-releasing hormone (CRH)** is secreted from the hypothalamus in response to diurnal variations (circadian), low serum cortisol levels, and stress. CRH stimulates release of stored ACTH from the anterior pituitary gland, which stimulates transport of free cholesterol into adrenal mitochondria, initiating steroid production.

Conversion of cholesterol to pregnenolone is both the first step and the rate-limiting step in steroid biosynthesis: six carbon atoms are removed from cholesterol by the enzyme cytochrome P450 (CYP450) present in the mitochondrial membrane (**Figure 16.3**). Newly synthesized pregnenolone is then returned to the cytosol for subsequent zonal conversion by microsomal enzymes in each layer by F-zone enzymes and/or androgens by enzymes in the R-zone (**Figure 16.4**).

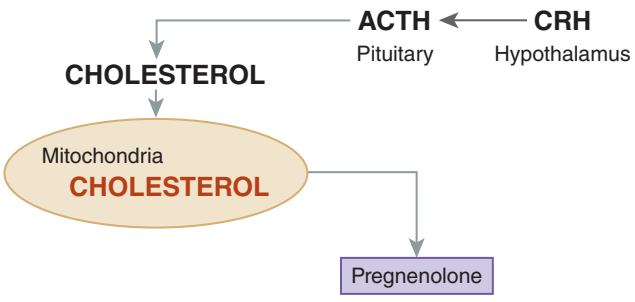


Figure 16.3 Conversion of cholesterol to pregnenolone. ACTH, adrenocorticotropic hormone; CRH, corticotropin-releasing hormone.

High levels of serum glucocorticoids suppress release of CRH and ACTH via a negative feedback loop. Cortisol is the primary feedback regulator of ACTH-stimulated hormone production in the adrenal cortex. ACTH generally does not impact G-zone aldosterone synthesis, although cortisol has mineralocorticoid action and in excess states can stimulate aldosterone.

Decreased activity of any enzyme required for biosynthesis can occur as an acquired or inherited (autosomal recessive) trait. Defects that decrease the production of cortisol cause increases in ACTH and CRH secretion in an attempt to stimulate cortisol levels and lead to adrenal hyperplasia or overproduction of androgens, depending on the affected enzyme.

Evaluation of adrenal function requires measuring relevant adrenal hormones, metabolites, and regulatory hormones. Diagnosis is then based on the correlation of clinical and laboratory findings.[4]

Congenital Adrenal Hyperplasia

Congenital adrenal hyperplasia refers to a group of clinical entities that arise from absent or diminished activity of enzymes involved in steroidogenesis. The mineralocorticoid, glucocorticoid, and androgen production pathways can be affected to varying degrees based on the enzyme affected. Blocks in one pathway result in upstream substrate buildup and potential upregulation of another pathway. The most common enzyme affected is 21-hydroxylase. Deficiency in this enzyme results in decreased glucocorticoid, in some cases mineralocorticoid and increased adrenal androgen production.[5] A very high serum concentration of 17-hydroxyprogesterone, the normal substrate for 21-hydroxylase, is diagnostic of classic 21-hydroxylase deficiency. The "classic" presentation is seen in infants and presents with features such as failure to thrive and low blood pressure. These

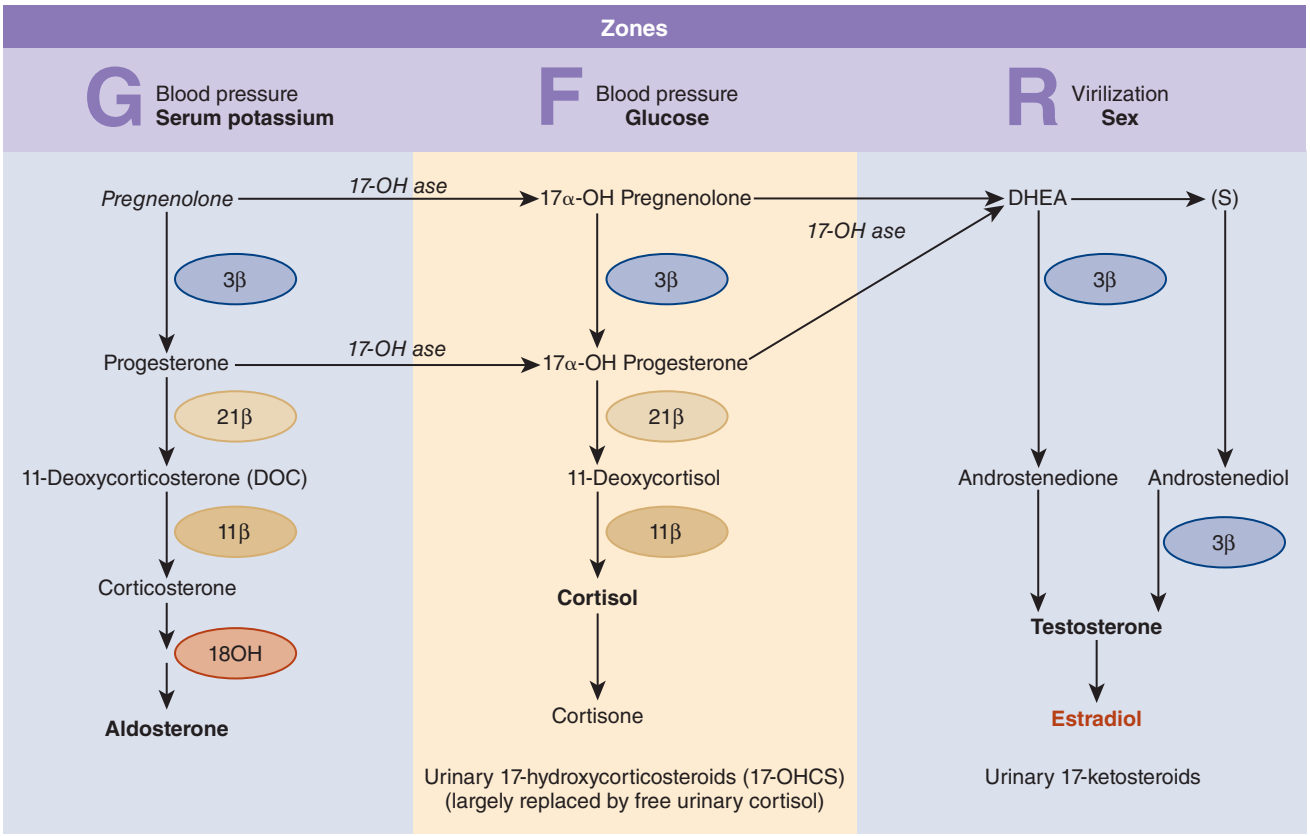


Figure 16.4 Adrenocortical hormone synthesis by zone.

infants need both glucocorticoid and mineralocorticoid replacement. A second "nonclassic" form is seen in adult women who present in their reproductive years with complaints of hirsutism, menstrual irregularities, and infertility. These individuals may need steroids during pregnancy. The different enzyme defects along with clinical and biochemical abnormalities have been summarized in **Figure 16.5**.

Primary Aldosteronism

The clinical entity in which excessive secretion of aldosterone cannot be suppressed with salt or volume replacement is termed **primary aldosteronism (PA)**. PA results in hypertension, hypokalemia, metabolic alkalosis, and increased risk of vascular disease such as stroke and is depicted in **Figure 16.6**. It is estimated that 5% to 10% of patients with hypertension have PA.[6] Endocrine Society guidelines on PA suggest that hypokalemia is found in 37% of patients with PA.[7] The guidelines recommend more aggressive screening may be appropriate in patients with the following signs: systolic blood pressure > 160, diastolic blood pressure > 100, drug-resistant hypertension, hypokalemia associated with hypertension, presence of adrenal mass, family history of early hypertension or stroke, and first-degree relatives of patients with PA.

Enzyme defect	New classification	HTN	Virilization	High lab value
3β-Hydroxysteroid dehydrogenase	3β-HSD II	N	Slight	DHEA
17α-Hydroxylase	CYP17	Y	No	Aldosterone
11β-Hydroxylase	CYP11B1GF	Y	Marked	11-DOC
21β-Hydroxylase	CYP21A2	N	Marked	17-OH-progesterone

Figure 16.5 Congenital adrenal hyperplasia syndromes.

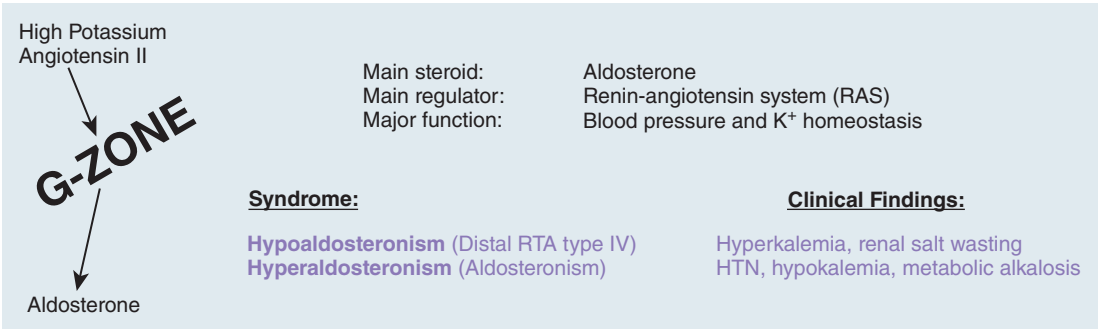


Figure 16.6 G-zone function and pathology. HTN, hypertension.

Etiology

The most common causes of PA are:

- Aldosterone-producing adrenal adenoma
- Unilateral or bilateral adrenal hyperplasia
- Familial hyperaldosteronism such as glucocorticoid remediable aldosteronism (GRA)
- Adrenocortical carcinomas that secrete aldosterone
- Ectopic aldosterone secretion

Diagnosis

Plasma aldosterone concentration (PAC) and plasma renin activity (PRA) are used in the diagnosis of PA.[8] Presence of both a PAC > 15 ng/dL and PAC/PRA ratio of 30 or greater is suggestive of PA. The 2008 Endocrine Society guidelines recommend the labs be drawn in the morning after the patient has been up for at least 2 hours. The patient should be seated for 5 to 15 minutes before the draw, and sodium intake should be unrestricted. Most antihypertensive medications do not need to be stopped prior to testing except mineralocorticoid antagonists. If clinical suspicion is high in a patient on an ACE inhibitor and testing is negative, the ACE inhibitor should be discontinued and testing repeated. If the initial screening is suggestive of PA, confirmatory tests include aldosterone measurement following either oral salt loading or IV saline infusion.

Biochemical evaluation is followed by adrenal imaging (CT, MRI). Aldosterone secretion may be due to bilateral hyperplasia. Adrenal venous sampling (AVS) is recommended to distinguish between hyperplasia and aldosterone-secreting adenoma.[9]

Treatment

Surgery is the treatment of choice for an aldosterone-producing adenoma, and hypertension is controlled in 30% to 60% of the patients treated surgically.[10] Mineralocorticoid antagonists such as spironolactone or eplerenone should be used for adrenal hyperplasia, and prednisone is the treatment of choice for those with familial hyperaldosteronism such as GRA.

Types of aldosteronism are diagrammed in **Figure 16.7** according to their PA (*y*-axis) and PRA (*x*-axis) values.

Adrenal Insufficiency

Adrenal insufficiency is a term that describes inadequate hormone secretion from the adrenal cortex and may be categorized according to the gland that is affected: primary (adrenal), secondary (pituitary), and tertiary (hypothalamic). A primary adrenal disease involves inadequate release of glucocorticoids (such as cortisol), mineralocorticoids (such as aldosterone), and adrenal androgens. In primary adrenal insufficiency, there is reduced production of adrenal hormones despite adequate stimulation.

In secondary adrenal insufficiency, adrenal gland function is preserved but the stimulus for hormone release is insufficient or absent. For example, this would occur if the pituitary gland failed to release adequate ACTH, which in turn does not stimulate proper cortisol levels. Tertiary adrenal insufficiency can be caused by any process that involves the hypothalamus and interferes with CRH secretion. The most common cause of tertiary adrenal insufficiency is chronic administration of synthetic glucocorticoids that suppress the hypothalamic–pituitary–adrenal (HPA) axis.[11]

The most common cause of primary adrenal insufficiency is autoimmune destruction of the adrenal gland.[12] Other causes of adrenal insufficiency include infections (tuberculosis and histoplasmosis), tumors, bilateral adrenal hemorrhage, etc.

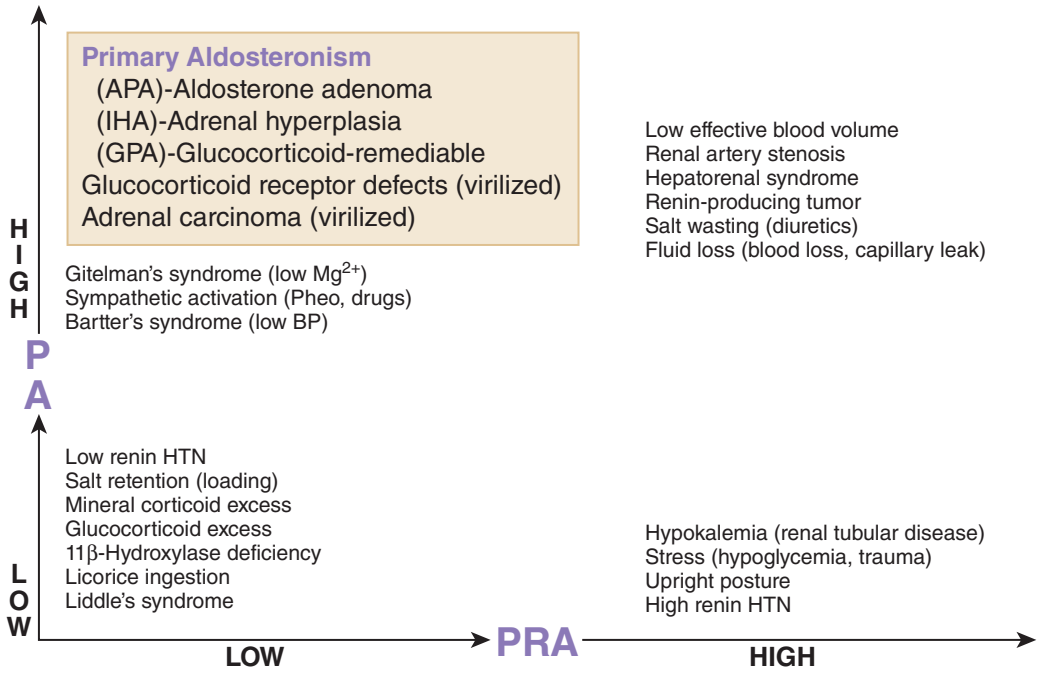


Figure 16.7 Types of aldosteronism according to plasma aldosterone: plasma renin activity (PRA) ratio. HTN, hypertension.

Symptoms

Clinical symptoms of chronic adrenal insufficiency can be nonspecific, with most patients complaining of fatigue, decreased appetite, weight loss, and nausea (**Figure 16.8**). Patients may also present with more severe findings such as low blood pressure, low blood sugar, low serum sodium, and high potassium levels.

Diagnosis

Since cortisol secretion is based on diurnal variation, samples must be drawn according to the peak (6:00–8:00 am; reference range: 7–25 µg/dL) and nadir (trough, which is usually midnight for cortisol based on a normal sleeping pattern reference range 2–14 µg/dL). The diagnosis of adrenal insufficiency can be suspected if the 8:00 am serum cortisol is low with concomitant elevation in ACTH levels. A cortisol level < 3 µg/dL in the morning is highly suggestive of adrenal insufficiency. The diagnosis of primary adrenal insufficiency is made by performing an ACTH stimulation test. The test is performed at 8:00 am in the fasting state.[13] Baseline cortisol and ACTH levels are obtained. The patient is then given 250 µg of cosyntropin (synthetic ACTH) intravenously, and cortisol level is drawn at 30 and 60 minutes post-ACTH administration.

CASE STUDY 16.1, PART 2

Remember Ciera. After review of Ciera's symptoms and history, her preliminary examination showed a low blood pressure. The physician orders a comprehensive metabolic panel, and the following laboratory tests were outside the normal reference intervals:

Na^+ = 122 mmol/L

K^+ = 5.9 mol/L

Glucose = 70 mg/dL (nonfasting)

1. If the physician suspects adrenal issues, what follow-up tests should be ordered?

CASE STUDY 16.1, PART 3

Remember Ciera. On a follow-up visit, Ciera indicates she is not feeling any better. In fact, she complains about more nausea in the last week. Her physician orders tests for cortisol and ACTH. The results are:

8 AM cortisol < 3 µg/dL (low) (Reference Range 7–25 µg/dL)

ACTH = elevated (Reference Range 7.2–63 pg; mL; morning collection only)

2. What could the results indicate?

Frequency	Symptoms	Signs
100%	Weakness Fatigue Anorexia	Weight loss
90%		Hyperpigmentation (primary adrenal insufficiency)
50%	Nausea Diarrhea	
10%	Pain	Adrenal calcification

Figure 16.8 Signs and symptoms of adrenal insufficiency.

A cortisol of 18 µg/dL (indicating an increase of at least 7µg/dL from baseline) or greater at either 30 or 60 minutes post-ACTH administration suggests normal adrenal function. In secondary adrenal insufficiency, the ACTH stimulation test may be normal or abnormal based on the duration of the disease. If secondary adrenal insufficiency is suspected, metyrapone suppression testing can be done but is currently only available for compassionate testing and rarely performed.[14] Metyrapone blocks certain enzymes (11β-hydroxylase) in the steroidogenesis pathway, preventing the conversion to cortisol. When metyrapone is administered orally at midnight in normal individuals, it will block 11β-hydroxylase, increasing 11-deoxycortisol (>7 µg/dL) and causing cortisol to decrease (<5 µg/dL). Secondary adrenal insufficiency is suggested in patients with a near-normal response to a 250-µg cosyntropin test but with an abnormal response to metyrapone. An alternative is to perform an insulin tolerance test (ITT). The ITT has long been considered the gold standard for assessing the HPA axis and is used mainly as a second-line measure to further evaluate those patients who had a borderline response to the ACTH stimulation test. Unlike the ACTH stimulation test, the ITT assesses the integrity of the entire HPA axis. If the HPA axis is intact, insulin-induced hypoglycemia stimulates the hypothalamus and the pituitary gland to secrete CRH and ACTH, respectively, which, in turn, leads to a rise in cortisol. Hypoglycemia is induced with insulin administration, and ACTH, cortisol, and growth hormone response to hypoglycemia is measured. The ITT is contraindicated in those with ischemic **cardiovascular disease**, in older adults (>70 years of age), during pregnancy, and in those with a history of seizures. **Figure 16.9** provides a summary of primary and secondary adrenal insufficiency.

Treatment

The treatment of primary adrenal insufficiency includes both glucocorticoid replacement (prednisone or hydrocortisone) and mineralocorticoid replacement (fludrocortisone). In secondary adrenal insufficiency, only glucocorticoid replacement is required since aldosterone production remains normal, as it is not under the control of the pituitary gland.

Isolated Hypoaldosteronism

Insufficient aldosterone secretion is seen with adrenal gland destruction, with chronic heparin therapy, following unilateral adrenalectomy (transient), and with G-zone enzyme deficiencies. Usually hypoaldosteronism occurs in patients with mild renal insufficiency—such as persons with diabetes who present with mild metabolic acidosis and have a high serum potassium, low urinary potassium excretion (urine K^+ < urine Na^+), and low renin. Treatment is with dietary changes and fludrocortisone, which enhances salt retention and the secretion of both potassium and hydrogen.

CASE STUDY 16.1, PART 4

Remember Ciera. In order to confirm her diagnosis, an ACTH stimulation test was performed. The results are as follows:

Post stimulation: Cortisol = 10 µg/dL; ACTH = elevated

3. What is Ciera's diagnosis?

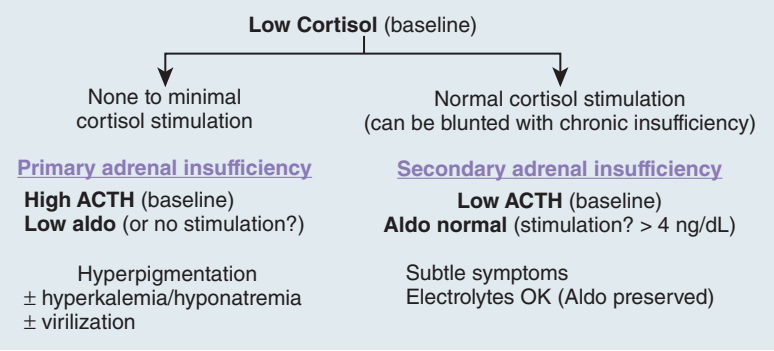


Figure 16.9 Differential diagnosis of low cortisol states. ACTH, adrenocorticotropic hormone.

Hypercortisolism (Cushing's Syndrome)

Excess cortisol production presents with a constellation of clinical manifestations. Each is not unique to cortisol excess but, when present together, may suggest hypercortisolism. The common findings are as follows: There is progressive central obesity with sparing of extremities. Dorsocervical fat pad deposition and fat accumulation in the cheeks result in the "buffalo hump" and "moon face," respectively. Dermatologic manifestations can be seen such as atrophy of skin, easy bruising, wide purple striae, and hyperpigmentation based on the etiology of cortisol excess. In women, menstrual irregularity, signs of androgen excess such as acne, and *hirsutism* can be seen. Proximal muscle weakness and decreased bone mineral density may be present. Patients with cortisol excess also have increased risk of cardiovascular disease, thromboembolic events such as disseminated intravascular coagulopathy, increased anxiety, and risk of infection including atypical fungal infections.[15]

Etiology

Cortisol excess may be due to adrenal tumors that produce excess cortisol, in which case pituitary release of ACTH levels remain low due to the feedback loop. Alternatively, cortisol production could be due to elevated ACTH levels, either from a pituitary tumor or from ectopic (nonpituitary) sources often associated with malignancy, such as lung cancer. The term *Cushing's disease* is used specifically when the source of elevated ACTH is the pituitary gland, while other forms of cortisol excess are labelled as *Cushing's syndrome*. Cushing's disease is the most common reason for cortisol excess, accounting for about 70% of cases (**Figures 16.10** and **16.11**).

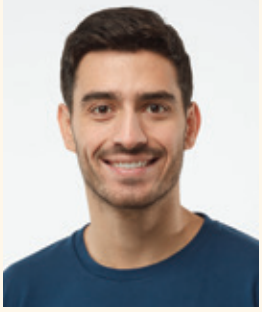
Diagnosis

Once exogenous cortisol use (oral, topical, injected, and/or inhaled) is ruled out, initial screen tests for cortisol excess include: (a) urine free cortisol collected over 24 hours (on two occasions to be considered positive for accuracy); (b) midnight salivary cortisol (two occasions); and (c) dexamethasone suppression

CASE STUDY 16.2, PART 2

Remember Jacob. During Jacob's physical examination, the physician notes central fat pads, plethora, thin skin, bruises, and the beginnings of a "buffalo hump." His physician suspects Cushing's syndrome.

1. What laboratory test can be used to confirm this assumption?



testing (1 mg overnight or 2 mg over 48 hours). A systematic review during the development of Endocrine Society guidelines in 2008 suggested that the three tests were similar in diagnostic accuracy.[16] The guidelines recommend demonstrating at least two positive tests for the diagnosis as this increases the diagnostic accuracy over any one test.

Urine free cortisol (UFC) measurement requires accurate collection and measurement of urine volume. Fluid intake > 5 L a day and urine volume > 3 L can lead to a false-positive test.[17] Depression and alcohol intake may also influence UFC, resulting in so-called pseudo-Cushing's.[18] A cortisol level three times above normal is suggestive of cortisol excess;

Stress

Infection

Severe obesity (visceral)
Polycystic ovary syndrome (up to 40% have slightly elevated urine cortisol)
Chronic alcoholism (cortisol normalizes with abstinence)
Depression (up to 80% have abnormal cortisol levels, disappears with remission)
Iatrogenic Cushing's (<1% inhaled, topical, oral glucocorticoid use)

Cushing's Syndrome

Symptoms of Cushing's Syndrome	
HTN	**(85–90%)**
Central obesity	**(90%)**
Glucose intolerance	**(80%)**
Plethoric faces	**(80%)**
Purple striae	**(65%)**
Hirsutism	**(65%)**
Abnormal menses	**(60%)**
Muscle weakness	**(60%)**

Figure 16.10 Conditions associated with hypercortisolism. HTN, hypertension.

Causes of Cushing's Syndrome (High Cortisol) by ACTH

High ACTH
Cushing's with hyperpigmentation

Low ACTH

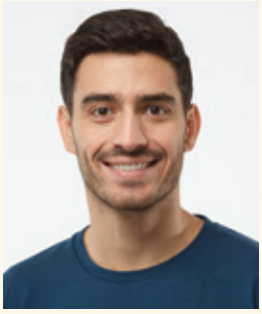
ACTH Dependent	ACTH Independent
Primary ACTH (pituitary disease)	Adrenal adenoma
Ectopic ACTH	Adrenal carcinoma
Ectopic CRF	Nodular adrenal hyperplasia
	Exogenous glucocorticoids

Figure 16.11 Differentiating source of adrenocorticotropic hormone (ACTH) secretion.

CASE STUDY 16.2, PART 3

Remember Jacob. The physician has ordered a 24-hour urine collection for cortisol testing.

2. What specific instructions should Jacob be following during the collection period?



© Damir Khabirov/Shutterstock.

however, this would need confirmation with another test. When liquid chromatography-tandem mass spectrometry is used, sensitivity of UFC is 97% and specificity was 91%. The reference range for UFC is 3.5–45 µg/24 h for adults.

The overnight dexamethasone suppression test (DST) involves administration of 1 mg dexamethasone between 11:00 pm and midnight, followed by an 8:00 am serum cortisol level.[19] The test is considered negative if the 8:00 am cortisol level is < 1.8 µg/dL. This is a stringent criterion and may have greater sensitivity but lower specificity. It is the test of choice when an adrenal nodule or mass is an incidental finding on abdominal imaging.

Late-night salivary cortisol level (LNSC) is an alternate screening test. It takes advantage of the fact that patients with true excess endogenous production of cortisol lose diurnal variability in cortisol production and so will have increased LNSC levels. On the other hand, those with pseudo-Cushing's maintain that variability. A pooled analysis of seven studies suggested sensitivity of 93% and specificity of 96% of LNSC[20] and is superior to plasma and UFC tests. These numbers may be lower in mild hypercortisolism and in pregnancy.

Once hypercortisolism is established, a reliable sandwich electrochemiluminescence immunoassay for ACTH is used to determine the cause for high cortisol levels. If the 8:00 am ACTH level is > 15 µg/dL, cortisol secretion is considered ACTH dependent. On the other hand, with ACTH < 5 µg/dL, cortisol secretion is thought to be ACTH independent. While levels between 5 and 15 µg/dL are deemed indeterminate, there are little data to support these cutoffs, and the recommendations come from clinical guidelines. In ACTH-independent Cushing's syndrome, the next step is to obtain CT or MRI imaging of the adrenal glands. The reference range for ACTH at 8:00 am is 7.2–63.0 pg/mL.

In ACTH-dependent hypercortisolism, the source of ACTH is most often the pituitary gland. The source can also be ectopic such as lung cancers, pancreatic tumors, etc. Pituitary adenomas secreting ACTH are somewhat resistant to negative feedback regulation by glucocorticoids, while most non-pituitary, ACTH-secreting malignant tumors are highly resistant to feedback suppression. Therefore, dynamic testing employing high-dose steroids is used to distinguish between ectopic and pituitary source of ACTH.

The high-dose dexamethasone suppression test can be performed in two ways: overnight[19,21] or over 48 hours.

In the overnight test, a baseline morning serum cortisol level is drawn, and 8 mg of dexamethasone is administered at 11:00 pm followed by an 8:00 am plasma ACTH and serum cortisol measurement. A cortisol level < 5 µg/dL is considered significant suppression. Alternatively, if the cortisol post-dexamethasone is suppressed by 50% or more compared to the baseline value, the source of ACTH is thought to be pituitary.

In the 2-day high-dose dexamethasone suppression test, baseline UFC from a 24-hour urine collection is obtained, sometimes with a baseline serum cortisol. Following this, 2 mg of dexamethasone is administered every 6 hours for 48 hours. At the end of the study, urine free cortisol (24 hours) or serum cortisol is measured. Low urinary cortisol and suppression of serum cortisol < 5 µg/dL would suggest pituitary ACTH production.

The gold standard for determining the source of ACTH production is by inferior petrosal sinus sampling (IPSS). This is an invasive procedure in which the ACTH levels in the inferior petrosal sinus and a peripheral vein are determined simultaneously. This study can also be done following stimulation with CRH. If the ratio of inferior petrosal sinus ACTH to peripheral ACTH is 2:1 or greater without CRH and is 3:1 or greater with use of CRH, it suggests that ACTH hypersecretion is coming from the pituitary itself.[22] **Figure 16.12** summarizes the diagnostic testing algorithm.

Treatment

Cushing's Disease

Transsphenoidal resection of pituitary tumor is the treatment of choice.[23,24] Cure rates of 80% to 90% can be achieved with microadenomas (<1 cm). Cure rates

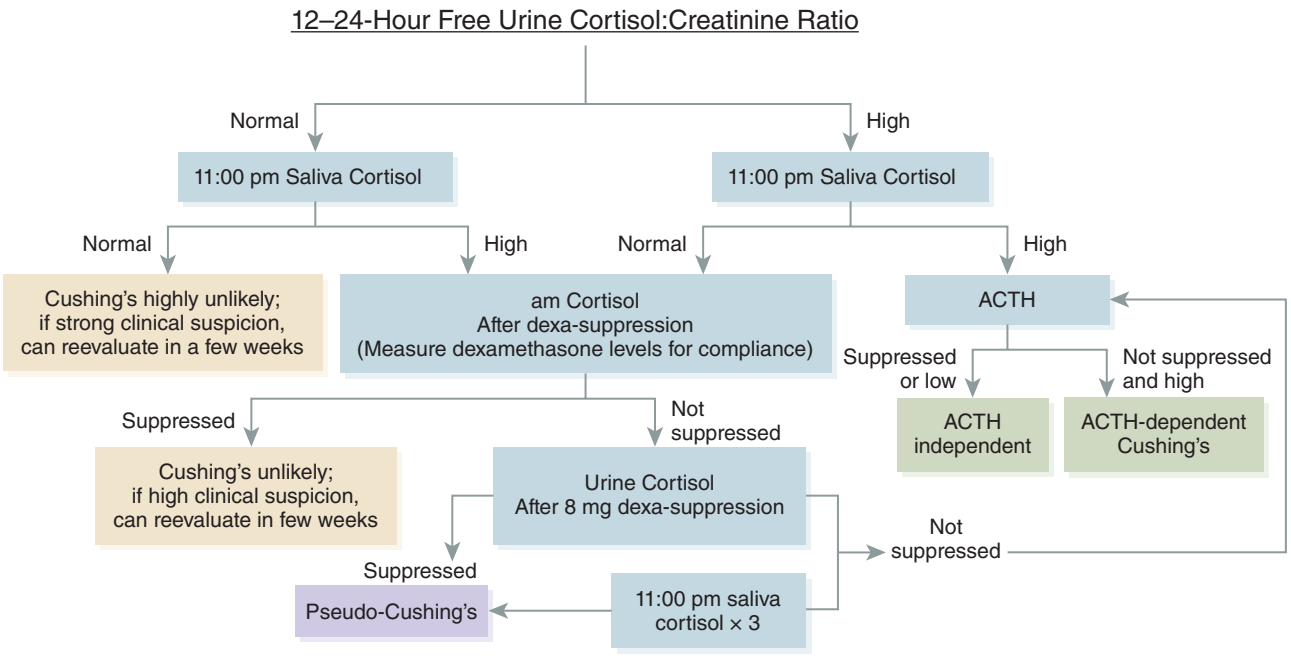


Figure 16.12 Cushing's syndrome workup. Note that dexamethasone serum levels (for standard test, 2 ng/mL; suppression test, 6.5 ng/mL) can be drawn at 8:00 am (or 6 hours after the last dose) to help clarify or determine compliance. Normal dexamethasone saliva levels have not been determined. ACTH, adrenocorticotropic hormone.

are around 60% or less when the pituitary tumor is > 1 cm (macroadenoma).

Cushing's Syndrome

Surgical resection (adrenal or ectopic) would be the first step in management. If hypercortisolism persists after surgical intervention, medical therapy can be used. Adrenal enzyme inhibitors such as ketoconazole, adrenolytic agents such as mitotane, pituitary ACTH secretion suppressors such as pasireotide or cabergoline, and finally glucocorticoid receptor blockers are all available for use. Bilateral adrenalectomy is occasionally necessary if medical therapy is unable to suppress cortisol levels and the patient has infectious or thrombotic complications from cortisol excess.

Adrenal Androgens

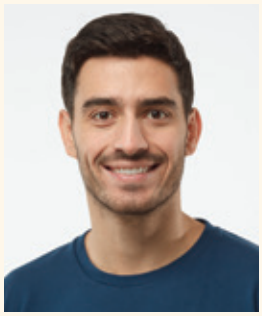
Androgens are produced as byproducts of cortisol synthesis that are regulated by ACTH. Although prolactin, proopiomelanocortin peptides, and T lymphocytes are known stimulators of androgens, regulatory mechanisms of R-zone biosynthesis remain uncertain (**Figure 16.13**). R-cells primarily produce DHEA and multiple 19-carbon steroids (androgens and estrogens) from 17α-hydroxylated pregnenolone and progesterone. DHEA is sulfated to DHEA-S by sulfotransferase, an adrenal enzyme, and secreted daily.

Both DHEA and DHEA-S are precursors to more active androgens (e.g., androstenedione, testosterone, and **5-dihydrotestosterone**) and estrogens (e.g., estradiol and estrone). Although DHEA and DHEA-S have minimal androgenic activity, adverse effects are caused by conversion to active androgens in the adrenal and peripheral tissue (e.g., hair follicles, sebaceous glands, genitalia, adipose, and prostate tissue). Although men derive less than 5% of

CASE STUDY 16.2, PART 4

Remember Jacob.The physician has ordered a 24-hour urine collection for cortisol.

3. What specific instructions should Jacob be following during the collection period?



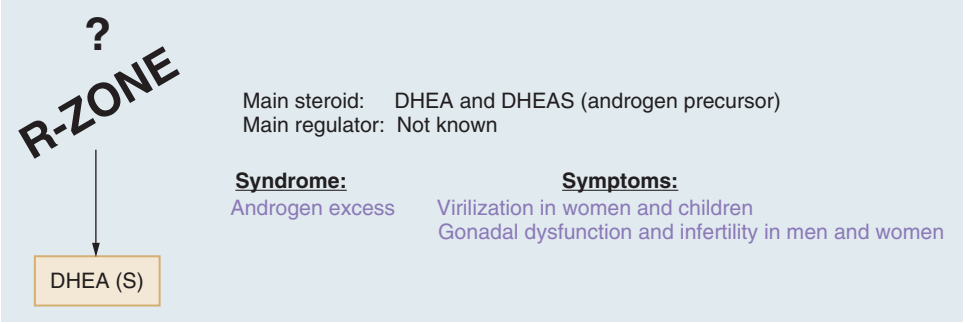


Figure 16.13 R-zone function and pathology. Manifestations of adrenal hyperandrogenism vary with age, onset, and gender. DHEA, dehydroepiandrosterone; DHEA-S, dehydroepiandrosterone sulfate.

their testosterone from adrenal or peripheral sources, women rely on the adrenals for 40% to 65% of their daily testosterone production.

Adrenal androgen production increases in both genders in late childhood and correlates with the onset of pubic hair (adrenarche); however, it peaks in young adults and progressively declines with age. See Chapter 17, *Gonadal Function*.

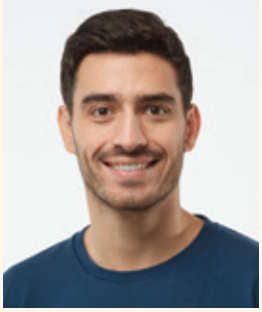
Androgen Excess

Androgen excess causes ambiguous genitalia in female infants and precocious puberty in children of both sexes. Androgens stimulate organ development, linear growth, and epiphyseal fusion. Excess androgens in boys includes premature sexual maturation and exaggerated male characteristics such as penile enlargement, increased hair growth on the face and body, baldness, and others. Increased androgen production in girls leads to virilization such as hirsutism, acne, deepening of voice, uterine atrophy, and clitoral hyperplasia. Excess androgens can also cause short stature by causing early epiphyseal fusion.

CASE STUDY 16.2, PART 5

Remember Jacob. The results of the 24-hour urine collection for cortisol shows elevated levels. Reference Range: 3.5–45 µg/24 h

4. Is this the only test needed for a definitive diagnosis?



In adult women, androgen overproduction causes similar effects as described above in girls and can lead to amenorrhea and infertility. In adult men, excess adrenal androgens that are converted to estrogen can also cause infertility with feminizing effects and inhibition of pituitary gonadotropins, which lowers testicular testosterone production. Despite overall adrenal androgen excess, males can experience hypogonadal symptoms including loss of muscle mass and decreased hair growth, decrease in testes size, low testicular testosterone production, and decreased spermatogenesis.

Diagnosis

Gonads produce less than 10% of DHEA-S and DHEA; therefore, high DHEA-S and DHEA production strongly suggests adrenal hyperandrogenism, whereas elevated testosterone values are seen with either adrenal or gonadal hyperandrogenism.

Plasma DHEA-S, DHEA, or urinary 17-ketosteroids can identify patients with adrenal causes of pathologic masculinization (females) and feminization (males). References ranges for DHEA-S and DHEA vary by age. The reference range for urinary 17-ketosteroids is 10–20 mg/24 h for males and 5–15 mg/24 h for females.

Treatment

Similar to previously described disorders of overproduction, differentiation between ACTH-dependent and ACTH-independent secretion is assessed by dexamethasone suppression tests followed by imaging studies (CT, MRI). Adenomas and carcinomas are surgically removed. Glucocorticoid-suppressible causes are treated accordingly. Exogenous sources are discontinued, and other nonadrenal conditions are treated. Drugs with antiandrogenic properties

INCREASE	DECREASE	No Change
IGF-1 DHEA(S) Testosterone Androstenedione	HDL (transient)	Estradiol Estrone
Sexuality Sebum production Hair growth Acne Mood	Depression Fatigue	Muscle mass Strength Hot flashes Sexuality

Figure 16.14 Effects of DHEA supplementation. DHEA, dehydroepiandrosterone; IGF-1, insulin growth factor 1; HDL, high-density lipoprotein.

(e.g., minoxidil, spironolactone, birth control pills) are occasionally used.

Exogenous DHEA is a popular nutritional supplement with numerous purported properties including vasodilatory, anti-inflammatory, antiaging, and antiatherosclerosis; however, only a few have been studied clinically.

The clinical relevance of DHEA(S) interactions with non-androgen/estrogen receptors remains unknown. DHEA (4 mg/d) and DHEA-S (7 to 15 mg/d) are secreted as the major components of adrenal androgens. There is no evidence that this molecule is required for health or that it contributes to disease. Steroid receptors for DHEA have not been clearly identified. DHEA actions are attributed to its downstream products, testosterone and estrogen.

In normal and compromised patients (e.g., adrenal insufficiency, glucocorticoid therapy, depression, older persons, and trained athletes), DHEA can increase the sense of well-being and raises or lowers a variety of serum markers (**Figure 16.14**), although the changes are small. Supplementation (50 to 100 mg/d) in some androgen-deficient patients (e.g., those with adrenal insufficiency, ACTH deficiency, and glucocorticoid therapy) may help ameliorate adverse effects of deficiency and may inhibit glucocorticoid-induced bone loss. However, DHEA can cause adverse androgenic effects in women, and the long term consequences of supplementation remain unknown.

The Adrenal Medulla

The adrenal medulla produces catecholamines from the amino acid tyrosine. The secreted catecholamines regulate the body's sympathetic response to stress. Catecholamine products serve as first responders to stress by acting within seconds (while cortisol takes 20 minutes) to promote the fight-or-flight response. This increases cardiac output and blood pressure, diverts blood toward muscle and brain, and mobilizes glucose from storage for fuel. Catecholamines have a half-life in the blood of 1 or 2 minutes.

Embryology

Sympathetic nervous system arises in the fetus from primordial neural crest stem cells. At about the fifth week of gestation, they migrate from the spinal ganglia in the thoracic region. At the sixth week of gestation, these primitive cell groups migrate to the adrenal cortex to form the adrenal medulla. These cells are referred to as sympathogonia and pheochromoblasts, which later differentiate into pheochromocytes. The paraganglia are formed by migration of pheochromoblasts and pheochromocytes on either side of the aorta.

Tumors that arise from either cell line share similar histologic and biochemical properties. Malignant neuroblastomas and benign ganglioneuromas arise from lymphoblasts, secrete **homovanillic acid (HVA)**, and are rarely seen after adolescence. In contrast, tumors of chromaffin cells (pheochromocytomas) maintain the capacity to synthesize and store catecholamines (**norepinephrine [NE]** and EPI) throughout life.[25]

Biosynthesis, Storage, and Secretion of Catecholamines

Catecholamines have a benzene ring with a catechol nucleus and two hydroxyl side groups linked to a side chain amine group, hence the name "catecholamine," and include **dopamine (DA)**, NE, and EPI. Biosynthesis begins with the sequential conversion of phenylalanine substrates in a tightly regulated, compartmentalized manner. All reactions take place in the cytoplasm, except for the production of NE, which occurs within lipid vesicles or outer mitochondrial membranes, as illustrated in **Figure 16.15**.

In the cytosol, NE is converted into EPI by a cortisol-dependent enzyme called **phenylethanolamine *N*-methyltransferase (PNMT)**. Any form of stress that increases cortisol levels stimulates EPI production.

In the adrenal medulla, EPI and NE are stored in neurosecretory vesicles but are in dynamic equilibrium with the surrounding cytoplasm. They

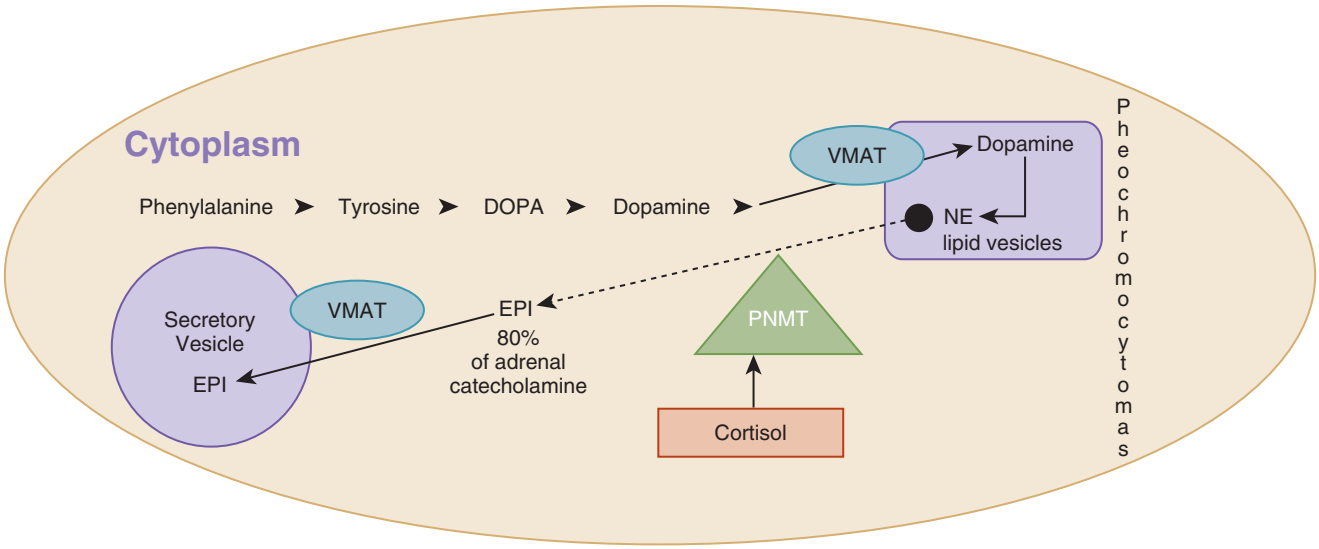


Figure 16.15 Biosynthesis and storage of catecholamines. VMAT, vesicular monoamine transporter; NE, norepinephrine; EPI, epinephrine; PNMT, phenylethanolamine *N*-methyltransferase.

diffuse into the cytoplasm and actively re-enter the vesicles via **vesicular monoamine transporters (VMATs)**. They enter the circulation via exocytosis of the vesicles. The cytoplasmic catecholamines may also be metabolized by the enzyme catecholamine-*O*-methyltransferase (COMT) into metanephrines and normetanephrines, which enter circulation via diffusion.

EPI is almost exclusively produced and secreted by the adrenal medulla. The ratio of NE/E in the medulla is about 1:4 and contains 80% EPI. However, because all three catecholamines are also synthesized within the central and sympathetic nervous systems, the peripheral NE/E ratio is closer to 9:1 (98% from postganglionic neurons, 2% from the medulla).[26] In adrenal insufficiency (low cortisol), that ratio increases to 45:1 in females and 24:1 in males.

In sympathetic neurons, cytoplasmic dopamine is sequestered into vesicles, converted into NE, and stored until nerve stimulation causes its release into a synapse by exocytosis.

Metabolism and Excretion of Catecholamines

All catecholamines are rapidly eliminated from target cells and the circulation by three mechanisms:

1. Reuptake into secretory vesicles
2. Uptake in nonneuronal cells (mostly liver)
3. Degradation

Degradation of free catecholamines relies on two enzymes—COMT (in nonneuronal tissues) and **monoamine oxidase (MAO)** (within neurons)—to produce the metabolites metanephrines and normetanephrines, which are then converted to vanillylmandelic acid (VMA). Metabolites and free catecholamines are eliminated by direct filtration into the urine and excreted as free NE (5%), conjugated NE (8%), metanephrines (20%), and VMA (30%) (**Figure 16.16**).

Pheochromocytoma and Paraganglioma

Pheochromocytomas are rare catecholamine-secreting tumors arising from chromaffin cells. The name was coined based on the dusky color of the tumor after staining with chromium salts (in Greek, *phios* = dusky, *chromo* = color, and *cytoma* = tumor). Chromaffin cells are found in the adrenal medulla and in neural ganglia. Most of the tumors referred to as **pheochromocytoma** (90%) arise from the adrenal medulla; the smaller number arising from the sympathetic ganglia are referred to as *paragangliomas* (simply indicating that it is an extra-adrenal pheochromocytoma). Although 90% of pheochromocytomas are benign,[26] it is important to distinguish between the two types because of implications related to neoplastic associations, risk of malignancy, and genetic testing; though rare, they are potentially lethal. In this chapter, **pheochromocytoma–paraganglioma (PPGL)** will be used to refer to both adrenal and extra-adrenal tumors.

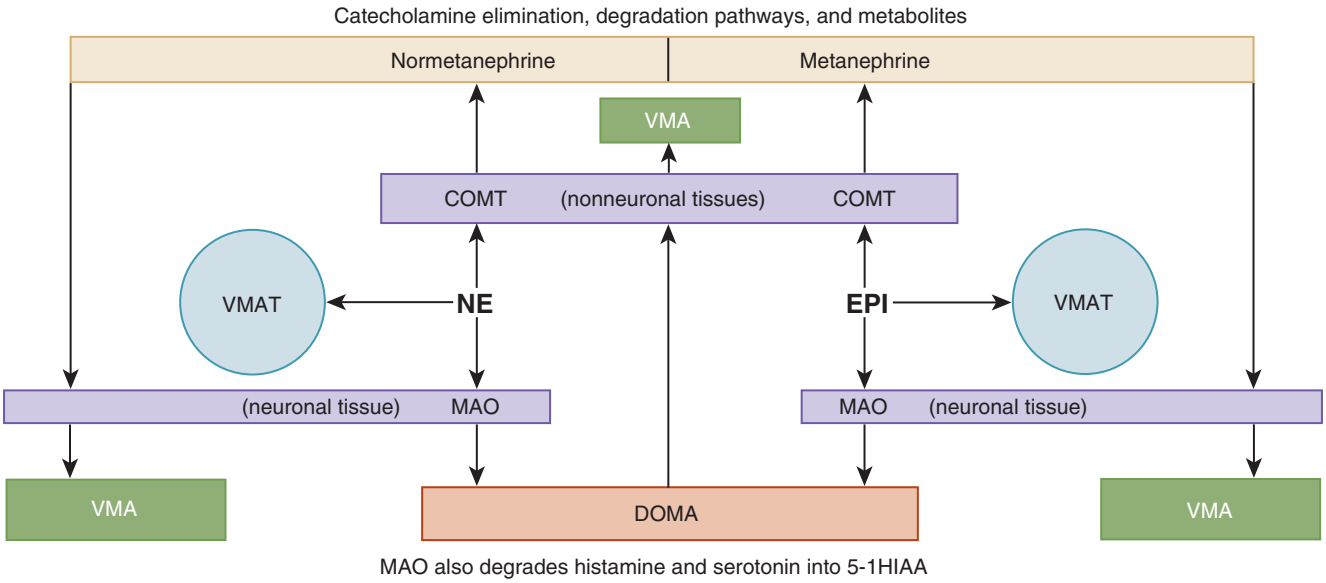


Figure 16.16 Catecholamine degradation. Free catecholamines (EPI and NE) are either sequestered into VMAT-containing vesicles or converted into metabolites, DOMA by neuronal MAOs and metanephrines by nonneuronal COMTs. These metabolites are ultimately degraded to VMA (DOMA by COMTs and metanephrines by MAOs) and excreted. EPI, epinephrine; NE, norepinephrine; VMAT, vesicular monoamine transporter; DOMA, 3,4-dihydroxymandelic acid; MAO, monoamine oxidase; COMT, catechol methyltransferases; VMA, vanillylmandelic acid.

Epidemiology

PPGLs occur in 0.2% to 0.6% of patients with hypertension, and their annual incidence is approximately estimated to be 0.8 per 100,000 patient-years. PPGLs most commonly present between fourth and fifth decade of life and do not have a sex predilection.[27]

About 98% of tumors are located within the abdomen, with 85% originating from the adrenal gland. Malignancy, defined by the occurrence of metastases, occurs in 10% of pheochromocytomas, with paragangliomas having a higher rate of malignancy (35%). Paragangliomas may be located in multiple sites throughout the body.

Clinical Presentation

PPGLs are diagnosed in four different clinical situations: (1) patient presents with symptoms, (2) during evaluation of an adrenal mass, (3) screening of patients with related syndromes, and (4) screening of family members diagnosed with PPGLs.[28]

The symptoms of PPGLs are due to paroxysmal secretion of catecholamines and occur in only 50% of patients. The classic triad includes spells of heart palpitations, headaches, and sweating. Half of these patients have paroxysmal hypertension, while the remaining have sustained hypertension or normal blood pressure. Other symptoms include pallor, hand tremors, shortness of breath, and generalized weakness/fatigue. Patients may also present with symptoms of panic attack, especially with tumors that produce EPI. Less common presentations include weight loss, dizziness, orthostatic hypotension, papilledema, cardiomyopathy, erythrosis, hyperglycemia, or psychiatric diagnoses. PPGLs should also be suspected if a spell occurs during diagnostic procedures like endoscopy, during induction of anesthesia or surgical procedures, and during exposure to foods rich in tyramine or drugs like MAO inhibitors or metoclopramide. In a retrospective study, 40% of patients evaluated for pheochromocytoma met the criteria for panic disorder, compared with 5% of control patients with hypertension.

Diagnosis

Evaluation for PPGLs should be based on clinical suspicion, and testing should be performed in the following scenarios:

1. Symptoms and signs of PPGL, especially paroxysmal
2. Early- or late-onset hypertension, resistant hypertension, or atypical presentation of diabetes mellitus
3. Symptoms of PPGL provoked by drugs

4. Incidentally discovered adrenal nodules (irrespective of hypertension status)
5. Syndromic presentation
6. Idiopathic cardiomyopathy
7. History of gastric stromal tumor or pulmonary chondromas (Carney triad)
8. Personal or family history of PPGL

Interfering Medications

Certain medications may lead to false-positive testing due to (1) direct interference with the measurement method and/or (2) interference with disposition of catecholamines. These medications are acetaminophen, labetalol, sotalol, α-methyldopa, tricyclic antidepressants, buspirone, phenoxybenzamine, MAO inhibitors, sympathomimetics, cocaine, sulfasalazine, and levodopa. A review of the patient's prescribed medications should be performed prior to testing to avoid false-positive test results.

Biochemical Testing

A high diagnostic sensitivity is of utmost importance to avoid missing a diagnosis of PPGL. Plasma-free metanephrines or 24-hour urine fractionated metanephrines and catecholamines may be used for initial testing. Some experts recommend using plasma levels for patients suspected to have a high risk of PPGL and a 24-hour urine collection for those with low risk of PPGL.

Plasma-Free Metanephrines. The guidelines on PPGL by The Endocrine Society recommend venipuncture for plasma-fractionated metanephrines to be measured with the patient in supine position for at least 30 minutes prior to blood sampling. Evidence suggests that the upright posture activates the sympathetic nervous system, stimulating the release of NE and subsequent metabolism to normetanephrines, leading to false-positive results. Most laboratories use high-performance liquid chromatography or radioimmunoassays. This yields a high sensitivity of 96% to 100% but a poor specificity of 85% to 89%, which is even lower in patients older than 60 years at 77%.[29,30]

24-Hour Urine Fractionated Metanephrines and Catecholamines. The collection of 24-hour urine should be performed appropriately, and patients must be counseled on the proper method of collection. Urine volume and creatinine can be quantified to verify adequacy of collection. Most laboratories use high-performance liquid chromatography (HPLC) with electrochemical detection or tandem mass spectroscopy (MS/MS). This yields a high sensitivity of 98% and specificity of 98%.

The 24-hour VMA and plasma chromogranin A (neuroendocrine protein) tests are not routinely measured since they have poor sensitivity and specificity unless they are ordered in conjunction with other tests.

Result Interpretation

Case Detection

If testing yields normal results, no further testing is indicated. An exception is for patients who present with spells: repeat testing using a 24-hour urine collection should be performed during a spell. If one test is equivocal (less than threefold elevation above the upper limit of normal), a different test should be performed if clinical suspicion remains high.

If results of the preceding tests are equivocal, a clonidine suppression test (92% accurate) should be performed to separate patients with pheochromocytoma from those without pheochromocytoma who are experiencing similar symptoms secondary to increased sympathetic outflow.[31] Clonidine, an antihypertensive agent, acts as an α2-adrenoceptor agonist and inhibits release of neuronal NE in patients without PPGL alone.

Sympatholytic drugs should be stopped at least 48 hours prior to testing. Baseline plasma normetanephrine is measured, oral clonidine is given at 300 μg/70 kg body weight, and plasma normetanephrines are measured 3 hours later. If the 3-hour plasma normetanephrine is elevated and decreases by < 40% compared to baseline level, the test is abnormal, indicating a PPGL.

In general, the metanephrines and catecholamines should be elevated to above threefold upper limit of normal for the particular lab value to proceed with radiological investigation. A retrospective review of 1896 patients reported a false-positive rate of 19% to 21% for both plasma-free and 24-hour urine fractionated metanephrines.[32] The cutoff values for a positive case detection with 24-hour urine fractionated metanephrine and catecholamine testing are:

Normetanephrine	> 900 μg/24 h or metanephrine > 400 μg/24 h
Norepinephrine	> 170 μg/24 h

Epinephrine	> 35 µg/24 h
Dopamine	> 700 µg/24 h

Plasma-Fractionated Metanephrines

The normal ranges for plasma metanephrines and normetanephrines depend upon the method used to obtain the blood sample. In a sample obtained following an overnight fast, via an indwelling cannula placed for 20 minutes before the blood draw, the ranges are:

Metanephrine	< 0.3 nmol/L
Normetanephrine	< 0.66 nmol/L

In a sample obtained via venipuncture in a seated, ambulant, nonfasting patient, the ranges are:

Metanephrine	< 0.5 nmol/L
Normetanephrine	< 0.9 nmol/L

Radiographic Localization

Biochemical confirmation of pheochromocytoma should be followed by imaging for tumor localization.[33] Ninety-five percent of the PPGLs are in the abdomen or pelvis. The most common sites for paragangliomas are the superior and inferior para-aortic areas (75%), bladder (10%), thorax (10%), and head, neck, and pelvis (5%).

Either a computed tomography (CT) or magnetic resonance imaging (MRI) of the abdomen/adrenals is a reasonable option; both have a high sensitivity of 98% to 100% to detect sporadic pheochromocytoma but a low specificity of 70% due to presence of adrenal incidentalomas. Although the choice of imaging depends upon multiple factors, such as radiation exposure, contrast agents, availability, and cost, CT is favored due to its excellent spatial resolution and is the first choice for imaging the thorax and pelvis, while MRI is preferred for skull base and neck paraganglioma.

Treatment

Management of PPGL involves surgical removal of the tumor after appropriate medication therapy. It is extremely important to control blood pressure, heart rate, and volume status prior to therapy.

Surgical excision is a high-risk procedure; patients with severe preoperative hypertension, highly secretory tumors, or those undergoing repeat intervention are at highest risk for complications. After a successful surgery, plasma catecholamines usually normalize within a week.

Outcome, Prognosis, and Follow-up

Surgical removal of a pheochromocytoma is the primary therapy; nevertheless, excision does not necessarily lead to long-term cure of pheochromocytoma or hypertension (even in patients with benign tumors). Patients with familial pheochromocytomas, right adrenal tumors, and extraadrenal tumors are more likely to have recurrence. In one retrospective series of 176 patients, pheochromocytoma recurred in 16% (52% of those were malignant). Hence, long-term monitoring is indicated in all patients, even those who seem cured.

Genetic Testing

Although most PPGLs are sporadic, up to 40% of them are familial and associated with a syndrome.[34] All the syndromes have autosomal dominant inheritance with frequencies (in parentheses) and mutations as described below:

- Multiple endocrine neoplasia 2 (50%): mutations in the *RET* proto-oncogene
- von Hippel-Lindau syndrome (10% to 20%): mutations in the *VHL* tumor suppressor gene
- Neurofibromatosis type 1 (0.1% to 5.7%): mutations in the *NF1* gene

It is recommended to proceed with genetic testing in the following scenarios:

- Paraganglioma
- Bilateral pheochromocytoma
- Unilateral pheochromocytoma and a family history of PPGL
- Unilateral pheochromocytoma onset at age < 45 years
- Clinical findings suggestive of one of the syndromes

Genetics is a rapidly changing field, and hence for the latest information, please refer to www.genetests.org.

Adrenal Incidentaloma

An adrenal incidentaloma is defined as an adrenal mass > 1 cm, which is discovered on imaging performed to evaluate a nonadrenal pathology.[35] This is the result of advancement in medical imaging technology and the widespread use of imaging studies. The prevalence on an abdominal CT is 4.4% and increases to 10% in the elderly, with 10% to 15% being bilateral nodules. Autopsies report a

Incidentaloma Function Evaluation		
	Clinical Features	**Screening Tests** **Negative Results**
Pheochromocytoma	**HTN** (paroxysmal) with spells (sweating, HA, or palpitations)	24-hour urine metanephrines <1 µg or 5.5 µmol/mg creatine
Cushing's syndrome	**HTN**, obesity (truncal) weakness	1 mg bedtime dexamethasone 8:00 am cortisol < 3.6 µg/dL, or 24-hour urine free cortisol normal
Primary aldosteronism	**HTN**, hypokalemia, weakness	Serum potassium, *if low* urine K^+ excretion (<30 mmol)
Adrenocarcinoma	Virilization (+ above)	Plasma renin:aldosterone ratio < 30 Plasma DHEAS (<9.2 µmol/L) Urine 17-ketosteroid < 20 mg

Figure 16.17 Brief functional screen for adrenal masses. HTN, hypertension; HA, hyperaldosteronism; DHEA-S, dehydroepiandrosterone sulfate.

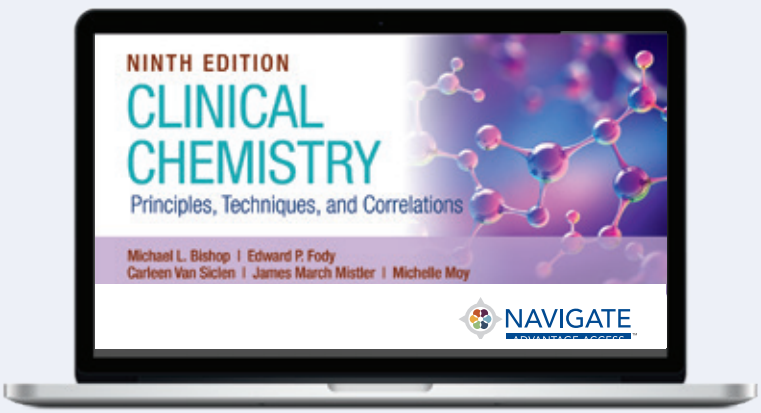
prevalence of 1% to 9%, with higher prevalence among people who have obesity, diabetes mellitus, and hypertension. Workup requires biochemical testing to evaluate for hyperfunction, hypofunction, and imaging to define tumor characteristics (benign or malignant).

An adrenal protocol abdominal CT or MRI is the imaging of choice to determine tumor characteristics, while fluorodeoxyglucose (FDG)-positron emission tomography (PET) scanning is used selectively. Size of the nodule is important since 90% of the adrenocortical carcinomas are > 4 cm and smaller size at diagnosis correlates to better prognosis. Fine needle aspiration cytology, if performed, should be done after ruling out a pheochromocytoma. It will help distinguish between adrenal and extra-adrenal malignancy but cannot distinguish an adrenal adenoma from a carcinoma. It can also help diagnose infiltrative disorders (amyloidosis, sarcoidosis) and infectious diseases (histoplasmosis, tuberculosis).

In case of bilateral lesions, in addition to routine workup, appropriate biochemical testing should be performed to evaluate for adrenal insufficiency. **Figure 16.17** illustrates a brief functional screen for adrenal masses, which assesses the function of all adrenal layers and serves as a clinical summary.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Kacsoh D. *Endocrine Physiology*. New York, NY: McGraw-Hill; 2000.
2. Clark O, Kebebew E, Clark OH. *Textbook of Endocrine Surgery*. Philadelphia, PA: Elsevier Saunders; 2005:557.
3. Miller W, Chrousos GP. The adrenal cortex. In: Felig P, Frohman LA, eds. *Endocrinology and Metabolism*. 4th ed. New York, NY: McGraw-Hill; 2001:387–493.
4. Lacroix A. In: Pinchera A, ed. *Endocrinology and Metabolism*. London, UK: McGraw-Hill; 2001:285–297.
5. Merke DP, Auchus RJ. Congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *N Engl J Med*. 2020 Sep 24;383(13):1248–1261. doi: 10.1056/NEJMra1909786. PMID: 32966723.
6. Puar, TH, Mok Y, Debajyoti R, Khoo J, How CH, Ng AK. Secondary hypertension in adults. *Singapore Med J*. 2016;57(5):228–232. doi:10.11622/smedj.2016087
7. Funder JW, Carey RM, Mantero F, et al. The management of primary aldosteronism: case detection, diagnosis, and treatment: an endocrine society clinical practice guideline. *J Clin Endocrinol Metab*. 2016;105(5):1889–1916. doi:10.1210/jc.2015-4061
8. Weinberger MH, Fineberg NS. The diagnosis of primary aldosteronism and separation of two major subtypes. *Arch Intern Med*. 1993;153(18):2125–2129.
9. England RW, Geer EB, Deipolyi AR. Role of venous sampling in the diagnosis of endocrine disorders. *J Clin Med*. 2018;7(5):114. doi:10.3390/jcm7050114
10. Meyer A, Brabant G, Behrend M. Long-term follow-up after adrenalectomy for primary aldosteronism. *World J Surg*. 2005;29(2):155–159. doi:10.1007/s00268-004-7496-z
11. Nicolaides NC, Chrousos GP, Charmandari E. Adrenal insufficiency. [Updated 2017 Oct 14]. In: Feingold KR, Anawalt B, Boyce A, et al., eds. *Endotext* [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279083/
12. Erichsen MM, Løvås K, Skinningsrud B, et al. Clinical, immunological, and genetic features of autoimmune primary adrenal insufficiency: observations from a Norwegian registry. *J Clin Endocrinol Metab*. 2009;94(12):4882–4890. doi:10.1210/jc.2009-1368
13. Dickstein G, Shechner C, Nicholson WE, et al. Adrenocorticotropin stimulation test: effects of basal cortisol level, time of day, and suggested new sensitive low dose test. *J Clin Endocrinol Metab*. 1991;72(4):773–778. doi:10.1210/jcem-72-4-773
14. Staub JJ, Noelpp B, Girard J, et al. The short metyrapone test: comparison of the plasma ACTH response to metyrapone and insulin-induced hypoglycaemia. *Clin Endocrinol (Oxf)*. 1979;10(6):595–601.
15. Nieman LK, Biller BM, Findling JW, et al. The diagnosis of Cushing's syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2008;93(5):1526–1540. doi:10.1210/jc.2008-0125
16. Elamin MB, Murad MH, Mullan R, et al. Accuracy of diagnostic tests for Cushing's syndrome: a systematic review and metaanalyses. *J Clin Endocrinol Metab*. 2008;93(5):1553–1562. doi:10.1210/jc.2008-0139
17. Mericq MV, Cutler GB. High fluid intake increases urine free cortisol excretion in normal subjects. *J Clin Endocrinol Metab*. 1998;83(2):682–684. doi:10.1210/jcem.83.2.4555
18. Carroll BJ, Curtis GC, Davies BM, et al. Urinary free cortisol excretion in depression. *Psychol Med*. 1976;6(1):43–50.
19. Tyrrell JB, Findling JW, Aron DC, et al. An overnight high-dose dexamethasone suppression test for rapid differential diagnosis of Cushing's syndrome. *Ann Intern Med*. 1986;104(2):180–186.
20. Carroll T, Raff H, Findling JW. Late-night salivary cortisol for the diagnosis of Cushing syndrome: a meta-analysis. *Endocr Pract*. 2009;15(4):335–342. doi:10.4158/EP09023OR
21. Bruno OD, Rossi MA, Contreras LN, et al. Nocturnal high-dose dexamethasone suppression test in the aetiological diagnosis of Cushing's syndrome. *Acta Endocrinol*. 1985;109(2):158–162.
22. Oldfield EH, Doppman JL, Nieman LK, et al. Petrosal sinus sampling with and without corticotropin-releasing hormone for the differential diagnosis of Cushing's syndrome. *N Engl J Med*. 1991;325(13):897–905. doi:10.1056/NEJM199109263251301
23. Bigos ST, Somma M, Rasio E, et al. Cushing's disease: management by transsphenoidal pituitary microsurgery. *J Clin Endocrinol Metab*. 1980;50(2):348–354. doi:10.1210/jcem-50-2-348
24. Nieman LK, Biller BM, Findling JW, et al. Treatment of Cushing's syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2015;100(8):2807–2831. doi:10.1210/jc.2015-1818
25. Fitzgerald PA. In: *Gardner D*, Shoback D, eds. *Greenspan's Basic & Clinical Endocrinology*. 9th ed. China: McGraw-Hill; 2011:345–356.
26. Pacak, K, Chrousos, G, Koch, C, Lenders, J, and Eisenhofer, G. Pheochromocytoma: progress in diagnosis, therapy, and genetics. In: Margioris AN, Chrousos GP, eds. *Adrenal Disorders. Contemporary Endocrinology*. Totowa, NJ: Humana Press; 2001:479–513.
27. Martucci VL, Pacak K. Pheochromocytoma and paraganglioma: diagnosis, genetics, management, and treatment. *Curr Probl Cancer*. 2014;38(1):7–41. doi:10.1016/j.currproblcancer.2014.01.001
28. Lenders JW, Duh QY, Eisenhofer G, et al. Pheochromocytoma and paraganglioma: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2014;99(6):1915–1942. doi:10.1210/jc.2014-1498
29. Lenders JW, Pacak K, Walther MM, et al. Biochemical diagnosis of pheochromocytoma: which test is best? *JAMA*. 2002;287(11):1427–1434.
30. Sawka AM, Jaeschke R, Singh RJ, et al. A comparison of biochemical tests for pheochromocytoma: measurement of fractionated plasma metanephrines compared with the combination of 24-hour urinary metanephrines and catecholamines. *J Clin Endocrinol Metab*. 2003;88(2):553–558. doi:10.1210/jc.2002-021251
31. Eisenhofer G, Goldstein DS, Walther MM, et al. Biochemical diagnosis of pheochromocytoma: how to distinguish true from false-positive test results. *J Clin Endocrinol Metab*. 2003;88(6):2656–2666. doi:10.1210/jc.2002-030005
32. Yu R, Wei M. False positive test results for pheochromocytoma from 2000 to 2008. *Exp Clin Endocrinol Diabetes*. 2010;118(9):577–585. doi:10.1055/s-0029-1237699

33. Allen BC, Francis IR. Adrenal imaging and intervention. *Radiol Clin North Am.* 2015;53(5):1021–1035. doi:10.1016/j.rcl.2015.05.004
34. Rana HQ, Rainville IR, Vaidya A. Genetic testing in the clinical care of patients with pheochromocytoma and paraganglioma. *Curr Opin Endocrinol Diabetes Obes.* 2014;21(3):166–176. doi:10.1097/MED.0000000000000059
35. Ioachimescu AG, Remer EM, Hamrahian AH. Adrenal incidentalomas: a disease of modern technology offering opportunities for improved patient care. *Endocrinol Metab Clin North Am.* 2015;44(2):335–354. doi:10.1016/j.ecl.2015.02.005

CHAPTER 17

Gonadal Function

Jane R. Semler

CHAPTER OUTLINE

KEY TERMS

Amenorrhea
Androgen
Anosmia
Corpus luteum
Dehydroepiandrosterone (DHEA)
Follicle-stimulating hormone (FSH)
Follicular phase
Graafian follicle
Gynecomastia
Hirsutism
Hypogonadism
Inhibin
Leydig cells
Luteal phase
Luteinizing hormone (LH)
Menorrhagia
Menopause
Oligomenorrhea
Ovulation
Sertoli cells
Virilization

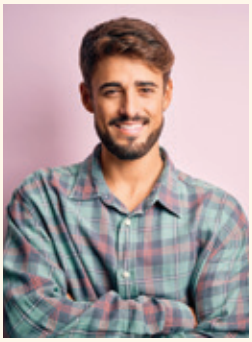
CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- State the principal androgen and its function.
- Name three hormones produced by the ovaries and their function.
- Diagram the menstrual cycle including the follicular phase and luteal phase.
- Discuss the biosynthesis, secretion, transport, and action of the sex steroids and gonadotropins discussed in this chapter.
- Identify the location of the pituitary gland, ovaries, and testes.
- Differentiate the hypothalamic–pituitary–ovarian and hypothalamic–pituitary–testicular axes and how they regulate sex steroid and gonadotropin hormone production.
- Explain the principles of each diagnostic test for pituitary–gonadal axes dysfunction discussed in this chapter.
- Interpret laboratory results with regard to suspected gonadal disorders, given a patient's presentation and clinical data.
- Select appropriate laboratory tests to effectively evaluate or monitor patients with suspected gonadal disease.
- Summarize causes of infertility in both males and females.

CASE STUDY 17.1, PART 1

Roger, a 23-year-old male, presented to his physician with a chief complaint of infertility. Roger and his wife have been attempting to conceive for 3 years without result. Roger's wife had a full OB-GYN workup that found normal ovarian function and regular menstrual cycles.



© Krakenimages.com/Shutterstock.

CASE STUDY 17.2, PART 1

Karen, a 17-year-old female, presented to her pediatrician for an annual well exam. Patient history reveals that the girl was born at term following an unremarkable pregnancy and vaginal delivery at home with a midwife.

© Klaus Vedfelt/DigitalVision/Getty Images.

The Testes

By the sixth week of development in both sexes, the primordial germ cells have migrated from their extraembryonic location to the gonadal ridges, where they are surrounded by the sex cords to form a pair of primitive gonads. Whether chromosomally 46,XX or 46,XY, the developing gonad at this stage of development is bipotential.

The ovarian pathway is followed unless a gene on the short arm of the Y chromosome, designated TDF (testis determined factor), also known as the sex-determining region Y (*SRY*) gene, makes a DNA-binding protein that acts as a switch, diverting development into the male pathway.[1] In the presence of the Y chromosome (due to the SRY protein), the medullary tissue starts forming typical testes with seminiferous tubules and **Leydig cells** by 7 weeks of gestation. Leydig cells under the stimulation of human chorionic gonadotropin (hCG) from the placenta become capable of **androgen** secretion by week 10. The spermatogonia, derived from the primordial germ cells, form the walls of the seminiferous tubules together with supporting Sertoli cells.

There are three important steps in sexual differentiation and the development of the normal male phenotype.[2] The first is the differentiation of the bipotential gonad primordium (identical in both XX and XY fetuses) into the testes. The testes will secrete testosterone, which is under the control of the SRY protein coded by the Y chromosome gene. The second is the development of the internal reproductive tract, initiated by fetal testicular androgen production. In male fetuses, this requires the presence of antimüllerian hormone (AMH), which causes involution of the müllerian ducts; these ducts are the rudimentary organs for the female reproductive tract.[3] The third is the development of the external genitalia that requires testosterone and, in some target tissues, its more potent metabolite 5α-dihydrotestosterone (DHT).

Functional Anatomy of the Male Reproductive Tract

Adult testes are paired, ovoid organs that hang from the inguinal canal by the spermatic cord, which comprises a neurovascular pedicle, vas deferens, and cremasteric muscle. The testes are located outside the body, encased by a muscular sac. Blood flow is governed by an intricate plexus of arterial and venous blood flow that, together with contraction of the muscles in the scrotal sac, regulates the temperature of the testicles to 2°C below core body temperature. This important function is vital to uninterrupted sperm production. Also encased in the muscular sheath is the spermatic cord, which has the ability to retract the testicles into the inguinal canal in instances of threatened injury. The testes themselves are comprised of two anatomical units: a network of tubules, known as the seminiferous tubules, and an interstitium. The tubules contain germ cells and Sertoli cells and are responsible for sperm production. The testes serve dual functions, which include (1) production of sperm and (2) production of reproductive steroid hormones.[1] In the embryonic stage, the dominant male sex hormone, testosterone (T), aids in development and differentiation of the primordial gonads.

Puberty marks the transition from a nonreproductive state into a reproductive state and is associated with adrenarche, characterized by an increase in adrenal androgens and gonadarche. Tanner staging of the pubertal and axillary hair changes

and genitalia is used in children to mark pubertal changes, which are characterized by growth spurt, increased muscle mass, psychological changes, and male pattern hair growth. The earliest sign of puberty in boys is testicular enlargement that results from rising **luteinizing hormone (LH)** and **follicle-stimulating hormone (FSH)**. After puberty, throughout adulthood, and until late in old age, testosterone helps with sperm production and maintains secondary sexual characteristics.

The sperm move sequentially through the tubuli recti; rete testes; ductuli efferentes testes; the head, body, and tail of the epididymis; and, finally, into the vas deferens. Various secretory products of the seminal vesicles and prostate mix with sperm to form the final product: semen. Seminal vesicle secretions are rich in vitamin C and fructose, important for the preservation of motility of the sperm.

Physiology of the Testicles

Spermatogenesis

Sperm are formed from stem cells called *spermatogonia*. The spermatogonia undergo mitosis and meiosis; finally, the haploid cells transform into mature sperm. The mature sperm has a head, body, and tail, which enables it to swim for the purpose of fertilizing the haploid ovum. Certain spermatogonia stagger division so that sperm production is uninterrupted and continuous. The **Sertoli cells** are polyfunctional cells that aid in the development and maturation of sperm.[4–6]

Hormonogenesis

The testes produce 95% of circulating testosterone in males. This production is regulated by the anterior pituitary hormones FSH and LH.[7] Because these hormones were first described in women, they are named in reference to the menstrual cycle. Both hormones are produced by a single group of cells in the anterior pituitary called *gonadotrophs*. FSH acts primarily on germinal stem cells,[8] and LH acts primarily on the Leydig cells[7]—located in the testicular interstitium—that synthesize testosterone. Gonadotropins (LH and FSH) are glycoproteins and share an α subunit with thyroid stimulating hormone (TSH) and hCG. The β subunit found on FSH and LH confers their biological specificity. The remaining ~5% of testosterone is produced in peripheral circulation from adrenal **dehydroepiandrosterone (DHEA)** and androstenedione. These testosterone precursors are produced in the zona reticularis of the adrenal glands.[9]

Hypothalamic–Pituitary–Testicular Axis

The hypothalamic–pituitary–testicular axis is depicted in **Figure 17.1**. The hypothalamus, located in the brain, releases gonadotropin-releasing hormone (GnRH) in a pulsatile pattern in conjunction with diurnal variation (circadian rhythm). A higher concentration is present in the morning due to a nocturnal surge of GnRH. The concentration wanes throughout the afternoon. GnRH is a peptide hormone synthesized in neurons situated in the arcuate nucleus and other nuclei of the hypothalamus, and is released into the portal hypophyseal system that, in turn, stimulates production of LH and FSH from the anterior pituitary gland. Impaired pulse generation of GnRH leads to inadequate production of LH and FSH, resulting in **hypogonadism**.[10] The first, and rate-limiting, step in the testicular steroidogenesis is the conversion of cholesterol to pregnenolone. This cholesterol is either trapped by endocytosis from the blood lipoproteins or synthesized within the Leydig cells. The LH binds to the glycoprotein receptor in the cell membrane and induces intracellular cyclic AMP production that, in turn, activates protein kinase A, which catalyzes protein phosphorylation. This latter step induces testosterone synthesis. The testicular steroidogenesis pathway is similar to the pathway in the adrenal cortex, and they share the same enzymatic systems. Testosterone is the principal androgen hormone in men. After puberty, the testes secrete 4 to 10 mg daily. About 50% of testosterone in circulation is bound to albumin while about 45% is bound to sex hormone–binding globulin (SHBG).

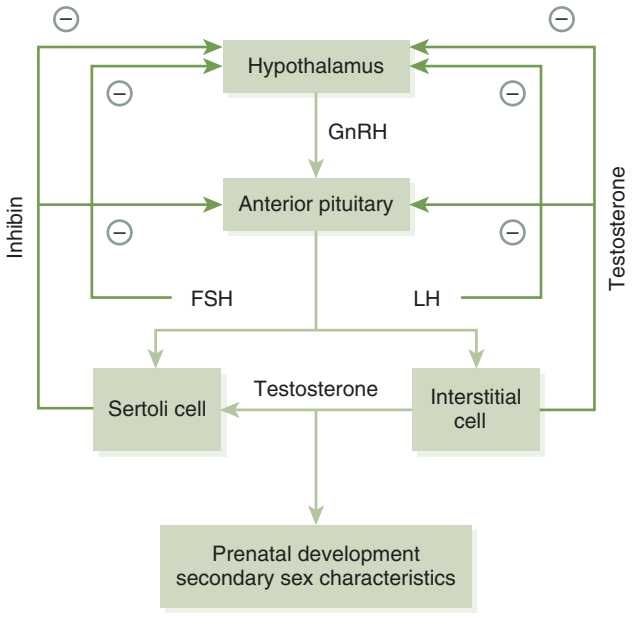


Figure 17.1 The hypothalamic–pituitary–testicular axis.

The remaining 2% to 3% of testosterone circulates free of a carrier protein. The concentration of binding proteins affect measurement of total testosterone but do not affect the laboratory estimation of free testosterone. Testosterone and **inhibin** are the two hormones secreted by the testes that provide feedback control to the hypothalamus and pituitary. Leydig cell steroidogenesis of testosterone is primarily stimulated by LH secretion and inhibited by testosterone and estradiol, which provide negative feedback to the hypothalamus and anterior pituitary. FSH acts on Sertoli cells to stimulate protein synthesis and release of inhibin and androgen-binding protein. The actions of both testosterone and FSH on Sertoli cells are synergistic, permitting completion of spermatogenesis. FSH stimulates production of an androgen receptor that makes the Sertoli cell responsive to androgen, and, in turn, androgens stimulate synthesis of FSH receptors. Inhibin provides negative feedback to the anterior pituitary to regulate secretion of FSH.[11]

Testosterone concentration fluctuates in a diurnal fashion, reflecting the parallel rhythms of GnRH, LH, and FSH secretion. This fact should be considered when interpreting serum levels of testosterone: the highest level is found at about 6:00 am and correlates with most laboratory reference ranges, and the lowest level is found at about 12:00 am in males with a normal sleep–wake cycle.

Cellular Mechanism of Testosterone Action

Once inside a cell, testosterone is metabolized to one of two active metabolites: dihydrotestosterone (DHT) or estradiol. In the androgen target tissues, primarily the skin and prostate, 5α-reductase converts testosterone to DHT. Both DHT and testosterone can bind to an intracellular androgen receptor protein. Binding of this complex to the nuclear receptor effects protein synthesis and cell growth. DHT binds to the androgen receptor with higher affinity and, thus, is a more potent metabolite as compared to testosterone. In non-androgen target tissues, testosterone and androstenedione are aromatized to estradiol. This occurs primarily in adipose tissue, which is rich in aromatase. The rate of estradiol formation in males increases in proportion to body fat. The waste products of testosterone metabolism, known as 17-ketosteroids, are excreted in the urine.[12]

Physiologic Actions of Testosterone

Prenatal Development. Early in development, embryos have primordial components of the genital tracts of both sexes. The primitive gonads become distinguishable at about the seventh week of embryonic stage. Both chorionic gonadotropins and fetal LH stimulate production of testosterone by the fetal Leydig cells. In the fetus, the hypothalmic-hypophyseal vascular connections are responsible for LH release by hypothalamic GnRH and are established between 11 and 12 weeks after conception, which is 3 weeks after testosterone production by the Leydig cells of the testis. hCG, which has LH-like action and is made by the placenta, accounts for this gap in stimulation of testosterone production by the fetal testes. Exposure of testosterone to the Wolffian duct leads to differentiation of the various components of the male genital tract. Sertoli cells produce AMH, which aids in regression of the female primordial genital tract. The scrotal skin is rich in 5α-reductase, which converts testosterone to DHT. Fetal exposure to drugs that block this enzyme and DHT formation leads to feminization of the genitalia of the male fetus.

Postnatal Development. Testicular function is reactivated during puberty after a period of quiescence to produce testosterone that results in development of secondary sex characteristics such as hair growth (face, chest, axilla, and pubis), enhanced linear skeletal growth, development of internal and external genitalia, increased upper body musculature, and development of larynx and vocal cords with deepening of the voice.[13–15] Possible mood changes and aggression are undesired effects that may occur during puberty. The linear growth effects of testosterone are finite, with epiphyseal closure when genetically determined height is achieved. Hypogonadism during puberty leads to imprecise closure of growth plates, leading to excessive height, long limbs, and disproportionate upper and lower body segments. Male secondary sexual characteristics can be staged by a system of development devised by Marshall and Tanner (**Table 17.1**).

Effect on Spermatogenesis. Stimulation of Leydig cells induces production of testosterone. Testosterone, acting with FSH, has paracrine effects on the seminiferous and Sertoli cells, inducing spermatogenesis. Exogenous overuse or abuse of testosterone, such as occurs with some athletes, will reduce the high intratesticular concentration of testosterone, leading to reduction of sperm production.

Effect on Secondary Sexual Effects. Testosterone has growth-promoting effects on various target tissues. The secondary sex characteristics that develop during puberty are maintained into late adulthood by testosterone.[17] The prostate enlarges

CASE STUDY 17.1, PART 2

Remember Roger. His history reveals that he was born at term following an unremarkable pregnancy and vaginal delivery. He has no siblings and is uncircumcised. He was diagnosed with micropenis (3.1 cm) as a child. He is unaware of his parents' medical history. Roger admits to undergoing two surgical procedures during childhood. The first surgery was to remove an undescended testicle, and the second surgery corrected excessive lacrimation (tearing of the eyes). His weight is 118 lbs., height 5 feet 5 inches, arm span 5 feet 11 inches, crown to pubis 2 feet 5 inches, pubis to heel 3 feet 2 inches.[16,17]

1. How does Roger's medical history contribute to the assessment of infertility?
2. Define eunuchoid.
3. Do Roger's characteristics meet the definition of eunuchoid status?

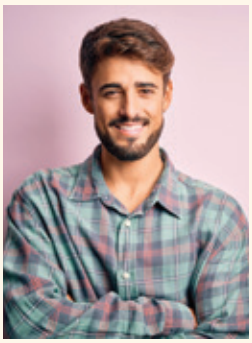


Table 17.1 Tanner Staging of Genital and Pubic Hair Development in Males

Stages of Genital Development
1. Prepubertal
2. Enlargement of the scrotum and testes
3. Increased length of the penis, further enlargement of the testes
4. Enlargement of the testes, scrotum, and penis with growth of glans; darkening of scrotal skin
5. Mature genitalia
Stages of Pubic Hair Development
1. Lanugo-type hair (prepubertal)
2. Dark terminal hair at the base of the penis
3. Darker terminal hair spreading over junction of pubis
4. Terminal hair covering pubic region, no spread to medial thighs
5. Mature stage with horizontal distribution of terminal hair to inner thighs

progressively during adulthood, while exposure of scalp hair results in regression of the hair follicles (temporal hairline recession). Failure to develop secondary sexual characteristics should prompt evaluation for hypogonadism or constitutional delay in boys. Loss of secondary sexual characteristics might occur gradually and should prompt evaluation for hypogonadism because, among other effects, low testosterone levels lead to loss of bone mass and development of osteoporosis in males at any age.

Disorders of Sexual Development and Testicular Hypofunction

Pubertal development could be premature (precocious) or delayed, even if development is normal at birth.[18,19] Precocious sexual development results from the premature exposure of sex steroids, which might arise from early gonadotropin secretion or over production by the adrenal glands or testes. Detailed descriptions of the sequence of hormonal pubertal abnormalities of hair, genitals, and breasts are beyond the scope of this text. The differential diagnosis of hypogonadism includes a diverse group of disorders affecting the testicles and the hypothalamic–pituitary regulation of the testes outlined in **Box 17.1**. Certain important disorders are explained in the following section.

Hypergonadotropic Hypogonadism

Hypergonadotropic hypogonadism is the result of primary gonadal insufficiency. This group of disorders is characterized by low testosterone, elevated FSH or LH, and impaired sperm production.

Klinefelter's Syndrome. Klinefelter's syndrome occurs in about 1 of 400 to 600 men and is caused by the presence of an extra chromosome; in fact, it is the most common human sex chromosome abnormality. The most common karyotype is 47,XXY.[21] Men with this disorder have small (<2.5 cm), firm testicles. **Gynecomastia** (enlargement of the male breast) is commonly present at the time of diagnosis. Due to reduced production of testosterone, LH levels are elevated. Underproduction of inhibin as a result of deficient seminiferous tubule mass causes elevation of FSH.[21] Patients with Klinefelter's syndrome also have azoospermia and are sterile. Patients with a mosaic

Box 17.1 Causes of Delayed Puberty

- Delayed puberty or hypogonadism, with increased gonadotropins (FSH and/or LH)
- Klinefelter's syndrome
- Bilateral gonadal failure
- Primary testicular failure
- Anorchia
- Vanishing testicles
- Chemotherapeutic agents
- Irradiation
- Trauma
- Infection (mumps orchitis)
- Delayed puberty with normal or low FSH and/or LH
- Constitutional delayed puberty[20]
- Hypothalamic dysfunction
- Malnutrition
- Chronic systemic illness
- Severe obesity
- Central nervous system tumors
- Hypopituitarism
- Panhypopituitarism
- Kallmann's syndrome (**anosmia**, cleft palate, and reduced FSH and LH levels)
- Isolated GH deficiency
- Hyperprolactinemia (prolactinoma or drug induced)
- Hypothyroidism
- Miscellaneous
- Prader-Willi syndrome
- Laurence-Moon syndrome
- LEOPARD syndrome
- Bloom syndrome
- Germ cell neoplasia
- Male pseudohermaphroditism
- Ataxia–telangiectasia
- Steroidogenic enzyme defects

form of Klinefelter's syndrome may be fertile. Because elevated levels of FSH and LH induce increased aromatase activity resulting in elevated estrogen levels, men with Klinefelter's syndrome may have reduced bone density and increased breast cancer risk.[22]

Androgen Insensitivity Syndrome. Androgen insensitivity syndrome (AIS) is the most severe form of androgen resistance syndrome. AIS results from mutations of the androgen receptor gene. Depending on the amount of androgen receptor present, AIS may be complete or partial. Complete AIS (CAIS), formerly known as testicular feminization syndrome, results in physical development of the female phenotype despite male karyotypic 46,XY status. Patients with CAIS have female external genitalia and characteristic female distribution of fat and hair. The testicles are often undescended, and prompt surgical removal of these organs is required to avoid malignant transformation. Müllerian-inhibiting hormone produced by Sertoli cells prevents development of the upper vagina, uterus, and fallopian tubes. Most patients present for evaluation of primary amenorrhea, at which time the lack of female internal genitalia becomes apparent. Biochemical evaluation reveals normal or elevated serum concentrations of testosterone with elevated FSH and LH levels. Administration of exogenous testosterone is not useful, as there is no response to it in these patients. There is wide variation in androgen insensitivity and in corresponding clinical deficits in male sexual development.

5α-Reductase Deficiency. Deficiency of 5α-reductase is a rare cause of androgen insensitivity and results in a mutation encoding the type 2 isoenzyme, maps to chromosome 2p23, and is expressed in XY males. A reduction in levels of the enzyme 5α-reductase results in decreased conversion of testosterone to DHT. Because DHT is essential for the development of the prostate and external genitalia during embryonic virilization, males are born with ambiguous genitalia. Physical development resembles the female phenotype until puberty, when the Wolffian ducts virilize in response to testosterone. The female internal genitalia are absent, and the male internal genitalia are well developed (epididymis, vas deferens, and seminal vesicle) in response to testosterone.

Myotonic Dystrophy. Myotonic dystrophy comes in two forms: type 1 is caused by a *DMPK* gene mutation, and type 2 is caused by a *CNBP* gene mutation. Both are inherited in an autosomal dominant fashion; an affected person almost always has one parent with the condition and presents with primary hypogonadism, frontal balding, diabetes, and muscle weakness, atrophy, and dystonia (an inability of the muscle to relax adequately after contraction). The incidence of myotonic dystrophy is 1 in 8000 worldwide. Testicular failure typically occurs when the patient is in their 20s or 30s; however, puberty progresses normally, and male secondary sexual characteristics, height, bone growth, etc., are attained normally during the course of pubertal development. Primary hypogonadism occurs with primarily germ cell compartment failure (i.e., oligozoospermia and infertility), with elevated serum FSH. In later stages, failure of Leydig cell compartment resulting in low serum testosterone,

elevated LH concentrations, and testicular atrophy occurs. Testicular failure happens more commonly in type 1 myotonic dystrophy than in type 2.

Testicular Injury and Infection. Mumps orchitis develops in one-third of post-pubertal males with *Mumps orthorubulavirus* infection (mumps) and is the most frequent extra-salivary manifestation of this highly contagious infection. Mumps epididymo-orchitis is usually unilateral (85% of the time); therefore, permanent sterility is rare (contrary to popular belief). Most cases of orchitis occur within 4 to 6 days of parotitis, but it may precede parotitis or be the only manifestation of mumps. Intermittently, mumps outbreaks have been reported in secondary schools, colleges, universities, and military groups; however, the absolute numbers are still quite low, at a few hundred cases reported in a year in the United States. Other infectious causes of orchitis like tuberculosis, chlamydia, gonorrhea, etc., are rare in the United States. Radiation and chemotherapy for cancer can result in long-term testicular damage.

Sertoli Cell–Only Syndrome. Sertoli cell–only syndrome (SCO syndrome), or germ cell aplasia, is characterized by a lack of germ cells. Men present with small testes, high FSH levels, azoospermia, and normal testosterone levels. Testicular biopsy, which is necessary for diagnosis, shows lack (or complete absence) of spermatozoa. This disorder may arise from Y chromosome microdeletions on Yq11 locus (AZF region: Azoospermia Factor). It is relatively rare, affecting 5% to 10% of all infertile men in the country.

Hypogonadotropic Hypogonadism

The hallmark of disorders of hypogonadotropic hypogonadism is the occurrence of low testosterone levels together with low or inappropriately normal GnRH, FSH, or LH levels.

Kallmann's Syndrome. Kallmann's syndrome is a result of an inherited, X-linked recessive trait that manifests as hypogonadism during puberty. The impaired secretion of GnRH has been elucidated in the X-linked form of congenital GnRH deficiency, which results from impaired migration of GnRH neurons and olfactory nerves to the ventral hypothalamus during embryogenesis.

Mutations in the *KAL1* gene (a gene found on the X chromosome) occur less often in sporadic cases (<10%). In general, the clinical phenotypes of idiopathic hypogonadotropic hypogonadism (IHH) subjects with *KAL1* mutations are characterized with a high incidence of microphallus, cryptorchidism, and small testes. X-linked *KAL* mutations have never been observed in families with normosmic IHH (nIHH) or in families with both anosmic and nonanosmic individuals.[23,24]

Hyperprolactinemia. Prolactin elevation resulting from any cause (drug-induced or prolactin-producing tumors of the pituitary) can result in hypogonadotropic hypogonadism[25,26] due to impairment of both frequency and amplitude of FSH and LH pulses, because of hypothalamic kisspeptin neuron–mediated disruption of GnRH pulsatile secretion.[27,28]

Type 2 Diabetes. Type 2 diabetes is also associated with hypogonadotropic hypogonadism in at least 25% to 50% of men.[29] It is characterized by low free or total serum concentrations of testosterone and inappropriately low LH.[30] This may stem from insulin resistance (insulin action seems to be important for LH release by gonadotropes), low SHBG levels, inflammation (hypogonadism in type 2 diabetic males was seen to be associated with high *C*-reactive protein levels, an inflammatory marker),[29] and even elevated estradiol levels (due to testosterone aromatization in adipose tissue). Of note, type 1 diabetes does not seem to be associated with hypogonadism, provided glycemic control has remained good and body mass index is within normal limits.[31]

Age. Males undergo a gradual reduction in testosterone after age 30, with an average decline of about 110 ng/dL every decade. The Baltimore Longitudinal Study of Aging revealed "hypogonadism" (reduced total testosterone concentrations) of 19% at age 60, 28% at age 70, and 49% at age 80,[32] with free testosterone concentrations much lower in these men. Age is also associated with elevation of SHBG by about 1% per year. Similar findings emerged from the Massachusetts Male Aging Study, which showed a 1.6% decline per year in total testosterone levels and a 2% to 3% decline per year in bioavailable testosterone (free and albumin-bound testosterone) levels.[33] Results of the Testosterone Trials showed significant improvement in sexual function in men older than 65 years given testosterone replacement for 1 year, but they saw no significant improvement in physical function or vitality.[34]

Pituitary Disease. Acquired hypogonadism can follow injury to the pituitary as a result of tumors, surgical or radiation-induced trauma, vascular injury,

autoimmune hypophysitis, or granulomatous or metastatic disease. Hemochromatosis is a rare cause of pituitary dysfunction.

Opioid Use. Long-term or continuous use of narcotics has been linked to severe hypogonadotropic hypogonadism due to μ-opioid receptor–mediated decreased GnRH pulsatile production. Opioid use is also linked to a decrease in male fertility due to decreased sperm motility, decreased sperm counts, and abnormal sperm morphology.[35]

Obstructive Sleep Apnea. It is unclear if sleep apnea leads to hypogonadotropic hypogonadism (due to hypoxemia and sleep deprivation),[36] or if the obesity that is often present in men with sleep apnea leads to decreased testosterone levels. It is to be noted that testosterone replacement therapy, especially high-dose testosterone, may worsen preexisting sleep apnea, possibly due to increased oxygen consumption. Therefore, the Endocrine Society guidelines[37] recommend starting men with untreated severe obstructive sleep apnea on continuous positive airway pressure (CPAP) before initiating testosterone replacement therapy.

Diagnosis of Hypogonadism

Both clinical and biochemical features must be met to make the diagnosis of hypogonadism (see **Figure 17.2**). Evaluation should begin with the physical examination including body mass index, external genitalia, virilization, breast tissue, neurologic examination for anosmia and visual abnormalities, and thyroid examination. Patient history, including reproductive history, sexual history (including sexually transmitted infections), prior illness, recreational and prescribed drugs, alcohol use, and exposure to toxins, also plays an important role in the diagnosis of hypogonadism. Semen analysis should be performed prior to hormone analysis. If semen analysis is normal, further laboratory testing for male infertility is not indicated. If oligospermia or azoospermia is found, serum testosterone, LH, and FSH should be measured. Testosterone concentrations have a diurnal rhythm, and the time of sampling must be considered. Morning samples between 8:00 and 10:00 am are recommended to match values from reference range studies. Multiple estimation of free and bound testosterone levels should be done on different days before a diagnosis of testosterone deficiency is confirmed.[37] The distinction between primary (disease or destruction of the testes) versus secondary (disease or destruction of the pituitary) or tertiary (hypothalamic) causes is relatively easy to make. FSH and/or LH[38] values are elevated in primary hypogonadism and are inappropriately normal or low with secondary or tertiary etiologies. Pituitary MRI should be done in secondary hypogonadism in young individuals. Older individuals often have secondary or tertiary dysfunction as a result of reduced hypothalamic pulse generator frequency, resulting in low or inappropriately normal FSH and/or LH levels.[39] Clinical signs and symptoms of hypogonadism (e.g., loss of secondary sexual characteristics, decreased muscle mass, osteoporosis, etc.) should be corroborated with low testosterone levels, particularly when testosterone replacement therapy is considered.[37]

Hyperprolactinemia. As previously mentioned, hyperprolactinemia is a cause of secondary hypogonadism. If prolactin is elevated, the patient should also be evaluated for hypothyroidism because low levels of thyrotropin releasing hormone can result in hyperprolactinemia.[40]

CASE STUDY 17.1, PART 3

Recall Roger. His latest physical examination reveals micropenis, monorchism, sparse pubic hair (Tanner stage 2), flat philtrum and thin upper lip, and absence of facial and body hair. Upon questioning, the patient reveals that he has no sense of smell.

4. How do the results of the physical examination contribute to the differential diagnosis for the cause of infertility?
5. State two possible causes of infertility that should be included in the differential diagnosis for this patient.

© Krakenimages.com/Shutterstock.

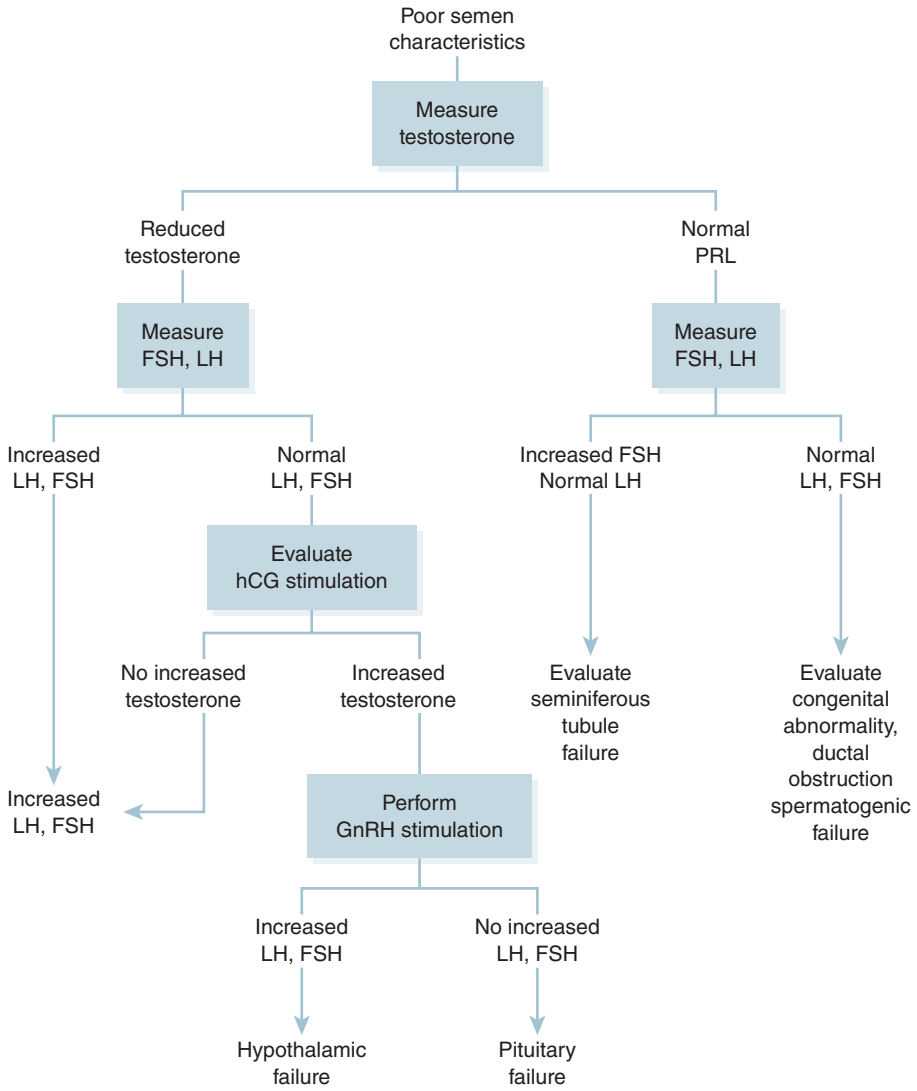


Figure 17.2 Clinical diagnostic evaluation of male hypogonadism. PRL, prolactin; FSH, follicle stimulating hormone; LH, luteinizing hormone; hCG, human chorionic gonadotropin; GnRH, gonadotropin-releasing hormone.

Hypergonadotropic hypogonadism. Elevated FSH in a patient with azoospermia is indicative of Sertoli-only syndrome or Klinefelter syndrome. Elevated FSH combined with decreased testosterone and oligospermia is consistent with primary testicular failure.[41]

Hypogonadotropic hypogonadism. GnRH stimulation may be used to distinguish between secondary and tertiary gonadotropic disorders. Serum LH is measured at 0, 15, 30, and 60 minutes after IV or subcutaneous injection of GnRH. Failure to reach an LH level 2.5 times the baseline (at 0 minutes) is indicative of secondary (pituitary) cause for hypogonadotropic hypogonadism.[42]

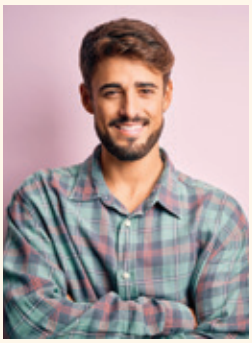
Testosterone Replacement Therapy

The following principles should guide testosterone therapy:

- Testosterone should be administered only to a man who is hypogonadal, as evidenced by clinical symptoms and signs consistent with androgen deficiency and a distinctly subnormal serum testosterone concentration.
- Treating symptoms of hypogonadism without corroborating biochemical evidence of testosterone deficiency is not recommended.
- Restoring testosterone to mid-normal concentrations can be achieved satisfactorily whether the testosterone deficiency is due to primary or secondary hypogonadism.

CASE STUDY 17.1, PART 4

Recall Roger. His doctor ordered an abdominal ultrasonography, which revealed no abnormalities. Scrotal ultrasound revealed the presence of one prepubertal testicle measuring 1.7 × 1.1 × 0.7 cm with a volume <4 mL. Karyotyping demonstrated 46,XY. MRI of the brain revealed normal pituitary gland, hypoplastic left olfactory bulb, and aplastic right olfactory bulb. Laboratory results for Roger are shown in the following table.



© Krakenimages.com/Shutterstock.

Analyte	Patient Value	Reference Range
FSH	0. 7 mIU/mL	1–10.5 mIU/mL
LH	0.4 mIU/mL	1.9–9.4 mIU/mL
Testosterone	1.2 nmol/L	9–38 nmol/L
Prolactin	6 ng/mL	5–25 ng/mL
TSH	1.2 mU/mL	0.3–5.0 mU/mL

6. Which test results are abnormal?
7. Based on the laboratory results, what is the cause for the patient's infertility?

The currently available modes[43] of testosterone administration in the United States are as follows:

1. *Parenteral testosterone.* This is the most widely available and cost-effective mode of administration. The cypionate and enanthate esters of testosterone are available for intramuscular injection. The peak level is achieved in 72 hours and the effect lasts for a period of 1 to 2 weeks. Recently, the FDA has approved extra-long-acting testosterone undecanoate injection, which can be injected every 10 weeks after the first two doses, which need to be injected in a 4-week interval, but this injection is available only under a special program,[44] due to possible risk of serious adverse effects of pulmonary oil microembolism (POME) and anaphylaxis.
2. *Transdermal testosterone patch therapy.* This mode of administration provides more physiologic levels of testosterone.
3. *Testosterone gel.* This hydroalcoholic gel preparation is applied to nongenital skin once daily.
4. *Testosterone buccal pellet.* This plastic tablet is placed along the gum line twice daily. Local discomfort (irritated gums) and the need for twice-daily dosing sometimes limit use.
5. *Subcutaneous testosterone pellet.* This therapeutic modality (Testopel) involves implantation of three to six 75-mg testosterone pellets under the skin.
6. *Nasal testosterone gel.* Testosterone gel (Natesto) is administered into each nostril by a metered-dose pump applicator three times daily.

Oral alkylated androgens are not recommended due to adverse effects of decreased high-density lipoprotein,[45] increase in low-density lipoprotein (LDL), and incidence of cholestatic jaundice and peliosis hepatitis.

Monitoring Testosterone Replacement Therapy

Prostate-specific antigen (PSA), blood counts (for hematocrit), and lipid levels should be checked 3 to 6 months after initiation of testosterone replacement and at least yearly thereafter. Routine clinical evaluation for leg edema, worsening of sleep apnea, and prostate enlargement is also recommended. Pharmacologic use of testosterone may also reduce sperm count by reducing the intratesticular testosterone concentration that is much higher than serum concentrations. If PSA elevation is noted after testosterone replacement, prostate evaluation with possible biopsy is recommended. Prostatic carcinoma is a contraindication to testosterone replacement.

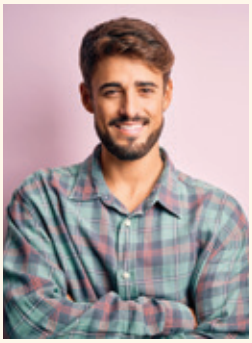
The Ovaries

Early Ovarian Development

If no Y chromosome or TDF is present, the gonad, by default, forms an ovary; the cortex develops, the medulla regresses, and oogonia begin to develop

CASE STUDY 17.1, PART 5

Recall Roger. He underwent GnRH stimulation testing. Serum LH was measured 30 minutes and60 minutes after administration of 100 mg GnRH. Results are shown in the following table.



© Krakenimages.com/Shutterstock.

GnRH Stimulation Test

Time (minutes)	LH	Reference Range
0, GnRH	<2 mU/mL	3–18 mU/mL post stimulation: 2.5 times baseline at some point
30	<2 mU/mL	
60	3 mU/mL	

Following one month of GnRH priming, the GnRH stimulation test was repeated. Results are shown in the following table.

GnRH Simulation After GnRH priming

Time (minutes)	LH	Reference Range
0, GnRH	<2 mU/mL	3–18 mU/mL post stimulation: 2.5 times baseline at some point
30	15 mU/mL	
60	13 mU/mL	

8. Based on the GnRH stimulation tests, what is the cause for the patient's infertility?
9. What do these results indicate regarding the possibility of restoring fertility for this patient?

within follicles. The oogonia are derived from the primitive germ cells by a series of about 30 mitoses, far fewer than the number required for spermatogenesis. Beginning at about the end of the third month, the oogonia enter meiosis I, but this process is arrested at a stage called dictyotene, in which the cell remains until **ovulation** occurs many years later during menstruation (this cell is called primary oocyte). Many of the oogonia degenerate before birth, and only about 400 mature into ova during the 30 years or so of sexual maturity of the female. It is evident that the gonads are paired bilateral structures; in very rare circumstances, the developmental processes may be different on either side. Estrogen formation in the fetal ovary begins in early development despite primordial follicles not having begun forming until the second trimester of pregnancy (at 16 weeks of gestation). The gonadotropins from the pituitary gradually take over the role of maternal placental hCG, and fetal pituitary LH and FSH peak near midgestation and then fall to low concentrations at birth. Postpartum, a smaller peak of LH and FSH occur, which stimulates steroid secretion leading to neonatal milk production from the breast.

Pubertal Changes of Ovarian Function

The onset of puberty is characterized by increasing secretion of LH and FSH that stimulates gonadal activity and is driven by increased activity of the hypothalamic GnRH neurons. In the ovary, both LH and FSH are involved in the control of steroidogenesis. The genes coding LH and FSH receptors are on chromosome 2p21. Inactivating mutations of the β-chain of the LH receptor have been described. Before the onset of puberty, both LH and FSH are secreted in small amounts, but as puberty approaches, the amplitude of LH and FSH pulsatile secretions increases and the nocturnal rise in LH secretion increases. The nocturnal rise disappears with adulthood.

The ovaries are paired organs and, like the male gonads, perform the dual functions of gamete (ovum) and steroid hormone production.[46–48] Unlike in the male, the primordial reproductive cells in the female typically produce a solitary gamete. Ovarian and menstrual events are carefully synchronized by a complex interplay of hormones among the hypothalamus, pituitary, and ovaries to prepare the uterus for implantation of an embryo. In the absence of implantation, the uterine lining is shed, resulting in menses.[49,50]

CASE STUDY 17.2, PART 2

Recall Karen. Karen had several bouts of otitis media throughout childhood and sustained mild hearing impairment. The patient's mother is concerned that her daughter has never had a menstrual period but states that her daughter is an avid runner. Karen appears lean and short in stature with height measuring 4'1" and weight 78 lbs. The patient is below the 5th percentile in height for her age. Her mother is also short and states that the father is 5'3". [51]

1. How does the patient's history contribute to assessment of the lack of pubertal development?
2. Define primary and secondary amenorrhea.
3. Does the patient meet the definition of either primary or secondary amenorrhea?

© Klaus Vedfelt/DigitalVision/Getty Images.

The length of the menstrual cycle is the time between any two consecutive cycles. The typical duration is 25 to 35 days, with average menstrual flow about 3 to 6 days.[50]

Functional Anatomy of the Ovaries

The ovaries are oval organs that lie in the pelvic fossa, formed by the posterior and lateral pelvic wall, and attach to the posterior surface of the broad ligament by the peritoneal fold, otherwise known as the mesovarium. They are positioned near the fimbrial end of the fallopian tubes, which are connected to the uterine cavity. An adult ovary averages 2 to 5 cm in length, weighs an average of 14 g, and typically contains 2 to 4 million primordial follicles.[52] These primordial follicles are present at birth; however, maturation is blocked until puberty. Following the onset of puberty, each ovarian cycle is marked by recruitment of a few primordial follicles for maturation. Typically, all but one of these follicles will then atrophy, in a process termed the **follicular phase**.

The single remaining follicle—known as the **graafian follicle**—is composed of an outer and inner layer (the theca externa and theca interna, respectively) encasing a central fluid-filled cavity and a layer of cells known as the granulosa layer.[53–55] The maturing ovum attaches to the inside of the follicle via cells derived from granulosa cells, called *cumulus cells*. During the **luteal phase** of the ovarian cycle, the graafian follicle releases its ovum in response to ovarian stimulation by LH. When the ovum is extruded, the graafian follicle undergoes a morphologic change with hypertrophy of the theca and granulosa cells to become the **corpus luteum**. This process is called *luteinization*. The corpus luteum is rich in cholesterol and acts as a substrate for continued production of progesterone and estrogen, maintaining the endometrium for conception. If conception or implantation fails to occur, the endometrium is shed and the corpus luteum atrophies to an atretic follicle.

Hypothalamic–Pituitary–Ovarian Axis

The hypothalamic–pituitary–ovarian axis is depicted in **Figure 17.3**. As in the male, GnRH stimulates the release of FSH and LH from the anterior pituitary. FSH targets the ovarian follicle to stimulate growth, while LH stimulates ovulation. LH further stimulates the corpus luteum to secrete progesterone following ovulation. Both negative and positive feedback

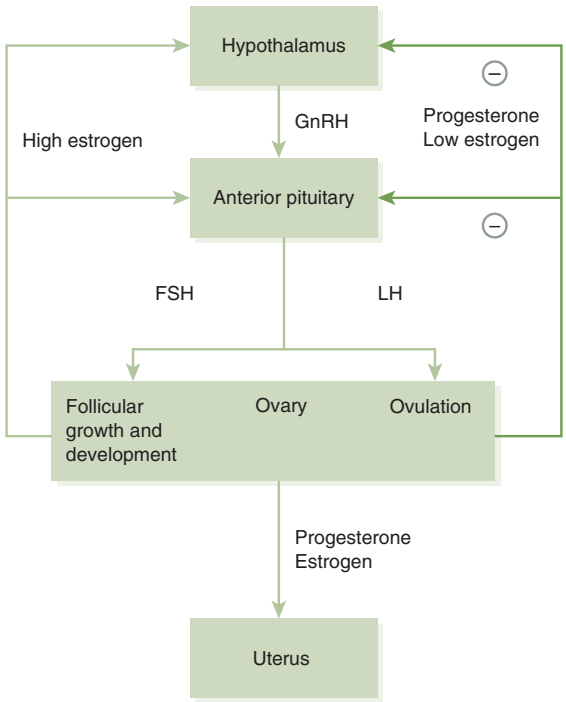


Figure 17.3 The hypothalamic–pituitary–ovarian axis.

© Jones & Bartlett Learning.

are imposed by estrogen and progesterone on the anterior pituitary and hypothalamus. Low estrogen concentration during the follicular phase results in negative feedback, while high concentration of estrogen midcycle results in positive feedback.

Hormonal Production by the Ovaries

As in the adrenal glands and the testes, the steroidogenic pathway and synthetic enzymes are present in the ovaries. Cholesterol is either synthesized from acetate or actively transported from the LDL particles in the blood and then used as a substrate for hormonal production.[56]

Estrogen

Naturally synthesized estrogens are carbon-18 compounds. The principal estrogen produced in the ovary is estradiol. Estrone and estriol are primarily metabolites of intraovarian and extraglandular conversion. Estrogens promote secondary characteristics such as breast, uterine, and vaginal development but also affect the skin, vascular smooth muscles, bone cells, and the central nervous system.[57] The lack of estrogen that naturally occurs with the onset of menopause leads to atrophic changes in these organs.[58] During the reproductive period, it is estrogen that is responsible for follicular phase changes in the uterus, with deficiency resulting in irregular and incomplete development of the endometrium.

Progesterone

Progesterone is a carbon-21 compound within the steroid family and is produced by the corpus luteum. Progesterone induces the secretory activity of those endometrial glands that have been primed by estrogen, readying the endometrium for embryo implantation. Other effects include thickening of the cervical mucus, reduction of uterine contractions, and thermogenic effect, in which basal body temperature rises after ovulation. This effect is of clinical use in marking the occurrence of ovulation. Progesterone is the dominant hormone responsible for the luteal phase, and deficiency results in failure of implantation of the embryo.[59]

Androgens

Ovaries produce the androgens androstenedione, dehydroandrostenedione, testosterone, and DHT, all of which are carbon-19 compounds. Excess production of ovarian androgens in women leads to excess hair growth (**hirsutism**), loss of female characteristics, and—in severe cases—development of overt male secondary sexual features (masculinization or (**virilization**).[59–62] Unlike estrogen, which is not produced in the ovary after menopause, ovarian androgen synthesis continues well into advanced age.[63]

Others

Inhibins A and B, which are produced by the ovaries, are hormones that inhibit FSH production.[63] Activin is a hormone that enhances FSH secretion and induces steroidogenesis.[64] Folliculostatin, relaxin,[63] follicle regulatory protein, oocyte maturation factor, and meiosis-inducing substance are hormones that appear to have important, yet not clearly characterized, functions.

The Menstrual Cycle

By convention, the menstrual cycle is considered to start on the first day of menses (day 1). The menstrual cycle consists of two phases of parallel events occurring at the ovaries and endometrium. Within the ovaries, these events are known as the follicular and luteal phases, while the concurrent endometrial events are known as the proliferative and secretory phases.[65]

The Follicular Phase

The follicular phase begins with the onset of menses and ends on the day of the LH surge. Early in the follicular phase, the ovary secretes very little estrogen or progesterone. A rise in FSH, however, stimulates estrogen production. The estrogen secreted by the developing follicle within the ovary stimulates uterine epithelial cells, blood vessel growth, and endometrial gland development to increase the thickness of the endometrium.

The Luteal Phase

Estrogen levels peak 1 day before ovulation, at which point a positive feedback system results in an LH surge. The start of the luteal phase is marked by the extrusion of the ovum approximately 36 hours after this LH surge, with subsequent luteinization of the graafian follicle to form the corpus luteum. The corpus luteum secretes progesterone to aid in the implantation of the embryo. In the absence of fertilization, with a gradual decline in the production of progesterone and estrogen by the corpus luteum, there is a loss of endometrial blood supply; this

results in shedding of the endometrium approximately 14 days after ovulation occurred. The typical duration of menstrual bleedingis 3 to 6 days, with blood loss averaging 50 mL. Onset of menses marks the end of the luteal phase.

Hormonal Control of Ovulation

The central control of FSH and LH secretion resides in the GnRH pulse generator of the arcuate nuclei and medial preoptic nuclei of the hypothalamus.[65] Positive and negative feedback responses exist among estrogen, progesterone, LH, and FSH production. It is because of the lack of estrogen after menopause that both FSH and LH levels rise.[66] During reproductive years, FSH levels are elevated early in the follicular phase. A midcycle surge in LH production stimulates a series of events that culminate in ovulation, with FSH levels falling after this event. Any injury to the hypothalamus or the presence of either psychosocial or physical stressors leads to changes in these hormonal cues and results in anovulation and amenorrhea.[67]

Pubertal Development in the Female

As with males, puberty in females consists of a sequence of hormonally mediated events resulting in the development of secondary sexual characteristics. Thelarche (development of breast tissue) is typically the earliest sign of sexual development, followed by development of pubic hair. Menarche, or initiation of menses, occurs an average of 2 to 3 years after the onset of puberty.

Precocious Sexual Development

Precocious sexual development occurs in response to premature exposure of tissues to sex steroids from any source and is distinguished from precocious puberty, which can be arrested by suppression of GnRH with analogs of GnRH that will suppress LH. If untreated, precocious puberty or sexual development can lead to short stature. Premature breast development is characterized by isolated breast development and occurs in response to earlier estrogen secretion, albeit at low circulating concentrations. Premature adrenarche may occur between 4 and 8 years of age and is characterized by pubic hair growth, and most commonly, precocious puberty does not ensue. Gonadotropins and gonadal steroid secretion help differentiate between them. DHEA and dehydroepiandrosterone sulfate (DHEAS) are elevated for age but match the child's bone age.

In girls with delayed puberty, sex steroids and gonadotropins are low, but they continue growing and may be tall because growth hormone (GH) and other pituitary hormones are unaffected unless pituitary or hypothalamic dysfunction is present.

Devised by Marshall and Tanner as a way to determine pubertal staging, the Tanner staging system, outlined in **Table 17.2**, is used to monitor and assess the growth stages of breast and pubic hair.[68]

Table 17.2 Tanner Staging of Breast and Pubic Hair Development in Females

Stages of Breast Development
1. Prepubertal
2. Elevation of breast bud and papilla, areolar enlargement
3. Elevation of breast tissue and papilla
4. Elevation of areola and papilla in secondary mound above the level of the breast
5. Mature stage: recession of areola into the breast with projection of papilla only
Stages of Pubic Hair Development
1. Lanugo-type hair (prepubertal)
2. Dark terminal hair on labia majora
3. Terminal hair covering labia majora and spreading to the mons pubis
4. Terminal hair fully covering the labia majora and mons pubis
5. Terminal hair covering the labia majora, mons pubis, and inner thighs

Menstrual Cycle Abnormalities

The menstrual cycle ranges from 25 to 35 days, with an average duration of 28 days. The average age of menopause in the United States is between 45 and 55 years, with the median at 53 years.[66]

Amenorrhea is defined as the absence of menses. *Primary amenorrhea* is used to describe a woman who has never menstruated by age 16 years, while *secondary amenorrhea* is used to describe a woman who has had at least one menstrual cycle followed by absence of menses for a minimum of 3 to 6 months.[67] Frequency of the different etiologies for

Table 17.3 Etiologies of Amenorrhea

	Primary (%)	Secondary (%)
Hypothalamus	27	38
Pituitary	2	15
Polycystic ovarian syndrome	7	30
Ovary	43	12
Uterus/outflow	19	7

amenorrhea, both primary and secondary, is listed in **Table 17.3**.

Oligomenorrhea refers to infrequent irregular menstrual bleeding, with cycle lengths in excess of 35 to 40 days. Uterine bleeding in excess of 7 days is dysfunctional and is termed **menorrhagia**. In a patient with infertility, the diagnosis of inadequate luteal phase is made when the luteal phase is less than 10 days or when an endometrial biopsy indicates the progression of endometrial changes is delayed or out of phase, resulting in implantation failure.[69] The multiple causes of male and female infertility are shown in **Table 17.4**.

Table 17.4 Causes of Infertility

Target	Result	Cause
Female		
Hypothalamus	Decreased GnRH	Drugs
		Increased stress
		Diet
Pituitary	Decreased FSH and LH	Destructive tumor or vesicular lesion
Ovaries	Decreased estradiol or progesterone	Organ failure
		Organ dysgenesis
		Antiovarian antibodies
		Malnourishment, very low weight, metabolic disease
Fallopian tubes and uterus	Inadequate endometrium	Low progesterone output
	Tubal scarring and closure	Pelvic inflammatory disease
	Decreased cervical mucus	Cervical infections
Conception	Immobilization and destruction of sperm	Antisperm antibodies
Male		
Hypothalamus and pituitary	Oligospermia to azoospermia (no sperm)	Primary defects in hypothalamic or pituitary glands
		Exogenous androgens
		Testicular dysfunction
Testes	Oligospermia to azoospermia	Orchitis
	Delayed or deficient sexual maturity	Testicular infections (mumps)
	Decreased testosterone	Alcoholism/substance abuse
		Chromosomal defects
Prostate	Decreased seminal fluid	Infections of prostate or seminal vesicles
Urethrogenital tract	Retrograde or absent ejaculation	Physical abnormalities
		Chronic diabetes

GnRH, gonadotropin-releasing hormone; FSH, follicle-stimulating hormone; LH, luteinizing hormone.

Box 17.2 Classification of Amenorrhea by WHO Guidelines

- Type 1: hypothalamic hypogonadism (low or normal FSH and/or LH)
- Hypothalamic amenorrhea (anorexia nervosa, idiopathic, exercise induced)
- Kallmann's syndrome
- Isolated gonadotropin deficiency
- Type 2: estrogenic chronic anovulation (normal FSH and LH)
- Polycystic ovarian syndrome (LH > FSH in some patients 2:1 or greater)
- Hyperthecosis
- Type 3: hyperthalamic hypogonadism (elevated FSH and/or LSH)
- Premature ovarian failure
- Turner's syndrome

The principles underlying the evaluation of disorders of normal menstrual functions, as outlined by the World Health Organization (WHO), are the same for ovarian and pituitary dysfunction (**Box 17.2**). A diagnostic approach to secondary amenorrhea is outlined in **Figure 17.4**.

Evaluation of Female Infertility

Evaluation of female infertility should begin with a physical examination including body mass index, external genitalia, hair growth patterns, pelvic examination, PAP smear, breast examination, neurologic examination for anosmia and visual abnormalities, and thyroid examination. Clinical history such as menstrual pattern, number of pregnancies (parity) and births, and sexual history, including sexually transmitted infections, also plays an important role in the evaluation of female infertility. Ovulation can be predicted by a surge in LH combined with an increase of 0.5–1.0°F in basal body temperature, which occur 24 to 36 hours prior to ovulation.[70] Laboratory evaluation of serum progesterone is also used as an indicator of ovulation. An increase in serum progesterone follows formation of the corpus luteum. However, it does not definitively confirm release of an oocyte. Following pregnancy, hCG and progesterone levels continue to rise. Midluteal serum progesterone concentration greater than 300 ng/dL indicates that ovulation occurred.[71]

Hypogonadotropic Hypogonadism

Hypogonadotropic hypogonadism, or gonadotropin (FSH and LH) deficiency resulting in decreased sex steroid production, is a common cause of secondary amenorrhea, and FSH and LH levels are usually <10 IU/L. There are many physiologic and pathologic causes of hypogonadotropic hypogonadism, including weight loss as associated with anorexia nervosa or various disease processes, intense physical exercise (commonly termed *runner's amenorrhea*), hypothyroidism, and pituitary tumors that disrupt secretion of FSH or LH.[67] Prolactin production by prolactinomas can have similar effects. Any secondary cause of chronic hypogonadism can induce pathologic bone loss, resulting in osteopenia or, if severe, osteoporosis. This is known as the "*athlete's triad*" consisting of amenorrhea, low energy (with or without an eating disorder), and osteoporosis.

Hypergonadotropic Hypogonadism

Hypergonadotropic hypogonadism is characterized by ovarian failure resulting in elevation of FSH concentrations, with or without LH elevations. Ovarian failure occurs naturally between 45 and 55 years of age in American women. When the depletion of oocytes and follicles occurs at the expected time, it is termed **menopause**. Menopause is a natural, inevitable event that results in elevation of FSH and LH levels, with low levels of estrogen.[66]

Premature ovarian failure is defined as primary hypogonadism occurring in a woman before the age of 40 and can be a result of congenital chromosomal abnormality (e.g., Turner's syndrome[72]) or premature menopause.[73] Patients with Turner's syndrome are missing all (45,X) or part of one X chromosome. Some patients have *mosaicism*, in which parts of the body have normal cells that contain two X chromosomes (46,XX), while other cells are missing one X chromosome (45,X). Turner's syndrome is not hereditary.[74] Patients with Turner's syndrome do not exhibit the same hot flashes experienced by patients with secondary hypergonadotropic hypogonadism. Premature menopause can occur in isolation or in association with other endocrine gland failure such as hypoparathyroidism, hypothyroidism, or hypoadrenalism.[73] Basal FSH that is repeatedly >30 IU/L or a single elevation >40 IU/L combined with estradiol <20 pg/mL indicates premature ovarian failure.[75]

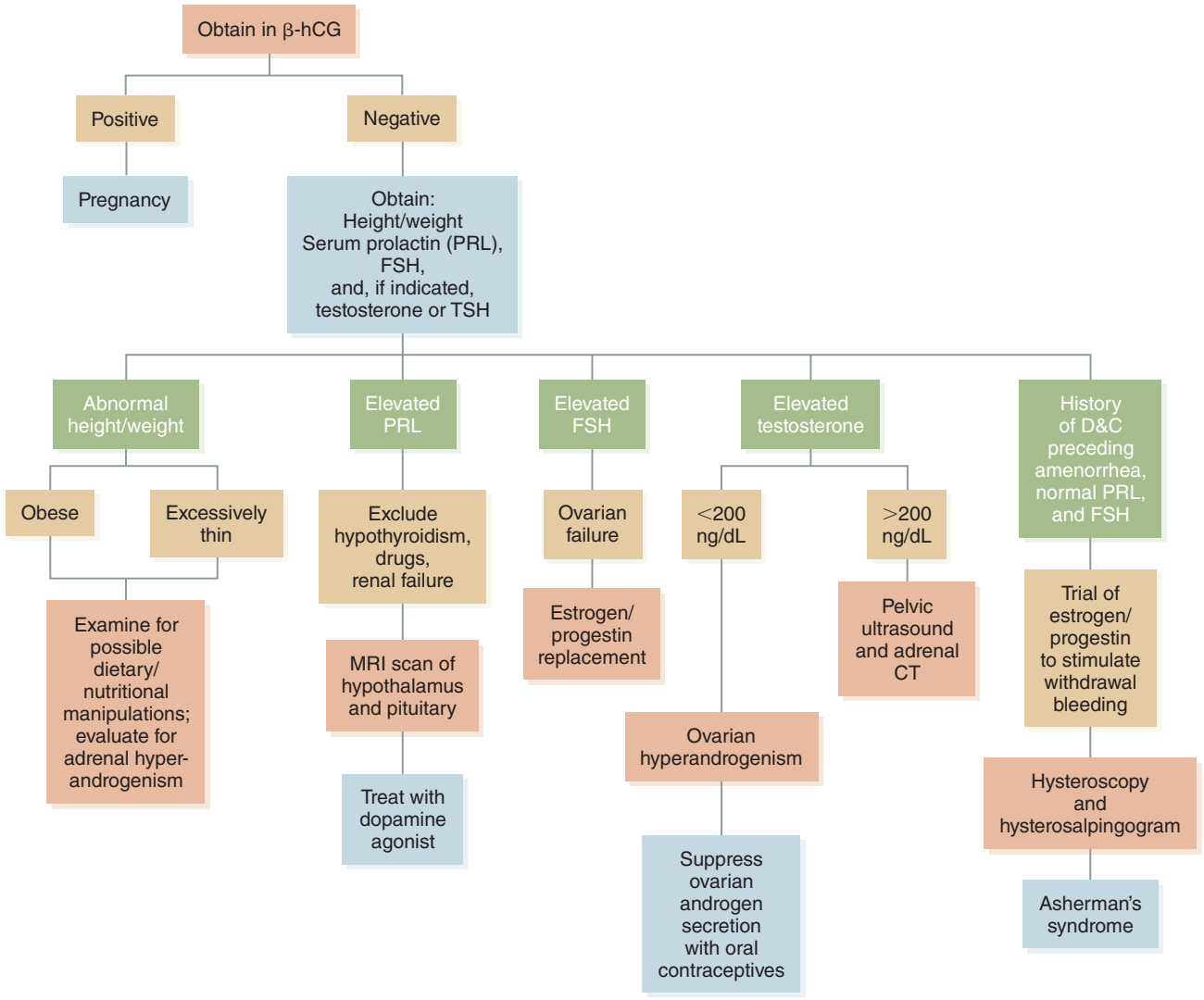


Figure 17.4 Diagnostic approach to secondary amenorrhea. PRL, prolactin; FSH, follicle-stimulating hormone; TSH, thyroid-stimulating hormone; hCG, human chorionic gonadotropin.

Polycystic Ovary Syndrome

This common disorder can present in many ways: infertility, hirsutism, chronic anovulation, glucose intolerance, hyperlipidemia or dyslipidemia, and hypertension.[59–61] The onset is often perimenarcheal, chronic, and notable for its slow progression. Investigations for this disorder involve estimation of free testosterone, SHBG, FSH, LH, fasting glucose, insulin, lipid levels, DHEAS, prolactin, and with or without AMH levels. Ovarian ultrasound may reveal multiple cysts in most patients, however, about 30% of patients do not have ovarian cysts. Most patients with this disorder are overweight; however, some patients with polycystic ovary syndrome (PCOS), especially outside the United States, are of normal weight. Most symptoms and laboratory abnormalities are reversed with weight loss and increased physical activity. The drug Glucophage (metformin), commonly used for the treatment of diabetes, is useful in this condition, even in the absence of diabetes. Although not approved for this use by the U.S. Food and Drug Administration, it reportedly can normalize menstrual cycles and improve conception rates.[76]

Hirsutism

Hirsutism is an abnormal, abundant, androgen-sensitive terminal hair growth in areas in which terminal hair follicles are sparsely distributed or not

Box 17.3 Classification of Hirsutism[51]

- Functional (normal androgen levels with excess hair growth) or true androgen excess (elevated androgens)
- Ovarian (LH mediated) or adrenal (adrenocorticotropin hormone mediated)
- Peripheral conversion of androgens (obesity)
- Tumoral hyperandrogenism (ovarian, adrenal)
- Chorionic gonadotropin mediated

Data from Escobar-Morreale HF, Carmina E, Dewailly D, et al. Epidemiology, diagnosis and management of hirsutism: a consensus statement by the Androgen Excess and Polycystic Ovary Syndrome Society. *Hum Reprod Update.* 2012;18(2):146–170.

normally found in women (**Box 17.3**). Typical causes of hirsutism are listed in **Table 17.5**.[60]

Hirsutism should only be considered in the context of a woman's ethnic origin. Women of Italian, eastern European, south Asian, and Irish descent possess more androgen-sensitive terminal hair than do most northern European women, making a careful elicitation of ethnic background important prior to initiation of an extensive laboratory evaluation in a woman born in the United States. It is estimated that about 5% to 10% of American women have hirsutism, which can be quantified using a measurement technique known as the modified Ferriman-Gallwey Scale that identifies nine areas (lip, chin, sideburn region, neck, chest, abdomen, upper and lower back, and thigh) for assessment and allots points on a scale of 1 to 4 based on hair thickness and pigmentation. A score of higher than 8 is consistent with a diagnosis of hirsutism.[77]

Table 17.5 Causes of Hirsutism

Common
Idiopathic
Polycystic ovary syndrome
Uncommon
Drugs: danazol, oral contraceptives with androgenic progestins
Congenital adrenal hyperplasia
Hyperprolactinemia
Cushing's syndrome
Adrenal tumors
Ovarian tumors

Hormonal abnormalities associated with hirsutism are summarized in **Table 17.6**.

Estrogen Replacement Therapy

Estrogen replacement, also known as hormone replacement therapy (HRT), remains a contentious issue. The Women's Health Initiative study enrolled 16,608 postmenopausal women who were placed on conventional hormone replacement combinations. The study showed an overall increased incidence of invasive breast cancer (hazard ratio, 1.26), stroke, venous clot formation, and coronary heart

CASE STUDY 17.2, PART 3

Recall Karen, the avid runner. Her physical examination reveals edema of the hands and feet, low-set ears, presence of redundant nuchal skin, and absence of secondary sexual characteristics (Tanner stage 1). Karen complains that she is the shortest one in her class and that some of the other students tease her because of her spider veins.

4. How do the results of the physical examination contribute to the differential diagnosis for the cause of delayed puberty?
5. State three possible causes of delayed puberty that should be included in the differential diagnosis for this patient.

CASE STUDY 17.2, PART 4

Recall Karen. Abdominal ultrasonography revealed horseshoe kidney. Pelvic ultrasound revealed small uterus and ovaries. Karyotyping demonstrated 45,X. Laboratory results for patient serum are shown in the following table.

© Klaus Vedfelt/DigitalVision/Getty Images.

Hormone	Patient Result	Reference Range[84]
FSH	100 IU/mL	0.1–11 IU/mL
LH	30 IU/mL	0.1–16.4 IU/mL
Estradiol	25 pmol/L	20–111 pmol/L
Inhibin B	Undetectable	2–80 pg/mL

6. Which test results are abnormal?
7. Based on the laboratory results, what is the cause for the patient's delayed puberty?
8. What is the prognosis for this patient?

Table 17.6 Androgen Levels in Hirsutism and Virilization

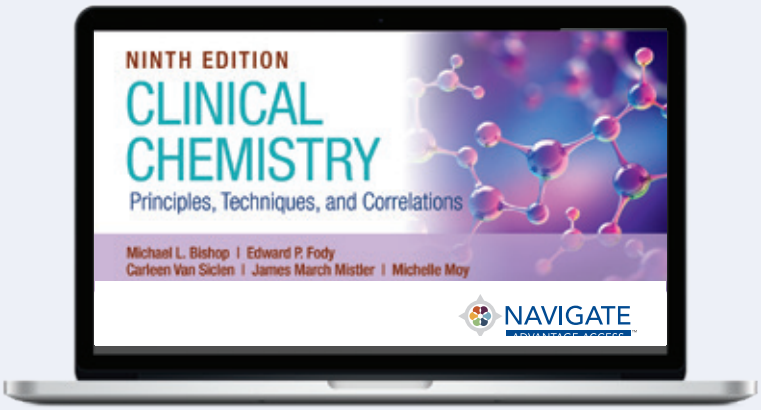
Analyte			
Condition	Total Testosterone	Free Testosterone	DHEAS
Idiopathic hirsutism[78]	↑	↑↑↑	↑
Polycystic ovary syndrome	↑	↑↑	↑
Congenital adrenal hyperplasia	↑↑	↑↑	↑↑↑
Virilizing Tumors			
Ovarian	↑↑↑	↑↑↑	↑
Adrenal	↑↑	↑↑	↑↑↑

Data from Demers LM. *Hirsutism and Virilization, News and Views.* Washington, DC: American Association for Clinical Chemistry; 1989.

disease (CHD) events, yet no benefit in slowing cognitive decline, with combined estrogen plus progestin therapy. However, *unopposed estrogen therapy* (i.e., only estrogen without progesterone) was not associated with an increase in either breast cancer or CHD, suggesting that these risks may be conferred by the progestin component. With HRT, reductions in bone loss, colon polyp formation, and menopausal symptoms (hot flashes and vaginal dryness) were noted. The published ELITE trial (Early versus Late Intervention Trial with Estradiol) results showed decreased progression of subclinical atherosclerosis in menopausal women who were started on HRT within 6 years of menopause, lending further credence to the "timing hypothesis;" that is, younger women may indeed benefit from HRT in terms of decreased CHD risk.[79–82] In large meta-analyses of randomized controlled trials, including the Women's Health Initiative, a significantly lower risk of CHD (hazard ratio, 0.68; 95% CI, 0.48 to 0.96) and of death from any cause (hazard ratio, 0.61; 95% CI, 0.39 to 0.95) was found with hormone therapy than with placebo among women who were younger than 60 years of age, less than 10 years past menopause, or both when they underwent randomization.[83]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Sinisi AA, Pasquali D, Notaro A, et al. Sexual differentiation. *J Endocrinol Invest.* 2003;26(3 suppl):23–28.
2. Dewing P, Bernard P, Vilain E. Disorders of gonadal development. *Semin Reprod Med.* 2002;20:189–198.
3. Gao F, Maiti S, Huff V. The Wilms tumor gene, Wt1, is required for Sox9 expression and maintenance of tubular architecture in the developing testis. *Proc Natl Acad Sci U S A.* 2006;103(32):11987–11992.
4. Nef S, Parada LF. Cryptorchidism in mice mutant for Insl3. *Nat Genet.* 1999;22:295–299.
5. Durlinger AL, Visser JA, Themmen AP. Regulation of ovarian function: the role of anti-Müllerian hormone. *Reproduction.* 2002;124:601–609.
6. Sharpe RM, McKinnell C, Kivlin C, et al. Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood. *Reproduction.* 2003;125:769–784.
7. Svechnikov K, Landreh L, Weisser J, et al. Origin, development and regulation of human Leydig cells. *Horm Res Paediatr.* 2010;73:93–101.
8. Rey RA, Grinspon RP. Normal male sexual differentiation and aetiology of disorders of sex development. *Best Pract Res Clin Endocrinol Metab.* 2011;25(2):221–238.
9. Havelock JC, Auchus RJ, Rainey WE. The rise in adrenal androgen biosynthesis: adrenarche. *Semin Reprod Med.* 2004;22(4):337–347.
10. Campbell RE, Suter KJ. Redefining the gonadotropin-releasing hormone neurone dendrite. *J Neuroendocrinol.* 2010;22(7):650–658.
11. Grover A, Smith CE, Gregory M, et al. Effects of FSH receptor deletion on epididymal tubules and sperm morphology, numbers, and motility. *Mol Reprod Dev.* 2005;72(2):135–144.
12. Schiffer L, Barnard L, Baranowski ES, et al. Human steroid biosynthesis, metabolism and excretion are differentially reflected by serum and urine steroid metabolomes: a comprehensive review. *J Steroid Biochem Mol Biol.* 2019;194:105439. doi:10.1016/j.jsbmb.2019.105439
13. Labrie F. Extragonadal synthesis of sex steroids: intracrinology. *Ann Endocrinol (Paris).* 2003;64:95–107.
14. Hiort O, Holterhus PM. Androgen insensitivity and male infertility. *Int J Androl.* 2003;26:16–20.
15. Arkoncel ML, Arkoncel FR, Lantion-Ang FL. A case of Kallmann syndrome. *BMJ Case Rep.* 2011;2011: bcr0120113727. doi:10.1136/bcr.01.2011.3727
16. Navarro NF, Sukster E, Feijó RB. An adolescent with Kallmann syndrome: a case report. *Resid Pediátr.* 2019;9(2): 173–175. doi:10.25060/residpediatr-2019.v9n2-18
17. Lee DK, Chang C. Molecular communication between androgen receptor and general transcription machinery. *J Steroid Biochem Mol Biol.* 2003;84:41–49.
18. Sultan C, Gobinet J, Terouanne B, et al. The androgen receptor: molecular pathology. *J Soc Biol.* 2002;196: 223–240.
19. Sultan C, Lumbroso S, Paris F, et al. Disorders of androgen action. *Semin Reprod Med.* 2002;20:217–228.
20. Segal TY, Mehta A, Anazodo A, et al. Role of gonadotropin releasing hormone and human chorionic gonadotropin stimulation tests in differentiating patients with hypogonadotropic hypogonadism from those with constitutional delay of growth and puberty. *J Clin Endocrinol Metab.* 2009;94(3):780–785.
21. Tong S, Wallace EM, Burger HG. Inhibins and activins: clinical advances in reproductive medicine. *Clin Endocrinol (Oxf).* 2003;58:115–127.
22. Brinton LA. Breast cancer risk among patients with Klinefelter syndrome. *Acta Paediatr.* 2011;100:814–818. https://doi.org/10.1111/j.1651-2227.2010.02131.x
23. Cheng CK, Leung PCK. Molecular biology of gonadotropinreleasing hormone (GnRH)-I, GnRH-II, and their receptors in humans. *Endocr Rev.* 2005;26(2):283–306.
24. Seminara SB, Hayes FJ, Crowley WF Jr. Gonadotropin releasing hormone deficiency in the human (idiopathic hypogonadotropic hypogonadism and Kallmann's syndrome): pathophysiological and genetic considerations. *Endocr Rev.* 1998;19:521–539.
25. De Rosa M, Zarrilli S, Di Sarno A, et al. Hyperprolactinemia in men: clinical and biochemical features and response to treatment. *Endocrine.* 2003;20:75–82.

26. Heaton JP, Morales A. Endocrine causes of impotence (nondiabetes). *Urol Clin North Am.* 2003;30:73–81.
27. Matsuzaki T, Azuma K, Irahara M, et al. Mechanism of anovulation in hyperprolactinemic amenorrhea determined by pulsatile gonadotropin-releasing hormone injection combined with human chorionic gonadotropin. *Fertil Steril.* 1994;62:1143–1149.
28. Kaiser U. Hyperprolactinemia and infertility: new insights. *J Clin Invest.* 2012;122:3467–3468.
29. Dandona P, Dhindsa S. Update: hypogonadotropic hypogonadism in type 2 diabetes and obesity. *J Clin Endocrinol Metab.* 2011;96(9):2643–2651.
30. Dhindsa S, Prabhakar S, Sethi M, et al. Frequent occurrence of hypogonadotropic hypogonadism in type 2 diabetes. *J Clin Endocrinol Metab.* 2004;89(11):5462–5468.
31. Holt SK, Lopushnyan N, Hotaling J, et al. Prevalence of low testosterone and predisposing risk factors in men with type 1 diabetes mellitus: findings from the DCCT/EDIC. *J Clin Endocrinol Metab.* 2014;99:E1655–E1660.
32. Harman SM, Metter EJ, Tobin JD, et al. Longitudinal effects of aging on serum total and free testosterone levels in healthy men. Baltimore Longitudinal Study of Aging. *J Clin Endocrinol Metab.* 2001;86:724–731.
33. Feldman HA, Longcope C, Derby CA, et al. Age trends in the level of serum testosterone and other hormones in middle-aged men: longitudinal results from the Massachusetts Male Aging Study. *J Clin Endocrinol Metab.* 2002;87(2):589–598.
34. Snyder PJ, Bhasin S, Cunningham GR, et al. Effects of testosterone treatment in older men. *N Engl J Med.* 2016;374(7):611–624. doi:10.1056/NEJMoa1506119
35. Agirregoitia E, Valdivia A, Carracedo A, et al. Expression and localization of delta-, kappa-, and mu-opioid receptors in human spermatozoa and implications for sperm motility. *J Clin Endocrinol Metab.* 2006;91:4969–4975.
36. Luboshitzky R, Aviv A, Hefetz A, et al. Decreased pituitary gonadal secretion in men with obstructive sleep apnea. *J Clin Endocrinol Metab.* 2002;87:3394–3398.
37. Bhasin S, Snyder PJ, Montori VM, et al. Testosterone therapy in adult men with androgen deficiency syndromes: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab.* 2006;91(6):1995–2010.
38. Dandona P, Rosenberg MT. A practical guide to male hypogonadism in the primary care setting. *Int J Clin Pract.* 2010;64(6):682–696.
39. Testosterone for 'late-onset hypogonadism' in men? *Drug Ther Bull.* 2010;48:69–72.
40. Fröhlich E, Wahl R. The forgotten effects of thyrotropin-releasing hormone: Metabolic functions and medical applications. *Front Neuroendocrinol.* 2019;52:29–43. https://doi.org/10.1016/j.yfrne.2018.06.006
41. Wosnitzer M, Goldstein M, Hardy MP. Review of azoospermia. *Spermatogenesis*. 2014;4:e28218. doi:10.4161/spmg.28218
42. Sun QH, Zheng Y, Zhang X-L, Mu Y-M. Role of gonadotropin-releasing hormone stimulation test in diagnosing gonadotropin deficiency in both males and females with delayed puberty. *Chin Med J (Engl).* 2015;128(18):2439–2443. doi:10.4103/0366-6999.164926
43. Srinivas-Shankar U, Wu FC. Drug insight: testosterone preparations. *Nat Rev Urol.* 2006;3:653–665.
44. AVEED® REMS (Risk Evaluation and Mitigation Strategy) Program [website]. https://www.aveedrems.com/AveedUI/rems/preHome.action.
45. Friedl KE, Hannan CJ Jr, Jones RE, et al. High-density lipoprotein cholesterol is not decreased if an aromatizable androgen is administered. *Metabolism.* 1990;39:69–74.
46. McGee EA, Hsueh AJW. Initial and cyclic recruitment of ovarian follicles. *Endocr Rev.* 2000;21(2):200–214.
47. Sherman BM, West JH, Korenman SG. The menopausal transition: analysis of LH, FSH, estradiol, and progesterone concentrations during menstrual cycles of older women. *J Clin Endocrinol Metab.* 1976;42:629–636.
48. Jost A, Vigier B, Prepin J, et al. Studies on sex differentiation in mammals. *Recent Prog Horm Res.* 1973;29:1–41.
49. Messinis IE. Ovarian feedback, mechanism of action and possible clinical implications. *Hum Reprod Update.* 2006;12(5):557–571.
50. Filicori M, Santoro N, Merriam GR, et al. Characterization of the physiological pattern of episodic gonadotropin secretion throughout the human menstrual cycle. *J Clin Endocrinol Metab.* 1986;62:1136–1144.
51. Pinsker JE. Turner syndrome: updating the paradigm of clinical care. *J Clin Endocrinol Metab.* 2012;97(6):E994–E1003. https://doi.org/10.1210/jc.2012-1245
52. Badouraki M, Christoforidis A, Economou I, et al. Sonographic assessment of uterine and ovarian development in normal girls aged 1 to 12 years. *J Clin Ultrasound.* 2008;36(9):539–544.
53. Canipari R, Cellini V, Cecconi S. The ovary feels fine when paracrine and autocrine networks cooperate with gonadotropins in the regulation of folliculogenesis. *Curr Pharm Des.* 2012;18(3):245–255.
54. Peters H, Byskov AG, Grinsted J. Follicular growth in fetal and prepubertal ovaries of humans and other primates. *Clin Endocrinol Metab.* 1978;7:469–485.
55. Gougeon A. Dynamics of follicular growth in the human: a model from preliminary results. *Hum Reprod.* 1986;1:81–87.
56. Fakheri RJ, Javitt NB. Autoregulation of cholesterol synthesis: physiologic and pathophysiologic consequences. *Steroids.* 2011;76(3):211–215.
57. Poutanen M. Understanding the diversity of sex steroid action. *J Endocrinol.* 2012;212:1–2.
58. te Velde ER, Pearson PL. The variability of female reproductive ageing. *Hum Reprod Update.* 2002;8(2):141–154.
59. Carmina E, Campagna AM, Lobo RA. A 20-year follow-up of young women with polycystic ovary syndrome. *Obstet Gynecol.* 2012;119(2):263–269.
60. Escobar-Morreale HF, Carmina E, Dewailly D, et al. Epidemiology, diagnosis and management of hirsutism: a consensus statement by the Androgen Excess and Polycystic Ovary Syndrome Society. *Hum Reprod Update.* 2012;18(2):146–170.
61. Panidis D, Tziomalos K, Misichronis G, et al. Insulin resistance and endocrine characteristics of the different phenotypes of polycystic ovary syndrome: a prospective study. *Hum Reprod.* 2012;27(2):541–549.
62. Sarfati J, Bachelot A, Coussieu C, et al.; Study Group Hyperandrogenism in Postmenopausal Women. Impact of clinical, hormonal, radiological, and immunohistochemical studies on the diagnosis of postmenopausal hyperandrogenism. *Eur J Endocrinol.* 2011;165:779–788.
63. Robertson DM. Inhibins and activins in blood: predictors of female reproductive health? *Mol Cell Endocrinol.* 2012:359(1–2):78–84. doi:10.1016/j.mce.2011.05.016

64. Sherwood OD. Relaxin's physiological roles and other diverse actions. *Endocr Rev.* 2004;25(2):205–234.
65. Jabbour HN, Kelly RW, Fraser HM, et al. Endocrine regulation of menstruation. *Endocr Rev.* 2006;27(1):17–46.
66. Broekmans FJ, Soules MR, Fauser BC. Ovarian aging: mechanisms and clinical consequences. *Endocr Rev.* 2009;30(5):465–493.
67. Santoro N. Update in hyper- and hypogonadotropic amenorrhea. *J Clin Endocrinol Metab.* 2011;96(11): 3281–3288.
68. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Arch Dis Child.* 1969;44:291–303.
69. Lessey BA. Assessment of endometrial receptivity. *Fertil Steril.* 2011;96:522–529.
70. Steward K, Raja A. Physiology, ovulation and basal body temperature. [Updated 2020 Jul 27]. *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan. Available from: https://www.ncbi.nlm.nih.gov/books/NBK546686/
71. Fritz MA. The modern infertility evaluation. *Clin Obstet Gynecol.* 2012;55(3):692–705. doi:10.1097/GRF.0b013e31825cab55
72. Elsheikh M, Dunger DB, Conway GS, et al. Turner's syndrome in adulthood. *Endocr Rev.* 2002;23(1):120–140.
73. Hoek A, Schoemaker J, Drexhage HA. Premature ovarian failure and ovarian autoimmunity. *Endocr Rev.* 1997;18(1):107–134.
74. Kasagani SK, Mutthineni RB, Jampani ND, Nutalapati R. Report of a case of Turner's syndrome with localized aggressive periodontitis. *J Indian Soc Periodontol*. 2011;15(2): 173–176. doi:10.4103/0972-124X.84389
75. Jankowska K. Premature ovarian failure. *Prz Menopauzalny*. 2017;16(2):51–56. doi:10.5114/pm.2017.68592
76. Palomba S, Falbo A, Zullo F, et al. Evidence-based and potential benefits of metformin in the polycystic ovary syndrome: a comprehensive review. *Endocr Rev.* 2009;30(1):1–50.
77. Cook H, Brennan K, Azziz R. Reanalyzing the modified Ferriman-Gallwey score: is there a simpler method for assessing the extent of hirsutism? *Fertil Steril.* 2011;96(5): 1266–1270.
78. Azziz R, Carmina E, Sawaya ME. Idiopathic hirsutism. *Endocr Rev.* 2000;21(4):347–362.
79. Chlebowski RT, Wactawski-Wende J, Ritenbaugh C, et al.; Women's Health Initiative Investigators. Estrogen plus progestin and colorectal cancer in postmenopausal women. *N Engl J Med.* 2004;350(10):991–1004.
80. Manson JE, Hsia J, Johnson KC, et al.; Women's Health Initiative Investigators. Estrogen plus progestin and the risk of coronary heart disease. *N Engl J Med.* 2003;349(6): 523–534.
81. Hays J, Ockene JK, Brunner RL, et al.; Women's Health Initiative Investigators. Effects of estrogen plus progestin on health-related quality of life. *N Engl J Med.* 2003;348(19):1839–1854.
82. Hodis HN, Mack W, Henderson VW, et al. Vascular effects of early versus late postmenopausal treatment with estradiol. *N Engl J Med.* 2016;374(13):1221–1231. doi:10.1056/NEJMoa1505241.
83. Boardman HM, Hartley L, Eisinga A, et al. Hormone therapy for preventing cardiovascular disease in post-menopausal women. *Cochrane Database Syst Rev.* 2015;(3):CD002229.
84. Soldin OP, Hoffman EG, Waring MA, Soldin SJ. Pediatric reference intervals for FSH, LH, estradiol, T3, free T3, cortisol, and growth hormone on the DPC IMMULITE 1000. *Clin Chim Acta.* 2005;355(1-2):205–210. doi:10.1016/j.cccn.2005.01.006

CHAPTER 18

Parathyroid Function

Rebbecca Silva

CHAPTER OUTLINE

KEY TERMS

1,25-dihydroxyvitamin D (calcitriol)
25-Hydroxyvitamin D
Bisphosphonates
Bone turnover
Calcium-sensing receptor (CSR)
Fibroblast growth factor 23 (FGF23)
Hypercalcemia
Hypocalcemia
Osteoblast
Osteoclast
Osteomalacia
Osteoporosis
Parathyroid hormone (PTH)
Parathyroid hormone–related protein (PTHrP)
Rickets
Vitamin D

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Identify the biologically active form of vitamin D and describe its production in the body.
- Describe the endocrine function and physiology of the parathyroid gland.
- Discuss the laboratory tools used to evaluate the function of the parathyroid gland.
- Correlate laboratory results to clinical disease states involving the parathyroid gland.
- Apply theoretical knowledge of parathyroid function to answer case study questions.

CASE STUDY 18.1, PART 1

Amanda, a 40-year-old woman, presents to her physician complaining of marked left flank pain that began the previous night. She reports that the pain is worse than that of giving birth. She also reports blood in her urine earlier on the day she came to see her doctor. She has felt more fatigued and as if her concentration has not been as good as normal for the last year or so, and she feels more forgetful.

CASE STUDY 18.2, PART 1

Robert, 58-year-old man, has been a smoker since childhood. He has been smoking three packs per day as long as he can remember but insists his cigarettes "don't hurt me none, doc." He has been feeling ill recently, however, with loss of appetite, malaise, and weight loss. His mental ability has been dulled recently, and he can't remember from "one minute to the next," especially notable during his work as a cowboy.

CASE STUDY 18.3, PART 1

Regan, a 6-year-old girl, is brought to a pediatrician by her parents, who report that her height is not progressing as they think it should (or like it did for her 8-year-old sister) and her legs look bowed.

Serum calcium level is maintained at a constant level for the optimal excitability of neural and muscular tissue and the coordinated functioning of various organ systems in the human body. Calcium levels are maintained within a narrow range for optimal nerve impulse transmission, muscular contraction, blood coagulation, hormone secretion, and intercellular adhesion. Imbalance of calcium can lead to serious clinical consequences. In this chapter, the regulation of blood calcium by hormonal control, as well as consequences of abnormal hormones, are discussed. For information on calcium as an electrolyte, see Chapter 11, *Electrolytes*.

Parathyroid Glands

Parathyroid Hormone

Parathyroid hormone (PTH) is secreted from four parathyroid glands that are located adjacent to the thyroid gland. The parathyroid glands were first described in 1852 by Sir Richard Owen,[1] when he performed dissection on an Indian rhinoceros that had died in London Zoo. Normal parathyroid glands are ovoid or bean shaped and measure approximately 3 mm in size. The superior parathyroid glands are smaller than the inferior pair. The parathyroid glands have an anatomically distinct vascular supply from that of the thyroid gland and are enveloped in a pad of anatomically distinct fibrofatty capsule.[2]

One human postmortem examination study revealed that four glands are found in 91% of the subjects, three glands in 5%, and five glands in 4%.[3] The locations of parathyroid glands vary widely due to the embryonic origination from the third and fourth pharyngeal pouches with eventual migration to the lower neck. The superior parathyroid glands develop in the fourth pharyngeal pouch and migrate caudally along with the ultimobranchial bodies, which give rise to parafollicular (C) cells of the thyroid gland. The superior pair is commonly located along the upper two-thirds of the posterior margin of the thyroid lobe. The superior parathyroid glands are relatively constant in their location along the posterior margin of the thyroid gland. The third bronchial pouches give rise to the inferior pair and the thymus gland, and together, they migrate to the lower neck; in some individuals, the inferior parathyroid glands can migrate into the thorax.

The parathyroid glands have specialized **calcium-sensing receptors (CSRs)** that respond to rising or falling calcium levels by increasing or decreasing PTH secretion, respectively. The primary targets for the PTH are the bone and kidneys. PTH mobilizes calcium from bone by increasing bone resorption, a process by which bone releases minerals from the bone matrix into the bloodstream. PTH has three effects on the kidneys:

1. Increase the reabsorption of renal tubular calcium
2. Increase phosphate excretion
3. Enhance 1α-hydroxylation of **25-hydroxyvitamin D**

Activated **1,25-dihydroxyvitamin D (calcitriol)** promotes meal-related intestinal absorption of calcium. In summary, low blood calcium is sensed by the parathyroid CSR, which in turn secretes PTH and sets into motion a cascade of events aimed at restoring normal blood calcium levels.

The multiple actions of PTH described above are mediated via PTH receptor, located on the cell membrane of target tissues. This receptor activates adenylate cyclase and the "second messenger" pathway involving cyclic AMP (cAMP), which modulates protein phosphorylation.

Vitamin D

Vitamin D is a steroid hormone that is synthesized in the skin from cholesterol following exposure to UVB rays from the sun. Following the photobiosynthesis of vitamin D from cholesterol, it undergoes further hydroxylation in the liver and kidneys (**Figure 18.1**).[4]

The hepatic enzyme, 25-hydroxylase, metabolizes vitamin D_3 to 25-hydroxyvitamin D. Serum 25-hydroxyvitamin D estimation is the best indicator of body stores; therefore, low 25-hydroxyvitamin D levels in the blood imply deficiency. The renal 1α-hydroxylase, under PTH regulation, completes the formation of the active metabolite, 1,25-dihydroxyvitamin D, also called $1,25(OH)_2D$, calcitriol, or simply vitamin D.

Endogenous synthesis of vitamin D is influenced by the amount of sun exposure, available sunlight (less in northern latitudes), skin covering (clothing and sun block), and age of the individual. Older individuals have less effective photobiosynthesis of vitamin D.

There are two forms of vitamin D: vitamin D_3 and vitamin D_2. Vitamin D_3 (cholecalciferol) is scarcely found in nature. The major dietary sources are internal organs such as liver and seafood. Vitamin D_2 (ergocalciferol) is found in edible mushrooms. Commercially available milk is fortified with vitamin D_3. Breastfed infants are at risk for vitamin D deficiency, since breast milk is a poor source of vitamin D. Multivitamins contain 400 units of vitamin D_3, about the same amount found in a quart of vitamin D fortified milk. Vitamin D_3 is more potent compared to vitamin D_2.

Vitamin D is a steroid hormone derived from cholesterol. The same cholesterol biosynthetic pathway that provides the precursors for vitamin D also provides the precursors for other steroid hormones from the adrenal gland and gonads. The steroid receptor supergene family of nuclear receptors carries out its physiologic regulation by directing transcription of specific vitamin D–responsive genes. $1,25(OH)_2D$ is the natural ligand for the vitamin D receptor. Synthetic analogs of vitamin D such as paricalcitol bind to vitamin D receptors with very high affinity and are used to treat secondary hyperparathyroidism due to chronic renal failure.

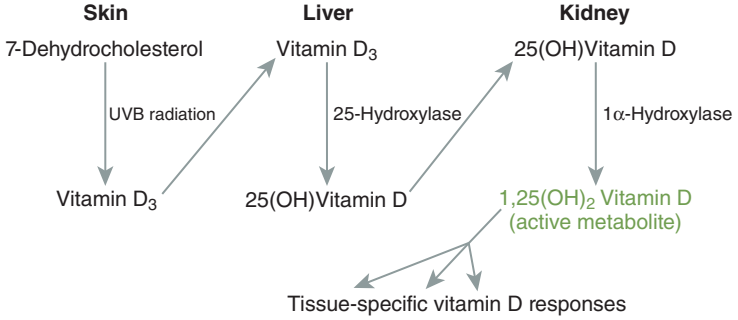


Figure 18.1 Vitamin D synthesis. Tissues involved in the synthesis of vitamin D and the steps that each tissue is responsible for. Also shown are enzymes responsible for the two enzymatically mediated steps (hepatic 25-hydroxylation and renal 1α-hydroxylation). The product of this pathway, $1,25(OH)_2D$, is responsible for the tissue-specific effects of vitamin D.

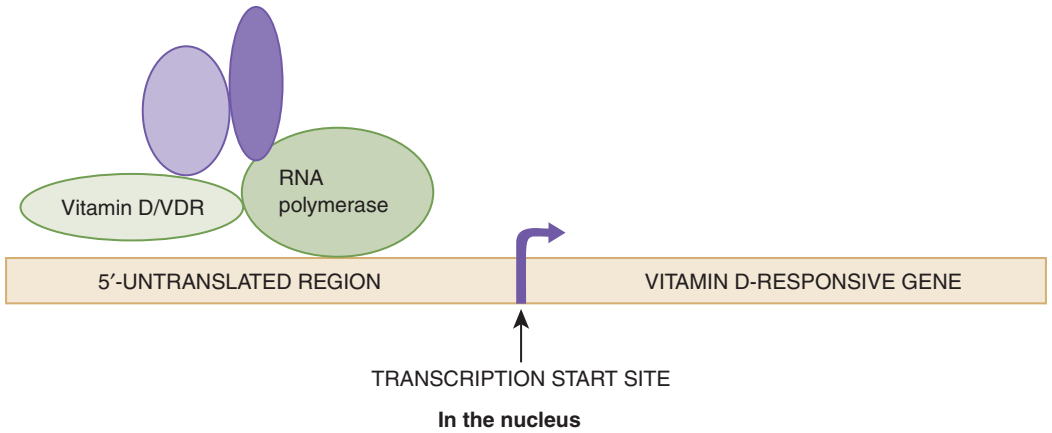


Figure 18.2 Vitamin D mechanism of action. DNA binding and interaction with other components of the transcriptional machinery of the vitamin D–vitamin D receptor complex are shown. Note the striking evolutionary similarity between this mechanism and that of other steroid and thyroid hormones. As prime examples, vitamin D inhibits transcription of the *PTH* gene in parathyroid tissue and stimulates transcription of the calcium transporter in the intestinal brush border epithelium.

The $1,25(OH)_2D$–vitamin D receptor complex binds to the vitamin D response element upstream (5′) of the transcription start site of vitamin D genes and induces gene transcription (**Figure 18.2**).

Activated vitamin D, $1,25(OH)_2D$, induces active absorption of calcium in the small intestinal cells, which is the dominant mechanism of calcium absorption in humans. Only about 5% to 10% of calcium is absorbed passively.[5,6] The lack of active intestinal absorption of calcium induced by activated vitamin D, $1,25(OH)_2D$, results in **hypocalcemia**. Even though activated $1,25(OH)_2D$ also stimulates absorption of phosphate, a larger fraction of phosphate absorption occurs passively and is much less dependent upon normal $1,25(OH)_2D$ levels. This mechanism is illustrated in chronic renal failure patients.

In the bone, $1,25(OH)_2D$ stimulates differentiation of **osteoclast** precursors to osteoclasts. $1,25(OH)_2D$ also stimulates **osteoblasts** to influence osteoclasts to mobilize bone calcium. $1,25(OH)_2D$ does not directly affect mature osteoclast function. $1,25(OH)_2D$ plays an important role in bone mineralization, a process of adding minerals to the extracellular matrix, and abnormal bone results when vitamin D is deficient (e.g., celiac disease and other malabsorption disorders) or its metabolism is defective (e.g., renal failure).

As noted earlier, $1,25(OH)_2D$ increases blood calcium by augmenting intestinal absorption of calcium. Blood calcium feeds back to parathyroid glands and regulates the synthesis and secretion of PTH. Also, $1,25(OH)_2D$ has direct control over PTH secretion, and high levels reduce secretion of PTH. Elevated serum phosphate levels reduce

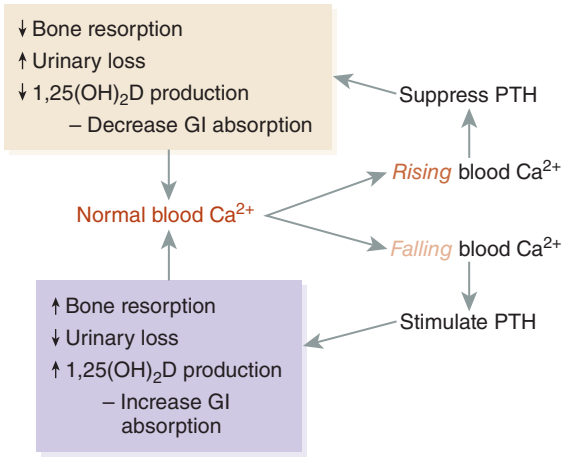


Figure 18.3 Calcium, PTH, and vitamin D feedforward and feedback loops. The endocrine response to shifts in blood calcium (rising or falling) is shown. A concerted hormone response helps restore blood calcium toward normal.

$1,25(OH)_2D$ formation (**Figure 18.3**). The reference range for 25-hydroxyvitamin D is 20–50 ng/mL and for $1,25(OH)_2D$ is 18–64 pg/mL for males and 18–64 pg/mL for females.

Effects of the Parathyroid Hormone

For calcium to be maintained at optimal levels, **vitamin D** and PTH work in conjunction with three organs: the gastrointestinal (GI) tract, the kidneys, and bones.

GI Regulation

Normal intestinal function is required for calcium absorption. Altered intestinal function may be seen with short bowel syndrome (following resection), gastric bypass surgery for weight loss, intestinal mucosal disease (celiac disease), genetic defects, and bowel fistula, which may affect calcium absorption. Adequate dietary calcium intake, the availability of normal amounts of vitamin D, and metabolism are all necessary for optimal calcium absorption. $1,25(OH)_2D$ controls calcium absorption from the small intestines. Blood calcium feeds back to parathyroid glands and regulates the synthesis and secretion of PTH.

Role of Kidneys

The kidneys play an essential role in calcium metabolism. This is exemplified in subjects who develop renal failure, leading to disturbances of calcium and phosphate metabolism.[7] Impaired hydroxylation of 25-hydroxyvitamin D to form the active $1,25(OH)_2D$ results in poor calcium absorption from the gut. Diseased kidneys fail to excrete phosphate, leading to elevated serum phosphate levels. Hyperphosphatemia is a powerful stimulus for PTH and **fibroblast growth factor 23 (FGF23)** secretion. FGF23 is a strong phosphaturic hormone secreted by the osteocytes that binds to the FGF receptor 1 and co-receptor Klotho in the kidney. FGF23 inhibits 1α-hydroxylation of inactive 25 vitamin D to active $1,25(OH)_2D$ but also stimulates 24α-hydroxylation of $1,25(OH)_2D$, causing inactivation. The net result is very low $1,25(OH)_2D$ levels, leading to hypocalcemia. Serum PTH levels can be elevated in chronic renal failure patients. Once a critical calcium/phosphate product is reached, precipitation of these minerals occurs within the tissues, leading to devastating consequences.

In physiological state, PTH increases renal tubular reabsorption of calcium, thereby reducing renal loss. **Hypercalcemia**, an increased calcium level in the blood, whether resulting from autonomous overproduction of PTH, termed *primary hyperparathyroidism (PHPT)*, or from other causes, increases the filtered load of calcium. Even though PTH stimulates tubular reabsorption of calcium, this process is overwhelmed, and the net calcium excretion is increased compared with normal state. This is the fundamental mechanism by which PHPT subjects form calcium kidney stones. See **Figure 18.4** for organ system integration of calcium homeostasis.

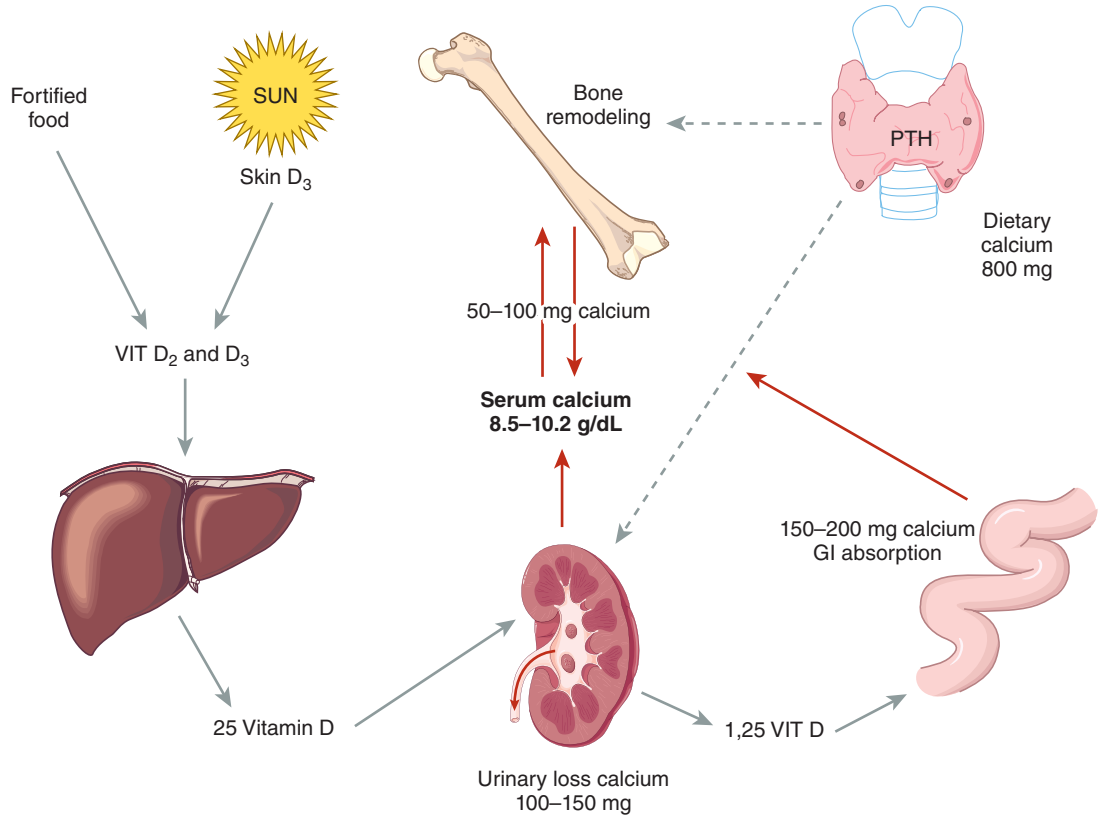


Figure 18.4 Organ system integration of calcium homeostasis.

Bone Physiology

Bone contains about 1 kg of calcium and serves as a repository for calcium, phosphate, and magnesium. **Bone turnover** or "remodeling" is a regulated and coupled process of simultaneous bone formation and breakdown that occurs to varying degrees throughout life and at multiple sites in the skeleton. The process of "coupling" ensures that neither resorption nor formation occurs in excess compared to the other, thereby preventing imbalance between formation and loss of bone. Too much resorption results in weaker bones, while too much bone formation can obliterate the red marrow space. Bone formation is mediated by osteoblasts, and bone breakdown, or resorption, is mediated by osteoclasts (a cell in the monocyte/macrophage lineage). Interestingly, although osteoclasts are required to mobilize calcium from bone, they do not appear to express receptors for either $1,25(OH)_2D$ or PTH. Rather, these hormones act directly on osteoblasts, which in turn produce a complex array of hormones that regulate the osteoclast activity by either activating or inactivating them. When the net rates of osteoblast-mediated bone formation and osteoclast-mediated bone resorption are mismatched, or uncoupled, such that resorption exceeds formation, bone mass will ultimately decrease. Decreased bone mass translates clinically into an increased risk of fracture, as seen in osteoporosis. As bone turns over, several proteins are released into the blood; these are eventually excreted in the urine, and collectively, they are viewed as bone turnover markers.[8] Bone turnover markers are useful in the monitoring of clinical response to osteoporosis therapies and for predicting fracture risks.

Bone resorption markers include the following:

- Hydroxyproline
- N-telopeptide/C-telopeptide
- Pyridinium Crosslinks
- Tartrate-resistant acid phosphatase (TRAP)

Bone formation markers include the following:

- Alkaline phosphatase (total and bone-specific ALP)
- Osteocalcin
- Procollagen N-terminal extension peptides

There are two main types of bones found in the skeleton: cortical bone and trabecular bone. Cortical bone is the predominant bone type found in the shaft of long bones, such as the femur. Cortical bone is very strong yet light in weight and is therefore well suited for the mechanical needs of the extremities. Trabecular bone is the primary bone type found in the axial skeleton, such as the vertebrae. It consists of numerous cross-hair or honeycomb-type connections called *trabeculae*. These interconnected trabeculae give this type of bone considerable strength for weight bearing. Although the femur is primarily cortical bone and the vertebrae is primarily trabecular bone, sites such as the femoral trochanter and proximal humerus are a mixture of both cortical and trabecular bones. The ends of long bones contain more trabecular bone. The shafts of long bones are predominantly cortical. Certain disease states induce preferential bone loss in specific bone types. In PHPT, cortical bone is preferentially lost, leading to weak bones and potential fractures.

In summary, normal calcium homeostasis is achieved by a complex and regulated interplay between vitamin D, PTH, and the target organs. Defects in the function and formation of hormones or organ system disorders can induce disease states, which will be reviewed in the rest of this chapter.

CASE STUDY 18.1, PART 2

Remember Amanda. Her physician reviews her medical history and finds no significant information. Amanda's family history contributes no pertinent information to the cause of her symptoms. She is not taking any medications or supplements. On physical examination, she appears to be in extreme pain. There is marked tenderness on very gentle percussion over the left costovertebral angle.

1. Given Amanda's signs and symptoms, what lab test should the physician order to determine the cause of her pain?

Parathyroid Disorders

Hypercalcemia results when blood calcium levels in a subject are above the expected normal range of a healthy population. The duration and the level of calcium elevation determine the severity of symptoms and the degree of end-organ damage. Some common symptoms of severe hypercalcemia characterized by critically high calcium results include lethargy, stupor, and coma. Patients with chronic, mild hypercalcemia are symptom free. Moderate calcium elevation is associated with intellectual weariness, personality changes, nausea, anorexia, polyuria, kidney stones, hypertension, and electrocardiogram (ECG) changes. In the outpatient setting, PHPT is the most common cause for hypercalcemia.

Primary Hyperparathyroidism

Primary hyperparathyroidism (PHPT) is the most common cause of hypercalcemia, with an incidence of 1 in 500 to 1000 in the general population, and women are affected three times more than men. The hallmark of this disorder is the autonomous overproduction of PTH, most commonly by a single adenoma in one of the parathyroid glands (>80%), less commonly by multiple gland hyperplasia (~5% to 10%), and very rarely by parathyroid cancer. Parathyroid carcinoma occurs more commonly in a familial syndrome called the hyperparathyroidism–jaw tumor syndrome. PHPT is also encountered in multiple endocrine neoplasia types 1 and 2 (MEN1 and MEN2a). The apparent increase in the incidence of PHPT was traced to the introduction and wide availability of multichannel analyzers in the 1970s. The incidence of PHPT increased from 7.8 cases per 100,000 to 51 cases per 100,000 in Olmsted County, Minnesota, United States, following the introduction of routine calcium testing.[8] Fuller Albright's description of the disease in the 1930s was notable for advanced bone and end-stage kidney disease at the time of diagnosis.[9] This presentation is very uncommon. The fortuitous detection of a large proportion of the prevalent subclinical PHPT subjects, along with early intervention and treatment, has changed the clinical presentation of PHPT. In countries where multichannel analyzers are not available, the disease presentation is severe and has catastrophic renal and skeletal consequences, as it was in the United States in the early 1900s. The dominant clinical phenotype of PHPT in the United States is asymptomatic hypercalcemia, invariably identified in patients undergoing yearly physical examination. The apparent increase in the incidence of PHPT was correlated to the wide availability and use of multichannel testing since the 1970s.[10] In a Kaiser Permanente study involving 3.5 million people, the incidence of PHPT varied from 34 to 124 per 100,000 person years. The incidence increases with advancing age. Women are affected threefold more frequently than men.[11] Along with the increase in the diagnosis of patients with PHPT, the surgical case volume increased an astounding 177% in a state-wide health planning study from California between 1999 and 2008.[12]

The diagnosis of PHPT is made by biochemical testing, and elevated serum calcium (or ionized calcium) level with inappropriately normal or elevated PTH level confirms the diagnosis. The biochemical differential diagnosis of hypercalcemia is listed in **Table 18.1**. The reference range for PTH is 11–54 pg/mL.

CASE STUDY 18.1, PART 3

Recall Amanda. Her physician orders a complete metabolic panel and urinalysis. The labs are drawn, and the only notable result is a calcium level of 11.2 mg/dL (reference range: 8.5–10.2 mg/dL). Renal function is normal with a BUN of 25 mg/dL (reference range: 6–20 mg/dL) and creatinine of 0.9 mg/dL (reference range: 0.6–1.1 mg/dL). Urinalysis is notable for blood and >50 red blood cells per high-power field (reference range: 0–2 RBC/HPF). This result prompts a 24-hour urine collection, which reveals elevated calcium of 483 mg/24 h (reference range: 100–250 mg/24 h). A CT scan of her abdomen is ordered, and kidney stones are noted.

2. Based on these results, what follow-up tests should the physician order?

Table 18.1 Biochemical Differential Diagnosis of Hypercalcemia

Causes	Sr CA	Sr PO_4	PTH	PTHrP	1,25 Vit D	Urine Calcium
PHPT	↑	↓	↑→	↓	↑	↑↑
Cancer PTHrP	↑↑	↓	↓	↑	↓	↑↑
CRF–3HPT Creatinine ↑	↑↓	↑↑	↑↑	↓	↓	↓
FHH Mg ↑	↑	→	↑→	↓	→	↓↓

PTH, parathyroid hormone; PTHrP, parathyroid hormone–related protein; PHPT, primary hyperparathyroidism; CRF, chronic renal failure; 3HPT, tertiary hyperparathyroidism; FHH, familial hypocalciuric hypercalcemia.

Biochemical Findings in Primary Hyperparathyroidism

In a patient with primary hyperparathyroidism (PHTP), the following biochemical findings are generally present:

1. Hypercalcemia
2. Hypophosphatemia: PTH induces phosphaturia
3. Elevated PTH relative to serum calcium (inappropriately normal PTH in the face of hypercalcemia)
4. Low-normal 25-OH-D and high-normal or elevated (30%) $1,25(OH)_2D$ due to PTH-enhancing renal hydroxylation
5. Elevated urinary calcium excretion
6. Metabolic hyperchloremic acidosis (chloride is exchanged for phosphate)
7. Elevated serum alkaline phosphatase in severe disease

PTH Assays

PTH is a polypeptide hormone with carboxy (C)-terminal and amino (N)-terminal ends composed of 84 amino acid residues. The integrity of 1 to 11 amino acid residues in the N-terminal end of PTH is required for the biologic activity of PTH. PTH 1 to 84 (sometimes refered to as intact PTH) refers to the active hormone sequence, which has a half-life of several minutes. *n*-Truncated PTH (7 to 84 PTH) are deficient in the first few residues of the N-terminal end, and C-terminal portions (34-84 PTH) are both biologically inactive and have half-lives of 20–36 hours. Common PTH assays use a sandwich technique in which two antibodies bind to specific sites on the PTH molecule: a solid phase antibody binds to the C-terminal end and a signal antibody binds the N-terminal ends, which enable precise detection of intact PTH. Electrochemiluminescent assays are also available for intact PTH measurement. A few commonly available assays include the following:

- Sandwich ELISA or electrochemiluminescent PTH (intact, whole): measures the entire 1 to 84 PTH sequence.
- CAP assay: cAMP inducible PTH. This assay detects biologically active PTH by its ability to induce formation of cAMP.
- PRHrP assays and PTH assays are exclusive to each other, by design. There is no cross-reactivity. Therefore, PTH levels are low in subjects with cancers producing PTHrP.
- PTH can also be estimated in needle biopsy specimens obtained from parathyroid tumors. Hook effect has not been observed despite massive elevation of PTH levels in these specimens obtained by direct puncture of enlarged parathyroid glands.[13]

Management of PHPT

After a diagnosis of PHTP is made, a treatment plan is initiated. The treatment for primary hyperparathyroidism is usually parathyroidectomy.

Surgical resection of the affected gland(s) easily results in cure. Only about 30% of patients with PHPT need surgery at the time of diagnosis. Therefore, the following consensus criteria were generated by a panel of experts, and some changes to these criteria were adopted later; the goal is for the early identification of subjects that are at risk of developing end-organ damage and to prevent progression. The

indications for surgery recommended by the expert panel are as follows:

Parameter	2002	2009	2014
Serum calcium	>1 mg	>1 mg	>1 mg
Urine calcium	>400 mg/d	Not recommended	>400 mg/d and nephrolithiasis or risk
Creatinine clearance	Reduction of 30% or more	<60 mL/min	<60 mL/min
BMD-DXA	T-score < –2.5 at any site	T-score < –2.5 same plus previous fracture	T-score < –2.5 at any site and vertebral fracture by other imaging
Age	<50 y	<50 y	<50 y

Surgical management of PHPT involves neck exploration with removal of the abnormal gland. Since most individuals have four parathyroid glands, minimally invasive surgery (MIS) utilizes presurgical imaging to locate the abnormal gland, thereby guiding the surgeon to the diseased gland. MIS results in smaller incision; high-resolution neck ultrasound and the shorter technetium-99m (Tc-99m) sestamibi scan are the two commonly used imaging tools. Parathyroid adenomas are rich in mitochondria, which trap the Tc-99m isotope. (This scan is also used in cardiac imaging, as the heart is another mitochondria-rich organ.) Also, with the availability of rapid PTH, also known as intraoperative PTH, levels are used to monitor the adequacy of parathyroid removal. Once the adenoma is removed, PTH levels plummet; since the half-life of PTH is <5 min, this enables biochemical confirmation of adenoma removal. A baseline PTH sample is taken before surgery and then another during the procedure but before excision of the adenoma. Once resection of the suspected abnormal parathyroid gland is done, an intraoperative PTH is drawn and run as a rapid sample. If the post-removal PTH specimen level drops 50% or more from baseline, the correct parathyroid glad was removed. If not, further exploration and resection of remaining glands is needed until a 50% reduction in PTH is realized.

Medical management of PHPT is practiced in poor surgical candidates and those with mild disease. **Bisphosphonates**, selective estrogen receptor modulators (SERMs), and estrogen are all useful to varying degrees in any given patient. An allosteric calcium receptor modulator, such as cinacalcet, that binds to CSR is used for controlling calcium levels in select patients with PHPT.

PHPT: Important Points

Insidious onset and gradual progression
Kidney stones in 20% to 30%
Bone loss—cortical > trabecular bone
Surgery is the only definitive management
Single gland tumor is common (85%)

CASE STUDY 18.1, PART 4

Recall Amanda. The crystal analysis from the urinalysis timed collection was reported to be composed of calcium oxalate, and the PTH result was 162 pg/mL (reference range: 11–54 pg/mL).

3. What is the diagnosis for this patient?
4. What treatment is indicated for this disease?

Secondary Hyperparathyroidism

Hypocalcemia refers to low blood calcium levels. It can result from a wide variety of conditions, from organ system dysfunction to a lack of hormone effect to acid–base disturbances. The causes of hypocalcemia will be discussed in the setting of both endocrine and organ system dysfunction. However, one key concept should be emphasized when considering hypocalcemia: when functioning properly, the parathyroid glands will not only correct falling blood calcium but also prevent it by increasing PTH secretion. The compensatory rise in PTH secretion in response to factors that would lower blood calcium is known as *secondary hyperparathyroidism*. Thus, an individual may have an elevated PTH level in response to low 25-hydroxyvitamin D and still maintain normal calcium levels. Secondary hyperparathyroidism is notable for the normal response of parathyroid glands with appropriate and vigorous secretion of PTH. The biochemical constellation includes low blood calcium, elevated PTH, low serum phosphate, elevated alkaline phosphatase, hypocalciuria, phosphaturia, and vitamin D deficiency or lack of vitamin D effect.[14] Treatment of secondary hyperparathyroidism is directed toward correcting the process that is inducing hypocalcemia and/or vitamin D deficiency.[15]

Secondary and Tertiary Hyperparathyroidism in Renal Failure

The kidney's central role in the regulation of bone and mineral metabolism was discussed earlier. Chronic kidney disease (CKD) results in striking bone mineral and skeletal changes. The diseased kidneys fail to excrete phosphate; this, along with the impaired formation of $1,25(OH)_2D$, leads to a vicious cycle of events resulting in parathyroid gland stimulation and hyperplasia. The bone mineral metabolic changes are progressive and proportional to the severity of renal dysfunction and eventually ubiquitous in patients with end-stage renal disease. In severe cases, the ensuing parathyroid hyperplasia results in autonomy of the parathyroid gland, leading to hypercalcemia; this is also referred to as tertiary hyperparathyroidism (although the term *tertiary* is not widely used nowadays). The difference between secondary and tertiary hyperparathyroidism is the development of sustained hypercalcemia. Patients with renal failure who develop transient, iatrogenic hypercalcemia should not be given the diagnosis of tertiary hyperparathyroidism. In early stages of CKD, in response to low blood calcium and elevation of phosphate levels, compensatory elevation of PTH and FGF23 maintains near-normal calcium and phosphorus levels.[14] As kidney disease becomes severe, the compensatory mechanisms are overwhelmed, leading to permanent abnormalities of calcium, phosphorus, PTH, vitamin D, bone mineralization, and vascular and soft tissue calcification. These biochemical changes completely resolve following renal transplantation. Renal replacement therapies such as hemodialysis and peritoneal dialysis only induce partial correction of these changes.[16]

One of the earliest changes in CKD is reduction of urinary phosphorus excretion, leading to an increase in serum phosphorus levels. Compensatory increase in the phosphaturic hormones such as FGF23 and PTH helps maintain near-normal serum phosphate levels for a period. When renal function declines to lower than 30 mL/min (the glomerular filtration rate), hyperphosphatemia is observed. Hyperphosphatemia also stimulates PTH secretion and inhibits $1,25(OH)_2D$ production; the low $1,25(OH)_2D$ leads to poor GI absorption of calcium. Hyperphosphatemia and elevation of FGF23 levels are independent predictors of survival; higher levels are associated with poorer outcomes.[14] Hyperphosphatemia is treated with a combination of dietary restriction of phosphorus and oral phosphate binders. Commonly used phosphate binders include calcium carbonate, calcium acetate, and noncalcium binders including sevelamer and lanthanum. Patients with CKD have low blood calcium levels. The imbalance in calcium and phosphate levels results in extracellular deposition of calcium, resulting in vascular and tissue calcification, a condition called calciphylaxis, which results in increased mortality.

Most patients with CKD also have low $1,25(OH)_2D$ levels. As phosphorus levels increase, FGF23 levels increase as a compensatory mechanism, which in turn inhibits 1α-hydroxylation of 25-hydroxyvitamin D. The decreased $1,25(OH)_2D$ leads to impaired calcium absorption, and the lower calcium levels stimulate PTH release. This triggers a cascade of bone mineral metabolic events, leading to poor bone mineralization.

PTH is secreted in response to hypocalcemia, hyperphosphatemia, and/or low $1,25(OH)_2D$ levels. PTH levels trend up with increasing severity of renal disease. Secondary hyperparathyroidism is treated with vitamin D analogs such as calcitriol, doxercalciferol, and paricalcitol.[15] Calcimimetics are allosteric activators of the extracellular CSR, sensitizing the parathyroid gland to extracellular calcium and decreasing PTH secretion independent of vitamin D. Calcimimetics such as cinacalcet have been shown to decrease PTH, calcium, and phosphorus.[17]

Familial Hypocalciuric Hypercalcemia

Familial hypocalciuric hypercalcemia (FHH) is a benign condition that results from germline mutation involving the CSR.[18,19] These inactivating mutations affecting the CSR upset their operating set point ranges (**Figure 18.5**). Unlike in PHPT, the PTH production and calcium elevation in this condition are not progressive and result in stable mild hypercalcemia since birth. The hallmarks of this disorder include mild hypercalcemia, hypocalciuria (renal tubular cell calcium sensing is upset, resulting in urine 24-hour calcium levels of <100 mg/d), mild PTH elevation (the parathyroid cellular calcium sensing is also abnormal), and mild elevation of magnesium. The hallmark is low urinary calcium excretion. Urinary fractional excretion in this condition is less than 1%, that is, < 0.01. Magnesium levels are mildly elevated. Inheritance is autosomal dominant with 100% penetrance among carriers, resulting in multiple members in the family being affected.

The following are the salient clinical and diagnostic features in FHH:

1. Serum calcium is elevated (mild to moderate elevation).
2. Serum magnesium is mildly elevated.
3. PTH is mildly elevated.
4. Urine calcium levels are typically < 100 mg in 24 hours, with fractional excretion of urine calcium of less than 1% (i.e., ≤0.01 is highly suggestive of a diagnosis of FHH).
5. Mutant CSR can be demonstrated.
6. End-organ dysfunction (bone and kidney) is uncommon; therefore, surgery should not be performed.

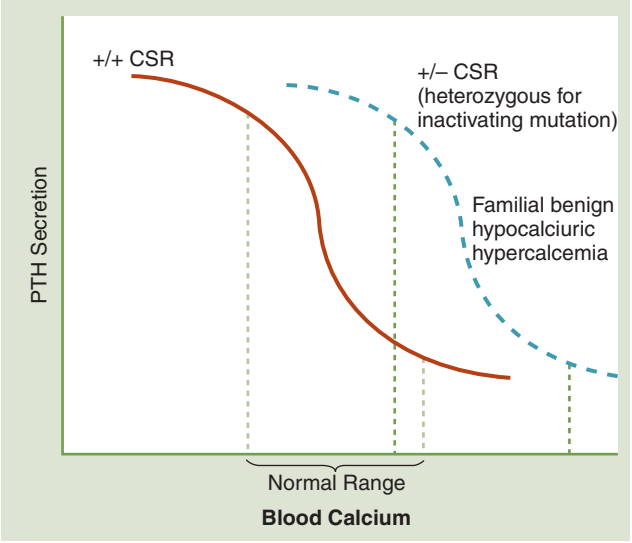


Figure 18.5 Calcium-sensing receptors' effect on PTH secretion. Shown is the response of parathyroid tissue (as demonstrated by PTH secretion) to blood calcium. The *set point* is determined by the parathyroid response mediated by the transmembrane CSR. The *normal curve* is for heterozygosity for the "wild-type" SCR (+/+). The *right-shifted curve* is shown for the case where there is heterozygosity for the receptor (familial benign hypocalciuric hypercalcemia): one wild-type copy of the gene and one inactivating mutation (+/–).

Antibodies to the CSR upset the set point and induce hypercalcemia; these patients have normal calcium receptors. Surgical removal of parathyroid gland(s) should not be performed in these patients. Hence, it is very important to differentiate this disorder from PHPT to avoid unnecessary and ineffective surgeries.

CASE STUDY 18.2, PART 2

Recall Robert. He noted his cigarettes have not been as enjoyable for him as they used to be. He complained of a worsening baseline cough, and he has noticed blood streaking his sputum when he clears his throat. He has no significant past medical history other than his tobacco use. Robert takes no medications. His family history is only notable for his father dying of lung cancer at age 63 and his mother dying of emphysema at age 68. On physical examination, he is a thin man who looks much older than his chronologic age and appears unwell. When he produces some sputum at the physician's request, it does indeed have a pink tinge and is streaked with blood. Chest examination reveals some scattered wheezing and some rales in the right upper lung region. He is diffusely weak on muscle strength testing.

1. Based on Robert's presentation, what lab tests should the physician order?

FHH: Important Points

Low urine calcium
No end-organ damage
Multiple family members are affected
Mild elevation of magnesium
Surgery is ineffective and should not be performed

Hyperthyroidism

Due to the direct effects of thyroxine on the bones, there is increased bone resorption and hypercalcemia in patients with hyperthyroidism. PTH levels are low in these patients.

Addison's Disease

Hypercalcemia is commonly seen in patients presenting with Addisonian crisis as PTH levels are low. Fluid and steroid replacement promptly corrects it. Hypercalcemia in this circumstance is thought to be a result of hypovolemia and dehydration; however, the exact mechanism is unknown.

Medications That Cause Hypercalcemia

Hydrochlorothiazide (HCTZ) and lithium are the most common drugs that induce hypercalcemia. HCTZ, a diuretic used to treat hypertension, enhances distal tubular calcium absorption, leading to hypercalcemia. Lithium, used in treating bipolar and other mental health disorders, affects the formation of intracellular inositol triphosphate, thereby upsetting the CSR function (similar to FHH).

Hypervitaminosis

Hypervitaminosis D is a condition that may result from excessive oral intake of vitamin D supplements. It may also result from aberrant production of $1,25(OH)_2D_3$ as a result of extrarenal 1α-hydroxylation of 25-hydroxyvitamin D.

Parathyroid Hormone–Related Protein

Parathyroid hormone–related protein (PTHrP) is a substance secreted by cancers that shares structural similarities to the N-terminal portion of human PTH molecule. Therefore, it retains functional features of PTH, with some critical differences. Both PTH and PTHrP bind to the same receptor in the kidney and bone, as well as a variety of other tissues. In humans, PTHrP functions in the paracrine regulation of cartilage, skin, brain, and lactating breast tissue. In healthy humans, circulating levels of PTHrP are very low or immeasurable. Normal lactating breast tissue is capable of producing PTHrP, which can result in hypercalcemia and resolves following the cessation of breast-feeding. In most clinical situations of hypercalcemia due to PTHrP secretion, an underlying cancer is the root cause.[20] PTHrP can be secreted by a variety of cancers. The secretion is not regulated by even very high blood calcium levels, which results in severe hypercalcemia (**Figure 18.6**). When humoral hypercalcemia of malignancy is suspected, PTHrP can be measured by specific immunoassay. Also, the intact PTH assay does not cross-react with circulating PTHrP. Normal parathyroid glands sense the elevated calcium levels and markedly reduce the secretion of PTH; therefore, PTH is very low. One salient difference in the biological functions between PTH and PTHrP is the inability of PTHrP to facilitate renal hydroxylation of 25-hydroxyvitamin D. This dichotomy in function between these otherwise functionally similar protein hormones remains unexplained.

Hypoparathyroidism

Because proper PTH secretion and action are necessary to maintain normocalcemia, any inadequacy of parathyroid gland function will cause hypocalcemia. The most common cause of hypoparathyroidism is neck surgery, especially, after thyroidectomy with lymph node dissection. This results from accidental removal of these glands due to their small size. The most common outcome is temporary

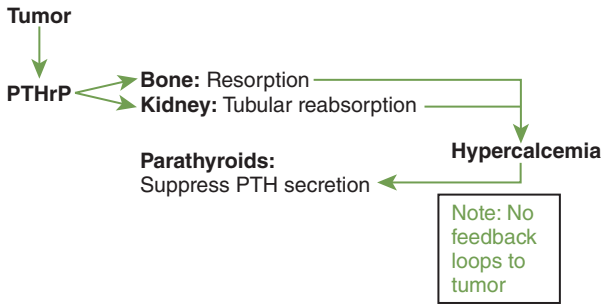


Figure 18.6 PTHrP endocrine pathophysiology. This demonstrates the effect of tumors that overproduce PTHrP. The pathophysiologic effect is via the same organ systems used by PTH to increase blood calcium. The difference between PTH and PTHrP is that PTH is subject to feedback regulation, whereas PTHrP is not subject to any feedback regulation by calcium (compare with Figure 18.4).

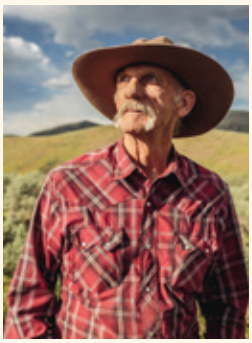
CASE STUDY 18.2, PART 3

Recall Robert. His physician orders a complete metabolic panel and a chest x-ray. The notable chemistry results are:

Calcium, 16.8 mg/dL (reference range: 8.5–10.2 mg/dL)
Albumin, 3.4 (reference range: 3.5–4.8 g/dL)
BUN, 27 mg/dL (reference range: 6–20 mg/dL)
Creatinine, 1.3 mg/dL (reference range: 0.6–1.1 mg/dL)

Robert's chest radiograph reveals a 3-cm proximal right hilar mass with distal streaking.

2. Given Robert has a mass in his lung as well as hypercalcemia, what follow up testing should be done?
3. What is the relationship between Robert's new diagnosis of lung cancer and hypercalcemia?



hypoparathyroidism following surgery, due to damage of the delicate parathyroid blood supply, which results in full recovery in most subjects. Parathyroid glands can be reimplanted into pouches created in skeletal muscles (such as the deltoid, sternomastoid, or forearm muscles). Parathyroid glands also survive careful cryopreservation and reimplantation back into their original owner. Additional causes of hypoparathyroidism include autoimmune destruction of parathyroid tissue. This condition is often associated with other autoimmune diseases such as type 1 diabetes, Hashimoto's thyroiditis, and Addison's disease. Magnesium deficiency can inhibit the secretion of PTH and also blunt its actions on target tissues. Following the correction of hypomagnesemia, PTH secretion and function are reestablished. Depending on the cause, hypoparathyroidism can usually be treated with relatively high doses of vitamin D and calcium. In the absence of PTH, even the small fraction of passively absorbed calcium may simply be excreted in the urine. Hypercalciuria increases the risk of development of kidney stones in these subjects. The use of HCTZ reduces urine calcium losses and elevates serum calcium levels in PTH-deficient subjects.

Pseudohypoparathyroidism is a inheritable disorder resulting in a lack of responsiveness to PTH in the target tissue. This results from uncoupling of the PTH receptor from adenylate cyclase, due to a mutant stimulatory G protein (G_s). PTH binds its receptor but cannot activate the second messenger, cAMP, and thus, there is no response. Hypocalcemia develops, although unlike other forms of hypoparathyroidism mentioned, those with pseudohypoparathyroidism have markedly elevated levels of PTH. This is an example of a hormone-resistance syndrome. Treatment is with calcium and vitamin D supplementation.

Hypovitaminosis D describes a collection of conditions, including low vitamin D availability, defective metabolism of vitamin D, or mutations in the vitamin D receptor, all of which predispose patients to hypocalcemia.

Organ System Causes of Hypocalcemia

A variety of intestinal disorders can result in malabsorption of calcium or vitamin D, resulting in hypocalcemia. Causes include short bowel syndrome, abdominal irradiation, weight-loss surgeries, celiac disease, and bowel fistulation. Treatment endpoints are normalization of urine calcium excretion and normalization of PTH.

Metabolic Bone Diseases

A variety of disease states can affect skeletal architecture, strength, and integrity. These include **rickets**, **osteomalacia**, and **osteoporosis**.

Rickets and Osteomalacia

Rickets and osteomalacia are diseases caused by abnormal mineralization of bone. They result from vitamin D deficiency. *Rickets* refers to the disease state affecting growing bones (in children) that result in permanent skeletal deformity. Rickets is associated with bone deformities due to the softening in the epiphyseal plate (growth tissue) at the ends of the child's bones; this causes the bones to bend under weight, leading to bowing. Epiphyseal plate closure occurs during the teenage years for girls and by age 24 in boys. Therefore, bone deformity is not seen in adults. *Osteomalacia* refers to the abnormal mineralization of bone after completion of skeletal maturation (adulthood). Both conditions are associated with similar

biochemical findings of secondary hyperparathyroidism. Fractures may result in either case because of poor bone structure. Hypocalcemia may be seen when the response of secondary hyperparathyroidism is inadequate to counteract the threat of hypocalcemia posed by the vitamin D deficiency. Foods fortified with vitamin D, primarily dairy, are intended to reduce the risk of developing vitamin D deficiency. Despite these attempts, due to poor consumption of dairy and avoidance of sun exposure, vitamin D deficiency states do occur. However, those of any age who live indoors with minimal or no sun exposure, or who lack other dietary sources of vitamin D, are at risk for developing this condition. As mentioned in the discussion on vitamin D physiology, adequacy of vitamin D in the body can be assessed by measuring the blood level of 25-hydroxyvitamin D. Because secondary hyperparathyroidism is also expected in the setting of rickets or osteomalacia, PTH and calcium levels should be obtained to further confirm the suspected diagnosis.

Rickets can, however, develop under conditions of adequate amounts of vitamin D. This unique situation may develop from genetic defects in vitamin D metabolism or in the vitamin D receptor. Although 25-hydroxyvitamin D level is often normal, $1,25(OH)_2D$ level may be low, normal, or high, depending on the genetic defect. Defects in vitamin D metabolism are best treated by supplying the metabolically active compound, $1,25(OH)_2D$. A wide variety of vitamin D receptor defects have been described, including abnormal ligand binding, abnormal DNA binding, and abnormal transactivation of transcriptional machinery at the regulatory site of vitamin D–responsive genes. The type of defect present determines how well the patient will respond to pharmacologic doses of calcitriol.

Osteoporosis

Osteoporosis is the most prevalent metabolic bone disease in adults. Osteoporosis, meaning "porous bone," is a bone disease in which bone density decreases, usually as a product of either bone malformation or excessive bone resorption caused by a wide range of nutritional, endocrine, and metabolic disorders. It affects an estimated 10.2 million Americans, with an estimated 4:1 female-to-male predominance.[21] Because decreased bone density can lead to bone fragility, bone fractures are common; osteoporosis is believed to cause approximately 2.3 million fractures annually in the United States.[22] The most devastating consequence of osteoporosis is a hip fracture. While as many as half of vertebral compressions may be asymptomatic, a hip fracture carries with it a significant morbidity as well as an increased mortality. Most hip fractures require surgery at the very least. Mortality from hip fracture is increased by about 22% in the first year following the fracture, and it is estimated that the number of deaths related to hip fracture is comparable to that of breast cancer.[23,24]

Multiple additional conditions have been identified as significant risk factors for reduced bone mass and a consequent increased risk of fracture, although risk factor assessment alone is generally not sufficient to characterize or quantify bone mass and diagnose osteoporosis. The following are validated risk factors for the prediction of fracture, independent of formal bone density evaluation: decreased bone mass due to previous fracture, advanced age, family history of osteoporosis or fracture, body weight less than 127 lb., long-term glucocorticoid therapy, cigarette smoking, or excess alcohol intake. Other conditions known to also alter calcium metabolism and increase

CASE STUDY 18.3, PART 2

Remember Regan. She likes chocolate milk and drinks it regularly. Other than her shorter stature and bowed legs, she has the normal characteristics of her 6-year-old friends. She takes no medications. The family history is notable for some cousins on the father's side with a similar problem back in the Appalachian hill country along Virginia/Tennessee border, where the family originates. The pediatrician orders a complete metabolic panel. Notable results are:

Calcium, 7.2 mg/dL (reference range: 8.5–10.2 mg/dL)
Albumin, 4.1 g/dL (reference range: 3.5–4.8 g/dL))

Lower extremity radiographs show bowing of the long bones and generalized demineralization.

1. Based on Regan's presentation, what condition do the preliminary lab tests indicate?
2. What follow up laboratory tests should the physician order?

CASE STUDY 18.3, PART 3

Remember Regan. Her follow-up test findings include:

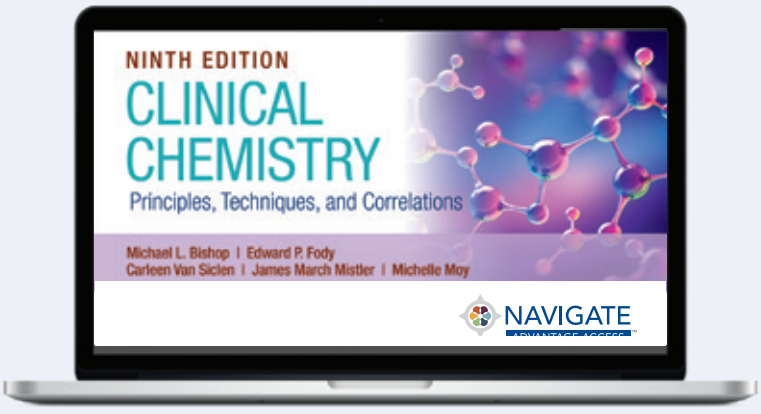
Intact PTH elevated: 866 pg/mL (reference range: 11–54 pg/mL)
25-hydroxyvitamin D: 35 ng/mL (reference range: 20–50 ng/mL)
$1,25(OH)_2D$: < 1 pg/mL (reference range: 18–64 pg/mL)

3. What condition do the preliminary lab tests indicate?
4. What is the significance of 25-hydroxyvitamin D and $1,25(OH)_2D$ levels in the follow-up laboratory tests?
5. Describe the inborn error of metabolism with this patient.

fracture risk include Cushing's syndrome, hyperparathyroidism, disorders of vitamin D metabolism, hyperthyroidism, and certain malignancies (such as mast cell disease). Several medications have negative effects on the skeleton, leading to low bone mass and an increase in the risk of fractures. The most notable are the glucocorticoids, which are widely used to treat a variety of inflammatory conditions such as asthma, rheumatoid arthritis, and lupus, as well as to prevent rejection after organ transplantation. "Glucocorticoid-induced osteoporosis" is a major source of morbidity associated with pharmacologic doses of glucocorticoids. Two bisphosphonates, alendronate and risedronate, are approved by the Food and Drug Administration (FDA) for the treatment of glucocorticoid-induced bone loss. Other medications that induce low bone mass include anticonvulsants (particularly phenytoin) and cyclosporin A.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Owen R. On the anatomy of the Indian Rhinoceros (*Rh. unicornis L.*). *Trans Zool Soc Lond.* 1862;4:31–58.
2. Gilmore JR. The gross anatomy of parathyroid glands. *J Pathol.* 1938;46:133.
3. Alveryd A. Parathyroid glands in thyroid surgery. *Acta Chir Scand.* 1968;389:1.
4. Norman AW, Roth J, Orci L. The vitamin D endocrine system: steroid metabolism, hormone receptors and biological response. *Endocrine Rev.* 1982;3:331–366.
5. Sheikh MD, Ramirez A, Emmett M, et al. Role of vitamin D dependent and vitamin D independent mechanisms in absorption of food calcium. *J Clin Invest.* 1988;81:126–132.
6. Gallagher JC, Riggs BL, Eisman J, et al. Intestinal calcium absorption and serum vitamin D metabolites in normal subjects and osteoporotic patients. *J Clin Invest.* 1979;64:729–736.
7. Sakou T. Bone morphogenetic proteins: from basic studies to clinical approaches. *Bone.* 1998;22:591–603.

8. Heath H, Hodgson SR, Kennedy MA. Primary hyperparathyroidism: incidence, morbidity and economic impact in a community. *N Eng J Med.* 1980;302:189–193.
9. Albright E, Reifenstein EC. *The Parathyroid Glands and Metabolic Bone Disease*. Baltimore, MD: Williams and Wilkins; 1948.
10. Yeh MW, Ituarte PHG, Zhou HC, et al. Incidence and prevalence of primary hyperparathyroidism in a racially mixed population. *J Clin Endocrinol Metab.* 2013;98(3): 1122–1129.
11. Abdulla AG, Ituarte P, Harari A, et al. Trends in the frequency and quality of parathyroid surgery: analysis of 17,082 cases over 10 years. *Ann Surg.* 2015;261(4):746–750.
12. Eastell R, Brandi ML, Costa AG, et al. Diagnosis of asymptomatic primary hyperparathyroidism: Proceedings of the Fourth International Workshop. *J Clin Endocrinol Metab.* 2014;99:3570–3579.
13. Abraham D, Sharma PK, Bentz J, et al. The utility of ultrasound guided FNA of parathyroid adenomas for pre-operative localization prior to minimally invasive parathyroidectomy. *Endocr Pract.* 2007;13(4):333–337.
14. Gutierrez OM, Mannstadt M, Isakova T, et al. Fibroblast growth factor 23 and mortality among patients undergoing hemodialysis. *N Engl J Med.* 2008;359(6):584–592.
15. Sprague SM, Coyne D. Control of secondary hyperparathyroidism by vitamin D receptor agonists in chronic kidney disease. *Clin J Am Soc Nephrol.* 2010;5(3):512–518.
16. Block GA, Klassen PS, Lazarus JM, et al. Mineral metabolism, mortality, and morbidity in maintenance hemodialysis. *J Am Soc Nephrol.* 2004;15(8):2208–2218.
17. Block GA, Martin KJ, de Francisco AL, et al. Cinacalcet for secondary hyperparathyroidism in patients receiving hemodialysis. *N Engl J Med.* 2004;350(15):1516–1525.
18. Brown EM, Gamba G, Riccardi D, et al. Cloning and characterization of an extracellular calcium sensing receptor from bovine parathyroid. *Nature.* 1993;366:575.
19. Hendy GN, D'Souza-Li L, Yang B, et al. Mutations of the calcium-sensing receptor (CASR) in familial hypocalciuric hypercalcemia, neonatal severe hyperparathyroidism, and autosomal dominant hypocalcemia. *Hum Mutat.* 2000;16:281.
20. Lips P, van Schoor NM, Bravenboer N. Vitamin D-related disorders. In: Rosen CJ, ed. *Primer on the Metabolic Bone Diseases and Disorders of Mineral Metabolism*. 7th ed. Washington, DC: American Society of Bone and Mineral Research; 2008:329.
21. International Osteoporosis Foundation. *Key statistics for North America: United States*. Accessed at https://www.osteoporosis.foundation/facts-statistics/key-statistic-for-north-america on August 20, 2021.
22. Whitaker Elam RE, Jackson NN, Machua W, Carbone LD. Osteoporosis. *Medscape*. January 20, 2021. Accessed at https://emedicine.medscape.com/article/330598-overview#a5 on August 20, 2021.
23. Sing C-W, Lin T-C, Bartholomew S, et al. Global epidemiology of hip fractures: a study protocol using a common analytical platform among multiple countries. *BMJ Open*. 2021;11(7):e047258.
24. Barclay L. Mortality greater for hip fracture than breast cancer in elderly women. *Medscape*. May 18, 2007. Accessed at https://www.medscape.com/viewarticle/556768 on August 20, 2021.

Organ Systems and Clinical Correlation

Liver Function

Janelle M. Chiasera

CHAPTER OUTLINE

KEY TERMS

Bile
Bilirubin
Cirrhosis
Conjugated bilirubin
Crigler-Najjar syndrome
Delta bilirubin
Hepatic jaundice
Hepatitis
Hepatocellular carcinoma
Hepatoma
Icterus
Jaundice
Kupffer cells
Posthepatic jaundice
Prehepatic jaundice
Unconjugated bilirubin
Uridine diphosphate glucuronosyltransferase
Urobilinogen

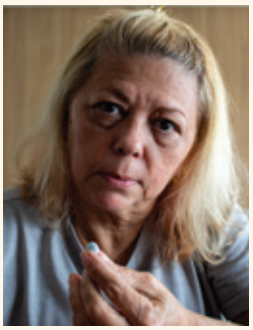
CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian will be able to do the following:

- Differentiate between unconjugated and conjugated bilirubin in the blood.
- Explain liver function including bile secretion, synthetic activity, and detoxification.
- Name two important cell types associated with the liver and state the function of each.
- Define jaundice and differentiate prehepatic, hepatic, and posthepatic jaundice.
- Discuss liver disorders and appropriate laboratory tests needed for diagnosis.
- Evaluate laboratory data and correlate with normal or pathological liver disorders.
- Compare and contrast how total and direct bilirubin measurements are performed.
- Describe the enzymes commonly used to assess hepatocellular and hepatobiliary disorders.
- Identify the various types of hepatitis to include cause, transmission, occurrence, alternate name, physiology, diagnosis, and treatment.
- Apply knowledge of liver function to answer case studies throughout the chapter.

CASE STUDY 19.1, PART 1

Meet Felicia, a 58-year-old female currently being evaluated for a liver transplant. She has severe jaundice, and her skin appears yellowish in color. She presents with right upper quadrant abdominal pain, fever, and chills.



© Victorhugosilvafotografo/Shutterstock.

CASE STUDY 19.2, PART 1

Meet Verlita, a 55-year-old female, who seeks medical treatment for jaundice and with the chief complaints of a recent, unexplained 10-lb. weight loss, nausea, and vomiting.

© stockimagesbank/Shutterstock.

CASE STUDY 19.3, PART 1

DeShawn, a 19-year-old male college student, visited the Student Health Service because of fatigue, bloodshot eyes, and lack of appetite. Upon physical examination, the sclera of his eyes appeared yellowish.

© Monkey Business Images/Shutterstock.

CASE STUDY 19.4, PART 1

José, a 36-year-old male, had laboratory testing completed as a part of a preinsurance examination. The nurse practitioner received the results and advised José to see his primary care physician.

© justoomm/Shutterstock.

The liver is a very large and complex organ responsible for performing vital tasks that impact all body systems. It has a myriad of functions including the metabolism of carbohydrates, lipids, proteins, and bilirubin; detoxification of harmful substances; storage of essential compounds; and clearing waste products into the bile or blood for excretion. The liver is unique in the sense that it is a relatively resilient organ that can regenerate cells that have been destroyed by some short-term injury or disease or have been removed. However, if the liver is damaged repeatedly over a long period of time, it may undergo irreversible changes that permanently interfere with its essential functions. If the liver becomes completely nonfunctional for any reason, death will occur within approximately 24 hours due to hypoglycemia. This chapter focuses on the normal structure and function of the liver, the pathology associated with it, and the laboratory tests used to aid in the diagnosis of liver disorders.

Anatomy

Gross Anatomy

Understanding the function and dysfunction of the liver depends on understanding its gross and microscopic structure. The liver is a large and complex organ weighing approximately 1.2 to 1.5 kg in the healthy adult. It is located beneath and attached to the diaphragm, protected by the lower rib cage, and held in place by ligamentous attachments. Despite the functional complexity of the liver, it is relatively simple in structure. It is divided unequally into two lobes by the falciform ligament. The right lobe is approximately six times larger than the left lobe. The lobes are functionally insignificant; however, communication flows freely between all areas of the liver (**Figure 19.1**).

Unlike most organs, which have a single blood supply, the liver is an extremely vascular organ that receives its blood supply from two major sources: the hepatic artery and the portal vein. The hepatic artery, a branch of the aorta, supplies oxygen-rich blood from the heart to the liver and is responsible for providing approximately 25% of the total blood supply to the liver. The portal vein supplies nutrient-rich blood, collected from the digestive tract as food is digested; this vein is responsible for providing approximately 75% of the total blood supply to the liver. The two blood supplies eventually merge into the hepatic sinusoid, which is lined with hepatocytes capable of removing potentially toxic substances from the blood. From the sinusoid, blood flows to the central canal (central vein) of each lobule. It is through the central canal

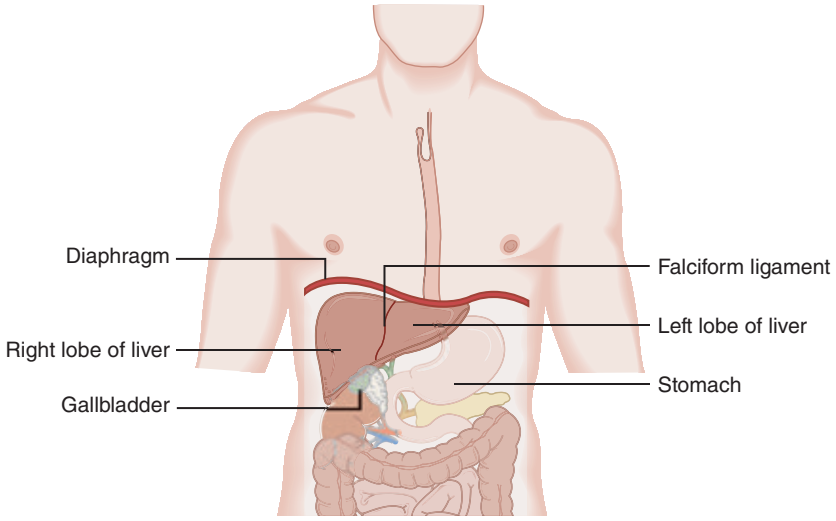


Figure 19.1 Gross anatomy of the liver.

that blood leaves the liver. Approximately 1500 mL of blood passes through the liver per minute. The liver is drained by a collecting system of veins that empties into the hepatic veins and ultimately into the inferior vena cava (**Figure 19.2**).

The excretory system of the liver begins at the bile canaliculi. The bile canaliculi are small spaces between the hepatocytes that form intrahepatic ducts, where excretory products of the cells can drain. The intrahepatic ducts join to form the right and left hepatic ducts, which drain the secretions from the liver. The right and left hepatic ducts merge to form the common hepatic duct, which is eventually joined with the cystic duct of the gallbladder to form the common bile duct. Combined digestive secretions are then expelled into the duodenum (**Figure 19.3**).

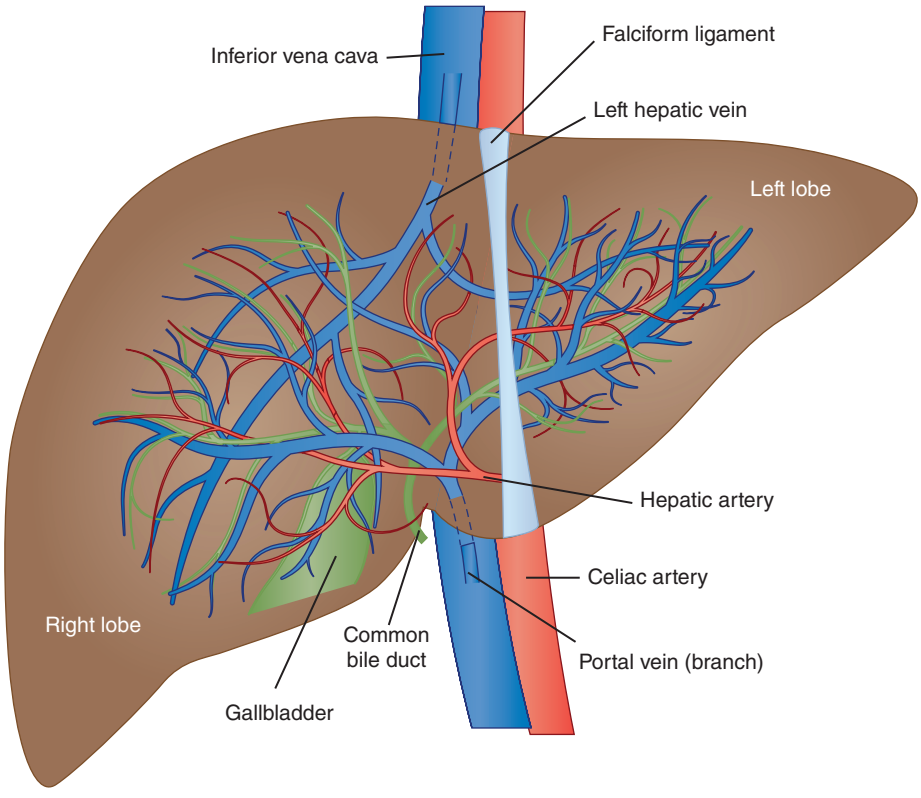


Figure 19.2 Blood supply to the liver.

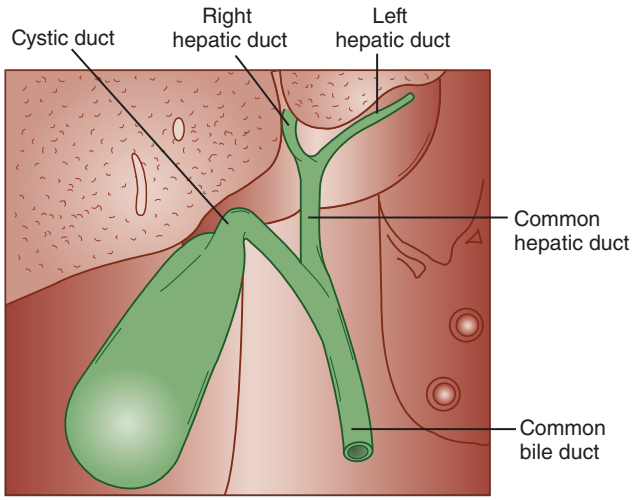


Figure 19.3 Excretory system of the liver.

Microscopic Anatomy

The liver is divided into microscopic units called lobules. The lobules are the functional units responsible for all metabolic and excretory functions performed by the liver. Each lobule is roughly a six-sided structure with a centrally located vein (called the *central vein*) with portal triads at each of the corners. Each portal triad contains a hepatic artery, a portal vein, and a bile duct surrounded by connective tissue. The liver contains two major cell types: hepatocytes and **Kupffer cells**. The hepatocytes, making up approximately 80% of the volume of the organ, are large cells that radiate outward in plates from the central vein to the periphery of the lobule. These cells perform the major functions associated with the liver and are responsible for the liver's regenerative properties. Kupffer cells are macrophages that line the **sinusoids** of the liver and act as active phagocytes capable of engulfing bacteria, debris, toxins, and other substances flowing through the sinusoids (**Figure 19.4**).

Biochemical Functions

The liver performs four major functions: excretion/secretion, metabolism, detoxification, and storage. The liver is so important that if the liver becomes nonfunctional, death will occur within 24 hours due to hypoglycemia. Although the liver is responsible for many other functions, this chapter focuses on these four major functions.

Excretory and Secretory Functions

The first major function of the liver is excretory and secretory processes. One of the most important functions is the processing and excretion of endogenous and exogenous substances into the **bile** or urine such

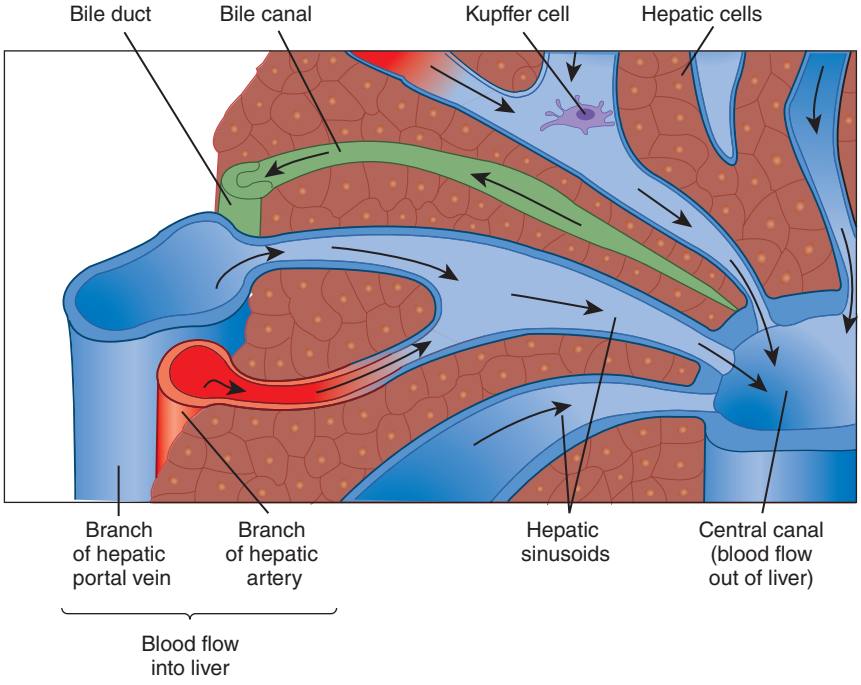


Figure 19.4 Microscopic anatomy of the liver.

as the major heme waste product **bilirubin**. The liver is the only organ that has the capacity to rid the body of heme waste products. Bile is made up of bile acids or salts, bile pigments, cholesterol, and other substances extracted from the blood. The body produces approximately 3 L of bile per day and excretes 1 L of what is produced. Bilirubin is the principal pigment in bile, and it is derived from the breakdown of red blood cells. The life span of a red blood cell is approximately 120 days, after which red blood cells are phagocytized and hemoglobin is released. Hemoglobin is degraded to heme, globin, and iron. The iron is bound by transferrin and is returned to iron stores in the liver or bone marrow for reuse. The globin is degraded to its constituent amino acids, which are reused by the body. The heme portion of hemoglobin is converted to bilirubin in 2 to 3 hours. Bilirubin is bound by albumin and transported to the liver. This form of bilirubin is referred to as *unconjugated* or *indirect bilirubin*. **Unconjugated bilirubin** is insoluble in water and cannot be removed from the body until it has been conjugated by the liver. Once at the liver cell, unconjugated bilirubin flows into the sinusoidal spaces and is released from albumin and then picked up by a carrier protein called *ligandin*. Ligandin, which is located inside the hepatocyte, is responsible for transporting unconjugated bilirubin to the endoplasmic reticulum, where it is conjugated. The conjugation (esterification) of bilirubin occurs in the presence of the enzyme **uridine diphosphate glucuronosyltransferase (UDPGT)**, which transfers a glucuronic acid molecule to each of the two propionic acid side chains of bilirubin to form bilirubin diglucuronide, also known as **conjugated bilirubin**. Conjugated bilirubin is water soluble and can be secreted from the hepatocyte into the bile canaliculi. Once in the hepatic duct, it combines with secretions from the gallbladder through the cystic duct and is expelled through the common bile duct to the intestines. Intestinal bacteria (especially the bacteria in the lower portion of the intestinal tract) work on conjugated bilirubin to produce mesobilirubin, which is reduced to form mesobilirubinogen and then **urobilinogen** (a colorless product). Most of the urobilinogen formed (roughly 80%) is oxidized to an orange-colored product called *urobilin* (stercobilin) and is excreted in the feces. The urobilin or stercobilin is what gives stool its brown color. There are two things that can happen to the remaining 20% of urobilinogen formed. The majority will be absorbed by extrahepatic circulation to be recycled through the liver and re-excreted. The other very small quantity left will enter systemic circulation and will subsequently be filtered by the kidney and excreted in the urine (**Figure 19.5**).[1]

Approximately 200 to 300 mg of bilirubin is produced per day, and it takes a normally functioning liver to process the bilirubin and eliminate it from the body. This, as stated earlier, requires that bilirubin be conjugated. Almost all the bilirubin formed is eliminated in the feces, and a small amount of the colorless product, urobilinogen, is excreted in the urine. The healthy adult has very low levels of total bilirubin (0.2 to 1.0 mg/dL) in the serum, and of this amount, the majority is in the unconjugated form.[2]

Metabolism

The second major function of the liver is metabolic processes. The liver has extensive capacity for metabolizing many biological compounds including carbohydrates, lipids, and proteins.

The metabolism of carbohydrates is one of the most important functions of the liver. When carbohydrates are ingested and absorbed, the liver can process them in one of three ways: (1) use the glucose for its own cellular energy requirements, (2) circulate the glucose for use at the peripheral tissues, or (3) store glucose as glycogen (principal storage form of glucose) within the liver itself or within other tissues. The liver is the major player in maintaining stable glucose concentrations due to its ability to store glucose as glycogen (glycogenesis) and degrade glycogen (glycogenolysis), depending on the body's needs. Under conditions of stress or in a fasting state when there is an increased requirement for glucose, the liver will break down stored glycogen (glycogenolysis), and when the supply of glycogen becomes depleted, the liver will create glucose from non-sugar carbon substrates like pyruvate, lactate, and amino acids (*gluconeogenesis*).

Lipids are metabolized in the liver under normal circumstances when nutrition is adequate and the demand for glucose is being met. The liver is responsible for metabolizing both lipids and the lipoproteins and is responsible for gathering free fatty acids from the diet, and those produced by the liver itself, and breaking them down to produce acetyl-CoA. Acetyl-CoA can then enter several pathways to form triglycerides, phospholipids, or cholesterol. Despite popular belief, the greatest source of cholesterol in the body comes from the cholesterol produced by the liver, not from dietary sources. In fact, approximately 70% of the daily production of cholesterol (roughly 1.5 to 2.0 g) is produced

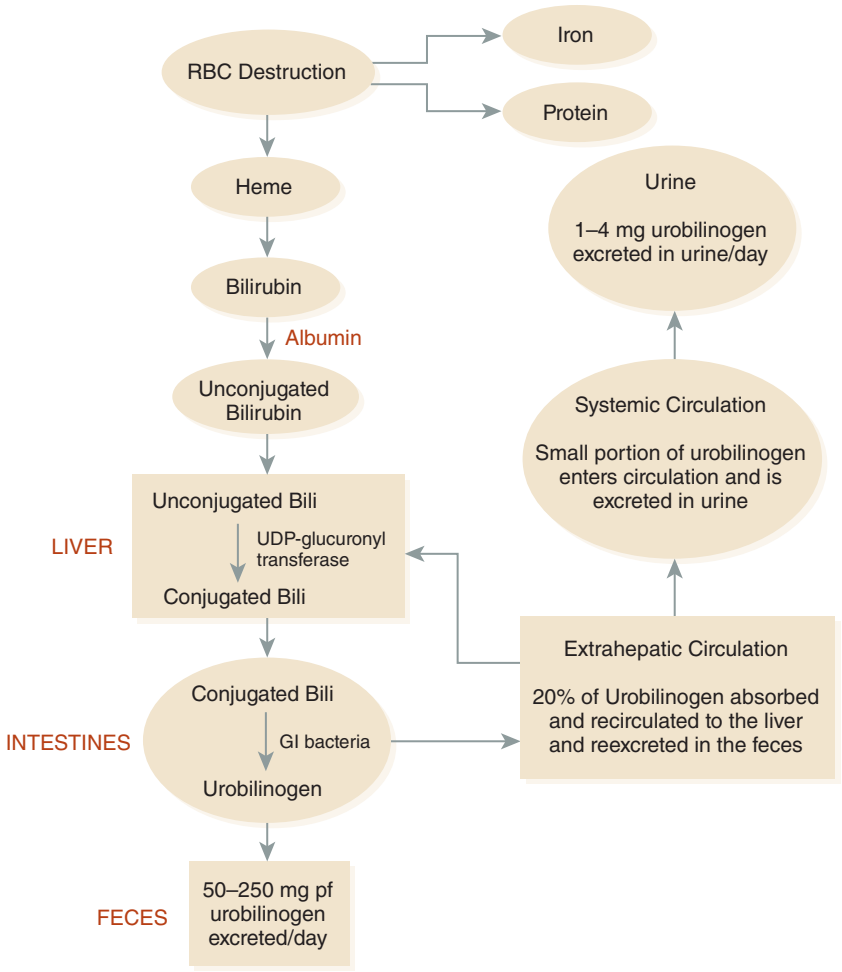


Figure 19.5 Metabolism of bilirubin.

by the liver.[3] A more thorough discussion of lipid metabolism may be found in Chapter 10, *Lipids and Lipoproteins*.

Almost all proteins are synthesized by the liver except for the immunoglobulins and adult hemoglobin. The liver plays an essential role in the development of hemoglobin in infants. One of the most important proteins synthesized by the liver is albumin, which carries with it a wide range of important functions. The liver is also responsible for synthesizing the positive and negative acute-phase reactants and coagulation proteins, and it also serves to store a pool of amino acids through protein degradation. The most critical aspect of protein metabolism includes transamination and deamination of amino acids. Transamination (via a transaminase) results in the exchange of an amino group on one acid with a ketone group on another acid. After transamination, deamination degrades them to produce ammonium ions that are consumed in the synthesis of urea, which is excreted by the kidneys.

Although it would seem logical that any damage to the liver would result in a loss of synthetic and metabolic functions of the liver, that is not the case. The liver must be extensively impaired before it loses its ability to perform these essential functions.

Detoxification and Drug Metabolism

The third major function of the liver is detoxification of waste products from physiological processes. The liver serves as a gatekeeper between substances absorbed by the gastrointestinal tract and those released into systemic circulation. Every substance that is absorbed in the gastrointestinal tract must pass through the liver before reaching the rest of the body; this is referred to as *first pass*. This is an important function of the liver because it can allow important substances to reach the systemic circulation yet can serve as a barrier to prevent toxic or harmful substances from reaching systemic circulation. The body

has two mechanisms for detoxification of foreign materials (drugs and poisons) and metabolic products (bilirubin and ammonia). It either binds the material reversibly to inactivate the compound, or it chemically modifies the compound so it can be excreted in its chemically modified form. The fourth major function of the liver is drug detoxification. This system is responsible for the detoxification of many drugs through oxidation, reduction, hydrolysis, hydroxylation, carboxylation, and demethylation. Many of these take place in the liver microsomes via the cytochrome P-450 (CYP450) isoenzymes.

Liver Function Alterations During Disease

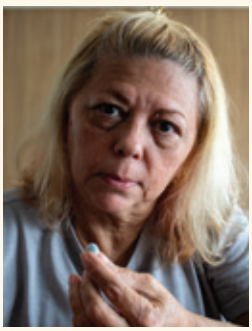
Jaundice

The word **jaundice** comes from the French word *jaune*, which means "yellow," and it is one of the oldest known pathologic conditions reported, having been described by Hippocratic physicians.[4] Jaundice is used to describe the yellow discoloration of the skin, eyes, and mucous membranes, most often resulting from the retention of bilirubin. Although the upper limit of normal for total bilirubin is 1.0 to 1.5 mg/dL, jaundice is usually not noticeable to the human eye (known as *overt jaundice*) until bilirubin levels reaches 3.0 to 5.0 mg/dL. The term **icterus** is used in the clinical laboratory to refer to a serum or plasma sample with a yellow discoloration due to an elevated bilirubin level. Jaundice is classified based on the site of the disorder: **prehepatic jaundice**, **hepatic jaundice**, and **posthepatic jaundice**. This classification is important because knowing the classification of jaundice will aid health care providers in formulating an appropriate treatment or management plan. *Prehepatic* and *posthepatic jaundice*, as the names imply, are caused by abnormalities outside the liver, either before, as in "prehepatic," or after, as in "posthepatic." In these conditions, liver function is normal, or it may be functioning at a maximum to compensate for abnormalities occurring elsewhere. This is not the case with hepatic jaundice, where the jaundice is due to a problem with the liver itself and may be caused by an intrinsic liver defect or disease.

Prehepatic jaundice occurs when the problem causing the jaundice occurs prior to liver metabolism. It is commonly caused by an increased amount of bilirubin being presented to the liver, as is seen in acute and chronic hemolytic anemias. Hemolytic anemia causes an increased amount of red blood cell destruction and the subsequent release of increased amounts of bilirubin presented to the liver for processing. The liver responds by functioning at maximum capacity. Individuals with prehepatic jaundice rarely have bilirubin levels that exceed 5.0 mg/dL because the liver is working efficiently to handle the increased workload. This type of jaundice may also be referred to as

CASE STUDY 19.1, PART 2

Remember Felicia, the 58-year-old female currently being evaluated for a liver transplant. Her liver function test results were sent to her physician. Assist her physician in evaluating these results.



© Victorhugosilvafotografo/Shutterstock.

Analyte	Patient Value	Reference Range
Alkaline phosphatase	420 U/L	42–98 U/L
Alanine aminotransferase	50 U/L	7–45 U/L
Aspartate aminotransferase	50 U/L	5–35 U/L
Total bilirubin	25.0 mg/dL	0.2–1.2 mg/dL
Direct bilirubin	19.9 mg/dL	0.0–0.3 mg/dL
Total protein	7.0 g/dL	6.5–8.3 g/dL
Albumin	4.4 g/dL	3.5–5.0 g/dL

1. What is the calculated indirect bilirubin result for Felicia?
2. What type of jaundice (prehepatic, hepatic, or posthepatic) does Felicia have?

unconjugated hyperbilirubinemia because the fraction of bilirubin increased in individuals with prehepatic jaundice is the unconjugated fraction. Unconjugated bilirubin is bound to albumin in the bloodstream and is not water soluble. This protein-bound bilirubin does not pass through the glomerular basement membrane and therefore is not detected in urine.

Hepatic jaundice occurs when the primary problem causing the jaundice resides in the liver (intrinsic liver defect or disease). This intrinsic liver defect or disease can be due to disorders of bilirubin metabolism and transport defects (Crigler-Najjar syndrome, Dubin-Johnson syndrome, Gilbert's disease, and neonatal physiologic jaundice of the newborn) or due to diseases resulting in hepatocellular injury or destruction. Gilbert's disease, Crigler-Najjar syndrome, and physiologic jaundice of the newborn are hepatic causes of jaundice that result in elevations in unconjugated bilirubin. Conditions such as Dubin-Johnson and Rotor syndrome are hepatic causes of jaundice that result in elevations in conjugated bilirubin.

Gilbert's syndrome, first described in the early twentieth century, is a benign autosomal recessive hereditary disorder that affects approximately 5% of the U.S. population.[5] Gilbert's syndrome results from a genetic mutation in the *UGT1A1* gene that produces the enzyme uridine diphosphate glucuronosyltransferase, one of the enzymes important for bilirubin metabolism. The *UGT1A1* gene is located on chromosome 2, and other mutations of this same gene produce Crigler-Najjar syndrome, a more severe and dangerous form of hyperbilirubinemia.[6]

Of the many causes of jaundice, Gilbert's syndrome is the most common cause, and interestingly, it carries no morbidity or mortality in the majority of those affected and carries generally no clinical consequences. It is characterized by intermittent unconjugated hyperbilirubinemia, underlying liver disease due to a defective conjugation system in the absence of hemolysis. The hyperbilirubinemia usually manifests during adolescence or early adulthood. Total serum bilirubin usually fluctuates between 1.5 and 3.0 mg/dL and it rarely exceeds 4.5 mg/dL. The molecular basis of Gilbert's syndrome is related to the UDPGT superfamily, which is responsible for encoding enzymes that catalyze the conjugation of bilirubin. The UGT1A1 (the hepatic 1A1 isoform of UDPGT) contributes substantially to the process of conjugating bilirubin. The UGT1A1 promoter contains the sequence (TA)6TAA. The insertion of an extra TA in the sequence, as seen in Gilbert's syndrome, reduces the expression of the *UGT1A1* gene to 20% to 30% of normal values. Thus, in Gilbert's syndrome, the liver's conjugation system is working at approximately 30% of normal.[7,8]

Crigler-Najjar syndrome was first described by Crigler and Najjar in 1952 as a syndrome of chronic nonhemolytic unconjugated hyperbilirubinemia.[9] Crigler-Najjar syndrome, like Gilbert's syndrome, is an inherited disorder of bilirubin metabolism resulting from a molecular defect within the gene involved with bilirubin conjugation. Crigler-Najjar syndrome may be divided into two types: type 1, where there is a complete absence of enzymatic bilirubin conjugation, and type 2, where there is a mutation causing a severe deficiency of the enzyme responsible for bilirubin conjugation. Unlike Gilbert's syndrome, Crigler-Najjar syndrome is rare, and it is a more serious disorder that may result in death.[10]

While Gilbert's disease and Crigler-Najjar syndrome are characterized as primarily unconjugated hyperbilirubinemias, Dubin-Johnson syndrome and Rotor syndrome are characterized as conjugated hyperbilirubinemias. Dubin-Johnson syndrome is a rare, autosomal recessive–inherited disorder caused by a deficiency of the canalicular multidrug resistance/multispecific organic anionic transporter protein (MDR2/cMOAT). In other words, the liver's ability to uptake and conjugate bilirubin is functional; however, the removal of conjugated bilirubin from the liver cell and the excretion into the bile are defective. This results in accumulation of conjugated and, to some extent, unconjugated bilirubin in the blood, leading to hyperbilirubinemia and bilirubinuria. Dubin-Johnson is a condition that is obstructive in nature, so much of the conjugated bilirubin circulates bound to albumin. This type of bilirubin (conjugated bilirubin bound to albumin) is referred to as **delta bilirubin**. An increase in delta bilirubin poses a problem in laboratory evaluation because the delta bilirubin fraction reacts as conjugated bilirubin in the laboratory method to measure conjugated or direct bilirubin. A distinguishing feature of Dubin-Johnson syndrome is the appearance of dark stained granules (thought to be pigmented lysosomes) on a liver biopsy sample. Usually, the total bilirubin concentration remains between 2 and 5 mg/dL, with more than 50% due to the conjugated fraction. This syndrome is relatively mild in nature, with an excellent prognosis. People with Dubin-Johnson syndrome have a normal life expectancy, so no treatment is necessary.[11,12]

Rotor syndrome (also known as Rotor type hyperbilirubinemia) was first described in 1948 and is an autosomal recessive disease caused by mutations in both the *SLCO1B1* and *SLCO1B3* genes on

chromosome 12. These genes provide instructions for making proteins found on liver cells that mediate the cellular uptake of compounds such as bilirubin, bile acids, and steroids. In Rotor syndrome, the proteins are abnormally short; therefore, bilirubin is less efficiently taken up by the liver and removed. It is a rare yet benign condition characterized by chronic jaundice. Therefore, it does not require treatment.

Although this is a benign condition, accurate diagnosis is required to aid in distinguishing it from more serious liver diseases that require treatment. Normal levels of alkaline phosphatase (ALP) and gamma-glutamyl transferase (GGT) seen in Rotor syndrome distinguish it from biliary obstruction, and abnormal levels of urinary coproporphyrin excretion with normal liver histology distinguish it from Dubin-Johnson syndrome.[13]

Physiologic jaundice, also referred to as neonatal hyperbilirubinemia, is a common clinical complication encountered during the neonatal period, especially during with the first week of life. Physiologic jaundice is caused by a deficiency in UDPGT, the enzyme responsible for bilirubin conjugation. It is one of the last liver functions to be activated in prenatal life, since the mother processes bilirubin until birth. Therefore, premature infants may be born with a UDPGT deficiency, resulting in the rapid buildup of unconjugated bilirubin, which can be life threatening. When unconjugated bilirubin builds up in the neonate, it is deposited in the brain, causing a form of brain injury called *kernicterus*. Kernicterus may result in permanent brain damage and even death when the neonatal bilirubin levels are critically high. This is the reason why all neonatal critical bilirubin results must be called to the healthcare provider so that treatment can be initiated.

Treatment options for physiologic jaundice include phototherapy (conventional, intensive), exchange transfusion, and pharmacological treatment with phenobarbitone, intravenous immunoglobulins (IVIG), and metalloporphyrins. Phototherapy has been effectively used as a relatively inexpensive and noninvasive method of treating neonatal hyperbilirubinemia through photo oxidation. Conventional phototherapy lowers serum bilirubin levels by using halogen or fluorescent lights to transform bilirubin into water-soluble isomers that can be eliminated without conjugation in the liver. During phototherapy, the baby's skin is exposed to the light, while the baby's eyes are covered to protect the nerve layer at the retina from the bright light. The bilirubin levels are monitored at least daily. Alternatively, intensive, fiberoptic phototherapy is also available via a blanket, called a bili-blanket, that consists of a pad of woven fibers used to transport light from a light source to the baby's back. The light generated through the bili-blanket breaks down the bilirubin through photo oxidation. The dose of phototherapy is a key factor in how quickly it works; the dose is determined by the wavelength of the light, the intensity of the light (irradiance), the distance between the light and the baby, and the body surface area exposed to the light. Commercially available phototherapy systems include those that deliver light via fluorescent bulbs, halogen quartz lamps, light-emitting diodes, and fiberoptic mattresses. In extreme cases of hyperbilirubinemia, an exchange transfusion may be used as the second-line treatment when phototherapy fails. An exchange transfusion involves removing aliquots of neonatal blood and replacing it with donor blood to remove the hyperbilirubinemia from circulation while maintaining adequate blood volume. Pharmacological treatment of neonatal jaundice aims to decrease plasma concentration of unconjugated bilirubin by inhibiting production, stimulating hepatic clearance, or interrupting the enterohepatic circulation of the pigment.[14]

Because hyperbilirubinemia is so serious in newborns, bilirubin levels are carefully and frequently monitored so the dangerously high levels of unconjugated bilirubin (approximately 20 mg/dL) can be detected and treated appropriately.[15]

Posthepatic jaundice results from biliary obstructive disease, usually from physical obstructions (gallstones or tumors) that prevent the flow of conjugated bilirubin into the bile canaliculi. Since the liver cell itself is functioning, bilirubin is effectively conjugated; however, it is unable to be properly excreted from the liver. Since bile is not being brought to the intestines, stool loses its source of normal pigmentation and becomes clay-colored. The laboratory findings for bilirubin and its metabolites in the abovementioned types of jaundice are summarized in **Table 19.1**. Mechanisms of hyperbilirubinemia may be found in **Figure 19.6**.

Cirrhosis

Cirrhosis is a clinical condition in which scar tissue replaces normal, healthy liver tissue. As the scar tissue replaces the normal liver tissue, it blocks the flow of blood through the organ and prevents the liver from functioning properly. Cirrhosis rarely causes signs and symptoms in its early stages, but as liver function deteriorates, signs and symptoms appear including fatigue, nausea, unintended weight loss, jaundice,

Table 19.1 Changes in Concentration of Bilirubin in Those with Jaundice

Type of Jaundice	Total Bilirubin	Conjugated Bilirubin	Unconjugated Bilirubin
Prehepatic	↑	↔	↑
Hepatic			
■ Gilbert's disease	↑	↔	↑
■ Crigler-Najjar syndrome	↑	↓	↑
■ Dubin-Johnson syndrome	↑	↑	↔
■ Rotor's syndrome	↑	↑	↔
■ Jaundice of newborn	↑	↔	↑
Posthepatic	↑	↑	↑

↑ Increased; ↓ Decreased; ↔ Increased, Normal, or Decreased

Adapted from Kaplan LA, Pesce AJ, Kazmierczak S. *Clinical Chemistry: Theory, Analysis, Correlation.* 4th ed. St. Louis, MO: Mosby; 2003:449.

bleeding from the gastrointestinal tract, intense itching, and swelling in the legs and abdomen. Although some patients with cirrhosis may have prolonged survival, they generally have a poor prognosis. Cirrhosis was the eleventh leading cause of death by disease in 2018, killing just over 42,800 people.[15]

In the United States, the most common cause of cirrhosis is chronic alcoholism. Other causes of cirrhosis include chronic infection with the hepatitis B (HBV), C (HCV), and D (HDV) viruses; autoimmune hepatitis; inherited disorders (e.g., α1-antitrypsin deficiency, Wilson's disease, hemochromatosis, and galactosemia); nonalcoholic steatohepatitis; blocked bile ducts; drugs; toxins; and infections.

Liver damage from cirrhosis cannot easily be reversed, but treatment can stop or delay further progression of the disorder. Treatment depends on the cause of cirrhosis and any complications a person is experiencing. For example, cirrhosis caused by alcohol abuse is treated by abstaining from alcohol. Treatment for hepatitis-related cirrhosis involves medications used to treat the different types of hepatitis, such as interferon for viral hepatitis and corticosteroids for autoimmune hepatitis.

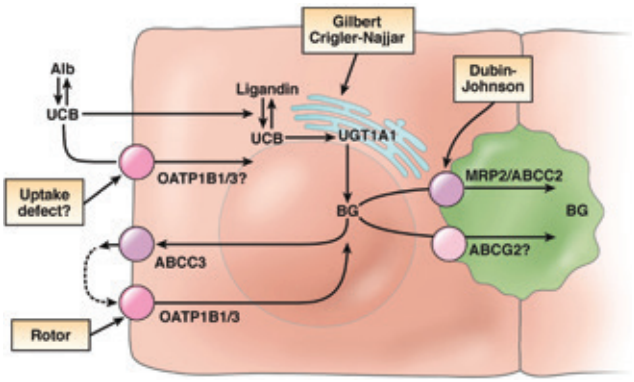


Figure 19.6 Mechanisms of hyperbilirubinemia: Gilbert's syndrome, Crigler-Najjar syndrome, Dubin-Johnson syndrome, Rotor's syndrome.

Reproduced from Erlinger S, Arias IM, Dhumeaux D. Inherited disorders of bilirubin transport and conjugation: new insights into molecular mechanisms and consequences. *Gastroenterology.* 2014 Jun;146(7):1625-38. doi: 10.1053/j.gastro.2014.03.047. Epub 2014 Apr 1.

Tumors

Cancers of the liver are classified as primary or metastatic. Primary liver cancer is cancer that begins in the liver cells, while metastatic cancer occurs when tumors from other parts of the body spread (metastasize) to the liver. Metastatic liver cancer is much more common than primary liver cancer; 90% to 95% of all hepatic malignancies are classified as metastatic. Cancers that commonly spread to the liver include colon, lung, and breast cancer. Tumors of the liver may also be classified as benign or malignant. The common benign tumors of the liver include hepatocellular adenoma (a condition occurring almost exclusively in females of childbearing age) and hemangiomas (masses of blood vessels with no known etiology). Malignant tumors of the liver include **hepatocellular carcinoma** (HCC) (also known as hepatocarcinoma, and **hepatoma**) and bile duct carcinoma. Hepatocellular carcinoma accounts for 90% of all primary liver cancers and it is the fourth leading cause of cancer-related deaths worldwide. Hepatoblastoma is an uncommon hepatic malignancy of children.

Hepatocellular carcinoma is the most common primary liver malignancy, is the leading cause of death worldwide, and it is the ninth leading cause of cancer deaths in the United States. Despite advances in

diagnosis, screening, prevention, and treatment, the incidence of HCC and the mortality associated with it continue to rise. Hepatocellular carcinoma develops with underlying cirrhosis and chronic liver inflammation, with the major risk factors for development being HBV/HCV infection, heavy alcohol consumption, aflatoxin B1 ingestion (mycotoxin), tobacco smoking, and nonalcoholic fatty liver disease.[16]

Chronic liver disease and cirrhosis remain the most important risk factors for the development of HCC; it occurs more often in males at a ratio of 2:1 to 4:1, with a higher incidence in regions where HBV is endemic, such as Southeast Asia and sub-Saharan Africa. Approximately 630,000 people are diagnosed annually with HCC worldwide, of which 42,230 new cases and 30,320 deaths were reported in the United States in 2021.[17] It is estimated that in the United States, chronic HBV and HCV infections account for approximately 30% and 40% of cases of HCC.[18]

Approximately 80% of cases worldwide are attributable to HBV and HCV; however, the mechanism by which the infection leads to HCC is not well understood. While surgical resection of HCC is sometimes possible, in people diagnosed with HCC with underlying cirrhosis who meet the Milan criteria (single tumor ≤5 cm in size or ≤3 tumors each ≤3 cm in size, and no macrovascular invasion), orthotopic liver transplantation is also currently available for HCC. The estimated 4-year survival rate is 85%, and the recurrence-free survival rate is 92%.[19]

Whether primary or metastatic, any malignant tumor in the liver is a serious finding and carries a poor prognosis, with survival times measured in months.

Reye's Syndrome

Reye's syndrome is a term used to describe a group of disorders caused by infectious, metabolic, toxic, or drug-induced disease found almost exclusively in children, although adult cases of Reye's syndrome have been reported.[20] Although the precise cause of Reye's syndrome is unknown, it is often preceded by a viral syndrome such as varicella, gastroenteritis, or an upper respiratory tract infection such as influenza.[21–23] Although the ingestion of aspirin during a viral infection has not been described as the precise cause of Reye's syndrome, studies have demonstrated a strong epidemiologic association between the ingestion of aspirin during a viral infection and the subsequent development of Reye's syndrome.[24,25] As a result, the Centers for Disease Control and Prevention (CDC) cautioned physicians and parents to avoid salicylate use in children with a viral syndrome, and the U.S. Surgeon General mandated that a warning label be added to all aspirin-containing medications beginning in 1986.[26,27] Reye's syndrome is an acute illness characterized by noninflammatory encephalopathy and fatty degeneration of the liver, with a clinical presentation of profuse vomiting accompanied with varying degrees of neurologic impairment such as fluctuating personality changes and deterioration in consciousness. The encephalopathy is characterized by a progression from mild confusion (stage 1) through progressive loss of neurologic function to loss of brain stem reflexes (stage 5). The degeneration of the liver is characterized by a mild hyperbilirubinemia and threefold increases in ammonia and the aminotransferases (aspartate aminotransferase [AST] and alanine aminotransferase [ALT]). Without treatment, rapid clinical deterioration leading to death may occur.[28,29]

Drug- and Alcohol-Related Disorders

Drug-induced liver disease is a major problem in the United States, accounting for one-third to one-half of all reported cases of acute liver failure. The liver is a primary target organ for adverse drug reactions because it plays a central role in drug metabolism. Many drugs are known to cause liver damage, ranging from very mild transient forms to fulminant liver failure. Drugs can cause liver injury by a variety of mechanisms, but the most common mechanism of toxicity is via an immune-mediated injury to the hepatocytes.[30] In this type of mechanism, the drug induces an adverse immune response directed against the liver itself and results in hepatic and/or cholestatic disease.[31]

Of all the drugs associated with hepatic toxicity, the most important is ethanol. In very small amounts, ethanol causes very mild, transient, and unnoticed injury to the liver; however, with heavier and prolonged consumption, it can lead to alcoholic cirrhosis. While the exact amount of alcohol needed to cause cirrhosis is unknown, a small minority of people with alcoholism develop this condition.[32] Approximately 90% of the alcohol absorbed from the stomach and small intestines is transported to the liver for metabolism. Within the liver, the elimination of alcohol requires the enzymes alcohol dehydrogenase and acetaldehyde dehydrogenase to convert alcohol to acetaldehyde and subsequently to acetate. The acetate can then be oxidized to water and carbon dioxide, or it may enter the citric acid cycle.

Long-term, excessive consumption of alcohol can result in a spectrum of liver abnormalities that may range from alcoholic fatty liver with inflammation (steatohepatitis) to scar tissue formation, as in hepatic fibrosis, to the destruction of normal liver structure seen in hepatic cirrhosis. Alcohol-induced liver injury may be categorized into three stages: alcoholic fatty liver, alcoholic hepatitis, and alcoholic cirrhosis. The risk for the development of cirrhosis increases proportionally with the consumption of more than 30 g (the equivalent of 3 to 4 drinks) of alcohol per day, with the highest degree of risk seen with the consumption of greater than 120 g (the equivalent of 12 to 16 drinks) per day.[33]

Fatty liver disease is very common and represents the mildest category where very few changes in liver function are measurable. This benign stage is characterized by slight elevations in AST, ALT, and γ-glutamyltransferase (GGT), and on biopsy, fatty infiltrates are noted in the vacuoles of the liver. This condition tends to affect middle-aged people with a history of obesity, diabetes, or moderate alcohol consumption. Treatment is aimed at reducing the risk factors and includes diet and exercise, diabetes management, and alcohol avoidance.

Alcoholic hepatitis presents with common signs and symptoms including fever, ascites, proximal muscle loss, and laboratory evidence of liver damage such as moderately elevated AST, ALT, GGT, ALP, and elevations in total bilirubin greater than 5 mg/dL. The elevations in AST are more than twice the upper reference of normal but rarely exceed 300 IU/mL. The elevations in ALT are comparatively lower than AST, resulting in an AST/ALT ratio (De Ritis ratio) greater than 2. Serum proteins, especially albumin, are decreased, and the prothrombin time, a screening test for coagulation disorders, is elevated. Prognosis is dependent on the type and severity of damage to the liver, and when serum creatinine levels begin to increase, it is a threatening sign, which may precede the onset of hepatorenal syndrome and death.[34] A variety of scoring systems have been used to assess the severity of alcoholic hepatitis and to guide treatment, including the Maddrey's discriminant function,[35] the Glasgow score,[36] and the Model for End-Stage Liver Disease (MELD) score.[37] All three scoring systems use bilirubin, INR, creatinine, age, white cell counts, blood urea nitrogen, and albumin levels to stage and guide treatment. The last and most severe stage is alcoholic cirrhosis. The prognosis associated with alcoholic cirrhosis is dependent on the nature and severity of associated conditions such as a gastrointestinal bleeding or ascites; however, the 5-year survival rate is 60% in those who abstain from alcohol and 30% in those who continue to drink. This condition appears to be more common in males than in females, and the symptoms tend to be nonspecific and include weight loss, weakness, hepatomegaly, splenomegaly, jaundice, ascites, fever, malnutrition, and edema. Laboratory abnormalities include increased liver function tests (AST, ALT, GGT, ALP, and total bilirubin), decreased albumin, and a prolonged prothrombin time. A liver biopsy is the only method by which a definitive diagnosis may be made.[38]

Other drugs, including tranquilizers, some antibiotics, antineoplastic agents, lipid-lowering medication, and anti-inflammatory drugs, may cause liver injury ranging from mild damage to massive hepatic failure and cirrhosis. One of the most common drugs associated with serious hepatic injury is acetaminophen. When acetaminophen is taken in massive doses, it is virtually certain to produce fatal hepatic necrosis unless rapid treatment is initiated.

Assessment of Liver Function/Liver Function Tests

Bilirubin

Analysis of Bilirubin: A Brief Review

The reaction of bilirubin with a diazotized sulfanilic acid solution to form a colored product was first described by Ehrlich in 1883 using urine samples. Since then, this type of reaction (bilirubin with a diazotized sulfanilic acid solution) has been referred to as the classic *diazo reaction*, a reaction on which all commonly used methods today are based. In 1913, van den Bergh found that the diazo reaction may be applied to serum samples but only in the presence of an accelerator (solubilizer). However, this methodology had errors associated with it. In 1937, Malloy and Evelyn developed the first clinically useful methodology for the quantitation of bilirubin in serum samples using the classic diazo reaction with a 50% methanol solution as an accelerator. In 1938, Jendrassik and Grof described a method using the diazo reaction with caffeine–benzoate–acetate as an accelerator. Today, all commonly used methods for measuring bilirubin and its fractions are modifications of the diazo reaction. Total bilirubin and direct bilirubin are measured, whereas indirect bilirubin (unconjugated) is calculated by subtracting conjugated bilirubin from total bilirubin (see **Figure 19.7**).

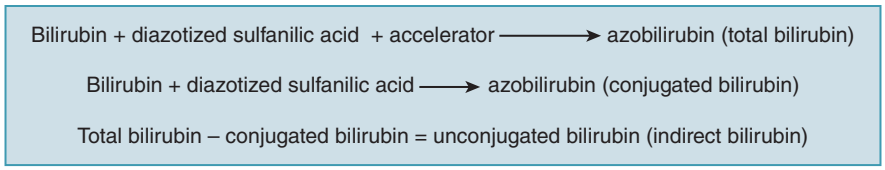


Figure 19.7 Determination of different fractions of bilirubin.

Bilirubin has also been quantified using bilirubinometry in the neonatal population. This methodology is only useful in the neonatal population because of the presence of carotenoid compounds in adult serum that cause strong positive interference in the adult population. Bilirubinometry involves the measurement of reflected light from the skin using two wavelengths that provide a numerical index based on spectral reflectance. Today's bilirubinometers use microspectrophotometers that determine the optical densities of bilirubin, hemoglobin, and melanin in the subcutaneous layers of the infant's skin. Mathematical isolation of hemoglobin and melanin allows measurement of the optical density created by bilirubin.[39]

When using the several methods described earlier, two of the three fractions of bilirubin were identified: conjugated (direct) and unconjugated (indirect) bilirubin. Unconjugated (indirect) bilirubin is a nonpolar and water-insoluble substance that is found in plasma bound to albumin. Because of these characteristics, unconjugated bilirubin will only react with the diazotized sulfanilic acid solution (diazo reagent) in the presence of an accelerator (solubilizer). Conjugated (direct) bilirubin is a polar and water-soluble compound that is found in plasma in the free state (not bound to any protein). This type of bilirubin will react with the diazotized sulfanilic acid solution directly (without an accelerator). Thus, conjugated and unconjugated bilirubin fractions have historically been differentiated by solubility of the fractions. Conjugated bilirubin reacts in the absence of an accelerator, whereas unconjugated bilirubin requires an accelerator. Bilirubin terminology may be challenging for some learners. A helpful tip may be to remember English grammar: vowels (indirect, unconjugated, insoluble) and consonants (direct, conjugated, water-soluble). While some studies reported direct and indirect terminology is outdated,[40] the truth of the matter is that this terminology is still commonly used in the clinical laboratory today.

CASE STUDY 19.2, PART 2

Remember Verlita, who sought medical treatment for jaundice, weight loss, nausea, and vomiting. Her physical examination revealed an enlarged liver. Her liver function results are in the following table.

	Laboratory Results	Reference Range
Total bilirubin	2.2 mg/dL	0.2–1.2 mg/dL
Direct bilirubin	1.2 mg/dL	0.0–0.3 mg/dL
ALP	120 U/L	42–98 U/L
AST	1050 U/L	5–35 U/L
ALT	1063 U/L	7–45 U/L
Albumin	3.5 g/dL	3.5–5.0 g/dL
Total Protein	8.0 g/dL	6.5–8.3 g/dL

1. What type of jaundice does Verlita have?
2. What additional testing is recommended?

The third fraction of bilirubin is delta bilirubin. Delta bilirubin is conjugated bilirubin that is covalently bound to albumin. This fraction of bilirubin is seen only when there is significant hepatic obstruction. Because the molecule is attached to albumin, it is too large to be filtered by the glomerulus and excreted in the urine. This fraction of bilirubin, when present, will react in most laboratory methods as conjugated bilirubin. Thus, total bilirubin is made up of three fractions: conjugated, unconjugated, and delta bilirubin. The three fractions together are known as *total bilirubin*.

Specimen Collection and Handling

Total bilirubin methods using a diazotized sulfanilic acid solution may be performed on either serum or plasma. Serum, however, is preferred for the Malloy-Evelyn procedure because the addition of the alcohol in the analysis can precipitate proteins and cause interference with the method. A fasting sample is preferred as the presence of lipemia will increase measured bilirubin concentrations. Hemolyzed samples should be avoided as they may decrease the reaction of bilirubin with the diazo reagent. Bilirubin is very sensitive to and is destroyed by light; therefore, specimens should be protected from light. If left unprotected from light, bilirubin values may reduce by 30% to 50% per hour. If serum or plasma is separated from the cells and stored in the dark, it is stable for 2 days at room temperature, 1 week at 4°C, and indefinitely at –20°C.[41]

Methods

There is no preferred reference method or standardization of bilirubin analysis; however, the American Association for Clinical Chemistry and the National Bureau of Standards have published a candidate reference method for total bilirubin, a modified Jendrassik-Grof procedure using caffeine–benzoate as a solubilizer.[42] Because they both have acceptable precision and are adapted to many automated instruments, the Jendrassik-Grof or Malloy-Evelyn procedure is the most frequently used method to measure bilirubin. The Jendrassik Grof method is slightly more complex, but it has the following advantages over the Malloy-Evelyn method:

- Not affected by pH changes
- Insensitive to a 50-fold variation in protein concentration of the sample
- Maintains optical sensitivity even at low bilirubin concentrations
- Has minimal turbidity and a relatively constant serum blank
- Is not affected by hemoglobin up to 750 mg/dL

Because this chapter does not allow for a detailed description of all previously mentioned bilirubin test methodologies, only the most widely used principles for measuring bilirubin in the adult and pediatric populations are covered.[43–45]

Malloy-Evelyn Procedure

Bilirubin pigments in serum or plasma are reacted with a diazo reagent. The diazotized sulfanilic acid reacts at the central methylene carbon of bilirubin to split the molecule, forming two molecules of azobilirubin. This method is typically performed at pH 1.2 where the azobilirubin produced is red-purple in color with a maximal absorption of 560 nm. The most commonly used accelerator to solubilize unconjugated bilirubin was methanol, although other chemicals have been used.[46]

Jendrassik-Grof Method for Total and Conjugated Bilirubin Determination Principle

Bilirubin pigments in serum or plasma are reacted with a diazo reagent (sulfanilic acid in hydrochloric acid and sodium nitrite), resulting in the production of the purple product, azobilirubin. The product azobilirubin may be measured spectrophotometrically. The individual fractions of bilirubin are determined by taking two aliquots of sample and reacting one aliquot with the diazo reagent only and the other aliquot with the diazo reagent and an accelerator (caffeine–benzoate). The addition of caffeine–benzoate will solubilize the water-insoluble fraction of bilirubin and will yield a total bilirubin value (all fractions). The reaction without the accelerator will yield conjugated bilirubin only. After a short period of time, the reaction of the aliquots with the diazo reagent is terminated by the addition of ascorbic acid. The ascorbic acid destroys the excess diazo reagent. The solution is then alkalinized using an alkaline tartrate solution, which shifts the absorbance spectrum of the azobilirubin to a more intense blue color that is less subject to interfering substances in the sample. The final blue product is measured at 600 nm, with the intensity of color produced directly proportional to bilirubin concentration. Indirect (unconjugated) bilirubin may be calculated by subtracting the conjugated bilirubin concentration from the total bilirubin concentration.

Sources of Error. Instruments should be frequently standardized to maintain reliable bilirubin results, and careful preparation of bilirubin standards is critical as these are subject to deterioration from exposure to light. Hemolysis and lipemia should be avoided as they will alter bilirubin concentrations. Serious loss of bilirubin occurs after exposure to fluorescent and indirect and direct sunlight; therefore, it is imperative that exposure of samples and standards to light be kept to a minimum. Specimens and standards should be refrigerated in the dark until testing can be performed.

Table 19.2 provides the reference ranges for bilirubin.

Urobilinogen in Urine and Feces

Urobilinogen is a colorless end product of bilirubin metabolism that is oxidized by intestinal bacteria to the brown pigment urobilin. In the normal individual, part of the urobilinogen is excreted in feces, and the remainder is reabsorbed into the portal blood and returned to the liver. A small portion that is not taken up by the hepatocytes is excreted by the kidney as urobilinogen. Increased levels of urinary urobilinogen are found in hemolytic disease and in defective liver cell function, such as that seen in **hepatitis**. Absence of urobilinogen from the urine and stool is most often seen with complete biliary obstruction. Fecal urobilinogen is also decreased in biliary obstruction, as well as in HCC.[47]

Table 19.2 Reference Ranges for Bilirubin in Adults and Infants

Population	Type of Bilirubin	Reference Range
Adults	Direct bilirubin	0.0–0.2 mg/dL
	Indirect bilirubin	0.2–0.8 mg/dL
	Total bilirubin	0.2–1.0 mg/dL
Premature infants	Total bilirubin at 24 hr.	1–6 mg/dL
	Total bilirubin at 48 hr.	6–8 mg/dL
	Total bilirubin 3–5 days	10–12 mg/dL
Full-term infants	Total bilirubin at 24 hr.	2–6 mg/dL
	Total bilirubin at 48 hr.	6–7 mg/dL
	Total bilirubin 3–5 days	4–6 mg/dL

Most quantitative methods for urobilinogen are based on a reaction first described by Ehrlich in 1901: the reaction of urobilinogen with *p*-dimethylaminobenzaldehyde (Ehrlich's reagent) to form a red color. Many modifications of this procedure have been made over the years to improve specificity. However, because the modifications did not completely recover urobilinogen from the urine, most laboratories use the less laborious, more rapid, semiquantitative method described next.

Determination of Urine Urobilinogen (Semiquantitative)

Principle. Urobilinogen reacts with *p*-dimethylaminobenzaldehyde (Ehrlich's reagent) to form a red color, which is then measured spectrophotometrically. Ascorbic acid is added as a reducing agent to maintain urobilinogen in the reduced state. The use of saturated sodium acetate stops the reaction and minimizes the combination of other chromogens with the Ehrlich's reagent.[48]

Specimen. A *fresh* 2-hour urine specimen is collected. This specimen should be kept cool and protected from light.

Comments and Sources of Error. Some important caveats about the test include the following:

1. The results of this test are reported in Ehrlich units rather than in milligrams of urobilinogen because substances other than urobilinogen account for some of the final color development.
2. Compounds, other than urobilinogen, that may be present in the urine and react with Ehrlich's reagent include porphobilinogen, sulfonamides, procaine, and 5-hydroxyindoleacetic acid. Bilirubin will form a green color and, therefore, must be removed, as previously described.
3. Fresh urine is necessary, and the test must be performed without delay to prevent oxidation of urobilinogen to urobilin. Similarly, the spectrophotometric readings should be made within 5 minutes after color production, because the urobilinogen–aldehyde color slowly decreases in intensity.

Reference Range. Urine urobilinogen, 0.1 to 1.0 Ehrlich units every 2 hours or 0.5 to 4.0 Ehrlich units per day (0.86 to 8 mmol/d); 1 Ehrlich unit is equivalent to approximately 1 mg of urobilinogen.

Fecal Urobilinogen

Visual inspection of the feces is usually sufficient to detect decreased urobilinogen. However, the semiquantitative determination of fecal urobilinogen is available and involves the same principle described earlier for the urine. It is carried out in an aqueous extract of fresh feces, and any urobilin present is reduced to urobilinogen by treatment with alkaline ferrous hydroxide before Ehrlich's reagent is added. A range of 75 to 275 Ehrlich units per 100 g of fresh feces or 75 to 400 Ehrlich units per 24-hour specimen is considered a normal reference range.[48]

Serum Bile Acids

Serum bile acid analysis is rarely performed because the methods required are very complex. These involve extraction with organic solvents, partition chromatography, gas chromatography-mass spectrometry, spectrophotometry, ultraviolet light absorption, fluorescence, radioimmunoassay, and enzyme immunoassay (EIA) methods. Although serum bile acid levels are elevated in liver disease, the total concentration is extremely variable and adds no diagnostic value to other tests of liver function. The variability of the type of bile acids present in serum, together with their existence in different conjugated forms, suggests that more relevant information of liver dysfunction may be gained by examining patterns of individual bile acids and their state of conjugation. For example, it has been suggested that the ratio of the trihydroxy to dihydroxy bile acids in serum will differentiate patients with obstructive jaundice from those with hepatocellular injury and that the diagnosis of primary biliary cirrhosis and extrahepatic cholestasis can be made based on the ratio of the cholic to chenodeoxycholic acids. However, the high cost of these tests, the time required to do them, and the current controversy concerning their clinical usefulness render this approach unsatisfactory for routine use.[49,50]

Enzymes

Liver enzymes play an important role in the assessment of liver function because injury to the liver resulting in cytolysis or necrosis will cause the release of enzymes into circulation. Enzymes also play an important role in differentiating hepatocellular (functional) from obstructive (mechanical) liver disease, which is an important clinical distinction because failure to identify an obstruction will result in liver failure if the obstruction is not rapidly treated. Although many enzymes have been identified as useful in the assessment of liver function, the most clinically useful include the aminotransferases (ALT and AST), the phosphatases (ALP and 5′-neucleotidase), GGT, and lactate dehydrogenase (LD).

The methods used to measure these enzymes, the normal reference ranges, and other general aspects of enzymology are discussed in Chapter 8, *Enzymes*. Discussion in this chapter focuses on the characteristic changes in serum enzyme levels seen in various hepatic disorders. It is important to note that the diagnosis of disease depends on a combination of patient history, physical examination, laboratory testing, and sometimes radiologic studies and biopsy, and therefore, abnormalities in liver enzymes alone are not diagnostic in and of themselves.[51,52]

Aminotransferases

The two most common aminotransferases measured in the clinical laboratory are AST (formerly referred to as serum glutamic oxaloacetic transaminase [SGOT]) and ALT (formerly referred to as serum glutamic pyruvic transaminase [SGPT]). The aminotransferases are responsible for catalyzing the conversion of aspartate and alanine to oxaloacetate and pyruvate, respectively. In the absence of acute necrosis or ischemia of other organs, these enzymes are most useful in the detection of hepatocellular (functional) damage to the liver. ALT is found mainly in the liver (lesser amounts in skeletal muscle and kidney), whereas AST is widely distributed in equal amounts in the heart, skeletal muscle, and liver, making ALT a more "liver-specific" marker than AST. Regardless, the serum activity of both transaminases rises rapidly in almost all diseases of the liver and may remain elevated for up to 2 to 6 weeks. The highest levels of AST and ALT are found in acute conditions such as acute viral hepatitis, drug- and toxin-induced liver necrosis, and hepatic ischemia. The increase in ALT activity is usually greater than that for AST. Only moderate increases are found in less severe conditions. AST and ALT are found to be normal or only mildly elevated in cases of obstructive liver damage. Because AST and ALT are present in other tissues besides the liver, elevations in these enzymes may be a result of other organ dysfunction or failure such as acute myocardial infarction, renal infarction, progressive muscular dystrophy, and those conditions that result in secondary liver disease such as infectious mononucleosis, diabetic ketoacidosis,

and hyperthyroidism. It is often helpful to conduct serial determinations of aminotransferases when following the course of a patient with acute or chronic hepatitis, and caution should be used in interpreting abnormal levels because serum transaminases may actually decrease in some patients with severe acute hepatitis, owing to the exhaustive release of hepatocellular enzymes.[51,52]

Phosphatases

Alkaline Phosphatase. The ALP family of enzymes are zinc metalloenzymes that are widely distributed in all tissues; however, highest activity is seen in the liver, bone, intestine, kidney, and placenta. The clinical utility of ALP lies in its ability to differentiate hepatobiliary disease from osteogenic bone disease. In the liver, the enzyme is localized to the microvilli of the bile canaliculi, and therefore, it serves as a marker of extrahepatic biliary obstruction, such as a stone in the common bile duct, or in intrahepatic cholestasis, such as drug cholestasis or primary biliary cirrhosis. ALP is found in very high concentrations in cases of extrahepatic obstruction, with only slight to moderate increases seen in those with hepatocellular disorders such as hepatitis and cirrhosis. Because bone is also a source of ALP, it may be elevated in bone-related disorders such as Paget's disease, bony metastases, diseases associated with an increase in osteoblastic activity, and rapid bone growth during puberty. ALP is also found elevated in pregnancy due to its release from the placenta, where it may remain elevated up to several weeks postdelivery. As a result, interpretation of ALP concentrations is difficult because enzyme activity of ALP can increase in the absence of liver damage.[51,52]

5′-Nucleotidase. 5′-Nucleotidase (5NT) is rarely ordered and is typically only performed in very large reference laboratories. It is responsible for catalyzing the hydrolysis of nucleoside-5′-phosphate esters. Although 5NT is found in a wide variety of cells, serum levels become significantly elevated in hepatobiliary disease. There is no bone source of 5NT, so it is useful in differentiating ALP elevations due to the liver from other conditions where ALP may be seen in increased concentrations (bone diseases, pregnancy, and childhood growth). Levels of both 5NT and ALP are elevated in liver disease, whereas in primary bone disease, ALP level is elevated but the 5NT level is usually normal or only slightly elevated. This enzyme is much more sensitive to metastatic liver disease than is ALP because, unlike ALP, its level is not significantly elevated in other conditions, such as in pregnancy or during childhood. In addition, some increase in enzyme activity may be noted after abdominal surgery.[51–54]

γ-Glutamyltransferase. GGT is a membrane-localized enzyme found in high concentrations in the kidney, liver, pancreas, intestine, and prostate but not in bone. The clinical utility of GGT is differentiating whether elevated levels of ALP are due to skeletal or hepatobiliary disease; the highest levels of GGT are seen in biliary obstruction. GGT is the most sensitive hepatic enzyme indicator for liver disease. Therefore, ingestion of alcohol or certain drugs (barbiturates, tricyclic antidepressants, and anticonvulsants) elevates GGT. It is a sensitive test for cholestasis caused by chronic alcohol or drug ingestion. Measurement of this enzyme is also useful if jaundice is absent for the confirmation of hepatic neoplasms.[51–55]

Lactate Dehydrogenase. Lactate dehydrogenase (LD) is an enzyme with a very wide distribution throughout the body. It is released into circulation when cells of the body are damaged or destroyed, serving as a general, nonspecific marker of cellular injury. Moderate elevations of total serum LD levels are common in acute viral hepatitis and in cirrhosis, whereas biliary tract disease may produce only slight elevations. High serum levels may be found in metastatic carcinoma of the liver. As a result of its wide distribution, LD measurements provide no additional clinical information above that which is provided by the previously mentioned enzymes. However, fractionation of LD into its five tissue-specific isoenzymes may give useful information about the site of origin of the LD elevation.

Tests Measuring Hepatic Synthetic Ability

A healthy functioning liver is required for the synthesis of serum proteins (except the immunoglobulins). The measurement of serum proteins, therefore, can be used to assess the synthetic ability of the liver. Although these tests are not sensitive to minimal liver damage, they may be useful in quantitating the severity of hepatic dysfunction.

A decreased serum albumin may be a result of decreased liver protein synthesis, and the albumin level correlates well with the severity of functional impairment and is found more often in chronic rather than in acute liver disease. The serum α-globulins also tend to decrease with chronic liver

disease. However, a low or absent α-globulin suggests α-antitrypsin deficiency as the cause of the chronic liver disease. Serum γ-globulin levels are transiently increased in acute liver disease and remain elevated in chronic liver disease. The highest elevations are found in chronic active hepatitis and postnecrotic cirrhosis. In particular, immunoglobulin G (IgG) and IgM levels are more consistently elevated in chronic active hepatitis; IgM in primary biliary cirrhosis; and IgA in alcoholic cirrhosis.

Prothrombin time is commonly increased in liver disease because the liver is unable to manufacture adequate amounts of clotting factor or because the disruption of bile flow results in inadequate absorption of vitamin K from the intestine. However, a prothrombin time is not routinely used to aid in the diagnosis of liver disease. Rather, serial measurements of prothrombin times may be useful in following the progression of disease and the assessment of the risk of bleeding. A marked prolongation of the prothrombin time indicates severe, diffuse liver disease and a poor prognosis.

Tests Measuring Nitrogen Metabolism

The liver plays a major role in removing ammonia from the bloodstream by converting it to urea so that it can be removed by the kidneys. A plasma ammonia level, therefore, is a reflection of the liver's ability to perform this conversion. In liver failure, ammonia and other toxins increase in the bloodstream and may ultimately cause hepatic coma. In this condition, the patient becomes increasingly disoriented and gradually lapses into unconsciousness. The cause of hepatic coma is not fully known, although ammonia is presumed to play a major role. However, the correlation between blood ammonia levels and the severity of the hepatic coma is poor, and ammonia levels are most useful when multiple measurements are made over time.

The most common laboratory determination of ammonia concentrations is based on the following reaction:

$$NH_4 + \alpha\text{-ketoglutarate} + NADH \xrightarrow{\text{Glutamate dehydrogenase (GLDH)}} \text{Glutamate} + NAD + H_2O \quad \text{(Eq. 19.1)}$$

The resulting decrease in absorbance at 340 nm is measured and is proportional to ammonia concentration. The sample of choice is plasma collected in ethylenediaminetetraacetic acid (EDTA), lithium heparin, or potassium oxalate, and the samples should be immediately placed on ice to prevent metabolism of other nitrogenous compounds to ammonia in the sample, leading to false elevations in ammonia. If analysis cannot be performed immediately, the plasma should be removed and placed on ice or frozen. Frozen (–70°C) samples are stable for several days. Hemolyzed samples should be rejected for analysis as red blood cells have a concentration of ammonia two to three times higher than that of plasma.[56] Lipemic samples and those with high bilirubin concentrations may be unsuitable for analysis in some systems. The glutaminase activity of GGT is a major contributor to the endogenous production of ammonia; therefore, concentrations may be artifactually increased in samples with raised GGT activity.[57]

Hepatitis

Hepatitis refers to an inflammatory condition of the liver. Infectious causes for the inflammation of liver include viral, bacterial, and parasitic infections, as well as noninfectious causes, such as radiation, drugs, chemicals, and autoimmune diseases and toxins. Viral infections account for most hepatitis cases observed in the clinical setting. Major hepatitis virus subtypes include HAV, HBV, HCV, HDV, and HEV. Infections with these viruses can lead to the onset of acute disease with symptoms including jaundice, dark urine, fatigue, nausea, vomiting, and abdominal pain. Some subtypes, such as HBV and HCV, can lead to the prolonged elevation of serum transaminase level (longer than 6 months), a condition termed *chronic hepatitis*. Routes of transmission vary from one viral subtype to another. HAV and HEV are typically caused by ingestion of contaminated food or water (fecal–oral route). HBV and HCV are common bloodborne pathogens, and HBV is a common sexually transmitted infection. Like HBV, the primary mode of transmission of HDV is parenteral contact with infected body fluids and sexual contact. **Table 19.3** provides details about the different subtypes of hepatitis virus.

Hepatitis A

HAV, also known as infectious hepatitis or short-incubation hepatitis, is the most common form of viral hepatitis worldwide. It is caused by a nonenveloped RNA virus of the picornavirus family. Approximately 1.5 million HAV infections occur annually, with the most common source of infection being the ingestion of contaminated food and drinking

Table 19.3 The Hepatitis Viruses

	Nucleotide	Incubation Period	Primary Mode of Transmission	Vaccine	Chronic Infection	Serological Diagnosis Available
Hepatitis A	RNA	2–6 wk	Fecal–oral	Yes	No	Yes
Hepatitis B	DNA	8–26 wk	Parenteral, sexual	Yes	Yes	Yes
Hepatitis C	RNA	2–15 wk	Parenteral, sexual	No	Yes	Yes
Hepatitis D	RNA	2–8 wk	Parenteral, sexual	No	Yes	Yes
Hepatitis E	RNA	3–6 wk	Fecal–oral	No	Yes	Yes

water.[58] Because HAV is excreted in bile and shed in the feces, which can contain up to 10^9 infectious virions per gram, the fecal–oral route is the primary means of HAV transmission.[59,60] Patients with HAV infection present with symptoms of fever, malaise, anorexia, nausea, abdominal discomfort, dark urine, and jaundice. Symptoms are generally self-limited and resolve within 3 weeks. However, in rare instances, patients develop fulminant liver failure. Chronic infection with HAV is not found, and there is no evidence of a carrier state or long-term sequelae in humans.

Clinical markers for the diagnosis and the progression of HAV infection are measured through the presence of serologic antibodies. IgM antibodies to HAV (IgM anti-HAV) are detectable at or prior to the onset of clinical illness and decline in 3 to 6 months, when it becomes undetectable by commercially available diagnostic tests.[61] IgG antibodies to HAV (IgG anti-HAV) appear soon after IgM, persist for years after infection, and confer lifelong immunity.[62] IgM anti-HAV has been used as the primary marker of acute infection.[63] The presence of elevated titers of IgG anti-HAV in the absence of IgM indicates a past infection. Another reliable method to detect acute infection in patients is assaying for the presence of viral antigen, which is shed in the feces. However, the antigen is no longer present soon after liver enzymes have reached their peak levels. Another method of detecting HAV infection is amplification of viral RNA by reverse transcription polymerase chain reaction (RT-PCR). Nucleic acid detection techniques are more sensitive than immunoassays for viral antigen to detect HAV in samples of different origins (e.g., clinical specimens, environmental samples, or food). Because of the high proportion of asymptomatic HAV infections, nucleic acid amplification techniques are useful to determine the extent to which unidentified infection occurs.[64]

The availability of vaccines to provide long-term immunity against HAV infection has the potential to significantly reduce the incidence of disease and possibly eliminate the transmission of this virus worldwide.[65–68] Since approval of the HAV vaccine for children in the United States in 2006, the Advisory Committee on Immunization Practices (ACIP) recommends that all children receive the 2-dose series of the HAV vaccine beginning at age 12 months. The use of this vaccine has significantly reduced the incidence of HAV in the United States and has therefore changed the epidemiology of this infection. In addition to children, the HAV vaccine is recommended for the following people: those traveling to countries where hepatitis A is common, those who are family and caregivers of adoptees from countries where hepatitis A is common, men who have sexual encounters with other men, those who are users of recreational drugs (injected or not), those with chronic or long-term liver disease (including hepatitis B or hepatitis C), and those with clotting factor disorders.[69]

Hepatitis B

Known previously as *serum hepatitis* or *long-incubation hepatitis*, HBV can cause both acute and chronic hepatitis and is the most ubiquitous of the hepatitis viruses. Two billion individuals are infected globally and 1.5 million new HBV infections are diagnosed annually. In the United States, 12 million people have been infected, with 2.4 million of those individuals estimated to be chronically infected. The highest incidence of acute HBV is among adults aged 25 to 45 years.[70] HBV is comparatively stable in the environment and remains viable for longer than 7 days on environmental surfaces at room temperature.[71] It is detected in virtually all body fluids, including blood, feces, urine, saliva, semen, tears, and breast milk; the three major routes of transmission are parenteral, perinatal, and sexual. Persons at high risk for infection in the United States include persons who engage

in high-risk sexual behaviors and the sharing of drug injection needles. Children born to mothers who are hepatitis B surface antigen (HBsAg) positive at the time of delivery, immigrants from endemic areas, and sexual partners and household contacts of patients who have HBV are high-risk groups for HBV infection. Although transmission of HBV by blood transfusion occurs, effective screening tests now make this transmission route rare. Healthcare workers, including laboratory personnel, may be at increased risk for developing HBV, depending on their degree of exposure to blood and body fluids.[72]

Serologic Markers of HBV Infection. HBV is a 42-nm DNA virus classified in the Hepadnaviridae family. The liver is the primary site of HBV replication. Following an HBV infection, the core of the antigen is synthesized in the nuclei of hepatocytes and then passed into the cytoplasm of the liver cell, where it is surrounded by the protein coat. An antigen present in the core of the virus (HBcAg) and a surface antigen present on the surface protein (HBsAg) have been identified by serologic studies. Another antigen, called the e antigen (HBeAg), has also been identified.[73]

Hepatitis B Surface Antigen (HBsAg). Previously known as the Australia antigen and hepatitis associated antigen, HBsAg is the antigen for which routine testing is performed on all donated units of blood. HBsAg is a useful serologic marker in patients before the onset of clinical symptoms because it is present during the prodrome of acute HBV. HBsAg is not infectious; however, its presence in the serum may indicate the presence of the hepatitis virus. Therefore, persons who chronically carry HBsAg in their serum must be considered potentially infectious because the presence of the intact virus cannot be excluded. HBsAg is the only serologic marker detected during the first 3 to 5 weeks after infection in newly infected patients. The average time from exposure to detection of HBsAg is 30 days (range 6 to 60 days).[74–76] Highly sensitive single-sample nucleic acid tests can detect HBV DNA in the serum of an infected person 10 to 20 days before detection of HBsAg.[76] HBsAg positivity has been reported for up to 18 days after HBV vaccination and is clinically insignificant.[77,78] Patients who achieve complete viral clearance develop the antibody to the HBsAg (anti-HBs), following the disappearance of the HBsAg (**Figure 19.8**).

Hepatitis B Surface Antibody (anti-HBs). The presence of anti-HBs antibody in patients is frequently observed in the general population, suggestive of past infection. Patients who have developed the antibody to the HBsAg are not susceptible to future reinfection with HBV.[72] It is useful for identifying previous exposure to hepatitis B virus and determining adequate immunity from hepatitis B vaccination. When interpreting anti-HBs results, a positive result indicates recovery from acute or chronic hepatitis B virus (HBV) infection or acquired immunity from HBV vaccination. This test does not differentiate between a vaccine-induced immune response and an HBV infection immune response. However, in conjunction with the presence of a positive total anti-HBc result, it would indicate that the anti-HBs response is due to past HBV infection. Positive results, as defined by the manufacturer, indicate

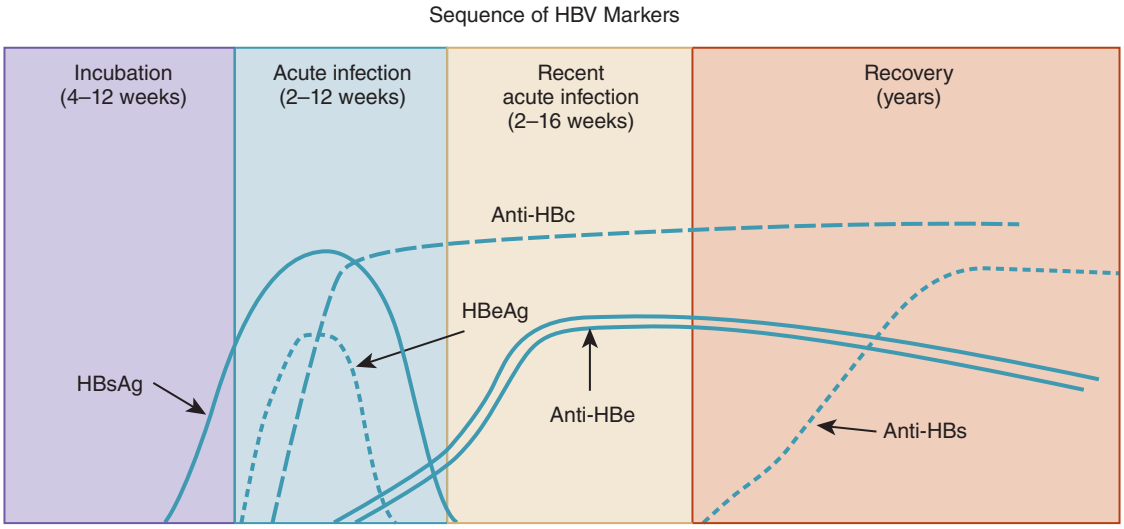
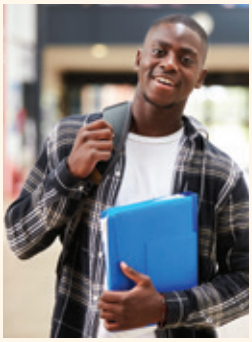


Figure 19.8 Serology of hepatitis B infection with recovery.

CASE STUDY 19.3, PART 2

Recall DeShawn. He returned to the Student Health Service to review his laboratory results with the advanced practice provider. The results are summarized in the following table.



© Monkey Business Images/Shutterstock.

	Laboratory Results
ALT (SGPT)	Elevated
AST (SGOT)	Elevated
Alkaline phosphatase	Minimally elevated
Lactate dehydrogenase	Elevated
Serum bilirubin	Increased
Urine bilirubin	Increased
Hepatitis A antibody (IgG)	Negative
Hepatitis A antibody (IgM)	Positive
Hepatitis B surface antigen	Negative
Hepatitis B surface antibody	Negative
Hepatitis C antibody	Negative

1. What is DeShawn's most likely diagnosis?
2. What is the most probable cause of his illness?
3. What are DeShawn's treatment and prognosis?

adequate immunity to hepatitis B from past hepatitis B or HBV vaccination. Per current CDC guidelines, individuals with anti-HBs levels greater than 10 mIU/mL after completing an HBV vaccination series are considered protected from hepatitis B. Negative results indicate a lack of recovery from acute or chronic hepatitis B or inadequate immune response to HBV vaccination. The U.S. Advisory Committee on Immunization Practices does not recommend more than two HBV vaccine series in nonresponders. Indeterminate results, defined as results below the positive threshold and above the negative threshold, indicate inability to determine if anti-HBs is present at levels consistent with recovery or immunity. Repeat testing is recommended in 1 to 3 months. The window period of hepatitis B infection is the time after the disappearance of the HBsAg and the appearance of the anti-HBs antibody.

Hepatitis B Core Antigen. Currently, there is no test assay to detect HBcAg in the plasma of patients or blood donors. This hepatitis B core antigen is present only in the nuclei of hepatocytes during an acute infection with HBV. The antibody to the HB core antigen, anti-HBc, usually develops earlier in the course of infection than the antibody to the surface antigen (Figure 19.8).

Hepatitis B Core Antibody, Total. This test is useful for the detection and differentiation between recent (past) and chronic HBV infection. It is used for the diagnosis of recent HBV infection during the "window period" when both the HBsAg and anti-HBs viral markers are negative. Positive anti-HBc (total) results may indicate recent (past) or chronic HBV infection. If positive, then testing for anti-HBc (IgM) is required to confirm the presence of acute or recent hepatitis B. A positive anti-HBc total result with a negative anti-HBc IgM result indicates past or chronic HBV infection. In chronic hepatitis B, the HBsAg will be detectable, thus differentiating the two conditions. A negative anti-HBc (total) result indicates the absence of recent, past, or chronic hepatitis B. An inconclusive result for anti-HBc (total) suggests presence of an interfering substance, which may or may not be resolved by retesting. Passively acquired maternal IgG antibodies may produce a positive anti-HBc (total) result in newborns. Confirmatory testing for anti-HBc (IgM) is necessary, to make a diagnosis of acute or recent hepatitis B in these infants.

Hepatitis B Core Antibody, IgM. A test for the IgM antibody to HBcAg was developed as a serologic marker for clinical use. The presence of this IgM

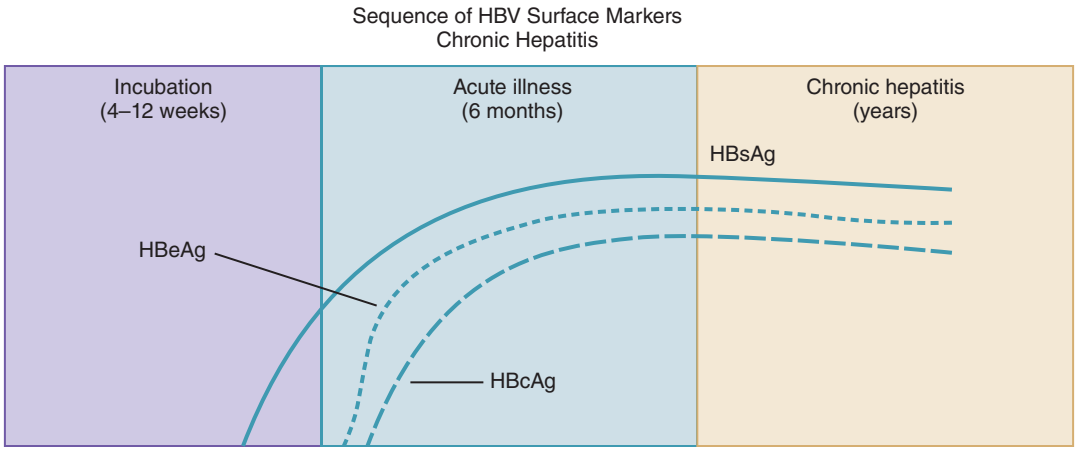


Figure 19.9 No antibody is formed against HBsAg. The persistence of HBeAg implies high infectivity and a generally poor prognosis. Without treatment, this patient would likely develop cirrhosis unless seroconversion occurs.

antibody is specific for acute HBV infection. In patients who have chronic HBV infection, the IgM anti-HBc antibody titer can persist during chronic viral replication at low levels that typically are not detectable by assays used in the United States. However, persons with exacerbation of chronic infection can test positive for IgM anti-HBc.[78] Another marker for acute infection is a viral DNA-dependent DNA polymerase that is closely associated with the presence of the core antigen. This viral enzyme is required for viral replication and is detectable in serum early in the course of viral hepatitis, during the phase of active viral replication.[79]

Hepatitis B Envelope Antigen (HBeAg). The hepatitis B envelope (HBe) antigen, an antigen closely associated with the core of the viral particle, is detected in the serum of persons with acute or chronic HBV infection. The presence of the HBe antigen appears to correlate well with both the number of infectious virus particles and the degree of infectivity of HBsAg-positive sera. The presence of HBeAg in HBsAg carriers is an unfavorable prognostic sign and predicts a severe course and chronic liver disease. Conversely, the presence of anti-HBe antibody in carriers indicates a low infectivity of the serum (**Figure 19.9**). The HBe antigen is detected in serum only when surface antigen is present (**Figure 19.10**).

Hepatitis B Envelope Antibody (Anti-HBe). During recovery from acute hepatitis B, the HBeAg level declines and becomes undetectable in the serum, and the anti-HBe becomes detectable in the serum. Anti-HBe remains detectable for several years after

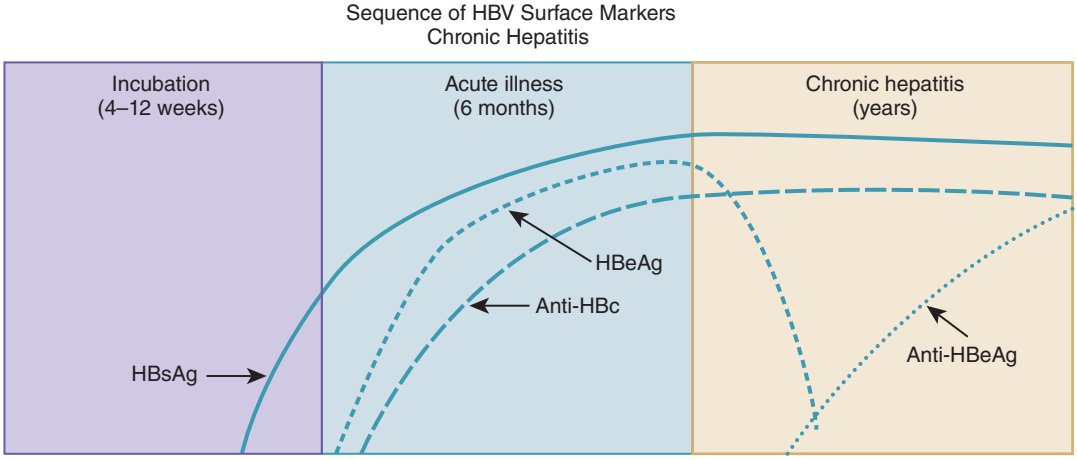


Figure 19.10 Serology of chronic hepatitis with formation of antibody to HBeAg. This is a favorable sign and suggests that the chronic hepatitis may resolve. Complete recovery would be heralded by the disappearance of HBsAg and formation of its corresponding antibody.

recovery from acute HBV infection. In chronic HBV infection, positive anti-HBe results indicate low infectivity of the patients and low risk of transmission. Positive anti-HBe results in the presence of detectable HBV DNA in serum indicate active viral replication. This test is useful for determining infectivity of HBV carriers and monitoring infection status and serologic response to antiviral therapy in individuals with chronic HBV infection. When interpreting results, the absence of HBeAg with the presence of anti-HBe is consistent with inactivity of the virus and loss of HBV infectivity. Although the detection of anti-HBe usually indicates resolution of chronic HBV infection, the HBV carrier state may persist.

In summary, the serologic markers of HBV infection typically used to differentiate among acute, resolving, and chronic infections are HBsAg, anti-HBc, and anti-HBs (**Table 19.4**). Persons who recover from natural infection typically will test positive for both anti-HBs and anti-HBc, whereas persons who respond to HBV vaccine have only anti-HBs. Persons who become chronically infected fail to develop antibody to the HBsAg, resulting in the persistent presence of HBsAg as well as the presence of anti-HBc in patient serum, typically for life.[80–83] HBeAg and anti-HBe screenings typically are used for the management of patients with chronic infection. Serologic assays are available commercially for all markers except HBcAg because no free HBcAg circulates in blood.[70]

Nucleic acid hybridization or PCR technique is used to detect HBV DNA in the blood and is another method used to measure disease progression. This technique provides a more sensitive measurement of infectivity and disease progression than serology. It may be used to monitor the effectiveness of antiviral therapy in patients with chronic HBV infection, but it supplements rather than replaces current HBV serologic assays.[84–86]

Chronic Infection with HBV. Approximately 90% of patients infected with HBV recover within 6 months. Recovery is accompanied by the development of the antibody to the HBsAg. However, about 10% of patients progress to a chronic hepatitis infection. The likelihood of developing chronic HBV infection is higher in individuals infected perinatally

Table 19.4 Typical Interpretation of Serologic Test Results for Hepatitis B Virus Infection

Serologic Marker				
HBsAg[a]	Total Anti-HBc[b]	IGM[c] Anti-HBc	Anti-HBs[d]	Interpretation
–[e]	–	–	–	Never infected
+[f,g]	–	–	–	Early acute infection; transient (up to 18 d) after vaccination
+	+	+	–	Acute infection
–	+	+	+ or –	Acute resolving infection
–	+	–	+	Recovered from past infection and immune
+	+	–	–	Chronic infection
–	+	–	–	False positive (i.e., susceptible); past infection, "low-level" chronic infection[h]; or passive transfer of anti-HBc to infant born to HBsAg-positive mother
–	–	–	+	Immune if concentration is ≥10 mIU/mL after vaccine series completion; passive transfer after hepatitis B immune globulin administration

[a]Hepatitis B surface antigen
[b]Antibody to hepatitis B core antigen.
[c]Immunoglobulin M
[d]Antibody to HBsAg
[e]Negative test result
[f]Positive test result
[g]To ensure that an HBsAg-positive test result is not a false positive, samples with reactive HBsAg results should be tested with a licensed neutralizing confirmatory test if recommended in the manufacturer's package insert.
[h]Persons positive only for anti-HBc are unlikely to be infectious except under unusual circumstances in which they are the source for direct percutaneous exposure of susceptible recipients to large quantities of virus (e.g., blood transfusion or organ transplant).

(90%) and during childhood (20% to 30%), when the immune system is thought to be less developed and unable to achieve efficient viral clearance, than in adult immunocompetent subjects (<1%).[87] Approximately 25% of persons who were chronically infected since childhood and 15% of those who were chronically infected since adulthood die prematurely from cirrhosis or liver cancer. In most cases, patients remain asymptomatic until the onset of cirrhosis.[88]

Patients with chronic HBV display the characteristic serologic profile as shown in Figure 19.9. The presence of HBsAg in chronically infected patients is an indication that they are infectious and at risk for developing complications, including cirrhosis and HCC. The natural course of chronic HBV infection is divided into four phases based on the virus–host interaction: immune tolerance, immune clearance (HBeAg-positive chronic hepatitis), low or nonreplication (inactive HBsAg carrier), and reactivation (HBeAg-negative chronic hepatitis).[89–91] Patients can be classified according to their serologic status as shown in **Table 19.5**.[87]

HBV Treatment and Prevention. Persons who have chronic HBV infection require medical evaluation and regular monitoring.[92–94] The FDA has approved several drugs for the treatment of chronic HBV that can achieve sustained suppression of HBV replication and remission of liver disease in certain persons.[93]

HBV vaccination is the most effective measure to prevent HBV infection and its consequences, including cirrhosis of the liver, liver cancer, liver failure, and death. HBsAg is the antigen used for HBV vaccination.[95,96] The vaccine is highly effective in stimulating the production of hepatitis B surface antibody and thereby rendering the recipient immune. As a result of the national program of childhood immunization, by 2005, a 98% decline in HBV infection was reported among children aged 13 years or younger, as well as a 97% decline among adolescents aged 12 to 19 years.[73] The Advisory Committee on Immunization Practices (ACIP) recommended universal HBV vaccination for all unvaccinated adults at risk for HBV infection and for all adults requesting protection from HBV infection.[73] All healthcare workers who handle blood products and are in close proximity to body fluids should have the hepatitis vaccine. The hepatitis B immune globulin (HBIG) provides passively acquired anti-HBs and temporary protection (i.e., 3 to 6 months) when administered in standard doses. HBIG typically is used as an adjunct to HBV vaccine for postexposure immunoprophylaxis to prevent HBV infection. For nonresponders to HBV vaccination, CDC recommends completing a second hepatitis B vaccine series.

Hepatitis C

HCV (originally "non-A non-B hepatitis") is caused by a virus with an RNA genome that is a member of the Flaviviridae family. HCV is transmitted parenterally. Although the sexual and fecal–oral routes as modes of transmission have been documented, the virus is transmitted primarily by blood transfusion of inappropriately screened blood products.[97–99] Approximately 2.5% of the world population is infected by the virus,[100] and most infections become chronic and may lead to cirrhosis, end-stage liver disease, HCC, and death. In the United States, an estimated 2.4 million people currently live with chronic hepatitis C, and in 2018, a total of 3620 cases were reported to the CDC, which is likely a substantial undercount of the true number of new infections.[101] HCV infection is present in about 0.84% of the noninstitutionalized adult population.[102]

Clinically, acute HCV infection presents only mild infection, and patients may remain completely asymptomatic. However, HCV infection has a high rate of progression to chronic hepatitis, cirrhosis, and liver carcinoma.[99,103] In addition, HCV infection is a leading cause of liver transplantation in this country.[104]

Laboratory Tests for Hepatitis C. The hepatitis C antibody (Anti-HCV) is usually not detected in the first few months of infection but will almost always be present in the later stages. The antibody is not protective against reinfection and sometimes disappears several years following resolution of the infection. Laboratory testing for the diagnosis of HCV infection in the clinical setting is relatively

Table 19.5 Serological Markers of Chronic Hepatitis B Virus Infection

	HBsAg	Anti-HBs	Anti-HBc (Total)	HBeAg	Anti-HBe
Chronic Hepatitis B (active replicating)	+	–	+	+	–
Chronic Hepatitis B (non- replicating)	+	–	+	–	+

straightforward. Currently, two laboratory tests are commonly used to diagnose HCV infection in clinical practice: anti-HCV detection by EIA and quantitative nucleic acid PCR assays for serum HCV RNA.[102,105–109] In clinical practice, the most common approach is initially to test a patient's serum for the presence of anti-HCV by EIA. If this test gives a positive result, the next step is to test for serum HCV RNA by PCR.[102] The HCV antibody test is designed to detect antibodies generated in response to HCV infection.[102,108] Although a positive HCV antibody test result generally indicates that the patient has been exposed to the HCV virus, this test cannot determine whether the patient is currently infected with HCV or has recovered from HCV infection.[105,108] Some patients with a positive HCV antibody test result have spontaneously cleared HCV. These patients (anti-HCV positive but HCV RNA negative) are recommended to retake the test for HCV RNA on a second occasion, 3 to 6 months after the first HCV RNA test.[108] Since HCV is an RNA virus, its detection and quantification is performed by real-time reverse transcription-PCR (RT-PCR) in serum specimens. This assay is replacing the need for serologic testing for HCV antibodies in the laboratory. New testing algorithms for screening and diagnosis use this RT-PCR molecular method.

Chronic HCV Infection. Most patients with HCV infection progress to chronic infection. Although patients with chronic HCV infection appear to be at high risk for liver cirrhosis, the role of HCV in the development of HCC is not clear. Most chronically infected patients are asymptomatic and manifest only mild elevations of liver function tests, especially transaminases. The degree of elevation in liver enzymes has little predictive value toward disease progression. About 80% of infected patients develop chronic hepatitis, although in most cases, the disease does not progress. The percentage of patients progressing to liver cirrhosis varies widely in different studies but has been estimated to be as high as 40% after 40 years. Alcohol consumption concomitant with chronic HCV infection significantly increases the risk of cirrhosis. Liver biopsies are performed periodically in these patients, with the degree of inflammation and fibrosis correlating with the risk of cirrhosis.[110–113]

Patients with chronic HCV infection are usually treated with pegylated interferon and ribavirin. Therapeutic efficacy is monitored by using PCR to determine the number of viral copies in serum.[114] A prototypic envelope peptide-based vaccine has been reported to induce antibodies in human subjects, but there is no evidence for the presence of a neutralizing antibody against HCV.[115] An effective and safe vaccine against HCV is unlikely to be available soon.

Hepatitis D

Hepatitis D virus (also known as delta hepatitis) is a unique subviral satellite virus. It is a small, defective RNA-containing virus that cannot replicate independently but rather requires the HBsAg of HBV for replication. Therefore, it is incapable of causing any illness in patients who do not have HBV infection. Infection with HBV and HDV can occur at the same time (known as coinfection), or infection with HDV can occur after first being infected with HBV (known as superinfection). Modes of transmission for HDV are similar to those of HBV. Approximately 5% of the global HBV carriers are coinfected with HDV, leading to a total of 10 to 15 million HDV carriers worldwide.[116,117] Each year, 7500 new cases of HDV infections are estimated to occur in the United States.[118] Chronic HDV infection is estimated to be responsible for more than 1000 deaths each year in the United States.[119]

HDV virions possess an outer envelope composed of HBsAg proteins and host membrane lipids and an inner nucleocapsid consisting of viral RNA and hepatitis delta antigen (HDAg). It is believed that HBsAg-mediated binding to a cellular receptor helps HDV penetrate the hepatocyte.[120] Nuclear localization signal domain on the HDAg facilitates transit of its genome into the nucleus, where viral replication takes place.[121] HDV infection can occur concurrently with HBV infection (coinfection) or in a patient with established HBV infection (superinfection) (**Table 19.6**). The rate of chronicity following coinfection with HBV and HDV is equal to that of HBV infection alone.[122] HDV superinfection is likely to become chronic simply because HBV infection is already chronic. In general, in the acute phase, HDV superinfected carriers may develop severe hepatitis, and around 70% to 90% will progress to chronicity.[123]

Table 19.6 Serological Markers of Hepatitis D Virus Infection

	Coinfection	Superinfection
HBsAg	+	+
Anti-HBc (IgM)	+	–
Anti-HBs	+	–
HDVAg	+	+
Anti-HDV (Total)	±	+

CASE STUDY 19.2, PART 3

Remember Verlita. Her laboratory results from the hepatitis profile are shown below:

Laboratory Results	
Hepatitis A virus antibody (IgM)	Negative
Hepatitis B surface antigen	Positive
Hepatitis B core antibody (IgM)	Positive
Hepatitis C virus antibody (IgM)	Negative

© stockimagesbank/Shutterstock.

3. What is Verlita's most likely diagnosis?
4. Does Verlita have an acute or chronic disease?
5. What complications may Verlita develop?

The diagnosis of HDV relies on detection of antibodies against HDAg and serum HDV RNA, as well as HBV markers.[124] Clinical symptoms of HDV cannot be distinguished from those of other hepatitis viruses. Accurate diagnosis is made by a negative test for IgM anti-HBc and confirmed by the detection of HDV markers.[125] Widespread use of the HBV vaccine has resulted in a decline in the incidence of HDV.[126] Interferon-α is currently the therapy used for treating chronic HDV infection. Compared with chronic HBV or HCV, chronic HDV treatment requires a higher dosage and a longer duration of treatment, and posttreatment relapses are common.[125]

Hepatitis E

The RNA-containing HEV, a nonenveloped RNA virus that is only 27 to 34 nm in diameter, is the sole member of the genus *Hepevirus* in the family Hepeviridae. After infection, the incubation period is short, generally between 21 and 42 days prior to the onset of symptoms. The virus may be detected in feces and bile by about 7 days after infection. HEV is transmitted primarily by the fecal–oral route, and waterborne epidemics are characteristic of HEV in many developing countries. However, in industrialized countries, several nonhuman mammals such as pigs, cows, and sheep are susceptible to infection with HEV, leading to the potential spread of the virus through zoonosis.[127] The clinical presentation of HEV is comparable to that of HAV. The severity of an HEV infection is generally greater than the severity of an HAV infection.[128] In general, HEV infection is mild, except in pregnant women, in whom it may be a devastating illness.[129]

Because cases of HEV are not clinically distinguishable from cases of other types of acute viral hepatitis, diagnosis is made by biochemical assessment of liver function. Acute HEV is diagnosed when the presence of IgM anti-HEV is detected.[128,130] The presence of a high or increasing anti-HEV IgG titer may support the diagnosis of acute HEV infection, and in such cases, acute HEV can be presumed even in the absence of IgM anti-HEV.[131] EIA and immunochromatography are most convenient for the detection of IgM and/or IgG anti-HEV. Additional testing by RT-PCR has a limited confirmatory role. Acute-phase HEV RNA can be detected in feces by PCR in approximately 50% of cases. In most instances, there is a very good positive correlation between the results assayed by RT-PCR and EIA.[128] HEV should be suspected in outbreaks of waterborne hepatitis occurring in developing countries, especially if the disease is more severe in pregnant women or if HAV has been excluded. If laboratory tests are not available, epidemiologic evidence can help in establishing a diagnosis.[132]

At present, no commercially available vaccines exist for the prevention of HEV. Experimental immune prophylaxis against HEV based on recombinant antigens appears to confer short-term protection and may be useful for pregnant women in endemic areas and travelers coming into these regions.

Other Forms of Hepatitis

Five forms of viral hepatitis (A, B, C, D, and E) are well recognized. Hepatitis F is an enteric agent that may be transmitted to primates. GB virus C, a member of the *Flaviviridae* family formerly known as hepatitis G, is known to infect humans but not cause human disease. Little is known about these diseases, and no diagnostic tests for them are commercially available at this time.[133–137] Cytomegalovirus, Epstein-Barr virus, SARS-CoV-2, and several other agents can also cause hepatitis.[138,139]

CASE STUDY 19.4, PART 2

Recall José. Based on the nurse practitioner's advice, José made a follow-up appointment with his primary care physician because of the abnormal laboratory test results from a preinsurance physical examination. The physician requested follow up testing and the results below are being reviewed.

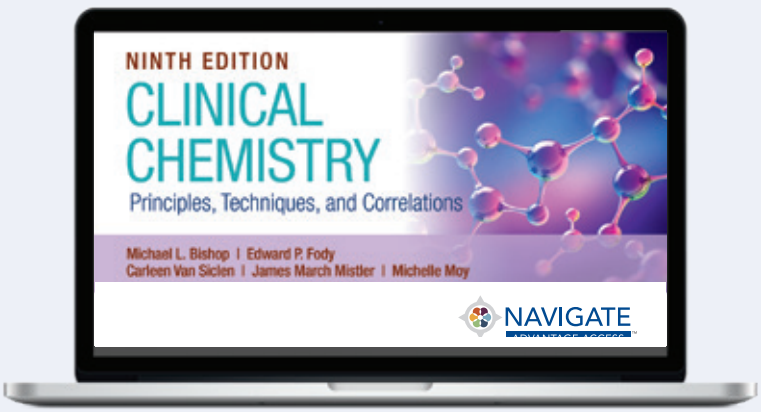
© justoomm/Shutterstock.

Laboratory Results	
Hepatitis A antibody (IgG)	Positive
Hepatitis A antibody (IgM)	Negative
Hepatitis B surface antigen	Positive
Hepatitis B surface antibody	Negative
Hepatitis B core antibody (IgM)	Positive
Hepatitis C antibody	Negative

1. What is José's most likely diagnosis?
2. Does José have an acute or chronic illness?
3. What complications may José develop over time?
4. What additional tests should be considered?

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Sticova E, Jirsa M. New insights in bilirubin metabolism and their clinical implications. *World J Gastroenterol*. 2013;19(38):6398–6407.
2. Ahlfors CE. Measurement of plasma unbound unconjugated bilirubin. *Anal Biochem*. 2000;279:130–135.
3. Canbay A, Bechmann L, Gerken G. Lipid metabolism in the liver. *Z Gastroenterol*. 2007;45(1):35–41.
4. Papavramidou N, Fee E, Christopoulou-Aletra H. Jaundice in the hippocratic corpus. *J Gastrointest Surg*. 2007;11:1728–1731.
5. Pashankar D, Schreiber RA. Jaundice in older children and adolescents. *Pediatr Rev*. 2001;22:219–226.
6. Gazzin S, Masutti F, Vitek L, Tirbelli C. The molecular basis of jaundice: an old symptom revisited. *Liver Int*. 2017;37:1094–1102.
7. Black M, Billing BH. Hepatic bilirubin UDP-glucuronyl transferase activity in liver disease and Gilbert's syndrome. *N Engl J Med*. 1969;280:1266–1271.
8. Kaplan M, Hammerman C. Bilirubin and the genome: the hereditary basis of unconjugated neonatal hyperbilirubinemia. *Curr Pharmacogenomics*. 2005;3:21–42.
9. Crigler JF Jr, Najjar VA. Congenital familial nonhemolytic jaundice with kernicterus. *Pediatrics*. 1952;10:169–180.

10. Berk PD, Noyer C. The familial unconjugated hyperbilirubinemia. *Semin Liver Dis.* 1994;14:346–351.
11. Berk PD. Bilirubin metabolism and the hereditary hyperbilirubinemias. *Semin Liver Dis.* 1994;14:321–322.
12. Berk PD, Noyer C. The familial conjugated hyperbilirubinemia. *Semin Liver Dis.* 1994;14:386–394.
13. Nisa AU, Ahmad Z. Dubin-Johnson syndrome. *J Coll Physicians Surg Pak.* 2008;18(3):188–189.
14. Cuperus FJ, Hafkamp AM, Hulzebos CV, Verkade HJ. Pharmacological therapies for unconjugated hyperbilirubinemia. *Curr Pharm Des*. 2009;15(25):2927–2938. doi: 10.2174/138161209789058219. PMID: 19754369
15. Erlinger S, Arias IM, Dhumeaux D. Inherited disorders of bilirubin transport and conjugation: new insights into molecular mechanisms and consequences. *Gastroenterology*. 2014;146(7):1625–1638.
16. Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: new trends. *J Hepatol*. 2020;72:250–261.
17. American Cancer Society. *Key Statistics About Liver Caner*. Available from: https://www.cancer.org/cancer/liver-cancer/about/what-is-key-statistics.html. Accessed September 14, 2021.
18. Murphy SL, Xu JQ, Kochanek KD, Arias E, Tejada-Vera B. *Deaths: Final data for 2018.* National Vital Statistics Reports; vol 69 no 13. Hyattsville, MD: National Center for Health Statistics. 2020.
19. Ferlay J, Soerjomataram I, Ervik M, et al. *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 11 [Internet]*. Lyon, France: International Agency for Research on Cancer; 2013. Available from: http://globocan.iarc.fr. Accessed May 1, 2016.
20. El-Serag HB. Hepatocellular carcinoma. *N Engl J Med.* 2011;365:1118–1127.
21. American Cancer Society. *Cancer Facts & Figures 2021*. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2021.html. Accessed April 2, 2021.
22. Makuuchi M. Surgical treatment for HCC—special reference to anatomical resection. *Int J Surg*. 2013;11(suppl 1): S47–S49.
23. Peters LJ, Wiener GJ, Gilliam J, et al. Reye's syndrome in adults. A case report and review of the literature. *Arch Intern Med.* 1986;146:2401–2403.
24. Preblud SR, Orenstein WA, Bart KJ. Varicella: clinical manifestations, epidemiology and health impact in children. *Pediatr Infect Dis.* 1984;3:505–509.
25. Edwards KM, Bennett SR, Garner WL, et al. Reye's syndrome associated with adenovirus infections in infants. *Am J Dis Child.* 1985;139:343–346.
26. Johnson GM, Scurletis TD, Carroll NB. A study of sixteen fatal cases of encephalitis. *NC Med J.* 1963;24:464–473.
27. Starko KM, Ray CG, Dominguez LB, et al. Reye's syndrome and salicylate use. *Pediatrics.* 1980;66:859–864.
28. Waldman RJ, Hall WN, McGee H, et al. Aspirin as a risk factor in Reye's syndrome. *JAMA.* 1982;247:3089–3094.
29. Surgeon General's advisory on the use of salicylates and Reye's syndrome. *MMWR Morb Mortal Wkly Rep.* 1982;31: 289–290.
30. Davis DL, Buffler P. Reduction of deaths after drug labelling for risk of Reye's syndrome. *Lancet.* 1992;340:1042.
31. Daugherty CC, Gartside PS, Heubi JE, et al. A morphometric study of Reye's syndrome. Correlation of reduced mitochondrial numbers and increased mitochondrial size with clinical manifestations. *Am J Pathol.* 1987;129:313–326.
32. Belay ED, Bresee JS, Holman RC, et al. Reye's syndrome in the United States from 1981 through 1997. *N Engl J Med.* 1999;340:1377–1382.
33. Liu ZX, Kaplowitz N. Immune-mediated drug-induced liver disease. *Clin Liver Dis.* 2002;6:755–774.
34. Shehu AI, Ma X, Venkataramanan R. Mechanisms of drug-induced hepatotoxicity. *Clin Liver Dis*. 2017;21(1):35–54.
35. Bissell DM, Gores GJ, Laskin DL, et al. Drug-induced liver injury: mechanisms and test systems. *Hepatology.* 2001;33:1009–1013.
36. Parker R, Neuberger JM. Alcohol, diet and drug use preceding alcoholic hepatitis. *Dig Dis*. 2018;36(4):298–305.
37. Aday AW, Mitchell MC, Casey LC. Alcoholic hepatitis: current trends in management. *Curr Opin Gastroenterol*. 2017;33(3):142–148.
38. Maddrey WC, Boitnott JK, Bedine MS, et al. Corticosteroid therapy of alcoholic hepatitis. *Gastroenterology.* 1978;75: 193–199.
39. Forrest EH, Evans CD, Stewart S, et al. Analysis of factors predictive of mortality in alcoholic hepatitis and derivation and validation of the Glasgow alcoholic hepatitis score. *Gut.* 2005;54:1174–1179.
40. Testino G, Leone S. Acute alcoholic hepatitis: a literature review and proposal of treatment. *Minerva Med*. 2018;109(4):290–299.
41. Nagata K, Suzuki H, Sakaguchi S. Common pathogenic mechanism in development progression of liver injury caused by non-alcoholic or alcoholic steatohepatitis. *J Toxicol Sci.* 2007;32:453–468.
42. Schmidt ET, Wheeler CA, Jackson GL, Engle WD. Evaluation of transcutaneous bilirubinometry in preterm neonates. *J Perinatol*. 2009;29:564–569.
43. Lott JA, Doumas BT. "Direct" and total bilirubin tests: contemporary problems. *Clin Chem.* 1993;39:641–647.
44. Wilding P, Zilva JF, Wilde CE. Transport of specimens for clinical chemistry analysis. *Ann Clin Biochem.* 1977;14:301–306.
45. Doumas BT, Kwok-Cheung PP, Perry BW, et al. Candidate reference method for determination of total bilirubin in serum: development and validation. *Clin Chem.* 1985;31:1779–1789.
46. Doumas BT, Wu TW. The measurement of bilirubin fractions in serum. *Crit Rev Clin Lab Sci.* 1991;28:415–445.
47. Ross JW, Miller WG, Myers GL, et al. The accuracy of laboratory measurements in clinical chemistry: a study of 11 routine chemistry analytes in the College of American Pathologists Chemistry Survey with fresh frozen serum, definitive methods, and reference methods. *Arch Pathol Lab Med.* 1998;122:587–608.
48. Slaughter MR. Sensitivity of conventional methods for bilirubin measurements. *J Lab Clin Med.* 1996;127:233.
49. Malloy HT, Evelyn KA. The determination of bilirubin with the photoelectric colorimeter. *J Biol Chem.* 1937;119: 481–490.
50. Hofmann AF. Bile acid secretion, bile flow and biliary lipid secretion in humans. *Hepatology.* 1990;12(3 pt 2):17S–22S.
51. Kotal P, Fevery J. Quantitation of urobilinogen in feces, urine, bile and serum by direct spectrophotometry of zinc complex. *Clin Chim Acta.* 1991;202:1–9.
52. Azer SA, Klaassen CD, Stacey NH. Biochemical assay of serum bile acids: methods and applications. *Br J Biomed Sci.* 1997;54:118–132.

53. Polkowska G, Polkowski W, Kudlicka A, et al. Range of serum bile acid concentrations in neonates, infants, older children, and in adults. *Med Sci Monit.* 2001;7(suppl 1):268–270.
54. Green RM, Flamm S. AGA technical review on the evaluation of liver chemistry tests. *Gastroenterology.* 2002;123: 1367–1384.
55. American Gastroenterological Association medical position statement: evaluation of liver chemistry tests. *Gastroenterology*. 2002;123:1364–1366.
56. Pagani F, Panteghini M. 5′-Neucleotidase in the detection of increased activity of the liver form of alkaline phosphatase in serum. *Clin Chem.* 2001;47:2046–2048.
57. Sunderman FW. The clinical biochemistry of 5′-nucleotidase. *Ann Clin Lab Sci.* 1990;20:123–139.
58. Jefferies M, Rauff B, Rashid H, Lam T, and Rafiq S. Update on global epidemiology of viral hepatitis and preventive strategies. *World J Clin Cases*. 2018;6(13):589–599.
59. Frank E. Nonprotein nitrogen compounds. In: Bishop ML, Fody EP, Schoeff, eds. *Clinical Chemistry: Principles, Techniques, Correlations*. Philadelphia, PA: Lippincott Williams & Wilkins; 2013:246–260.
60. da Fonseca-Wollheim F. Deamidation of glutamine by increased plasma gamma-glutamyltransferase is a source of rapid ammonia formation in blood and plasma specimens. *Clin Chem.* 1990;36:1479–1482.
61. Koff RS. Review article: vaccination and viral hepatitis—current status and future prospects. *Aliment Pharmacol Ther.* 2007;26:1285–1292.
62. Skinhoj P, Mathiesen LR, Kiryger P, et al. Faecal excretion of hepatitis A virus in patients with symptomatic hepatitis A infection. *Scand J Gastroenterol.* 1981;16:1057–1059.
63. Tassopoulos NC, Papaevangelou GJ, Ticehurst JR, et al. Fecal excretion of Greek strains of hepatitis A virus in patients with hepatitis A and in experimentally infected chimpanzees. *J Infect Dis.* 1986;154:231–237.
64. Kao HW, Ashcavai M, Redeker AG. The persistence of hepatitis A IgM antibody after acute clinical hepatitis A. *Hepatology.* 1984;4:933–936.
65. Skinhoj P, Mikkelsen F, Hollinger FB. Hepatitis A in Greenland: importance of specific antibody testing in epidemiologic surveillance. *Am J Epidemiol.* 1977;105: 140–147.
66. Cuthbert JA. Hepatitis A: old and new. *Clin Microbiol Rev.* 2001;14:38–58.
67. Nainan OV, Xia G, Vaughan G, et al. Diagnosis of hepatitis A virus infection: a molecular approach. *Clin Microbiol Rev.* 2006;19:63–79.
68. Bell BP, Feinstone SM. Hepatitis A vaccine. In: Plotkin SA, Orenstein WA, Offit PA, eds. *Vaccine*. 4th ed. Philadelphia, PA: WB Saunders; 2004:269–297.
69. Nelson NP, Weng MK, Hofmeister MG, et al. Prevention of hepatitis A virus infection in the United States: Recommendations of the Advisory Committee on Immunization Practices, 2020. *MMWR Recomm Rep* 2020;69(No. RR-5):1–38.
70. Hollinger FB, Martin A. Hepatitis A virus. In: Knipe SM, Howley PM, eds. *Fields Virology*. 6th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2013: 550–581.
71. Buffington J, Mast E. Viral hepatitis. In: Wallace RB, ed. *Maxcy-Rosenau-Last Public Health and Preventive Medicine*. 15th ed. McGraw-Hill Medical; 2008: 211–227.
72. Wodi AP, Ault K, Hunter P, McNally V, Szilagyi PG, Bernstein H. Advisory Committee on Immunization Practices Recommended Immunization Schedule for Children and Adolescents Aged 18 Years or Younger — United States, 2021. *MMWR Morb Mortal Wkly Rep.* 2021;70:189–192.
73. Schillie S, Vellozzi C, Reingold A, et al. Prevention of Hepatitis B Virus Infection in the United States: Recommendations of the Advisory Committee on Immunization Practices. *MMWR Recomm Rep.* 2018;67(RR-1):1–36.
74. Bond WW, Favero MS, Petersen NJ, et al. Survival of hepatitis B virus after drying and storage for one week. *Lancet.* 1981;1:550–551.
75. Aspinall EJ, Hawkins G, Fraser A, Hutchinson SJ, Goldberg D. Hepatitis B prevention, diagnosis, treatment and care: a review. *Occup Med (Lond)*. 2011;61(8):531–540.
76. Kwo PY, Cohen SM, Lim JK. ACG Clinical Guideline: Evaluation of Abnormal Liver Chemistries. *Am J Gastroenterol*. 2017 Jan;112(1):18–35.
77. Krugman S, Overby LR, Mushahwar IK, et al. Viral hepatitis, type B. Studies on natural history and prevention re-examined. *N Engl J Med.* 1979;300:101–106.
78. Hoofnagle JH, DiBisceglie AM. Serologic diagnosis of acute and chronic viral hepatitis. *Semin Liver Dis.* 1991;11: 73–83.
79. Biswas R, Tabor E, Hsia CC, et al. Comparative sensitivity of HBV NATs and HBsAg assays for detection of acute HBV infection. *Transfusion*. 2003;43:788–798.
80. Lunn ER, Hoggarth BJ, Cook WJ. Prolonged hepatitis B surface antigenemia after vaccination. *Pediatrics.* 2000;105:E81.
81. Kao JH, Chen PJ, Lai MY, et al. Acute exacerbations of chronic hepatitis B are rarely associated with superinfection of hepatitis B virus. *Hepatology.* 2001;34(4 pt 1):817–823.
82. Kurstak C, Kurstak E, Hossain A. Progress in diagnosis of viral hepatitis A, B, C, D and E. *Acta Virol.* 1996;40: 107–115.
83. Alward WL, McMahon BJ, Hall DB, et al. The long-term serological course of asymptomatic hepatitis B virus carriers and the development of primary hepatocellular carcinoma. *J Infect Dis.* 1985;151:604–609.
84. Liaw YF, Sheen IS, Chen TJ, et al. Incidence, determinants and significance of delayed clearance of serum HBsAg in chronic hepatitis B virus infection: a prospective study. *Hepatology.* 1991;13:627–631.
85. Adachi H, Kaneko S, Matsushita E, et al. Clearance of HBsAg in seven patients with chronic hepatitis B. *Hepatology.* 1992;16:1334–1337.
86. McMahon BJ, Holck P, Bulkow L, et al. Serologic and clinical outcomes of 1536 Alaska Natives chronically infected with hepatitis B virus. *Ann Intern Med.* 2001;135:759–768.
87. Jalava T, Ranki M, Bengtström M, et al. A rapid and quantitative solution hybridization method for detection of HBV DNA in serum. *J Virol Methods.* 1992;36:171–180.
88. Malavé Lara C, Gorriño MT, Campelo C, et al. Detection of hepatitis B virus DNA in chronic carriers by the polymerase chain reaction. *Eur J Clin Microbiol Infect Dis.* 1992;11: 740–744.
89. Aspinall S, Steele AD, Peenze I, et al. Detection and quantitation of hepatitis B virus DNA: comparison of two commercial hybridization assays with polymerase chain reaction. *J Viral Hepat.* 1995;2:107–111.
90. Fattovich G, Bortolotti F, Donato F. Natural history of chronic hepatitis B: special emphasis on disease progression and prognostic factors. *J Hepatol.* 2008;48:335–352.

91. Hadziyannis SJ, Papatheodoridis GV. Hepatitis B e antigen-negative chronic hepatitis B: natural history and treatment. *Semin Liver Dis.* 2006;26:130–141.
92. EASL International Consensus Conference on Hepatitis B. Consensus statement. *J Hepatol.* 2003;38:533–540.
93. Perz JF, Armstrong GL, Farrington LA, et al. The contribution of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *J Hepatol.* 2006;45:529–538.
94. Lok AS, McMahon BJ. Chronic hepatitis B. *Hepatology.* 2007;45:507–539.
95. Lok AS, McMahon BJ; Practice Guidelines Committee, American Association for the Study of Liver Disease (AASLD). Chronic hepatitis B. *Hepatology.* 2001;34: 1225–1241.
96. Lok AS, McMahon BJ; Practice Guidelines Committee, American Association for the Study of Liver Disease (AASLD). Chronic hepatitis B: update of recommendations. *Hepatology.* 2004;39:857–861.
97. Bruix J, Sherman M; Practice Guidelines Committee, American Association for the Study of Liver Disease (AASLD). Management of hepatocellular carcinoma. *Hepatology.* 2005;42:1208–1236.
98. Purcell RH, Gerin JL. Hepatitis B subunit vaccine: a preliminary report of safety and efficacy tests in chimpanzees. *Am J Med Sci.* 1975;270:395–399.
99. Hilleman MR, McAleer WJ, Buynak EB, et al. Quality and safety of human hepatitis B vaccine. *Dev Biol Stand.* 1983;54:3–12.
100. Petruzziello A, Marigliano S, Loquercio G, et al. Global epidemiology of hepatitis C virus infection: an up-date of the distribution and circulation of hepatitis C virus genotypes. *World J Gastroenterol.* 2016;22(34):7824–7840.
101. Hepatitis C questions and answers for health professionals. Centers for Disease Control and Prevention. Last reviewed August 7, 2020. Accessed August 23, 2021. https://www.cdc.gov/hepatitis/hcv/hcvfaq.htm#section1
102. Rosenberg ES, Rosenthal EM, Hall EW, et al. Prevalence of hepatitis C virus infection in the US states and the District of Columbia, 2013 to 2016. *JAMA Netw Open.* 2018;1(8):e186371.
103. Yen T, Keeffe EB, Ahmed A. The epidemiology of hepatitis C virus infections. *J Clin Gastroenterol.* 2003;36: 47–53.
104. Lingala S, Ghany MG. Natural history of hepatitis C. *Gastroenterol Clin North Am.* 2015;44(4):717–734.
105. Centers for Disease Control and Prevention. *Viral Hepatitis Surveillance — United States, 2018*. https://www.cdc.gov/hepatitis/statistics/SurveillanceRpts.htm. Published July 2020. Accessed [April 2, 2021].
106. Holland PV. Post-transfusion hepatitis: current risks and causes. *Vox Sang.* 1998;74(suppl 2):135–141.
107. Strader DB, Wright T, Thomas DL, et al. American Association for the Study of Liver Diseases practice guideline: diagnosis, management, and treatment of hepatitis C. *Hepatology.* 2004;39:1147–1171.
108. Dal Molin G, Tiribelli C, Campello C. A rational use of laboratory tests in the diagnosis and management of hepatitis C virus infection. *Ann Hepatol.* 2003;2:76–83.
109. Pawlotsky J-M. Diagnostic testing in hepatitis C virus infection: viral kinetics and genomics. *Semin Liver Dis.* 2003;23(suppl 1):3–11.
110. Pawlotsky JM, Lonjon I, Hezode C, et al. What strategy should be used for diagnosis of hepatitis C virus infection in clinical laboratories? *Hepatology.* 1998;27:1700–1702.
111. Alter MJ, Kuhnert WL, Finelli L. Guidelines for laboratory testing and result reporting of antibody to hepatitis C virus. *MMWR Recomm Rep.* 2003;52(RR-03):1–16.
112. Schille S, Wester C, Osborne M, Wesolowski L, and Blythe Ryerson A. CDC Recommendations for hepatitis C screening among adults—United States, 2020. *MMWR Recomm Rep.* 2020;69(2): 1–17.
113. Corrao G, Arico S. Independent and combined action of hepatitis C virus infection and alcohol consumption on the risk of symptomatic liver cirrhosis. *Hepatology.* 1994;20:1115–1120.
114. Seeff LB. The natural history of chronic hepatitis C virus infection. *Clin Liver Dis.* 1997;1:587–602.
115. Poynard T, Ratziu V, Charlotte F, et al. Rates and risk factors of liver fibrosis progression in patients with chronic hepatitis C. *J Hepatol.* 2001;34:730–739.
116. Forns X, Ampurdanès S, Sanchez-Tapias JM, et al. Long-term follow-up of chronic hepatitis C in patients diagnosed at a tertiary-care center. *J Hepatol.* 2001;35:265–271.
117. Mast EE, Kuramoto IK, Favorov MO, et al. Prevalence of and risk factors for antibody to hepatitis E virus seroreactivity among blood donors in Northern California. *J Infect Dis.* 1997;176:34–40.
118. Di Bisceglie AM, Frey S, Gorse GJ. A phase I safety and immunogenicity trial of a novel E1E2/MF59C.1 hepatitis C vaccine candidate in healthy HCV-negative adults [abstract]. *Hepatology.* 2005;42(suppl 1):750A.
119. Hadziyannis SJ. Decreasing prevalence of hepatitis D virus infection. *J Gastroenterol Hepatol.* 1997;12:745–746.
120. Farci P. Delta hepatitis: an update. *J Hepatol.* 2003;39(suppl 1):S212–S219.
121. Centers for Disease Control and Prevention. Hepatitis surveillance report. *MMWR Morb Mortal Wkly Rep.* 1990;53:23.
122. di Bisceglie AM. Interferon therapy for chronic viral hepatitis. *N Engl J Med* 1994;330:137–138.
123. Sureau C, Moriarty AM, Thornton GB, et al. Production of infectious hepatitis delta virus in vitro and neutralization with antibodies directed against hepatitis B virus pre-S antigens. *J Virol.* 1992;66:1241–1245.
124. Hughes SA, Wedemeyer H, Harrison PM. Hepatitis delta virus. *Lancet.* 2011, Jul. 2;378(9785):73–85.
125. Caredda F, d'Arminio Monforte A, Rossi E, et al. Prospective study of epidemic delta infection in drug addicts. *Prog Clin Biol Res.* 1983;143:245–250.
126. Wu JC, Chen TZ, Huang YS, et al. Natural history of hepatitis D viral superinfection: significance of viremia detected by polymerase chain reaction. *Gastroenterology.* 1995;108:796–802.
127. Macagno S, Smedile A, Caredda F, et al. Monomeric (7S) immunoglobulin M antibodies to hepatitis delta virus in hepatitis type D. *Gastroenterology.* 1990;98: 1582–1586.
128. Koh C, Heller T, Glenn JS. Pathogenesis of and New Therapies for Hepatitis D. *Gastroenterology.* 2019;156(2): 461–476.e1.
129. Mele A, Mariano A, Tosti ME, et al. Acute hepatitis delta virus infection in Italy: incidence and risk factors after the introduction of the universal anti-hepatitis B vaccination campaign. *Clin Infect Dis.* 2007;44:e17–e24.

130. Primadharsini PP, Nagashima S, Okamoto H. Genetic variability and evolution of hepatitis E virus. *Viruses*. 2019;11(5):456.
131. Emerson SU, Purcell RH. Hepatitis E virus. In: Knipe DM, Howley PM, eds. *Fields Virology*. 6th ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins; 2013:2242–2258.
132. Terrault NA, Levy MT, Cheung KW, Jourdain G. Viral hepatitis and pregnancy. *Nat Rev Gastroenterol Hepatol*. 2021;18(2):117–130.
133. Anderson D, Counihan N. Hepatitis A and E Viruses. In: Jorgensen JH, Pfaller MA, Carroll KC, et al., eds. *Manual of Clinical Microbiology*. 11th ed. Washington, DC: American Society for Microbiology Press; 2015:1584–1598.
134. Panda SK, Thakral D, Rehman S. Hepatitis E virus. *Rev Med Virol*. 2007;17:151–180.
135. Nimgaonkar I, Ding Q, Schwartz RE, Ploss A. Hepatitis E virus: advances and challenges. *Nat Rev Gastroenterol Hepatol*. 2018;15(2):96–110.
136. Frider B, Sookoian S, Castaño G, et al. Detection of hepatitis G virus RNA in patients with acute non-A-E hepatitis. *J Viral Hepat*. 1998;5:161–164.
137. Reshetnyak VI, Karlovich TI, Ilchenko LU. Hepatitis G virus. *World J Gastroenterol*. 2008;14(30):4725–4734.
138. Doo EC, Lian TJ, Schiff ER, et al. The hepatitis viruses. In: Schiff ER, Sorrell MF, Maddrey WC, eds. *Schiff's Diseases of the Liver*, 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003:917–940.
139. Alter HJ, Umemura T, Tanaka Y. Putative new hepatitis virus. In: Schiff ER, Sorrell MF, Maddrey WC, eds. *Schiff's Diseases of the Liver*, 9th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003:891–905.

CHAPTER 20

Cardiac Function

Teresa A. Mortier

CHAPTER OUTLINE

KEY TERMS

Acute coronary syndrome (ACS)
Angina pectoris
Atherosclerosis
Cardiac markers
Congestive heart failure
C-reactive protein
Creatine kinase (CK)
D-dimer
Homocysteine
Ischemia
Myocardial infarction (MI)
Myocarditis
Myoglobin
Troponin C
Troponin I
Troponin T

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Discuss the etiology and physiologic effects of the following cardiac conditions:
  - Congenital heart disease
  - Hypertensive heart disease
  - Infectious heart diseases
  - Coronary heart disease
  - Congestive heart failure
  - Pulmonary embolism
- Identify risk factors for coronary heart disease.
- Explain the characteristics of an ideal cardiovascular marker.
- Describe three novel markers of inflammation currently under investigation.
- Compare and contrast the specificity and sensitivity of commonly used cardiac markers.
- Correlate laboratory results with cardiac conditions, including myocardial infarction.
- Analyze the role of the laboratory in the assessment of a patient with cardiac disease.
- Calculate delta value for serial troponin measurements given patient data.
- Interpret laboratory results given patient data.
- Apply knowledge of cardiac markers to solve case study problems within the chapter.

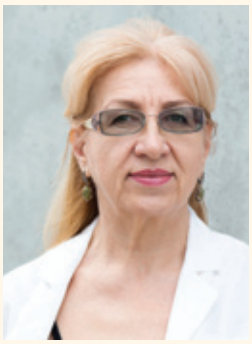
CASE STUDY 20.1, PART 1

Shaniece, a 15-month-old with a heart murmur since birth, was evaluated for repeated pulmonary infections, failure to grow, cyanosis, and mild clubbing of fingers and toes.

© pixelheadphoto/iStock/Getty Images Plus/ Getty Images.

CASE STUDY 20.2, PART 1

Frieda, a 58-year-old woman, came to the emergency department on Monday complaining of heartburn, abdominal discomfort, nausea, and vomiting after a family reunion on Friday.



© A StockStudio/Shutterstock.

CASE STUDY 20.3, PART 1

Carlos, an 83-year-old man with known severe coronary artery disease, diffuse small vessel disease, and significant stenosis distal to a vein graft from previous coronary artery bypass graft surgery (CABG), was admitted to the hospital when his healthcare provider referred him after a routine office visit.

© Fotoluminate LLC/Shutterstock.

Cardiovascular disease (CVD) remains the leading cause of death globally has increased by more than 2 million since 2009, evidenced by the 9 million deaths reported in 2019.[1] The four major types of CVD are based on the location in which it occurs: (1) coronary heart disease (CHD), (2) cerebrovascular disease, (3) peripheral arterial disease, and (4) aortic atherosclerotic disease. CHD manifests as **angina pectoris** (chest pain), **myocardial infarction (MI)** (heart attack), and heart failure. Cerebrovascular disease, where the blood supply is cut off to the brain, manifests as a stroke or transient ischemic attack (TIA), also called a mini-stroke. Peripheral arterial disease (PAD) manifests by blockage in the arteries to the extremities, usually the legs. Both PAD and deep vein thrombosis (DVT) are forms of peripheral vascular disease. The difference is that PAD only affects the arteries and not the veins. Aortic atherosclerosis manifests as aneurysms (abnormal widening of an artery) or dissection (tears in thoracic or abdominal aorta), which are acute, life-threatening events.

Cardiovascular disease, with damage to the heart muscle and heart valves, may present at birth as congenital heart disease or begin in childhood as a result of rheumatic fever. Other forms of CVD develop over time, manifesting in adulthood as a vascular disease. For example, in deep vein thrombosis (DVT) and pulmonary embolism (PE), a blood clot in the peripheral vasculature (often the legs) becomes dislodged and moves to the heart and lungs, where it can block blood flow and cause tissue damage and even death. **Atherosclerosis**, a chronic inflammatory disease in which the blood vessel endothelium becomes exposed and lipids infiltrate the endothelial space, is also common in adults. When the vessel lumen narrows due to lipid plaque formation, the likelihood of **ischemia** (lack of blood supply) can become localized in different parts of the body, including the heart. CHD in which atherosclerosis and ischemia are localized to the vasculature of the heart accounts for one-third of the total deaths related to CVD in Western countries.[2] The critical role of laboratory testing in diagnosing CHD is the focus of this chapter.

Pathophysiology of Atherosclerosis

Atherosclerosis is a chronic disease process that occurs over a number of years and contributes to approximately 50% of all deaths in modern Western

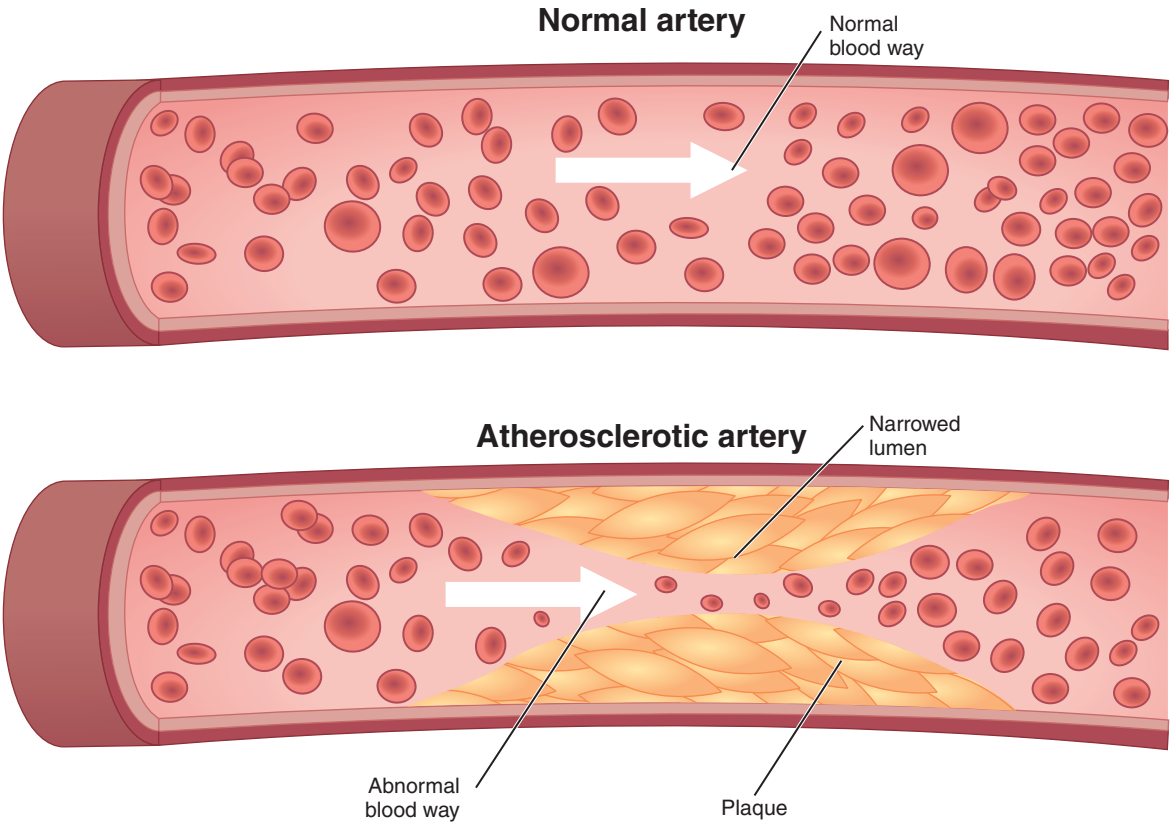


Figure 20.1 Comparison of normal and atherosclerotic arteries. Narrowing the arterial lumen due to atherosclerotic plaque leads to abnormal blood flow, contributing to progression of atherosclerosis.

societies.[3] Evidence of atherosclerosis can often be found in the human aorta before the age of 10,[4] and atherosclerosis becomes pathologic with the development of atherosclerotic plaques (atheromas), which predispose the vasculature to thrombosis, leading to organ ischemia and infarction (**Figure 20.1**). The pathophysiology of atherosclerosis is gradual and complicated, involving a progressive accumulation of lipids, smooth muscle cells, macrophages, and connective tissue within the intima (inner lining) of large and medium-sized arteries, ultimately causing luminal narrowing and decreased perfusion (Figure 20.1). Although the exact etiology of atherosclerosis remains unclear, the *reaction to injury* hypothesis is strongly favored by current evidence, proposing that atherosclerosis is due to a chronic inflammatory response to an accumulation of subtle vascular wall injuries.[5]

Pathohistological evidence from human and experimental studies demonstrates that endothelial and inflammatory cells interact with chemical and inflammatory mediators to promote the development of atherosclerotic plaques. This process begins with vascular injury, which is initiated when endothelial cells are damaged or rendered dysfunctional by vascular abnormalities, such as turbulent blood flow, hyperlipidemia, and hyperhomocysteinemia. A damaged vascular endothelium has an increased permeability to circulating lipids, so having high cholesterol (hypercholesterolemia) in the wake of a damaged endothelium favors the accumulation of lipoproteins, predominantly low-density lipoprotein (LDL) and very low density lipoprotein (VLDL), within the arterial intima (Figure 20.1).[6] In addition, apolipoprotein B- (Apo B-) containing LDL has high affinity for arterial wall proteoglycans.[7] Retention of lipids in arterial walls due to hypercholesterolemia and/or elevated levels of Apo B-containing lipoproteins is thus a crucial step in the pathogenesis of atherosclerotic lesions. The central role of cholesterol accumulation (i.e., LDL and VLDL) in the progression of atherosclerosis is the reason treating high cholesterol is a priority in preventing and attenuating heart disease.

Once vessel lesion initiation has begun, LDL deposited within the intima is oxidized by endothelial cells, lipoxygenase, and free radicals generated by the auto-oxidation of **homocysteine**.[4] The formation of oxidized LDL is central to lesion progression. Oxidized LDL is toxic to endothelial cells, causing

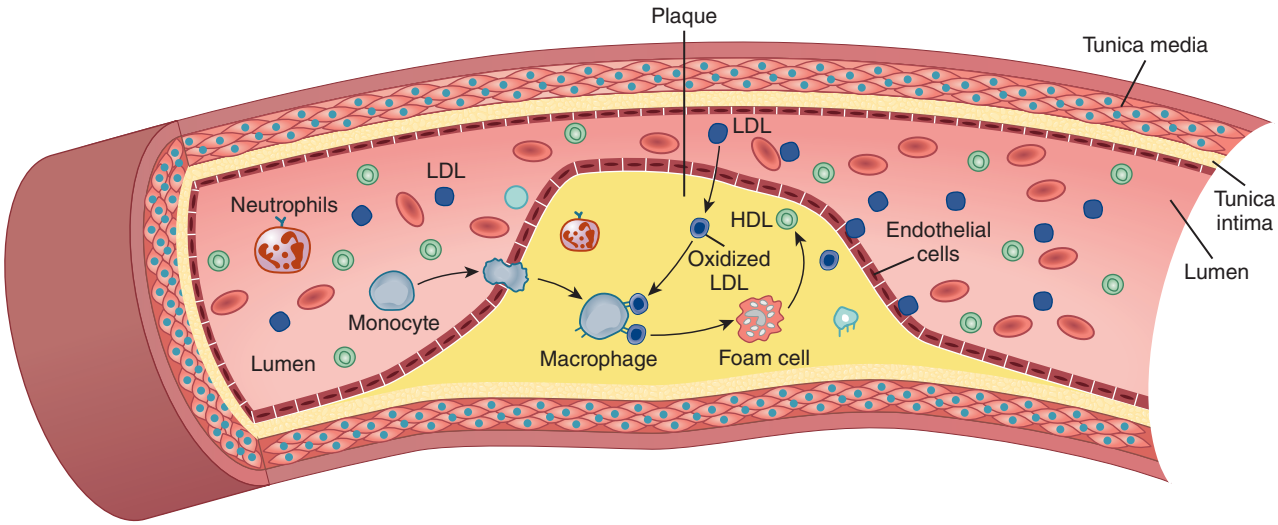


Figure 20.2 LDL deposition and oxidation within the vessel wall leads to monocyte recruitment and differentiation into activated macrophages that phagocytize oxidized LDL to become foam cells. Foam cells release HDL and proinflammatory mediators, and their rupture contributes to lesion progression.

additional intimal damage and subsequent retention of cholesterol-rich lipoproteins, and it elicits an inflammatory response that releases proinflammatory cytokines and recruits inflammatory cells to the early lesion. Among the earliest recruited leukocytes are neutrophils and monocytes, which play critical roles in early atherogenesis by maintaining a proinflammatory state around the initial lesion (**Figure 20.2**).[3]

Maturation of monocytes into activated macrophages is a key step in lesion progression. Macrophage scavenger receptors recognize oxidized LDL, but not native LDL, and activated macrophages rapidly phagocytose cholesterol-rich lipoproteins that have been oxidized within the vessel wall (Figure 20.2).[3] Excessive uptake of oxidized LDL transforms macrophages into bloated, cholesterol-filled cells called *foam cells*.[4,8] Filled with cytoplasmic lipid droplets, foam cells exhibit a variety of functions that both promote lesion progression, including the production of proinflammatory signals, and counter lesion progression, such as the secretion of HDL (Figure 20.2). Foam cells also display numerous metabolic abnormalities, including activation of inflammasomes and the nuclear factor-kappa B (NF-κB) pathway, and it is generally agreed that foam cells play a harmful role in lesion progression.[4] In addition, rupture of foam cells and release of their contents cause further damage to the vascular endothelium, stimulating more inflammation. Deposition and subsequent oxidation of LDL and its effects on monocyte differentiation are thus of key importance in the early development of an atherosclerotic lesion.

As the cycle of endothelial cell damage progresses, additional cell types are recruited to the plaque, in particular T and B lymphocytes and macrophages. These cells are activated by a stream of cytokines, such as interleukin (IL)-1 and tumor necrosis factor-α, that are released by endothelial cells within the plaque.[3] Interactions between T cells and foam cells promote a chronic inflammatory state and help recruit smooth muscle cells into the intima.[3] Additionally, growth factors, such as platelet-derived growth factor, fibroblast growth factor, and tissue growth factor-α, are released by lymphocytes and endothelial cells, further stimulating smooth muscle cell migration and activation.[9]

Once smooth muscle cells migrate into the core of the atheroma, they proliferate and deposit extracellular matrix components that give stability and strength to the plaque.[3] Cytokines released from T cells, such as interferon-γ, exert a number of both pro- and antiatherogenic effects on macrophages and smooth muscle cells, but the net effect of this cytokine release is proinflammatory and proatherogenic. Smooth muscle cells are induced to secrete collagen, elastin, and proteoglycans that form a fibrous backbone and outer shell that fix the plaque firmly in the vessel wall. As the atheroma grows, its core becomes increasingly isolated from surrounding blood supply, leading to hypoxia, which stimulates release of proangiogenic cytokines. This causes aberrant neovascularization around the periphery of the plaque, which predisposes to hemorrhage and deposition of erythrocyte components, such

as hemosiderin and lipids, within the plaque core.[10] Microvasculature hemorrhage and subsequent deposition of additional debris provokes further inflammation, leukocyte recruitment, and remodeling of the plaque.

The steps described above, beginning with lipid and cellular infiltration following vascular injury and progressing to chronic inflammation and fibrosis, continue in a cycle that ultimately culminates in complete vessel occlusion, thrombosis, plaque rupture, or some combination of the three. Each of these possible outcomes predisposes to organ ischemia. In the context of the heart, coronary artery ischemia and resultant hypoxia pose an immediate risk of cellular injury due to the high metabolic activity and oxygen demand of the myocardium. The ultimate outcome of such injury is ischemic heart disease, varying in severity from angina to MI, the symptoms of which are dependent on the extent of coronary artery atherosclerosis. The most common locations for symptomatic cardiac atheroma formation are the proximal left anterior descending, the proximal left main coronary, and the entire right coronary arteries. Coronary atherosclerosis typically becomes symptomatic only after atherosclerotic plaques obstruct approximately 75% (or >50% in the left main) of the cross-sectional area of the vessel.[3]

Cardiac Ischemia, Angina, and Heart Attacks

Acute coronary syndrome (ACS) resulting from CHD leads patients to seek treatment in the emergency department. ACS typically includes symptoms of chest pain, which signals ischemic events ranging from angina (no cell death) to acute myocardial infarction (AMI) (cell death). There are several noncardiac reasons for having chest pain, including common maladies such as esophageal reflux ("heartburn") due to stomach acid moving upward into the esophagus. It is critical to differentiate cardiac from noncardiac causes of chest pain, and making a differential diagnosis in these patients is challenging. This is evidenced by the fact that the diagnosis of ACS is missed in approximately 1% of patients.[11] Delay in treatment for ACS can result in substantial morbidity and mortality. Such patients benefit from rapid initiation of therapy directed at ACS. Patients without elevated **cardiac markers** are typically at lower risk and may undergo further evaluation of the cause of their symptoms, including nuclear cardiac imaging commonly known as a nuclear stress test, computerized tomography (CT) scan, and positron emission tomography (PET) scan.

The lack of an adequate blood supply to the heart leads to ischemia. Since the heart depends upon a constant supply of oxygen and nutrients supplied by the coronary arteries, any disruption in the balance of oxygen supply and cardiac demand puts the heart at serious risk. Adequate supply of blood to the heart is compromised by a number of factors, most commonly the formation of a thrombus in the context of underlying atherosclerosis, which is the underlying cause for MI (narrowing of the coronary arteries). The "classic" manifestation of cardiac ischemia is angina, described as a squeezing of the chest, heavy chest pressure, a burning feeling, or difficulty breathing.[12] This "classic" manifestation often radiates to the left shoulder, neck, or arm and typically increases in intensity over a period of minutes and gets worse with either physical or psychological stress. However, the symptoms often arise in the absence of situations that precipitate it. Nonclassical symptoms, more commonly experienced by women, include a more stabbing, pulsating, or sharp chest pain rather than the typical viselike pressure, nausea, shortness of breath, or abdominal pain.[13] In addition to these heart-related symptoms, the American Heart Association (AHA) and the American College of Cardiology (ACC) have also defined chest-related symptoms that are not indicative of cardiac involvement (see **Table 20.1**).

Patients with chest pain are initially evaluated by a physical examination, an electrocardiogram (ECG or EKG), and a chest x-ray, in addition to laboratory cardiac markers. Several advances in risk-stratifying patients with chest pain have been made, including

Table 20.1 Symptoms Not Related to Cardiac Ischemia-Induced Angina (Chest Pain)

Pleuritic pain (deep cough, sharp/knifelike pain)
Epigastric pain (primary/sole location in the abdomen)
Indigestion
Increasing dyspnea in the absence of chest pain
Psychiatric disorders (e.g., panic attack, anxiety)
Severe, noncardiac conditions such as sepsis, burns, respiratory failure, acute neurological diseases, and drug toxicity (including cancer chemotherapy): Pain that can mimic ACS.

Data from Barstow C, Rice M, McDivitt JD. Acute Coronary Syndrome: Diagnostic Evaluation. *Am Fam Physician.* 2017 Feb 1;95(3):170-177.

Table 20.2 Classification of MI According to the Universal Definition of Myocardial Infarction Released by the Joint ESC/ACC/AHA/WHF Task Force (2018)

Type	Description
Type 1	Coronary artery disease with plaque rupture and coronary thrombosis (STEMI and NSTEMI)
Type 2	Imbalance between oxygen supply and myocardial demand without thrombosis
Type 3	Myocardial infarction resulting in death when biomarker values are unavailable
Type 4a	Myocardial infarction related to PCI
Type 4b	Myocardial infarction related to stent thrombosis
Type 4c	Myocardial infarction due to restenosis ⩾50% after an initially successful PCI
Type 5	Myocardial infarction related to coronary artery bypass grafting

NSTEMI, non-ST elevation MI; PCI, percutaneous coronary intervention; STEMI, ST-elevation MI.

Data from Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al; Executive Group on behalf of the Joint European Society of Cardiology (ESC)/American College of Cardiology (ACC)/American Heart Association (AHA)/World Heart Federation (WHF) Task Force for the Universal Definition of Myocardial Infarction. Fourth Universal Definition of Myocardial Infarction (2018). *J Am Coll Cardiol*. 2018;72:2231-2264.

better laboratory markers of cardiac injury, which allows for rapid identification of those patients whose chest pain is due to CHD.[11] Risk factors for atherosclerosis should also be solicited, including advanced age, gender, diabetes, and any previous history of MI. The physical examination identifies possible causes of precipitating myocardial ischemia (high blood pressure, pulmonary disease, and heart failure or cardiac valve disease)[14] and may identify a noncardiac cause for the symptom. A normal ECG, however, does not rule out the presence of ACS. Chest radiography can also assist in the evaluation of a patient with chest pain by identifying a noncardiac source of the pain (e.g., pneumonia, pneumothorax, and aortic dissection) or sequelae of underlying cardiac causes (e.g., pulmonary edema caused by cardiac dysfunction).[14]

ACS is divided into categories according to ECG segments (waves) including ST elevation myocardial infarction (STEMI) and non-ST elevation myocardial infarction (NSTEMI), which includes unstable angina without myocardial infarction.[14] Biomarkers such as cardiac troponin I (cTnI) are critical in the evaluation of acute chest pain to identify ACS and assess tissue damage. Patients should also undergo a 12-lead ECG within 10 minutes of presenting with chest pain. If the results of the ECG reveal an ST segment elevation and the cTnI is increased, then it is recommended that percutaneous coronary intervention or fibrinolytic procedures be performed.[15] Patients without ST elevation (NSTEMI) and a negative cTnI level should be further evaluated by a treadmill test, cardiac perfusion study, or a stress echocardiography as an inpatient or within 72 hours if patient is discharged.[14]

The fourth universal definition of MI was published in 2018 outlining the consensus reached by The Executive Group on behalf of the Joint European Society of Cardiology (ESC)/American College of Cardiology (ACC)/American Heart Association (AHA)/World Heart Federation (WHF) Task Force (**Table 20.2**).[16]

These guidelines emphasized distinguishing myocardial injury from myocardial infarction. Myocardial injury has occurred if there are one or more cardiac troponin (cTn) values above the 99th percentile of upper reference limit (URL). It is termed *acute* if there is a rise and fall of cTn values. The term *myocardial infarction* (types 1, 2, or 3) should be used when the myocardial injury is acute and cTn values demonstrate a rise and/or fall above the 99th percentile of upper reference limit (URL) *and* one of the following:

- Type 1 designation includes:
  1. Ischemic symptoms
  2. Development of pathologic Q waves on ECG
  3. New ECG changes indicative of ischemia (ST-segment changes)
  4. Evidence of atherothrombosis by imaging or autopsy
  5. Imaging evidence of damage to viable myocardium
- Type 2 designation include the criteria of Type 1; however, imbalance of oxygen supply and demand is unrelated to cardiac atherothrombosis
- Type 3 criteria includes sudden, unexplained cardiac death before blood samples can be obtained

CASE STUDY 20.1, PART 2

Remember Shaniece, the 15-month-old with a heart murmur since birth. She has been on digitalis therapy initiated by the referring physician. Her radiograph showed a moderately enlarged heart and an enlarged pulmonary artery. Pertinent laboratory data were obtained.

Analyte	Patient Value	Reference Range
Total protein	5.4	(6.0–8.3 g/dL)
Albumin	3.0	(3.5–5.2 g/dL)
Hemoglobin	19.2	(14–18 g/dL)
Hematocrit	59	(40%–54%)

© pixelheadphoto/iStock/Getty Images Plus/ Getty Images

A cardiac catheterization is performed, and a large ventricular septal defect (VSD) is found.

1. How does this congenital defect affect the body's circulation?
2. Why are the red cell measurements increased in this patient?
3. What treatment will be suggested for this patient?
4. What is this patient's prognosis?

or before biomarkers can appear in the blood, and accompanied by evidence of symptoms, ECG, coronary angiogram, or autopsy.

- Myocardial injury with elevated troponins caused by coronary procedures is classified as either Type 4, 4a, 4b, or 5 based on elevated pre-procedural cTn values, whether stable (< 20% variation) or falling. In addition, patients must meet the criteria for a more than five- or tenfold increase and manifest a change from the baseline value of >20%.
  - Coronary procedure PCI patients are designated Type 4a, while stent placement patients are Type 4b.
  - Type 5 is designated for MI related to the CABG procedure.

Markers of Cardiac Damage

Initial Markers of Cardiac Damage

The first biochemical markers of cardiac damage were discovered in 1954. Today, cardiac biomarkers have become the centerpiece of evaluation and management of patients presenting with chest pain. The underlying principle behind serum biomarkers of cardiac damage relies on the fact that cell death releases intracellular proteins from the myocardium into the circulation. Detection of cardiac proteins provides insight into the occurrence, extent, and timing of MI, all of which are critical for proper medical management. Characteristics of a useful cardiac marker are (1) it is released rapidly from the heart into the circulation and provides specific and sensitive information about myocardial injury, (2) it should persist in circulation for several days to be detectable when patients present late, and (3) the assays should be designed to detect the biomarker at a low concentration.[17]

The first cardiac markers to be used extensively in clinical practice were enzymes, aspartate transaminase (AST) and lactate dehydrogenase (LD). However, both AST and LD were found to be nonspecific to cardiac myocytes, as they are elevated in numerous other conditions. One use for LD may be in the evaluation of acute MI versus subacute MI in patients with elevated troponin levels and normal levels of another marker, the enzyme creatine kinase (CK), particularly CK-MB (an isoenzyme found in the heart).[18]

Creatine kinase (CK) is found in nearly all cells in the body. In patients with AMI, CK levels exceed the reference range within 6 to 8 hours, reach peak levels by 24 hours, and return back within the reference range after 3 to 4 days (**Figure 20.3**).[19] A CK plasma concentration greater than two times normal was shown to correlate with MI,[19] and an elevation in serum CK, AST, and LD, was used as the primary means of enzymatic detection of MI for many years. The cytoplasmic CK isoenzymes are dimers composed of combinations of M and B subunits (where

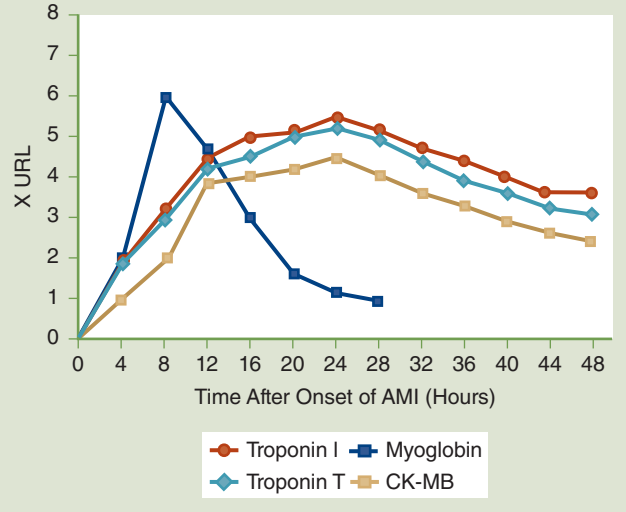


Figure 20.3 Temporal elevations of important markers of myocardial damage post-MI. URL, upper reference limit.

Data from French JK, White HD. Clinical implications of the new definition of myocardial infarction. *Heart*. 2004;90(1):99–106.

M is for muscle and B is for brain). The isoenzymes of total CK, CK-2 (CK-MB) and CK-3 (CK-MM) can be found in striated muscle from the heart and skeletal muscle, respectively. Elevation of CK-MB became an indicator of MI, since CK-MB is 10% to 20% of the total CK activity in myocardium. Because of its higher specificity and its rapid elevation post-MI, CK-MB was long considered the most reliable serum marker of MI. Total CK and CK-MB can be elevated in chronic muscle diseases, end-stage renal disease, and in the healthy population due to extreme exercise. However, in these conditions, cTn levels will be relatively normal if the myocardium is not damaged.[17]

As discussed, the use of enzymes (AST, LD, and CK) was the cornerstone of post-MI management for decades. Each had its individual applications: CK for the early presentation, LD for the late presentation, and AST for the intermediate presentation.[20] But these enzymes' lack of specificity to the myocardium and to myocardial injury fueled the search for better markers. Extensive research led to the discovery that muscle cells express structural and regulatory proteins, such as troponin, in a pattern that is tissue-specific. The subsequent development of rapid and sensitive techniques for detection of the tissue-specific forms of troponin quickly made troponin a marker of cardiac injury.

Cardiac Troponins

Troponin is a complex of three proteins (**Figure 20.4**) that regulate the calcium-dependent interactions of myosin heads with actin filaments during striated

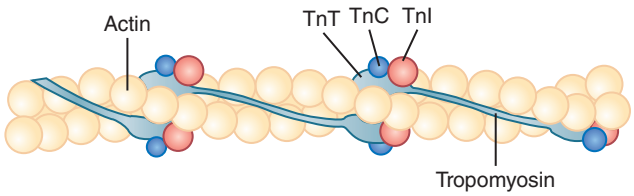


Figure 20.4 Troponin complex consisting of three subunits: TnT, TnC, and TnI.

muscle contraction. **Troponin T** (TnT) binds to tropomyosin, **troponin I** (TnI) inhibits the binding of actin and myosin, and **troponin C** (TnC) binds to calcium to reverse the inhibitory activity of troponin I (TnI).[21] The troponin complex is responsible for transmitting the calcium signal that triggers muscle contraction. Several properties of troponin made it attractive as a marker of myocardial damage.

In contrast to other cardiac markers, troponins were found to have tissue-specific isoforms in slow skeletal muscle, fast skeletal muscle, and cardiac muscle that could be used to detect damage to the heart. Although the same isoform of cTnC is expressed in slow-twitch (type 2) and cardiac muscle, unique isoforms of cTnI and cTnT are expressed in fast-twitch (type 1), slow-twitch, and cardiac muscle.[21] Each cTnI, cTnT, and cTnC isoform is encoded by a separate gene and expressed in a muscle-type–specific manner.[22] This allows for highly specific detection of troponin of cardiac origin. Further increasing the sensitivity of troponin detection was the demonstration that cardiac troponins are very tightly complexed to the contractile apparatus, such that circulating levels of cardiac troponins are normally extremely low.[21] Early studies demonstrated that normal levels of circulating cTnI are below 10 ng/mL, but patients with acute MI can have serum cTnI levels well over 100 ng/mL, demonstrating the specificity of cardiac troponins as indicators of myocardial cell damage.

The development of sensitive detection methods was the next step toward integrating cardiac troponins into the diagnostic workup of MI. Because the skeletal and cardiac isoforms of TnT and TnI contain differing amino acid sequences, monoclonal antibodies can be raised against cardiac-specific epitopes. This allowed for the development of sensitive immunoassays for cardiac troponins that have a far greater specificity than other markers. For example, detection of cTnI was specific for cardiac injury even in acute muscle disease, chronic muscle disease, chronic renal failure, and following intense exercise, such as running a marathon.

Additionally, measurement of plasma troponin levels allows for the detection of myocardial cell injury in syndromes, such as acute ischemic syndrome or **myocarditis**, that is undetectable using other markers.[23]

Finally, troponins offer additional advantages due to the timing of their release. CK-MB levels elevate rapidly post-MI (4 to 6 hours) but return to baseline after 2 to 4 days; CK-MB can thus be used only within a short window of time after a suspected MI. Conversely, LD levels remain elevated for up to 1 week but are not detectable until 24 to 48 hours post-MI (Figure 20.3). Cardiac troponins, on the other hand, are detectable in the plasma at 3 to 12 hours after myocardial injury, peaking at 12 to 24 hours and remaining elevated for more than 1 week: 8 to 21 days for TnT and 7 to 14 days for TnI (Figure 20.3).[19] Cardiac troponins thus offer the widest window for detection post-MI and with the highest sensitivity and specificity.

The characteristics of specific, sensitive, and rapid detection propelled cardiac troponins to the center of MI diagnosis. Cardiac troponins cTnI and cTnT are currently the preferred biomarker for myocardial necrosis, as they are recognized to have nearly absolute myocardial tissue specificity as well as high clinical sensitivity, detecting even minor cardiac damage (as utilized in the universal definition of MI).[16] An elevated plasma cardiac troponin level is defined as a measurement greater than the 99th percentile of a normal reference value range, specified as the upper reference limit (URL) (Figure 20.3).

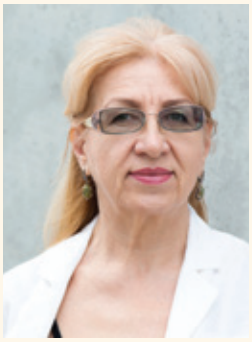
Specificity and Sensitivity of Troponins

Cardiac troponins have become the biomarker of choice to triage patients with chest pain and NSTEMI due to their specificity for cardiac tissue, but due to their relatively slow release, they may not be detected when a patient presents early with chest pain (< 2 hours). Detection of troponins is a key to diagnosing myocardial injury, and the subsequent rise and fall of troponin levels allows the injury to be considered acute. If myocardial ischemia is not present clinically, then elevated cTn levels without rise and fall may be indicative of acute myocardial injury and therefore related to an ongoing, more chronic injury. Clinicians needed a more sensitive test due to the slow release of troponins (< 2 hours). The fifth-generation high sensitivity troponin (hs-cTn) assays were introduced and met strict criteria: the assays must (1) have a coefficient of variation (CV) ≤ 10% at the upper limit (99th percentile) of the normal reference range (URL), and (2) have a low detection limit where troponin is quantitated in at least 50% of healthy individuals.[18] The hs-cTn assays can confirm myocardial injury at lower concentrations and thereby discern small changes in concentration in the normal reference range, i.e., above the limit of detection of the assay but below the URL. These increases (delta values) are associated with a higher probability of subsequent rises above the URL. The emergency department (ED) strives to identify patients who require treatment in the most efficient, cost-effective, and

CASE STUDY 20.2, PART 2

Remember Frieda, who came to the emergency department on Monday. She denied having any chest pain or radiating pain to her extremities. She takes daily medication for hypertension and high cholesterol.

Laboratory Test	Reference Range	4:00 pm	6:00 pm	10:00 pm
CK, Total	54–186 U/L	150 U/L		
CK-MB	0–5 ng/mL	3.5 ng/mL		
Myoglobin	<70 µg/L	pending	45 µg/L	
Troponin T	≤10 ng/L	15 ng/L	24 ng/L	33 ng/L



© A StockStudio/Shutterstock.

1. Do the signs and symptoms suggest Frieda had a heart attack?
2. Interpret Frieda's 4:00 pm laboratory results as normal, increased, or decreased.
3. If Frieda's signs and symptoms indicate an MI, then explain why the CK and CK-MB results are normal.

timely manner. Therefore, rule-out strategies in the ED aimed at accurately excluding myocardial injury as early as possible have led to serial measurements of hs-cTn at earlier times, adding criteria for delta changes, and defining low decision limits at the early measurements.[16]

This increase in sensitivity does create a challenge when chronic unhealthy patients present with chest pain. Single sample rule-out strategy using a very low value (in many cases the LoD of the assay) has high sensitivity for myocardial injury and therefore high negative predictive value to exclude MI. This strategy should not be used in those who present early, i.e., < 2 hours after the onset of chest discomfort. Some studies indicate that the single sample approach provides optimal sensitivity and negative predictive accuracy in patients otherwise at low risk and those with a normal ECG. However, one concern about very short rule-out periods is that the precision of the assays may not permit small differences to be distinguished. These criteria have not, and should not, be applied to patients with hs-cTn elevations.

Procedures and other diseases that may cause myocardial injury and/or MI should be considered when hs-cTn is elevated. Patients with suspected ACS in whom MI is ruled out with normal cardiac biomarker values below the URL may have unstable angina. These patients need further evaluation that can occur in an outpatient setting. Stable concentrations of hs-Tn in the detectable range below the URL are associated with structural heart disease, atherosclerosis risk factors, and a higher risk of future cardiovascular events. Consequently, nondetectable or very low hs-Tn concentrations, well below the URL, identify patients with lower cardiovascular risk.

Considerations in Patients with Kidney Disease

Prognosis in Chronic Kidney Disease. Increases in cTnI/cTnT are associated with increases in short-term cardiac outcomes in patients with chronic kidney disease diagnosed with acute MI.[24] Increased serum troponin levels predict worse long-term cardiovascular outcomes and worse survival in asymptomatic CKD patients in the absence of acute MI.[25]

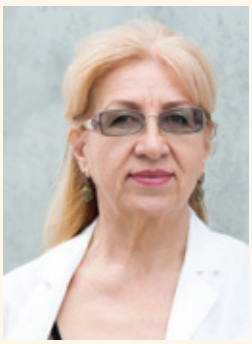
Dialysis Patients. In a recent meta-analysis of 11 studies, elevated cTnT was associated with an increase in all-cause mortality (95% confidence interval [CI] 2.4 to 43).[26] Few studies here used the high-sensitivity assay; cTnT measured with high-sensitive assay has also been reported as an independent predictor of all-cause mortality in dialysis patients.[27,28]

Effect of Hemodialysis. In kidney failure, the need for dialysis is common. Hemodialysis involves filtering the blood to remove the built-up toxins and results in altering the concentration of cardiac enzymes. In studies, dialysis minimally changed CK-MB and did not change cTnI concentrations.[29] CTnI and cTnT levels decreased 27% to 37% following dialysis with a high-flux membrane (but not low-flux membrane).[30] cTnT rose after dialysis in patients with CVD, suggesting the rise in cTnT resulting from cardiac injury.[31] While the cut-offs for the diagnosis for AMI should be the same for dialysis patients, caution should be exercised when interpreting troponin results in patients with chronic kidney disease.

CASE STUDY 20.2, PART 3

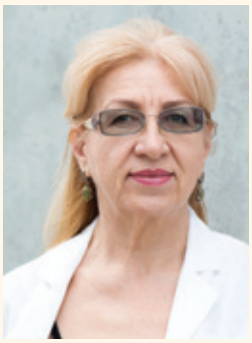
Remember Frieda. Let's look at Frieda's troponin results again.

4. What is the delta calculation between the baseline and 2 hours?
5. What is the delta calculation between the baseline and 6 hours?
6. How should the troponin results be interpreted?



CASE STUDY 20.2, PART 4

7. What is the usefulness of the myoglobin test in Frieda's case management?



© A StockStudio/Shutterstock.

CASE STUDY 20.3, PART 2

Remember Carlos, the 83-year-old man admitted to the hospital when his physician referred him. His symptoms included 3+ pedal edema, jugular vein distention, and heart sound abnormalities. Significant laboratory data obtained on admission were as follows:

© Fotoluminate LLC/Shutterstock.

Analyte	Patient Value	Reference Range
Urea nitrogen	53	(6–24 mg/dL)
Creatinine	2.2	(0.5–1.4 mg/dL)
Glucose	312	(60–110 mg/dL)
Total CK	134	(54–186 U/L)
CK-MB	4	(0–5 ng/L)
Myoglobin	62	(<70 µg/L)
Troponin T	0.2	(<15 ng/L)

1. Do the symptoms of this patient suggest acute MI?
2. Does the laboratory data correlate with a diagnosis of acute MI? Why or why not?
3. Based on the laboratory data, are there other organ system abnormalities present?
4. What are the indicators of these organ system abnormalities?
5. What laboratory test might indicate congestive heart failure in this patient?

Myoglobin

Myoglobin is an iron- and oxygen-binding protein found exclusively in the muscle and normally absent from the circulation. The myoglobin found in the muscle forms pigments, giving the muscle its characteristic red color. The exact color of the muscle depends on the oxidation state of the iron atom in myoglobin and the amount of oxygen attached to it. Myoglobin is a small protein, which is released quickly when the muscle is damaged (see Figure 20.3 for serum release pattern). It has a very short half-life of 9 minutes. Because myoglobin is released so quickly, it has been proposed as an adjunct marker for troponin or CK-MB in the early diagnosis of MI.[32] There are limitations to using myoglobin as a marker of cardiac damage. It is not specific for the heart, being elevated with any cause of skeletal muscle damage, and is cleared in an irregular pattern.[33] Recent studies have also shown that high-sensitivity troponin assays can detect elevated troponin levels prior to elevations in myoglobin,[34] making myoglobin's use less attractive.

Cardiac Injury

The damage that occurs in the context of MI is due to cardiac muscle cell (myocyte) death. A spectrum of cell death occurs in myocytes, ranging from apoptosis to necrosis to a mixture between the two. Apoptotic cell death is characterized by decreasing cell size, membrane blebbing, nuclear aggregation of chromosomal DNA, and purposeful degradation of DNA. The resulting apoptotic bodies formed from the process of apoptosis are recognized by macrophages, which then remove the debris in an effort to minimize the immune response.[35] Apoptosis requires energy, so as ischemia progresses, the apoptotic process stalls, and necrosis occurs if energy supply fails to meet the energy demand, which occurs rapidly in the heart. Thus, a rapid clinical response to myocardial ischemia is essential to limiting the extent of necrosis that occurs and is a central tenet of treating MI. The faster cardiac interventions are performed to restore perfusion (along with oxygen) to the heart, the greater the reduction in infarct size. The strategies for reperfusion include percutaneous interventions using

intravascular balloons and stents to clear coronary artery blockages and prevent their recurrence, coronary artery bypass grafting (CABG or "bypass surgery"), and the use of chemical thrombolytics, such as tissue plasminogen activator (tPA or tenecteplase).

A number of other causes of myocardial injury result in increased biomarker release. Like MI, these disease processes also have myocyte apoptosis and necrosis as part of their pathophysiology, which explains why these biomarkers may be elevated in these patients. Cell death, in the form of apoptosis and/or necrosis, is a key feature of cardiomyopathies due to genetic causes and/or stress-related causes and contributes to the worsening function that occurs with the disease (**Figure 20.5**).[36] Cell death also occurs in decompensated heart failure and volume overload.[37] In the cardiac hypertrophy "pre–heart failure" state, angiotensin II, catecholamine production, and cytokines release play a role in inducing cardiomyocyte apoptosis (Figure 20.5),[38] which can lead to low-level release of cardiac biomarkers. Cardiomyocyte cell death is also induced in myocarditis, whereby viruses and bacteria induce an autoimmune response, which recognizes cardiac-specific proteins as foreign and induces cardiac cells to die.[39] Similarly, bacterial toxins in sepsis can cause apoptosis and myocardial depression. Certain drugs are also capable of causing apoptosis and myocardial dysfunction, including doxorubicin and cyclophosphamide (Figure 20.5),[40,41] as well as alcohol, cocaine, and methamphetamines.

Cardiac Biomarkers in Heart Failure

Heart failure is a pathological state in which the heart fails to adequately supply the metabolic needs of the body, typically due to a decrease in pumping function. The clinical manifestations of heart failure result largely from the retention of fluid, which is one of the body's maladaptive responses to decreased cardiac output. The most common symptoms of heart failure are shortness of breath, fatigue, and lower extremity edema.

The ACC/AHA/Heart Failure Society of America and the European Society of Cardiology have recommended specific laboratory and clinical tests for patients with suspected heart failure (see **Table 20.3**).[42–44] Many of these tests may be familiar as routine laboratory tests; others are more specific to the atherosclerotic process and compensatory cardiac function mechanisms.

The ACC/AHA recommendations also suggest determining a cTnI or cTnT if the clinical presentation is suggestive of an ACS; elevated troponin levels may also indicate the severity of the heart failure (discussed in the next section).[42–44] A chest x-ray, 2D echocardiographic analysis, and Doppler flow studies to identify ventricular and/or valvular abnormalities

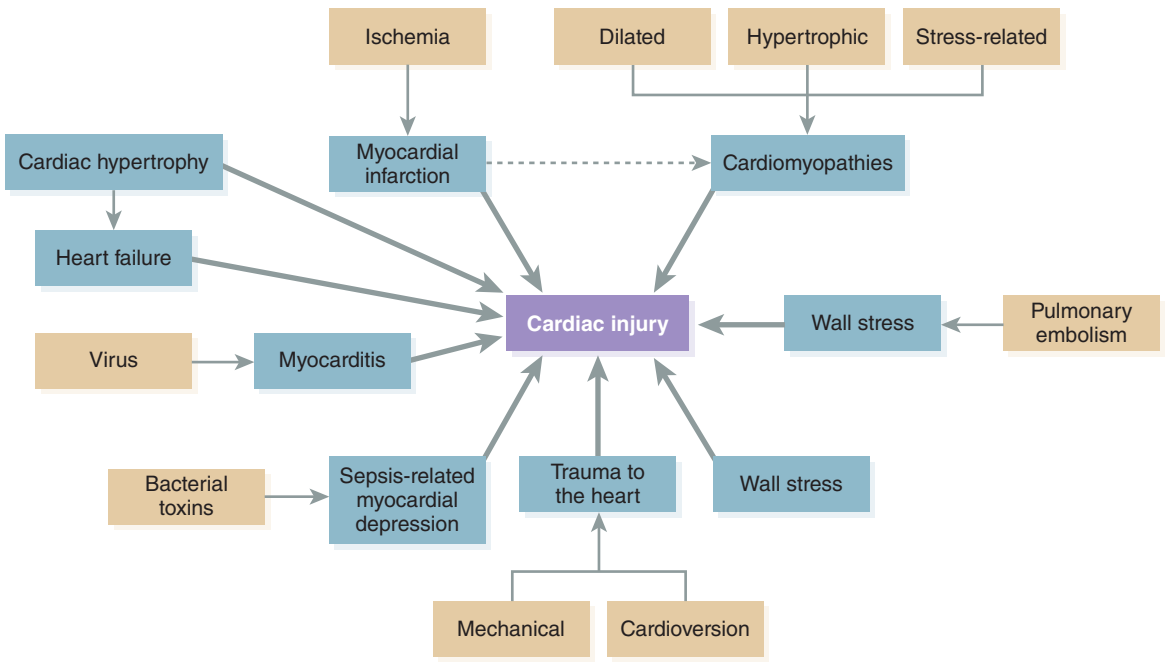


Figure 20.5 Common cardiac insults that result in cardiac injury and elevated cardiac biomarkers.

Data from: Asiri YA. Probucol attenuates cyclophosphamide-induced oxidative apoptosis, p53 and Bax signal expression in rat cardiac tissues. *Oxid Med Cell Longev.* 2010;3(5):308–316; Cooper LT. Myocarditis. *N Engl J Med.* 2009;360(15):1526–1538; Lai HC, Yeh YC, Ting CT, et al. Doxycycline suppresses doxorubicin-induced oxidative stress and cellular apoptosis in mouse hearts. *Eur J Pharmacol.* 2010;644(1–3):176–187; and Yaoita H, Maruyama Y. Intervention for apoptosis in cardiomyopathy. *Heart Fail Rev.* 2008;13(2):181–191.

Table 20.3 ACC/AHA/ESC Recommended Laboratory and Clinical Tests for Patients with Suspected Heart Failure

Complete blood count. A complete blood count can help determine if anemia or infection may be the cause of heart failure.
Serum electrolyte levels. An imbalance of electrolytes can cause fluid retention, and it may play a role in the severity of heart failure.
BUN/creatinine. Determination of BUN/creatinine can determine if damage to the kidneys, due to hypoperfusion, may be occurring secondary to heart failure.
Fasting glucose. Increased glucose levels in both diabetic patients and nondiabetic patients put them at an increased risk for heart failure.[45]
Liver function tests. Elevated liver enzymes (AST, ALT, and LD) may indicate if liver function is affected due to heart failure; congestion of the liver can occur secondary to the inefficiency with which the heart is able to move the venous circulation.
BNP/NT-proBNP levels. Both BNP and its precursor NT-proBNP are elevated after being released in the heart secondary to the stretch induced by heart failure. Their levels correlate closely with the severity of heart failure and can be used to differentiate cardiac causes of shortness of breath from primary lung disease.
ECG. A 12-lead ECG can detect ischemia, infarction, and arrhythmias as a cause for heart failure.
Urinalysis. Recent studies have described a relationship between the amount of protein excreted in urine and cardiovascular risk.[46] The specific relationship between urine protein and CV risk may be due to the common manifestation of vascular endothelial cell dysfunction.[47]
Lipid profile. The lipid profile, including total cholesterol, LDL, HDL, and triglyceride determination, should be performed to determine the risk of CHD.
TSH. Thyroid hormone function tests, in particular TSH levels, should be measured as both hyperthyroidism and hypothyroidism can be a primary cause of heart failure or can contribute to its severity.[42]

BUN, blood urea nitrogen; AST, aspartate aminotransferase; ALT, alanine aminotransferase; LD, lactate dehydrogenase; BNP, B-type natriuretic peptide; NT-proBNP, N-terminal pro-B–type natriuretic peptide; ECG, electrocardiogram; CV, cardiovascular; LDL, low-density lipoprotein; HDL, high-density lipoprotein; TSH, thyroid-stimulating hormone.

are also recommended in evaluating patients with suspected heart failure.[48] As CHD is the most common cause of heart failure, some evaluation for ischemic burden is indicated. To this end, coronary angiography or exercise stress testing may be used depending on the clinical scenario. Pulmonary function testing is generally not helpful in the diagnosis of heart failure as cardiogenic pulmonary edema and intrinsic lung disease can result in similar patterns of abnormalities.[44]

Diagnosis and Risk Stratification of Heart Failure

Though shortness of breath is the most common presentation of heart failure, it is a relatively nonspecific symptom. Measurement of circulating B-type natriuretic peptide (BNP), or its precursor N-terminal pro-B–type natriuretic peptide (NT-proBNP), can be particularly helpful in distinguishing cardiac from noncardiac causes of dyspnea and is widely used in EDs and other clinical settings.[49]

Natriuretic peptides are secreted from the heart in response to increased pressure and volume load. They play an important role in reducing intravascular volume by promoting natriuresis, diuresis, and vasodilation and inhibiting sympathetic nervous system signaling.[50] The NT-proBNP is released by myocardial cells in response to increased volume, increased pressure, and cardiac hypertrophy; this precursor is cleaved by the protease enzyme furin into the active BNP and the inactive NT-proBNP. Both NT-proBNP and BNP are elevated in patients with ventricular dysfunction and strongly predict morbidity and mortality in patients with heart failure.[51] Multiple clinical studies have assessed the utility of both NT-proBNP and BNP for ruling out heart failure as a cause of dyspnea (shortness of breath) in the acute clinical setting.[52] Meta-analysis of these studies has found that the pooled estimates of sensitivity and specificity are equivalent for NT-proBNP and BNP. However, the optimum cutoff value for each peptide remains difficult to determine across all populations.[52]

Vasodilation and natriuresis are beneficial in the setting of heart failure, and thus, BNP was pursued as a target for drug development. Scios (later Johnson & Johnson) developed recombinant BNP as nesiritide. It is the first in a new class of therapies designed to treat heart failure, acting as a neurohormonal suppressor just as endogenous BNP.[53] The importance of mentioning that recombinant BNP is used therapeutically is that it can be picked up by laboratory tests for BNP, but not NT-proBNP. Therefore, there are clinical situations in which NT-proBNP may be the more appropriate test to use to follow heart failure patients.

Cardiac Troponins

While used primarily for the diagnosis of myocardial injury caused by ischemia, elevations of cTnT and cTnI levels were recognized in heart failure patients more than a decade ago.[54] The precise reason for elevated troponins in the non-ACS setting of heart failure is not clear. Ongoing cell death including apoptosis and necrosis may be one of the reasons due to increased myocardial wall stress,[55] resulting in subendocardial ischemia as well as increased myocardial oxygen demand. Diminished cardiac perfusion and oxygen delivery to the heart itself and impaired renal clearance of troponins may also contribute.[55]

While the detection of troponins in heart failure is not diagnostic, it has been reported to add prognostic value. In acute heart failure patients without ACS, elevations in cTnI occur more often than cTnT elevations, although increases in either were related to increased mortality.[55] Other studies have described a 2.6-fold increased risk of in-hospital mortality for heart failure patients with elevated troponin levels at the time of admission.[56] Elevated cardiac troponin has also been associated with lower systolic blood pressure and lower left ventricular ejection fraction at the time of admission,[56] both of which are markers for worse outcome. In outpatients with more severe chronic heart failure (New York Heart Failure Class III or IV), the presence of elevated cardiac troponin was also associated with lower ejection fraction and deteriorating clinical course.[57] Troponin levels were one of the strongest predictors of mortality, particularly if used in conjunction with BNP levels.[58] In general, concomitant elevations in multiple markers (cardiac troponin, high-sensitivity **C-reactive protein** [hs-CRP], along with NT-proBNP) are associated with escalating risks of adverse events.[59] Few studies, however, have investigated how troponin levels may be helpful in the initial diagnosis of heart failure, so their current role in heart failure patients is limited to risk stratification.

Markers of Coronary Heart Disease Risk and Plaque Instability

Myeloperoxidase (MPO) and C-reactive protein (CRP) may indicate the presence of inflammation to predict mortality and provide independent prognostic information apart from clinical risk factors and high-sensitive cTnT.[60] MPO is released when neutrophils convene in blood vessels. It is elevated when patients present with ACS and chronic CAD. CRP is a marker of inflammation that was originally used as a marker of infection. With the development of hs-CRP, measurements above healthy values but lower than in infection are used as a prognostic marker of atherosclerotic processes. Lipoprotein a [Lp(a)] contains a small portion of total cholesterol; however, it has been demonstrated to be pro-atherogenic. Lp(a) has been utilized to predict premature cardiovascular disease. It is difficult to standardize the Lp(a) assay, but generally a level of >30 mg/dL has been used as a cutoff. The exact mechanisms leading to plaque instability, plaque fissure, and subsequent rupture are not fully understood, but all are expressed in atherosclerotic plaques and contribute to vessel inflammation.[61]

C-Reactive Protein

Inflammation plays an important role in the development and progression of atherosclerosis and CHD. C-reactive protein (CRP) is an acute marker of inflammation that is currently used clinically in the evaluation of CVD risk. CRP is a pentameric protein consisting of five identical subunits that bind to specific ligands, such as LDL cholesterol, in a calcium-dependent manner. CRP is normally present in human plasma at levels less than 10 mg/L, but its rapid synthesis in the liver after stimulus from a variety of inflammatory cytokines may increase plasma levels by 1000-fold, thus serving as a sensitive biomarker of systemic inflammation.[62] Because atherosclerosis and CHD derive largely from an inflammatory etiology, CRP has long been targeted as a biomarker for CHD, but it has recently gained widest acceptance as a marker of CHD risk.

CRP was first described when physicians studying the serum from patients with pneumonia observed

high seroreactivity with pneumococcal bacterial extracts. After separating the extracts into discrete fractions, only one fraction—arbitrarily designated fraction C—was found to react heavily with serum from acutely ill patients. This fraction contained "nonprotein material" that appeared "to be a carbohydrate common to the *Pneumococcus* species," which was later called C polysaccharide. Only serum from acutely ill patients reacted with the pneumococcal C polysaccharide, and this reactivity disappeared after resolution of their illness. In the early 1940s, a protein requiring calcium ions was identified to be present exclusively in the serum of acutely ill patients. This protein, called "reactive protein," was responsible for reactivity with C polysaccharide.

Although first observed in patients afflicted with pneumococcal pneumonia, it became clear that serum CRP was present in many more pathological conditions, and elevated CRP became associated with systemic inflammation. As a biomarker, CRP originally gained widespread use as an index of acute rheumatic fever. Its association with heart disease was first demonstrated in 1947 when the serum of patients with **congestive heart failure** was found to contain detectable CRP. In the 1950s, the discovery of elevated CRP after MI supported the hypothesis that MI is associated with systemic inflammation, as did the subsequent finding of elevated CRP in patients with CHD. Despite these findings, interest in CRP as a marker of CVD did not develop until the 1980s, when it was shown that CRP levels correlated remarkably well not only with serum CK-MB post-MI but also with the symptoms of cardiac disease, such as chest pain, unstable angina, and chronic atherothrombotic disease. All of these findings helped support the long-held hypothesis that vascular injury and inflammation play important roles in CVD, but CRP failed to provide significant benefit over the other clinically used markers of CVD or MI.

In the 1990s, extensive analysis of epidemiological data revealed that CRP could be applied clinically in a prognostic manner, rather than as a serum biomarker post-MI. This prognostic value was evaluated through several prospective cohort studies that compared baseline CRP levels in healthy individuals with CRP levels after cardiac events.[63] The most influential data were extracted from the Physicians Health Study (PHS), which found that baseline CRP levels were significantly higher in individuals that eventually experienced MI than those who did not.[64] These results showed that baseline plasma levels of CRP in apparently healthy individuals could help predict the risk of first MI, thus demonstrating a novel and substantial application of CRP as a prognostic marker. CRP data from the PHS also demonstrated that the use of anti-inflammatory medication (aspirin) reduced the risk of vascular events, which supported the hypothesis that chronic inflammation contributes to atherosclerosis and CHD.[63] In the late 1990s, the Cholesterol and Recurrent Events (CARE) trial showed that lipid-lowering drugs, such as statins, reduce CRP levels in a largely LDL-dependent manner, suggesting that CRP evaluation may help determine the efficacy of pharmacologic interventions used to treat CVD.[65]

These studies demonstrated that CRP had immense clinical value as a potential novel biomarker for both cardiovascular risk and cardiovascular therapy management. Importantly, the CRP levels in these and later studies were baseline values that were orders of magnitude less than the CRP levels typically present during acute inflammation. Whereas CRP levels present during acute inflammation are readily detectible using common clinical laboratory methods, the levels of CRP reported in these studies were far below the threshold of detection in most standard clinical assays. Because these studies demonstrated substantial prognostic value of monitoring CRP values at baseline levels, it was thus necessary to develop more sensitive methods of CRP measurement. Specifically, it was necessary to measure CRP levels with extremely high sensitivity (hs), in the range of 0.15 and 10 mg/L. In the early 2000s, such methods for hs-CRP measurement were developed and validated, and the first set of clinical guidelines for the use of hs-CRP as a marker of cardiovascular risk prediction was published by the AHA and the Centers for Disease Control and Prevention in early 2003.[66] These guidelines recommended that hs-CRP be the inflammatory marker of choice in the evaluation of cardiac heart disease risk, and they stated that hs-CRP concentrations of less than 1, 1 to 3, and greater than 3 correspond clinically to low, moderate, and high relative risk of CVD, respectively.[66]

Homocysteine

Homocysteine is a sulfur-containing amino acid formed in plasma from the metabolic demethylation of methionine, which is derived from dietary protein.[67] Plasma homocysteine circulates in four forms: (1) free thiol (1%), (2) disulfide (5% to 10%), (3) mixed disulfide (5% to 10%), and (4) protein-bound (80% to 90%). Total plasma homocysteine refers to

the combined pool of all forms of homocysteine. Normal total plasma homocysteine ranges from 5 to 15 µmol/L, moderate is 16 to 30 µmol/L, intermediate is 31 to 100 µmol/L, and severe hyperhomocystinemia is greater than 100 µmol/L.

Connections between homocysteine and vascular disease were made when a high incidence of vascular anomalies and arterial thromboses was observed in patients with homocystinuria. Five years later, a physician studying vascular abnormalities in homocystinuria found an association between homocysteine and atherosclerosis, concluding that "elevated concentration of homocysteine, homocystine, or a derivaive of homocysteine is the common factor leading to arterial damage." Later studies demonstrated that premature vascular disease is extremely common in patients with homocystinuria, such that advanced atherosclerosis is frequently found in children with homocystinuria, and approximately 50% of patients experience thromboembolic events in their lifetime.[68]

These early studies demonstrated a clear link between extremely high levels of plasma homocysteine (>100 µmol/L [>13.5 mg/L]) and CVD, but mildly elevated homocysteine was later shown to pose a risk of CVD as well. A 1976 study found that "a reduced ability to metabolize homocysteine" may contribute to premature coronary artery disease, and numerous cross-sectional, case–control, and prospective cohort studies further evaluated this relationship. Most, but not all, of these epidemiologic studies indicated that hyperhomocystinemia increases the risk of CVD, but the results varied significantly between studies and study type. Whereas cross-sectional and case–control studies consistently found that hyperhomocystinemia increases the risk of CVD, most prospective studies demonstrated little or no increased risk.[69] One meta-analysis, for example, found that case–control studies estimate approximately an 80% risk of developing CVD due to hyperhomocystinemia, whereas prospective cohort studies estimate only 20% risk.[70] Such variation in clinical and epidemiological data raised questions about what role, if any, homocysteine actually plays in the development of CVD. Many of these questions have been addressed through a wide body of basic research investigating the mechanisms through which homocysteine may contribute to CVD. Potential mechanisms that have been proposed include homocysteine-induced damage to vascular endothelium,[71] accelerated thrombin formation,[72] promotion of lipid peroxidation,[73] vascular smooth muscle proliferation,[74] and attraction of monocytes to the vascular endothelium.[75] Animal models have shown that mild hyperhomocystinemia contributes to atherosclerotic lesion development and to early lipid accumulation in vascular endothelium.[76] And studies in rats have shown that hyperhomocystinemia stimulates the expression of vascular adhesion molecules, such as monocyte chemoattractant protein (MCP-1), vascular cell adhesion molecule 1, and E-selectin, which increases the binding of monocytes to the endothelium.[77] Treatment of cultured human endothelial and smooth muscle cells with homocysteine also induces the expression of MCP-1, as well as expression of IL-8, a T lymphocyte and neutrophil chemoattractant.[78] Homocysteine-induced expression of these chemokines promotes a proinflammatory state that may contribute to general vascular inflammation that drives atherosclerosis.[79] Together, evidence from in vitro and animal models supports the hypothesis that hyperhomocystinemia promotes atherosclerotic lesion development, thereby increasing the risk of CVD.

Collectively, there is a growing body of evidence implicating hyperhomocysteinemia as an independent risk factor of CVD. Clinically, it has been estimated that up to 40% of patients diagnosed with premature coronary artery disease, peripheral vascular disease, or recurrent venous thrombosis exhibit some extent of hyperhomocystinemia.[79] Several recent meta-analyses found that for every 5 µmol/L (0.7 mg/L) increase in serum homocysteine concentration, the risk of ischemic heart disease increased 20% to 30%[80] and that decreasing plasma homocysteine by 3 µmol/L (0.4 mg/L) (through folate supplementation) can reduce the risk of ischemic heart disease by 16%, DVT by 25%, and stroke by 24%.[81] The clinical, epidemiological, and biochemical data support a role for homocysteine in the development of atherosclerosis and CVD, but further research will be necessary to delineate the exact mechanisms through which it exerts this effect.

Markers of Pulmonary Embolism

An embolus is a circulating mass of solid, liquid, or gas, and pulmonary embolism (PE) is an acute and serious condition in which an embolus becomes lodged within the pulmonary arteries, impairing

blood flow through the pulmonary vasculature and increasing right ventricular pressure. The extent of pulmonary vascular occlusion and subsequent symptoms is a function of the size and location of the embolus. Although most pulmonary emboli involve a pulmonary vessel of second, third, or fourth order yielding mild or no clinical symptoms, extremely large emboli can lodge at the bifurcation of the main pulmonary artery to form *saddle emboli* that can rapidly block pulmonary circulation.[82] Saddle emboli and other emboli that occlude over 60% of the pulmonary circulation greatly increase the risk of right heart failure, cardiovascular collapse, and sudden death. The coincidence of DVT and PE is quite high; approximately half of venous thromboemboli will develop into pulmonary emboli,[83] and approximately 95% of pulmonary emboli originate from deep veins of the legs (**Figure 20.6**, **Table 20.4**).[84]

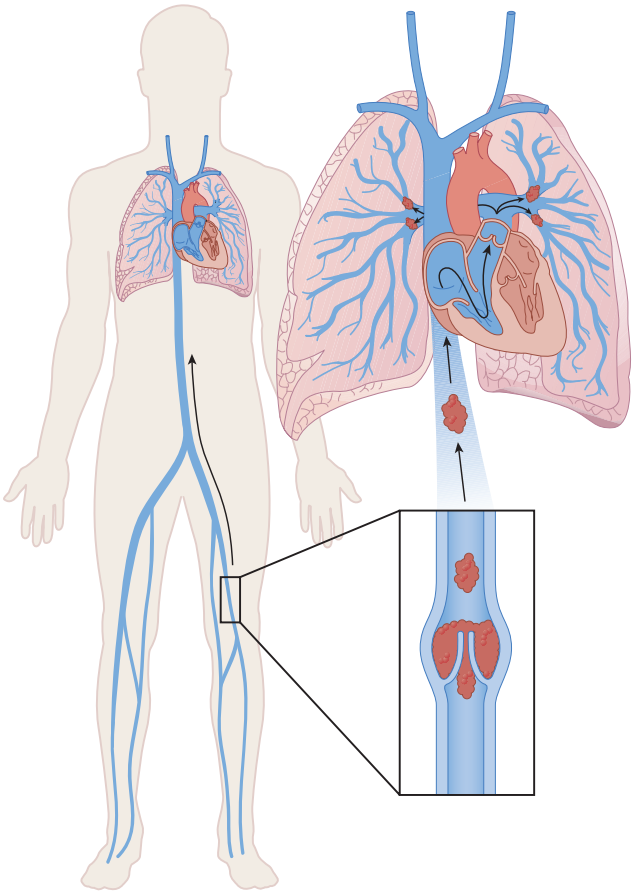


Figure 20.6 Thromboemboli originating from the deep veins of the legs travel through venous circulation through the heart and into the pulmonary vasculature, where they lodge in vessels of decreasing diameter and occlude blood flow.

Based on Douma RA, Kamphuisen PW, Büller HR. Acute pulmonary embolism. Part 1: epidemiology and diagnosis. *Nat Rev Cardiol.* 2010;7(10):585–596.

Table 20.4 Annual Number of Non-Fatal and Fatal, Community, and Hospital-Acquired Symptomatic Venous Thromboembolism Events in the US

Event Type	Community	Hospital	Total
Non-Fatal			
DVT	108,240	268,125	376,365
PE	85,358	151,700	237,058
Total	193,598	419,825	613,423
Fatal			
DVT	649	1,609	2,258
PE	105,902	188,210	294,112
Total	106,550	189,819	296,370

Reproduced with permission from Heit, John A., Cohen, Alexander T., and Anderson, Frederick A. Jr. *Science Direct.* November 16, 2005. Volume 106, Issue 11, pg. 910. Accessed October 15, 2021 at https://www.sciencedirect.com/science/article/pii/S0006497119758005

The incidence of PE increases almost exponentially with age, ranging from approximately 5 per 100,000 in childhood to nearly 600 per 100,000 in persons over 75 years old.[85] Women of reproductive age are at a greater risk for PE because of the associations between venous thromboembolism and pregnancy and the use of oral contraceptives. If left untreated, PE-related mortality can exceed 25%, but adequate treatment in the form of anticoagulation decreases this risk to approximately 5%.[86] Initial therapy after diagnoses of PE involves low molecular weight heparin, unfractionated heparin, or fondaparinux, and long-term treatment includes oral vitamin K antagonists.[87]

Diagnosis of PE is inherently challenging because of the similarity of its symptoms to other more common conditions, such as ACS, and because signs and symptoms are frequently not present.[88] The classical presentation of a patient with PE includes chest pain, dyspnea, tachycardia, tachypnea, and coughing.[85] Unilateral leg swelling and redness may indicate DVT and increases the likelihood of PE. Syncope due to circulatory collapse is present in approximately 15% of patients with a large PE, and crackles or decreased breath sounds are common.[85] Vital signs may reveal tachycardia or mild hypoxia. Evidence of increased venous pressure, such as neck vein distension, or increased right ventricular pressure, such as a loud P2 (pulmonic valve closure sound), increases the diagnostic suspicion for PE.[85] Distinction between PE and

ACS is particularly difficult because of their similar presentation, in particular chest pain, dyspnea, and ECG abnormalities.

Use of D-Dimer Detection in PE

The first step in the diagnostic workup of patients with suspected PE is determining the pretest clinical probability of PE using one of several decision rules.[85] The most widely used set of decision rules is the Wells score, which considers seven clinical variables obtained solely from medical history and physical examination, as well as the physician's judgment on the likelihood of PE versus other diagnoses.[89] When the pretest probability of PE is low or intermediate, it is reasonable to order a **D-dimer** blood test.[85] D-Dimer is a product of plasmin-mediated fibrin degradation that consists of two D-domains from adjacent fibrin monomers that are cross-linked by activated factor XIII. Because D-dimer is derived from cross-linked fibrin, not fibrinogen, the presence of D-dimer in the bloodstream is indicative of current or recent coagulation and subsequent fibrinolysis. D-Dimer thus serves as an indirect marker of coagulation and fibrinolysis.[85] The choice of D-dimer assay is important as the sensitivity of various tests varies greatly. Enzyme-linked fluorescent assay, enzyme-linked immunosorbent assay (ELISA), and latex quantitative assay are highly sensitive quantitative assays for circulating D-dimer levels and are the tests of choice in the workup patients with suspected PE.[90]

D-Dimer levels are abnormal in approximately 90% of patients with PE,[91] and numerous studies have shown that a normal D-dimer results can rule out PE safely in patients with low or intermediate clinical probabilities. However, D-dimer levels are normal in only 40% to 68% of patients without PE. Abnormal levels are often seen in patients with malignancy, recent surgery, renal dysfunction, or increased age.[91] The low specificity of D-dimer testing results in a poor positive predictive value and limited utility in patients with a high clinical probability of PE,[90] for whom CT and ventilation–perfusion scanning are reasonable initial diagnostic tests. However, the high sensitivity of the test translates to a valuable negative predictive value, meaning that D-dimer testing is most useful for excluding PE rather than diagnosing it. One study found that the sensitivity of the D-dimer (by high-sensitivity ELISA) for acute PE was 96.4% and the negative predictive value was 99.6%.[82] Thus further evaluation of PE is not indicated for most patients with normal D-dimer levels.[82]

Value of Assaying Troponin and BNP in Acute PE

A recent meta-analysis was performed to determine the prognostic value of elevated troponin levels in patients with acute PE.[92] Based on publications from January 1998 to November 2006, 122 of 618 patients were found to have elevated troponin levels (19.7%) compared with 51 of 1367 patients with normal troponin levels.[90] An elevated troponin level was associated significantly with a short-term mortality (odds ratio 5.24), resulting from PE (odds ratio 9.44) and with adverse outcome events (odds ratio 7.03).[90] These results were consistent between TnI and TnT and in both prospective and retrospective studies. Therefore, an elevated troponin measurement in patients presenting with PE does appear to have utility in determining their short-term mortality outcome. This information can drive the clinician's clinical management as to whether a more aggressive management approach is necessary, such as the use of thrombolysis.

Similarly, BNP has been used as a predictor of adverse outcome in patients with PE. In a study of 110 consecutive patients with PE, the positive and negative predictive values of BNP levels were determined.[93] The risk of death related to PE if BNP $>$ 21.7 pmol/L was 17%; the negative predictive value for uneventful outcome of a BNP $<$ 21.7 pmol/L was 99%.[93] A larger meta-analysis of 12 studies, including 868 patients with acute PE, determined that elevated BNP levels were significantly associated with an increase in short-term mortality from all causes (odds ratio 6.57) and with death resulting from PE (odds ratio 6.10) or serious adverse events (odds ratio 7.47) with positive predictive values of 14% and negative predictive values of 95%.[93] Together these studies suggest a role for elevated BNP in helping to identify patients with acute PE at high risk for an adverse outcome, in order to drive a more aggressive management approach such as thrombolysis. The high negative predictive value of a normal BNP is also a particularly useful piece of data used to select patients with an uneventful clinical course.

Coronary Heart Disease

CHD is an extremely common condition that causes substantial morbidity and death worldwide. CHD is due in large part to atherosclerosis, which is best

Table 20.5 Comparison of Past and Present Biomarkers for Cardiac Damage and Function with a Summary of Their Current Clinical Utility

Marker	Currently Used to Diagnose ACS/Acute MI	Currently Used to Differentiate Heart Failure from Lung Disease	Currently Used to Diagnose Pulmonary Embolism/Risk Stratification	Used for Cardiovascular Risk Stratification	No Longer Used, Used Less Commonly, or Experimental
LD					X
CK-MB	X				
Myoglobin					X
TnI, TnT	X		X	X	
hs-Tn	X				
hs-CRP				X	
Homocysteine				X	
BNP		X	X		
NT-proBNP		X	X		
D-Dimer			X		

LD, lactate dehydrogenase; CK-MB, creatine kinase MB; TnI, troponin I; TnT, troponin T; hs-Tn, high-sensitive troponin; BNP, B-type natriuretic peptide; NT-proBNP, N-terminal pro-B–type natriuretic peptide.

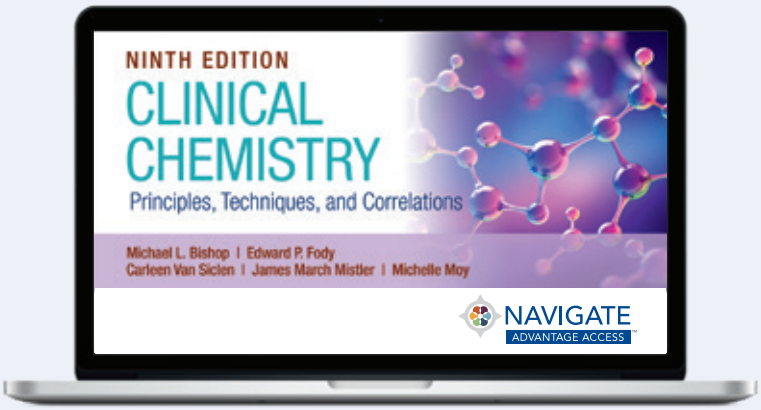
defined as a proinflammatory process in which cells, lipids, and connective tissue cause intimal thickening within large and medium-sized arteries. This process impairs normal blood flow and may progress to occlude the entire vessel diameter. This narrowing of the coronary arteries leads to cardiac ischemia, which may manifest as activity-induced chest pain (stable angina) or ACS. Severe occlusion of coronary vessels causes complete ischemia and subsequent necrosis of surrounding tissue, which is known as MI. The extent of MI and subsequent morbidity can vary tremendously, ranging from undetected infarction with little consequence to sudden death. Because of the physiological and pathological responses that take place within the myocardium immediately after infarction, rapid diagnosis is extremely important. Methods of rapidly diagnosing MI with a high sensitivity and specificity are thus extremely important for the management and potentially the survival of a patient that presents with symptoms of MI.

Plasma biomarkers have become the centerpiece of evaluation and diagnosis of such patients (**Table 20.5**). AST, LD, and CK-MB were widely used biomarkers in the diagnosis of MI but have largely been replaced by high-sensitivity troponin assays. cTnI and cTnT assays have extremely high specificity to cardiac tissue, and detection methods are sensitive enough to pick up even very minor cardiac tissue damage. Current guidelines therefore recommend measurement of cardiac troponins in the circulation as soon as possible after symptoms of MI. Other experimental markers, such as myoglobin, heart-type fatty acid–binding protein, and ischemia-modified albumin, lack the specificity of troponin testing but may be useful when assayed together with troponins.

Plasma biomarkers of cardiac disease risk are also an important factor in the management of patients at risk for CHD, in particular CRP and homocysteine, both of which indicate systemic inflammation and correlate with the elevated risk of CHD and MI. Similarly, the management and diagnosis of heart failure is facilitated by the analysis of circulating biomarkers, most importantly cardiac troponins and BNP.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. World Health Organization. Fact Sheet: The top 10 causes of death. 9 Dec 2020. Accessed at https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death on August 26, 2021.
2. Sanchis-Gomar F, Perez-Quilis C, Leischik R, Lucia A. Epidemiology of coronary heart disease and acute coronary syndrome. *Ann Transl Med*. 2016;4(13):256. doi:10.21037/atm.2016.06.33
3. Glass CK, Witztum JL. Atherosclerosis. the road ahead. *Cell*. 2001;104(4):503–516.
4. Lusis AJ. Atherosclerosis. *Nature*. 2000;407(6801):233–241.
5. Ross R. The pathogenesis of atherosclerosis—an update. *N Engl J Med*. 1986;314(8):488–500.
6. Nakashima Y, Wight TN, Sueishi K. Early atherosclerosis in humans: role of diffuse intimal thickening and extracellular matrix proteoglycans. *Cardiovasc Res*. 2008;79(1):14–23.
7. Skålén K, Gustafsson M, Rydberg EK, et al. Subendothelial retention of atherogenic lipoproteins in early atherosclerosis. *Nature*. 2002;417(6890):750–754.
8. Moore KJ, Tabas I. Macrophages in the pathogenesis of atherosclerosis. *Cell*. 2011;145(3):341–355.
9. Libby P. Changing concepts of atherogenesis. *J Intern Med*. 2000;247(3):349–358.
10. Virmani R, Kolodgie FD, Burke AP, et al. Atherosclerotic plaque progression and vulnerability to rupture: angiogenesis as a source of intraplaque hemorrhage. *Arterioscler Thromb Vasc Biol*. 2005;25(10):2054–2061.
11. Moy E, Barrett M, Coffey R, Hines A, Newman-Toker D. Missed diagnoses of acute myocardial infarction in the emergency department: variation by patient and facility characteristics. *Diagnosis*. 2015;2(1):29–40. https://doi.org/10.1515/dx-2014-0053
12. Runge M, Ohman M, Stouffer G. The history and physical exam. In: Runge M, Stouffer G, Patterson C, eds. *Netter's Cardiology*. 2nd ed. Philadelphia, PA: Saunders Elsevier; 2010.
13. Kudenchuk PJ, Maynard C, Martin JS, et al. Comparison of presentation, treatment, and outcome of acute myocardial infarction in men versus women (the Myocardial Infarction Triage and Intervention Registry). *Am J Cardiol*. 1996;78(1):9–14.
14. Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 AHA/ACC guideline for the management of patients with non-ST-elevation acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2014;130(25):e344–e426.
15. Barstow C, Rice M, McDivitt JD. Acute coronary syndrome: Diagnostic evaluation. *Am Fam Physician*. 2017;95(3):170–177. PMID: 28145667.
16. Thygesen K, Alpert J, White H, et al. Universal definition of myocardial infarction. *J Am Coll Cardiol*. 2018;72(18):2231–2264.
17. Burtis CA, Ashwood ER, Tietz NW. Cardiovascular diseases. In Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Fundamentals of Clinical Chemistry and Molecular Diagnostics*. St. Louis, MO: Elsevier; 2015: 632–650.
18. Alaour B, Liew F, Kaier TE. Cardiac Troponin - diagnostic problems and impact on cardiovascular disease. *Annals of Medicine*. 2018;50(8):655-665. doi:10.1080/07853890.2018.1530450
19. French JK, White HD. Clinical implications of the new definition of myocardial infarction. *Heart*. 2004;90(1):99–106.
20. Goldberg DM. Clinical enzymology: an autobiographical history. *Clin Chim Acta*. 2005;357(2):93–112.
21. Adams JE, Abendschein DR, Jaffe AS. Biochemical markers of myocardial injury. Is MB creatine kinase the choice for the 1990s? *Circulation*. 1993;88(2):750–763.
22. MacGeoch C, Barton PJ, Vallins WJ, et al. The human cardiac troponin I locus: assignment to chromosome 19p13.219q13.2. *Hum Genet*. 1991;88(1):101–104.
23. Coudrey L. The troponins. *Arch Intern Med*. 1998;158(11):1173–1180.
24. Stacy SR, Suarez-Cuervo C, Berger Z, et al. Role of troponin in patients with chronic kidney disease and suspected acute coronary syndrome: a systematic review. *Ann Intern Med*. 2014;161(7):502–512.
25. Khan NA, Hemmelgarn BR, Tonelli M, et al. Prognostic value of troponin T and I among asymptomatic patients

with end-stage renal disease: a meta-analysis. *Circulation.* 2005;112(20):3088–3096.
26. Michos ED, Wilson LM, Yeh HC, et al. Prognostic value of cardiac troponin in patients with chronic kidney disease without suspected acute coronary syndrome: a systematic review and meta-analysis. *Ann Intern Med.* 2014;161(7):491–501.
27. Hassan HC, Howlin K, Jefferys A, et al. High-sensitivity troponin as a predictor of cardiac events and mortality in the stable dialysis population. *Clin Chem.* 2014;60(2): 389–398.
28. McGill D, Talaulikar G, Potter JM, et al. Over time, high sensitivity TnT replaces NT-proBNP as the most powerful predictor of death in patients with dialysis-dependent chronic renal failure. *Clin Chim Acta.* 2010;411(13–14):936–939.
29. Tun A, Khan IA, Win MT, et al. Specificity of cardiac troponin I and creatine kinase-MB isoenzyme in asymptomatic long-term hemodialysis patients and effect of hemodialysis on these cardiac markers. *Cardiology.* 1998;90(4):280–285.
30. Lippi G, Tessitore N, Montagnana M, et al. Influence of sampling time and ultrafiltration coefficient of the dialysis membrane on cardiac troponin I and T. *Arch Pathol Lab Med.* 2008;132(1):72–76.
31. Wayand D, Baum H, Schätzle G, et al. Cardiac troponin T and I in end-stage renal failure. *Clin Chem.* 2000;46(9):1345–1350.
32. McCord J, Nowak RM, McCullough PA, et al. Ninety-minute exclusion of acute myocardial infarction by use of quantitative point-of-care testing of myoglobin and troponin I. *Circulation.* 2001;104(13):1483–1488.
33. Eggers KM, Oldgren J, Nordenskjöld A, et al. Diagnostic value of serial measurement of cardiac markers in patients with chest pain: limited value of adding myoglobin to troponin I for exclusion of myocardial infarction. *Am Heart J.* 2004;148(4):574–581.
34. Ilva T, Eriksson S, Lund J, et al. Improved early risk stratification and diagnosis of myocardial infarction, using a novel troponin I assay concept. *Eur J Clin Invest.* 2005;35(2):112–116.
35. Krysko DV, Denecker G, Festjens N, et al. Macrophages use different internalization mechanisms to clear apoptotic and necrotic cells. *Cell Death Differ.* 2006;13(12):2011–2022.
36. Yaoita H, Maruyama Y. Intervention for apoptosis in cardiomyopathy. *Heart Fail Rev.* 2008;13(2):181–191.
37. Dent MR, Das S, Dhalla NS. Alterations in both death and survival signals for apoptosis in heart failure due to volume overload. *J Mol Cell Cardiol.* 2007;43(6):726–732.
38. Leri A, Claudio PP, Li Q, et al. Stretch-mediated release of angiotensin II induces myocyte apoptosis by activating p53 that enhances the local renin-angiotensin system and decreases the Bcl-2-to-Bax protein ratio in the cell. *J Clin Invest.* 1998;101(7):1326–1342.
39. Cooper LT. Myocarditis. *N Engl J Med.* 2009;360(15): 1526–1538.
40. Lai HC, Yeh YC, Ting CT, et al. Doxycycline suppresses doxorubicin-induced oxidative stress and cellular apoptosis in mouse hearts. *Eur J Pharmacol.* 2010;644(1–3):176–187.
41. Asiri YA. Probucol attenuates cyclophosphamide-induced oxidative apoptosis, p53 and Bax signal expression in rat cardiac tissues. *Oxid Med Cell Longev.* 2010;3(5): 308–316.
42. Hunt SA, Abraham WT, Chin MH, et al. 2009 focused update incorporated into the ACC/AHA 2005 Guidelines for the Diagnosis and Management of Heart Failure in Adults: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines: developed in collaboration with the International Society for Heart and Lung Transplantation. *J Am Coll Cardiol.* 2009;53(15):e1–e90.
43. Lindenfeld J, Albert NM, Boehmer JP, et al. HFSA 2010 Comprehensive Heart Failure Practice Guideline. *J Card Fail.* 2010;16(6):e1–e194.
44. Dickstein K, Cohen-Solal A, Filippatos G, et al. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). *Eur Heart J.* 2008;29(19):2388–2442.
45. Haffner SJ, Cassells H. Hyperglycemia as a cardiovascular risk factor. *Am J Med.* 2003;115(suppl 8A):6S–11S.
46. Buckalew VM, Freedman BI. Effects of race on albuminuria and risk of cardiovascular and kidney disease. *Expert Rev Cardiovasc Ther.* 2011;9(2):245–249.
47. Ochodnicky P, Henning RH, van Dokkum RP, et al. Microalbuminuria and endothelial dysfunction: emerging targets for primary prevention of end-organ damage. *J Cardiovasc Pharmacol.* 2006;47(suppl 2):S151–S162; discussion S172–S156.
48. Temporelli PL, Scapellato F, Eleuteri E, et al. Doppler echocardiography in advanced systolic heart failure: a noninvasive alternative to Swan-Ganz catheter. *Circ Heart Fail.* 2010;3(3):387–394.
49. Januzzi JL, Camargo CA, Anwaruddin S, et al. The N-terminal Pro-BNP investigation of dyspnea in the emergency department (PRIDE) study. *Am J Cardiol.* 2005;95(8):948–954.
50. McFarlane SI, Winer N, Sowers JR. Role of the natriuretic peptide system in cardiorenal protection. *Arch Intern Med.* 2003;163(22):2696–2704.
51. Bibbins-Domingo K, Gupta R, Na B, et al. N-terminal fragment of the prohormone brain-type natriuretic peptide (NT-proBNP), cardiovascular events, and mortality in patients with stable coronary heart disease. *JAMA.* 2007;297(2):169–176.
52. Worster A, Balion CM, Hill SA, et al. Diagnostic accuracy of BNP and NT-proBNP in patients presenting to acute care settings with dyspnea: a systematic review. *Clin Biochem.* 2008;41(4–5):250–259.
53. Fonarow GC. B-type natriuretic peptide: spectrum of application. Nesiritide (recombinant BNP) for heart failure. *Heart Fail Rev.* 2003;8(4):321–325.
54. Güler N, Bilge M, Eryonucu B, et al. Cardiac troponin I levels in patients with left heart failure and cor pulmonale. *Angiology.* 2001;52(5):317–322.
55. Nagarajan V, Tang WH. Biomarkers in advanced heart failure: diagnostic and therapeutic insights. *Congest Heart Fail.* 2011;17(4):169–174.
56. Peacock WF, De Marco T, Fonarow GC, et al. Cardiac troponin and outcome in acute heart failure. *N Engl J Med.* 2008;358(20):2117–2126.
57. La Vecchia L, Mezzena G, Zanolla L, et al. Cardiac troponin I as diagnostic and prognostic marker in severe heart failure. *J Heart Lung Transplant.* 2000;19(7):644–652.
58. Horwich TB, Patel J, MacLellan WR, et al. Cardiac troponin I is associated with impaired hemodynamics, progressive left ventricular dysfunction, and increased mortality rates in advanced heart failure. *Circulation.* 2003;108(7): 833–838.

59. Yin WH, Chen JW, Feng AN, et al. Multimarker approach to risk stratification among patients with advanced chronic heart failure. *Clin Cardiol.* 2007;30(8):397–402.
60. Nicholls SJ, Hazen SL. Myeloperoxidase and cardiovascular disease. *Arterioscler Thromb Vasc Biol.* 2005;25(6):1102–1111.
61. Schaub N, Reichlin T, Meune C, et al. Markers of plaque instability in the early diagnosis and risk stratification of acute myocardial infarction. *Clin Chem.* 2012;58(1): 246–256.
62. Ridker PM. C-reactive protein: eighty years from discovery to emergence as a major risk marker for cardiovascular disease. *Clin Chem.* 2009;55(2):209–215.
63. Ridker PM, Cushman M, Stampfer MJ, et al. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N Engl J Med.* 1997;336(14): 973–979.
64. Ridker PM, Rifai N, Pfeffer MA, et al. Inflammation, pravastatin, and the risk of coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events (CARE) Investigators. *Circulation.* 1998;98(9):839–844.
65. Pearson TA, Mensah GA, Alexander RW, et al. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: a statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. *Circulation.* 2003;107(3):499–511.
66. Abraham JM, Cho L. The homocysteine hypothesis: still relevant to the prevention and treatment of cardiovascular disease? *Cleve Clin J Med.* 2010;77(12):911–918.
67. Nygård O, Nordrehaug JE, Refsum H, et al. Plasma homocysteine levels and mortality in patients with coronary artery disease. *N Engl J Med.* 1997;337(4):230–236.
68. Christen WG, Ajani UA, Glynn RJ, et al. Blood levels of homocysteine and increased risks of cardiovascular disease: causal or casual? *Arch Intern Med.* 2000;160(4):422–434.
69. Splaver A, Lamas GA, Hennekens CH. Homocysteine and cardiovascular disease: biological mechanisms, observational epidemiology, and the need for randomized trials. *Am Heart J.* 2004;148(1):34–40.
70. Blundell G, Jones BG, Rose FA, et al. Homocysteine mediated endothelial cell toxicity and its amelioration. *Atherosclerosis.* 1996;122(2):163–172.
71. Loscalzo J. Homocysteine-mediated thrombosis and angiostasis in vascular pathobiology. *J Clin Invest.* 2009;119(11): 3203–3205.
72. Heinecke JW. Biochemical evidence for a link between elevated levels of homocysteine and lipid peroxidation in vivo. *Curr Atheroscler Rep.* 1999;1(2):87–89.
73. Tsai JC, Perrella MA, Yoshizumi M, et al. Promotion of vascular smooth muscle cell growth by homocysteine: a link to atherosclerosis. *Proc Natl Acad Sci USA.* 1994;91(14): 6369–6373.
74. Sung FL, Slow YL, Wang G, et al, Homocysteine stimulates the expression of monocyte chemoattractant protein-1 in endothelial cells leading to enhanced monocyte chemotaxis. *Mol Cell Biochem.* 2001;216(1–2):121–128.
75. Chen Z, Karaplis AC, Ackerman SL, et al. Mice deficient in methylenetetrahydrofolate reductase exhibit hyperhomocysteinemia and decreased methylation capacity, with neuropathology and aortic lipid deposition. *Hum Mol Genet.* 2001;10(5):433–443.
76. Wang G, Woo CW, Sung FL, et al. Increased monocyte adhesion to aortic endothelium in rats with hyperhomocysteinemia: role of chemokine and adhesion molecules. *Arterioscler Thromb Vasc Biol.* 2002;22(11):1777–1783.
77. Poddar R, Sivasubramanian N, DiBello PM, et al. Homocysteine induces expression and secretion of monocyte chemoattractant protein-1 and interleukin-8 in human aortic endothelial cells: implications for vascular disease. *Circulation.* 2001;103(22):2717–2723.
78. Austin RC, Lentz SR, Werstuck GH. Role of hyperhomocysteinemia in endothelial dysfunction and atherothrombotic disease. *Cell Death Differ.* 2004;11(suppl 1): S56–S64.
79. Humphrey LL, Fu R, Rogers K, et al. Homocysteine level and coronary heart disease incidence: a systematic review and meta-analysis. *Mayo Clin Proc.* 2008;83(11):1203–1212.
80. Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *BMJ.* 2002;325(7374):1202.
81. Goldhaber SZ. Pulmonary embolism. *Lancet.* 2004;363(9417): 1295–1305.
82. Silverstein MD, Heit JA, Mohr DN, et al. Trends in the incidence of deep vein thrombosis and pulmonary embolism: a 25-year population-based study. *Arch Intern Med.* 1998;158(6):585–593.
83. Kumar V, Fausto N, Abbas A. *Robbins and Cotran Pathologic Basis of Disease*. 7th ed. Philadelphia, PA: W.B. Saunders Company; 2004.
84. Douma RA, Kamphuisen PW, Büller HR. Acute pulmonary embolism. Part 1: epidemiology and diagnosis. *Nat Rev Cardiol.* 2010;7(10):585–596.
85. Douketis JD, Kearon C, Bates S, et al. Risk of fatal pulmonary embolism in patients with treated venous thromboembolism. *JAMA.* 1998;279(6):458–462.
86. van Es J, Douma RA, Gerdes VE, et al. Acute pulmonary embolism. Part 2: treatment. *Nat Rev Cardiol.* 2010;7(11): 613–622.
87. Coutance G, Le Page O, Lo T, et al. Prognostic value of brain natriuretic peptide in acute pulmonary embolism. *Crit Care.* 2008;12(4):R109.
88. Wells PS, Anderson DR, Rodger M, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer. *Ann Intern Med.* 2001;135(2):98–107.
89. Agnelli G, Becattini C. Acute pulmonary embolism. *N Engl J Med.* 2010;363(3):266–274.
90. Stein PD, Hull RD, Patel KC, et al. D-Dimer for the exclusion of acute venous thrombosis and pulmonary embolism: a systematic review. *Ann Intern Med.* 2004;140(8):589–602.
91. Becattini C, Vedovati MC, Agnelli G. Prognostic value of troponins in acute pulmonary embolism: a meta-analysis. *Circulation.* 2007;116(4):427–433.
92. ten Wolde M, Tulevski II, Mulder JW, et al. Brain natriuretic peptide as a predictor of adverse outcome in patients with pulmonary embolism. *Circulation.* 2003;107(16):2082–2084.
93. Klok FA, Mos IC, Huisman MV. Brain-type natriuretic peptide levels in the prediction of adverse outcomes in patients with pulmonary embolism: a systematic review and meta-analysis. *Am J Respir Critical Csre Med.* 2008 Aug 15;178(4): 425-30. DOI: 10.1164/rccm.200803-459OC. Epub 2008 Jun 2. PMID: 1855626.

Renal Function

Peter O'Brien

CHAPTER OUTLINE

KEY TERMS

Acute kidney injury (AKI)
Albuminuria
Aldosterone
Arginine vasopressin (AVP)
β_2-Microglobulin (β_2-M)
Chronic kidney disease (CKD)
Countercurrent multiplier system
Creatinine clearance
Cystatin C
Diabetes mellitus
Erythropoietin
Estimated glomerular filtration rate (eGFR)
Glomerular filtration rate (GFR)
Glomerulonephritis
Glomerulus
Hemodialysis
Hemofiltration
Loop of Henle
Myoglobin
Nephrons
Nephrotic syndrome
Prostaglandin
Renin
Rhabdomyolysis
Tubular reabsorption
Tubular secretion
Tubule
Vitamin D

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Diagram and identify the major components of the nephron.
- Describe the physiologic role of each part of the nephron: glomerulus, proximal tubule, loop of Henle, distal tubule, and collecting duct.
- Describe the mechanisms by which the kidney maintains fluid and electrolyte balance in conjunction with hormones.
- State the usefullness of glomerular filtration rate and estimated glomerular filtration rate.

- Relate the clinical significance of total urine proteins, albuminuria, myoglobin clearance, serum β_2-microglobulin, and cystatin C.
- Explain the principle and clinical significance of each analyte in a routine urinalysis.
- Correlate laboratory test results with glomerulus and tubules disorders.
- Distinguish between acute kidney injury and chronic kidney disease.
- Evaluate chronic renal failure treatment options, including dialysis and transplantation.
- Apply knowledge of renal physiology and disease to answer case study questions throughout the chapter.

The kidneys are vital organs that perform a variety of important functions (**Table 21.1**). The most prominent functions are removal of unwanted substances from plasma (both waste and surplus); homeostasis (maintenance of equilibrium) of the body's water, electrolyte, and acid–base status; and participation in hormonal regulation. In the clinical laboratory, kidney function tests are used in the assessment of renal disease, water balance, and acid–base disorders and in situations of trauma, head injury, surgery, and infectious disease. This chapter focuses on renal anatomy and physiology and the analytic procedures available to diagnose, monitor, and treat kidney dysfunction.

Renal Anatomy

The kidneys are paired, bean-shaped organs located retroperitoneally on either side of the spinal column. Macroscopically, a fibrous capsule of connective tissue encloses each kidney. When dissected longitudinally, two regions can be clearly discerned—an outer region called the cortex and an inner region called the medulla (**Figure 21.1**). The pelvis can also be seen. It is a basin-like cavity at the upper end of the ureter into which newly formed urine passes. The bilateral ureters are thick-walled canals, connecting the kidneys to the urinary bladder. Urine is temporarily stored in the bladder until voided from

CASE STUDY 21.1, PART 1

George, a 52-year-old male with a history of AIDS, hypertension, diabetes mellitus, and alcohol abuse, was found unconscious in his home by his roommate.

© Simon Winnall/The Image Bank/Getty Images.

CASE STUDY 21.2, PART 1

Arthur, a 42-year-old man, presented to the hospital stating he had ingested about 3 cups of antifreeze and a rodenticide product (rat poison) 10 hours ago. Upon arrival to the hospital, the patient was somnolent and afebrile and vitals were within normal limits.

© brusinski/iStock/Getty Images Plus/Getty Images.

CASE STUDY 21.3, PART 1

Geraldine, a 78-year-old female with a history of hypertension, aortic thoracic graft, and esophageal reflux disease, complained of fever (100°F) and weakness. She was brought to the emergency department by her daughter.

© cobalt/iStock/Getty Images Plus/Getty Images.

Table 21.1 Kidney Functions

Urine formation
Fluid and electrolyte balance
Regulation of acid–base balance
Excretion of the waste products of protein metabolism
Excretion of drugs and toxins
Secretion of hormones
Renin
Erythropoietin
1,25-Dihydroxyvitamin D_3
Prostaglandins

the body by way of the urethra. The highlighted section in Figure 21.1 shows the arrangement of **nephrons** in the kidney; nephrons are functional units of the kidney that can only be seen microscopically. Each kidney contains approximately 1 million nephrons. Each nephron is a complex apparatus composed of five basic parts as shown in **Figure 21.2**.

These five parts are:

- The **glomerulus**—a capillary tuft surrounded by the expanded end of a renal **tubule** known as Bowman's capsule. Each glomerulus has an afferent arteriole that carries the blood in and an efferent arteriole carrying the blood out. The efferent arteriole branches into peritubular capillaries that supply the tubule.
- The proximal convoluted tubule—located in the cortex.
- The long loop of Henle—composed of the thin descending limb, which spans the medulla, and the ascending limb, which is located in both the medulla and the cortex, composed of a region that is thin and then thick.
- The distal convoluted tubule—located in the cortex.

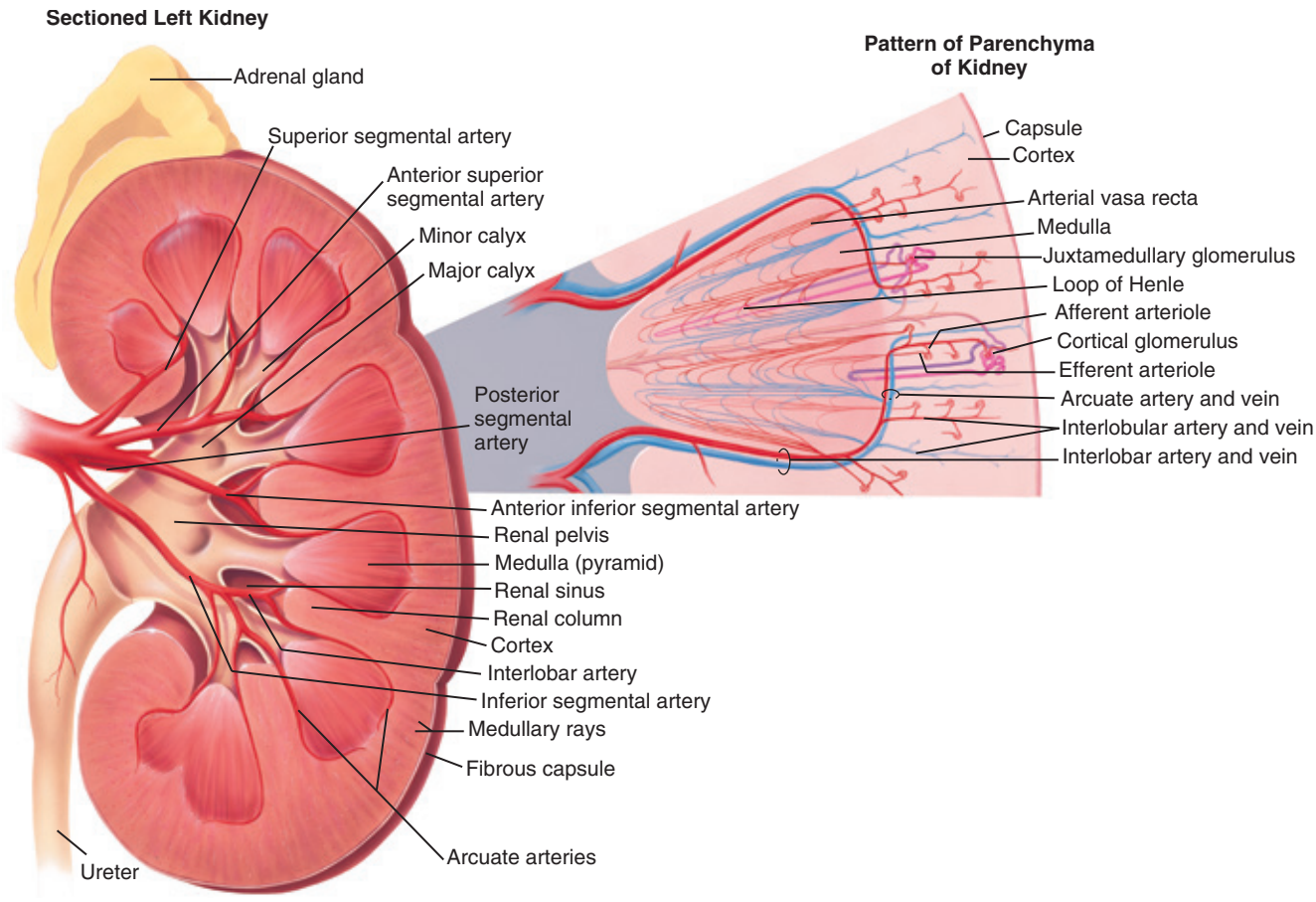


Figure 21.1 Anatomy of the kidney.

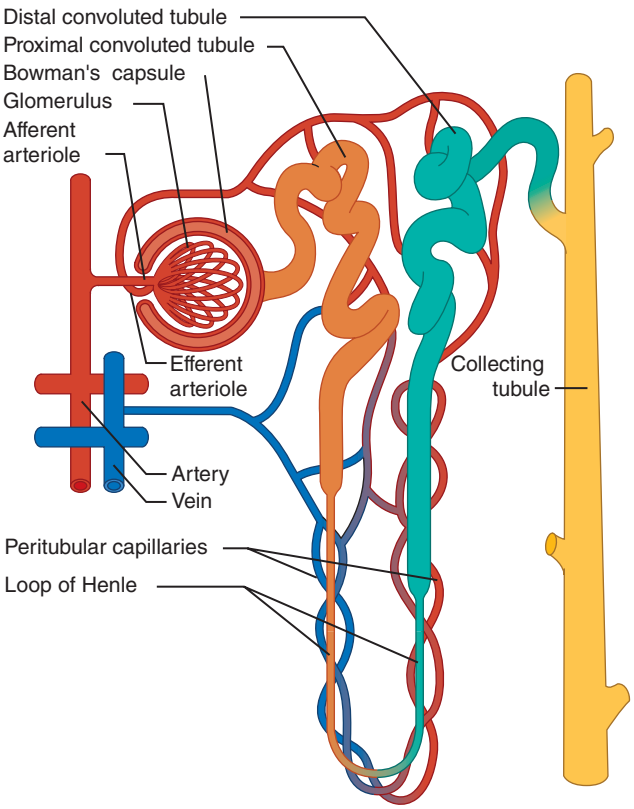


Figure 21.2 Representation of a nephron and its blood supply.

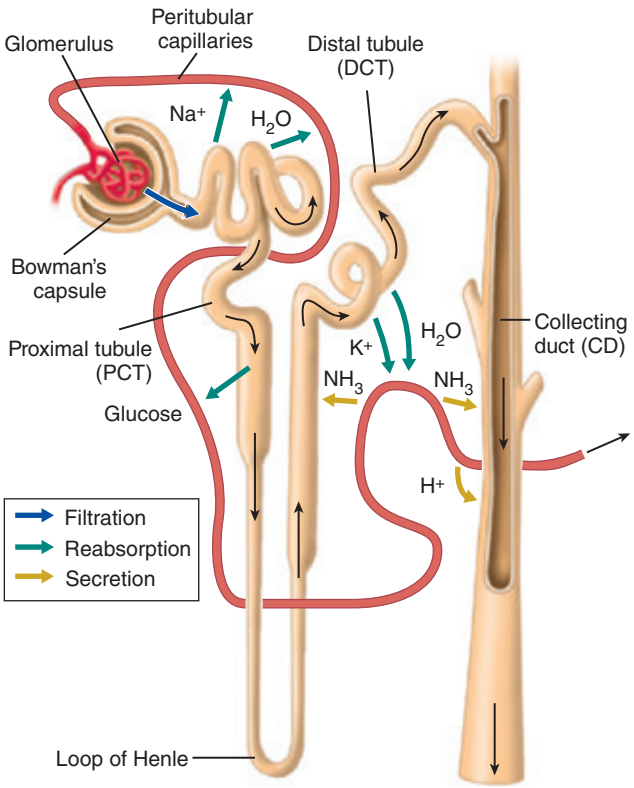


Figure 21.3 Renal processes of filtration, reabsorption, and secretion.

- The collecting duct—formed by two or more distal convoluted tubules as they pass back down through the cortex and the medulla to collect the urine that drains from each nephron. Collecting ducts eventually merge and empty their contents into the renal pelvis.

The following section describes how each part of the nephron normally functions.

Renal Physiology

There are three basic renal processes:

1. Glomerular filtration
2. Tubular reabsorption
3. Tubular secretion

Figure 21.3 illustrates three basic renal processes: filtration, reabsorption, and secretion by the nephron. NH_3 is filtered and secreted, but not reabsorbed; Glucose is filtered and a portion reabsorbed; and Amino Acids are filtered and completely reabsorbed (not shown).[1] The following is a description of how specific substances are regulated in this manner to maintain homeostasis.

Glomerular Filtration

The glomerulus is the first part of the nephron and functions to filter incoming blood. Several factors facilitate filtration. One factor is the unusually high pressure in the glomerular capillaries, which is a result of their position between two arterioles. This sets up a steep pressure difference across the walls. Another factor is the semipermeable glomerular basement membrane, which has a molecular size cutoff value of approximately 66,000 Da, about the molecular size of albumin. This means that water, electrolytes, and small dissolved solutes, such as glucose, amino acids, low-molecular-weight proteins, urea, and creatinine, pass freely through the basement membrane and enter the proximal convoluted tubule. Other blood constituents, such as albumin, many plasma proteins, cellular elements, and protein-bound substances such as lipids and bilirubin, are too large to be filtered. In addition, because the basement membrane is negatively charged, negatively charged molecules, such as proteins, are repelled. Of the 1200–1500 mL of blood that the kidneys receive each minute (approximately one-quarter of the total cardiac output), the glomerulus filters out 125 to 130 mL

CASE STUDY 21.1, PART 2

Remember George. His roommate called 911, and the EMTs transported him to the local trauma center. In the emergency department, he was hypotensive (103/60 mmHg), febrile (temperature 101°F), and unresponsive. Computed tomography scan of the abdomen showed cholecystitis and gallstones. Laboratory tests were ordered, and the results are:

© Simon Winnall/The Image Bank/Getty Images.

Drugs of Abuse	Patient Result (Reference Range)	Urinalysis	Patient Result (Reference Range)
Serum ethanol	84 mg/dL (Negative)	Blood	Positive (Negative)
		WBC	4 High Power Field (HPF) (0–4)
		RBC	2 HPF (0–4)
CK	3308 U/L (24–204)	Blood urea nitrogen (BUN)	71 mg/dL (8–21)
CK-MB	15 ng/mL (0–7.5)	Creatinine	4.1 mg/dL (0.9–1.5)
Troponin T	<0.01 ng/mL (0–0.4)	Alkaline phosphatase	443 U/L (45–122)
pH	7.50 (7.350–7.450)	Aspartate aminotransferase	305 U/L (9–45)
pCO_2	27 mmHg (35-45 mmHg)	Alanine aminotransferase	78 U/L (8–63)
Total CO_2	15 mmol/L (23-29 mmol/L)	Gamma glutamyl transpeptidase	724 U/L (11–50)
		Total bilirubin	2.7 mg/dL (0.2–1.0)
		Direct bilirubin	2.4 mg/dL (0–0.2)

1. What is the significance of the patient's elevated CK?
2. Explain why the physician ordered a CK-MB and troponin level. What can you conclude about the patient's cardiac status?

of an essentially protein-free, cell-free fluid, called *glomerular filtrate*. The volume of blood filtered per minute is the **glomerular filtration rate (GFR)**, and its determination is essential in evaluating renal function, as discussed in the section on Analytic Procedures.

Tubular Function

Proximal Convoluted Tubule

The proximal tubule is the next part of the nephron to receive the now cell-free and essentially protein-free blood. This filtrate contains waste products, which are toxic to the body above a certain concentration, and substances that are valuable to the body. One function of the proximal tubule is to return the bulk of each valuable substance back to the blood circulation. Thus, 75% of the water, sodium, and chloride; 100% of the glucose (up to the renal threshold); almost all the amino acids, vitamins, and proteins; and varying amounts of urea, uric acid, and ions, such as magnesium, calcium, potassium, and bicarbonate, are reabsorbed. Almost all (98% to 100%) of uric acid, a waste product, is actively reabsorbed, only to be secreted at the distal end of the proximal tubule.

When the substances move from the tubular lumen to the peritubular capillary plasma, the process is called **tubular reabsorption**. With the exception of water and chloride ions, the process is active; that is, the tubular epithelial cells use energy to bind and transport the substances across the plasma membrane to the blood. The transport processes that are involved normally have sufficient reserve for efficient reabsorption, but they are saturable. When the concentration of the filtered substance exceeds the capacity of the transport system, the substance is then excreted in the urine. The plasma concentration above which the substance appears in urine is known as the renal threshold, and its determination is useful in assessing both tubular function and nonrenal disease states. A renal threshold does not exist for water because it is always transported passively through diffusion down a concentration gradient. Chloride ions in this instance diffuse in the wake of sodium.

A second function of the proximal tubule is to secrete products of kidney tubular cell metabolism, such as hydrogen ions, and drugs, such as penicillin. The term **tubular secretion** is used in two ways: (1) tubular secretion describes the movement of substances from peritubular capillary plasma to the tubular lumen, and (2) tubular secretion also describes when tubule cells secrete products of their own cellular metabolism into the filtrate in the tubular lumen. Transport across the membrane of the cell is again either active or passive.

Loop of Henle

The osmolality in the medulla in this portion of the nephron increases steadily from the corticomedullary junction inward and facilitates the reabsorption of water, sodium, and chloride. The hyperosmolality that develops in the medulla is continuously maintained by the **Loop of Henle**, a hairpin-like loop between the proximal tubule and the distal convoluted tubule. The opposing flows in the loop—the downward flow in the descending limb and the upward flow in the ascending limb—are termed a *countercurrent flow*. To understand how the hyperosmolality is maintained in the medulla, it is best to look first at what happens in the ascending limb. Sodium and chloride are actively and passively reabsorbed into the medullary interstitial fluid along the entire length of the ascending limb. Because the ascending limb is relatively impermeable to water, little water follows, and the medullary interstitial fluid becomes hyperosmotic compared with the fluid in the ascending limb. The fluid in the ascending limb becomes hypotonic or dilute as sodium and chloride ions are reabsorbed without the loss of water, so the ascending limb is often called the diluting segment. The descending limb, in contrast to the ascending limb, is highly permeable to water and does not reabsorb sodium and chloride. The high osmolality of the surrounding interstitial medulla fluid is the physical force that accelerates the reabsorption of water from the filtrate in the descending limb. Interstitial hyperosmolality is maintained because the ascending limb continues to pump sodium and chloride ions into it. This interaction of water leaving the descending loop and sodium and chloride leaving the ascending loop to maintain a high osmolality within the kidney medulla produces hypo-osmolal urine as it leaves the loop. This process of creating an osmotic gradient and is called the **countercurrent multiplier system**.[2]

Distal Convoluted Tubule

The distal convoluted tubule is much shorter than the proximal tubule, with two or three coils that connect to a collecting duct. The filtrate entering this section of the nephron is close to its final composition. About 95% of the sodium and chloride ions and 90% of water have already been reabsorbed from the original glomerular filtrate. The function of the distal tubule is to effect small adjustments to achieve electrolyte and acid–base homeostasis. These adjustments occur under the hormonal control of both **aldosterone** and **arginine vasopressin (AVP)**, also called antidiuretic hormone (ADH) or just vasopressin. **Figure 21.4** describes the action of these hormones.

Arginine Vasopressin. AVP is a peptide hormone secreted by the posterior pituitary, mainly in response to increased blood osmolality; AVP is also released when blood volume decreases by more than 5% to 10%. Large decreases of blood volume will stimulate AVP secretion even when plasma osmolality is decreased. AVP stimulates water reabsorption. The walls of the distal collecting tubules are normally impermeable to water (like the ascending loop of Henle), but they become permeable to water in AVP. Water diffuses passively from the lumen of the tubules, resulting in more concentrated urine and decreased plasma osmolality.

Aldosterone. This hormone is produced by the adrenal cortex under the influence of the **renin**–angiotensin mechanism. Its secretion is triggered by decreased blood flow or blood pressure in the afferent renal arteriole and by decreased plasma sodium. Aldosterone stimulates sodium reabsorption in the distal tubules and potassium and hydrogen ion secretion. Hydrogen ion secretion is linked to bicarbonate regeneration and ammonia secretion, which also occur here. In addition to these ions, small amounts of chloride ions are reabsorbed.

Collecting Duct. The collecting ducts are the final site for either concentrating or diluting urine. The hormones AVP and aldosterone act on this segment of the nephron to control reabsorption of water and sodium. Chloride and urea are also reabsorbed here. Urea plays an important role in maintaining the hyperosmolality of the renal medulla. Because the collecting ducts in the medulla are highly permeable to urea, urea diffuses down its concentration gradient

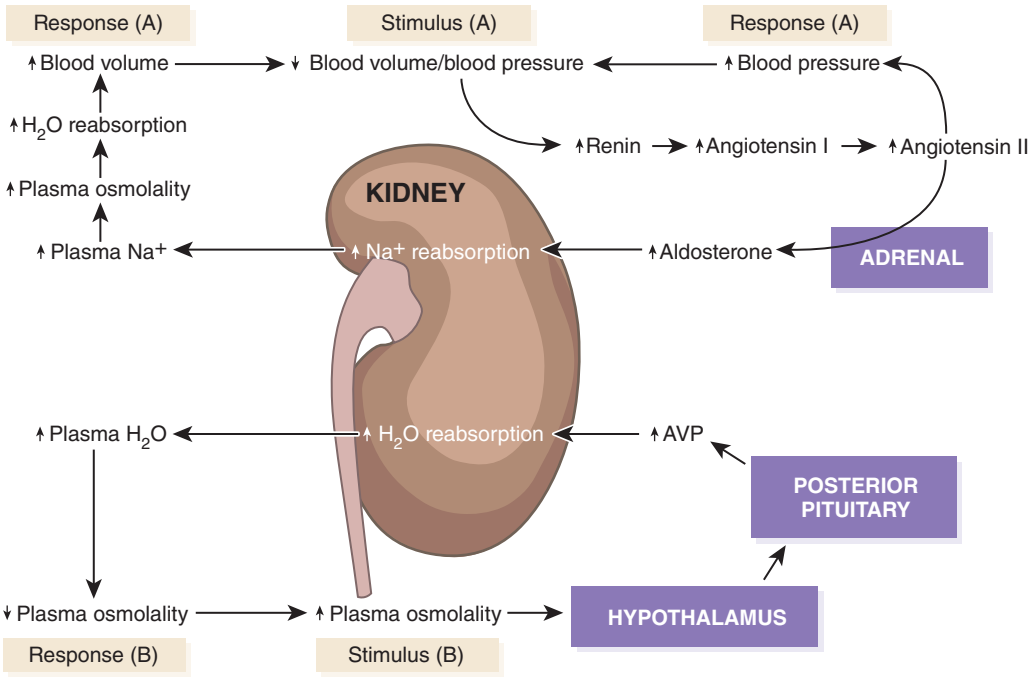


Figure 21.4 Arginine Vasopressin (AVP)/Antidiuretic hormone (ADH) and aldosterone control of the renal reabsorption of water and sodium.

Reprinted with permission from Kaplan A, et al. The kidney and tests of renal function. In: Kaplan A, Jack R, Orpheum KE, et al., eds. *Clinical Chemistry: Interpretation and Techniques.* 4th ed. Baltimore, MD: Williams & Wilkins; 1995:158, Figure 6.2.

out of the tubule and into the medulla interstitium, increasing its osmolality.[3]

Elimination of Nonprotein Nitrogen Compounds

Nonprotein nitrogen compounds (NPNs) are waste products formed in the body as a result of the degradative metabolism of nucleic acids, amino acids, and proteins. Excretion of these compounds is an important function of the kidneys. The three principal compounds are urea, creatinine, and uric acid.[4,5] For a more detailed treatment of their biochemistry and disease correlations, see Chapter 7, *Nonprotein Nitrogen Compounds*.

Urea

Urea makes up the majority (more than 75%) of the NPN waste excreted daily as a result of the oxidative catabolism of protein. Urea synthesis occurs in the liver. Proteins are broken down into amino acids, which are then deaminated to form ammonia. Ammonia is readily converted to urea, avoiding toxicity. The kidney is the only significant route of excretion for urea. It has a molecular weight of 60 Da and, therefore, is readily filtered by the glomerulus. In the collecting ducts, 40% to 60% of urea is reabsorbed. The reabsorbed urea contributes to the high osmolality in the medulla, which is one of the processes of urinary concentration mentioned earlier (see Loop of Henle).

Creatinine

Muscle contains creatine phosphate, a high-energy compound for the rapid formation of adenosine triphosphate (ATP). This reaction is catalyzed by creatine kinase (CK) and is the first source of metabolic fuel used in muscle contraction. Every day, up to 20% of total muscle creatine (and its phosphate) spontaneously dehydrates and cycles to form the waste product creatinine. Therefore, creatinine levels are a function of muscle mass and remain approximately the same in an individual from day to day unless muscle mass or renal function changes. Creatinine has a molecular weight of 113 Da and is, therefore, readily filtered by the glomerulus. Unlike urea, creatinine is not reabsorbed by the tubules. However, a small amount of creatinine is secreted by the kidney tubules at high serum concentrations.

Uric Acid

Uric acid is the primary waste product of purine metabolism. The purines, adenine and guanine, are precursors of nucleic acids ATP and guanosine

triphosphate, respectively. Uric acid has a molecular weight of 168 Da. Like creatinine, it is readily filtered by the glomerulus, but it then undergoes a complex cycle of reabsorption and secretion as it courses through the nephron. Only 6% to 12% of the original filtered uric acid is finally excreted. Uric acid exists in its ionized and more soluble form, usually sodium urate, at urinary pH > 5.75 (the first pK_a of uric acid). At pH < 5.75, it is undissociated. This fact has clinical significance in the development of urolithiasis (formation of calculi) and gout.

Water, Electrolyte, and Acid–Base Homeostasis

Water Balance

The kidney's contribution to water balance in the body is through water loss or water conservation, which is regulated by the hormone AVP. AVP responds primarily to changes in osmolality and intravascular volume. Increased plasma osmolality or decreased intravascular volume stimulates secretion of AVP from the posterior pituitary. AVP then increases the permeability of the distal convoluted tubules and collecting ducts to water, resulting in increased water reabsorption and excretion of more concentrated urine. In contrast, the major system regulating water intake is thirst, which appears to be triggered by the same stimuli that trigger AVP secretion.

In states of dehydration, the renal tubules reabsorb water at their maximal rate, resulting in production of a small amount of maximally concentrated urine (high urine osmolality, 1200 mOsm/L).[6] In states of water excess, the tubules reabsorb water at only a minimal rate, resulting in excretion of a large volume of extremely dilute urine (low urine osmolality, down to 50 mOsm/L).[7,8] The continuous fine-tuning possible between these two extreme states results in the precise control of fluid balance in the body (**Figure 21.5**).

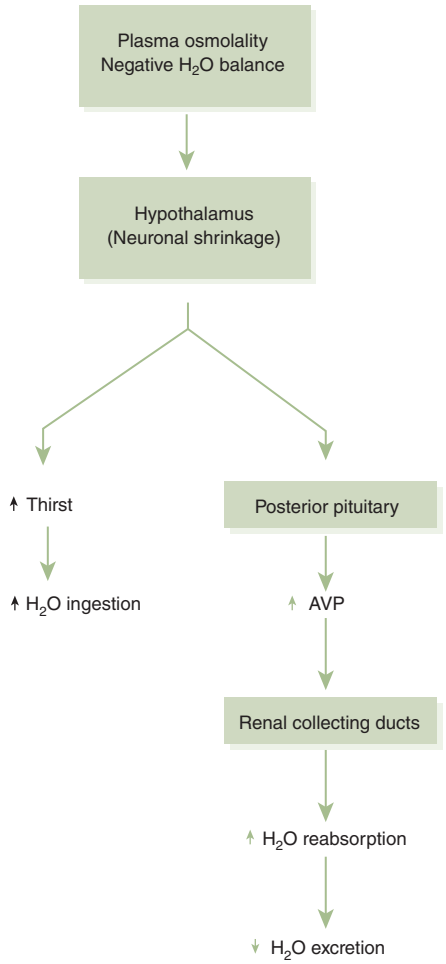


Figure 21.5 Arginine vasopressin (AVP) control of thirst mechanism.

Electrolyte Balance

The following is a brief overview of the notable ions involved in maintenance of electrolyte balance within the body. For a more comprehensive treatment of this subject, refer to Chapter 11, *Electrolytes*.

Sodium. Sodium is the primary extracellular cation in the human body and is excreted principally through the kidneys. Sodium balance in the body is controlled only through excretion. The renin–angiotensin–aldosterone hormonal system is the major mechanism for the control of sodium balance.

Potassium. Potassium is the main intracellular cation in the body. The precise regulation of its concentration is of extreme importance to cellular metabolism and is controlled chiefly by renal means. Like sodium, it is freely filtered by the glomerulus and then actively reabsorbed throughout the entire nephron (except for the descending limb of the Loop of Henle). Both the distal convoluted tubule and the collecting ducts can reabsorb and excrete potassium, and this excretion is controlled by aldosterone. Potassium ions can compete with hydrogen ions in their exchange with sodium (in the proximal convoluted tubule); this process is used by the body to conserve hydrogen ions and, thereby, compensate in states of metabolic alkalosis.

Chloride. Chloride is the principal extracellular anion and is involved in the maintenance of extracellular fluid balance. It is readily filtered by the glomerulus and is passively reabsorbed as a counterion when sodium is reabsorbed in the proximal

convoluted tubule. In the ascending limb of the Loop of Henle, potassium is actively reabsorbed by a distinct chloride "pump," which also reabsorbs sodium. This pump can be inhibited by loop diuretics, such as furosemide. As expected, the regulation of chloride is controlled by the same forces that regulate sodium.[6,8]

Phosphate, Calcium, and Magnesium. The phosphate ion occurs in higher concentrations in the intracellular than in the extracellular fluid environments. It exists as either a protein-bound or a non-protein-bound form; homeostatic balance is chiefly determined by proximal tubular reabsorption under the control of parathyroid hormone (PTH). Calcium, the second most predominant intracellular cation, is the most important inorganic messenger in the cell. It also exists in protein-bound and non-protein-bound states. Calcium in the non-protein bound form is either ionized and physiologically active or nonionized and complexed to small, diffusible ions, such as phosphate and bicarbonate. The ionized form is freely filtered by the glomerulus and reabsorbed in the tubules under the control of PTH. However, renal control of calcium concentration is not the major means of regulation. PTH and calcitonin-controlled regulation of calcium absorption from the gut and bone stores is more important than renal secretion or reabsorption. Magnesium, a major intracellular cation, is important as an enzymatic cofactor. Like phosphate and calcium, it exists in both protein-bound and ionized states. The ionized fraction is easily filtered by the glomerulus and reabsorbed in the tubules under the influence of PTH. See Chapter 18, *Parathyroid Function*, for more detailed information.

Acid–Base Balance

Many nonvolatile acidic waste products are formed by normal body metabolism each day. Carbonic acid, lactic acid, ketoacids, and others must be continually transported in the plasma and excreted from the body, causing only minor alterations in physiologic pH. The renal system constitutes one of three means by which constant control of overall body pH is accomplished. The other two strategies involved in this regulation are the respiratory system and the acid–base buffering system.[9]

The kidneys manage their share of the responsibility for controlling body pH by dual means: conserving bicarbonate ions and removing metabolic acids. For a more in-depth examination of these processes, refer to Chapter 12, *Blood Gasses, pH, and Buffer Systems*.

Regeneration of Bicarbonate Ions. In a complicated process, bicarbonate ions are first filtered out of the plasma by the glomerulus. In the lumen of the renal tubules, this bicarbonate combines with hydrogen ions to form carbonic acid, which subsequently degrades to carbon dioxide (CO_2) and water. This CO_2 then diffuses into the brush border of the proximal tubular cells, where it is reconverted by carbonic anhydrase to carbonic acid and then degrades back to hydrogen ions and regenerated bicarbonate ions. This regenerated bicarbonate is transported into the blood to replace what was depleted by metabolism; the accompanying hydrogen ions are secreted back into the tubular lumen, and from there, they enter the urine. Filtered bicarbonate is "reabsorbed" into the circulation, helping to return blood pH to its optimal level and effectively functioning as another buffering system.

Excretion of Metabolic Acids. Hydrogen ions are manufactured in the renal tubules as part of the regeneration mechanism for bicarbonate. These hydrogen ions, as well as others that are dissociated from nonvolatile organic acids, are disposed of by several different reactions with buffer bases.

Reaction with Ammonia (NH_3). The glomerulus does not filter NH_3. However, this substance is formed in the renal tubules when the amino acid glutamine is deaminated by glutaminase. This NH_3 then reacts with secreted hydrogen ions to form ammonium ions (NH^+), which are unable to readily diffuse out of the tubular lumen and, therefore, are excreted into the urine. This mode of acid excretion is the primary means by which the kidneys compensate for states of metabolic acidosis.

Reaction with Monohydrogen Phosphate (HPO_4^{2-}). Phosphate ions filtered by the glomerulus can exist in the tubular fluid as disodium hydrogen phosphate (Na_2HPO_4) (dibasic). This compound can react with hydrogen ions to yield dihydrogen phosphate (monobasic), which is then excreted. The released sodium then combines with bicarbonate to yield sodium bicarbonate and is reabsorbed. These mechanisms can excrete increasing amounts of metabolic acid until a maximum urine pH of approximately 4.4 is reached. After this, renal compensation is unable to adjust to any further decreases in blood pH, and metabolic acidosis ensues. Few free hydrogen ions are excreted directly in the urine.

Endocrine Function

In addition to numerous excretory and regulatory functions, the kidney has endocrine functions as well. It is both a primary endocrine site, as the producer of its own hormones, and a secondary site, as the target locus for hormones manufactured by other endocrine organs. The kidneys synthesize renin, **erythropoietin**, 1,25-dihydroxyvitamin D_3, and the **prostaglandins**.

Renin

Renin is the initial component of the renin–angiotensin–aldosterone system. Renin is produced by the juxtaglomerular cells of the renal medulla when extracellular fluid volume or blood pressure decreases. It catalyzes the synthesis of angiotensin by cleavage of the circulating plasma precursor angiotensinogen. Angiotensin is converted to angiotensin II by angiotensin-converting enzyme. Angiotensin II is a powerful vasoconstrictor that increases blood pressure and stimulates release of aldosterone from the adrenal cortex. Aldosterone, in turn, promotes sodium reabsorption and water conservation.[7,8] For a more detailed look at the complexities of this feedback loop, see Chapter 16, *Adrenal Function*.

Erythropoietin

Erythropoietin is a single-chain polypeptide produced by cells close to the proximal tubules, and its production is regulated by blood oxygen levels. Hypoxia produces increased serum concentrations within 2 hours. Erythropoietin acts on the erythroid progenitor cells in the bone marrow, increasing the number of red blood cells (RBCs). In chronic renal insufficiency, erythropoietin production is significantly reduced. The routine administration of recombinant human erythropoietin is common in chronic renal failure patients. Before this therapy was available, anemia was a clinical reality in these patients.[4,7] Erythropoietin concentrations in blood can be measured by immunoassays. Recombinant human erythropoietin has also been used in sports doping to stimulate erythrocyte production and increase the oxygen-carrying capacity in the blood of endurance athletes. Assays capable of detecting posttranslational modifications on erythropoietin have been produced and are capable of distinguishing exogenous from endogenous erythropoietin.

1,25-Dihydroxyvitamin D_3

The kidneys are the sites of formation of the active form of **vitamin D**, 1,25-$(OH)_2$ vitamin D_3. This form of vitamin D is one of three major hormones that determine phosphate and calcium balance and bone calcification in the human body. Chronic renal insufficiency is, therefore, often associated with osteomalacia (inadequate bone calcification, the adult form of rickets), owing to the continual distortion of normal vitamin D metabolism.

Prostaglandins

The prostaglandins are a group of potent cyclic fatty acids formed from essential (dietary) fatty acids, primarily arachidonic acid. They are formed in almost all tissue, and their actions are diverse. The prostaglandins produced by the kidneys increase renal blood flow, sodium and water excretion, and renin release. They act to oppose renal vasoconstriction due to angiotensin and norepinephrine.

Analytic Procedures

All laboratory methods used for the evaluation of renal function rely on the measurement of waste products in blood, usually urea and creatinine, which accumulate when the kidneys begin to fail. Renal failure must be advanced, with only about 20% to 30% of the nephrons still functioning, before the concentration of either substance begins to increase in the blood. The rate at which creatinine and urea are removed or cleared from the blood into the urine is termed *clearance*. Clearance is defined as that volume of plasma from which a measured amount of substance can be completely eliminated into the urine per unit of time expressed in milliliters per minute.[5] Calculation of creatinine clearance has become the standard laboratory method for determining the GFR. Urea clearance was one of the first clearance tests performed; however, it is no longer widely used since it does not accurately provide a full clearance assessment. Older tests used administration of inulin, or *p*-aminohippurate to assess glomerular filtration or tubular secretion. These tests are difficult to administer and are no longer performed. Iothalamate is used to evaluate potential living kidney transplant donors.

Creatinine Clearance

Creatinine is a nearly ideal substance for the measurement of clearance. It is an endogenous metabolic product synthesized at a constant rate for a given individual and cleared essentially only by glomerular filtration. It is not reabsorbed and is only slightly secreted by the proximal tubule. Serum creatinine levels are higher

in males than in females due to the direct correlation with muscle mass. Analysis of creatinine is simple and inexpensive using colorimetric assays; however, different methods for assaying plasma creatinine, such as kinetic or enzymatic assays, have varying degrees of accuracy and imprecision (see Chapter 7, *Nonprotein Nitrogen Compounds*).

Creatinine clearance is derived by mathematically relating the serum creatinine concentration to the urine creatinine concentration excreted during a period of time, usually 24 hours. Specimen collection, therefore, must include both a 24-hour urine specimen and a serum creatinine value, ideally collected at the midpoint of the 24-hour urine collection. The urine container must be kept refrigerated throughout the duration of both the collection procedure and the subsequent storage period until laboratory analysis can be performed. The concentration of creatinine in both serum and urine is measured by the applicable methods discussed in Chapter 7, *Nonprotein Nitrogen Compounds*. The total volume of urine is carefully measured, and the creatinine clearance is calculated using the following formula:

$$\frac{U_{Cr}\ (\text{mg/dL})\ V_{Ur}\ (\text{mL/24 hours})}{P_{Cr}\ (\text{mg/dL})\ 1440\ \text{minutes/24 hours}} \times \frac{1.73}{A} \quad \textbf{(Eq. 21.1)}$$

where Cr is creatinine clearance, U_{Cr} is urine creatinine clearance, V_{Ur} is urine volume excreted in 24 hours, P_{Cr} is serum creatinine concentration, and $1.73/A$ is normalization factor for body surface area (1.73 is the generally accepted average body surface in square meters, and A is the actual body surface area of the individual determined from height and weight). If the patient's body surface area varies greatly from the average (e.g., obese or pediatric patients), this correction for body mass must be included in the formula. The reference range for creatinine clearance is lower in females compared with males and normally decreases with age.

Estimated GFR

The National Kidney Foundation recommends that **estimated glomerular filtration rate (eGFR)** be calculated each time a serum creatinine level is reported. (Additional information is available at the National Kidney Foundation website at www.kidney.org.) The equation used to predict GFR is based on serum creatinine, age, body size, gender, and race, without the need for a 24-hour urine collection, which when improperly collected causes inaccurate test results. Because the calculation does not require a timed urine collection, it should be used more often than the traditional creatinine clearance and result in earlier detection of **chronic kidney disease (CKD)**. There are a number of formulas that can be used to estimate GFR on the basis of serum creatinine levels.

Cockcroft-Gault Formula

The Cockcroft-Gault formula is one of the first formulas used to estimate GFR. This formula predicts creatinine clearance, and the results are not corrected for body surface area. This equation assumes that women will have a 15% lower creatinine clearance than men at the same level of serum creatinine.

$$\text{GFR (mL/min)} = \frac{(140 - \text{Age}) \times \text{Weight (kg)}}{72 \times S_{Cr}\ (\text{mg/dL})} \times (0.85 \text{ if female}) \quad \textbf{(Eq. 21.2)}$$

Modification of Diet in Renal Disease Formula

The modification of diet in renal disease (MDRD) formula was developed in the Modification of Diet in Renal Disease Study of chronic renal insufficiency. The study showed that the MDRD formula provided a more accurate assessment of GFR than the Cockcroft-Gault formula. The MDRD formula was validated in a large population that included both Black and White individuals. It does not require patient weight and is corrected for body surface area. The MDRD formula is known to underestimate the GFR in healthy patients with GFRs over 60 mL/min and to overestimate GFR in underweight patients. The four-variable MDRD equation includes age, race, gender, and serum creatinine as variables.

$$\text{GFR (mL/min/1.7 m}^2) = 186 \times S_{Cr}\ (\text{mg/dL})^{-1.154} \times \text{Age}^{-0.203} \times (1.212 \text{ if Black}) \times (0.742 \text{ if female}) \quad \textbf{(Eq. 21.3)}$$

CKD-EPI Formula

The CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration) formula was published in 2009.[10] It was developed in an effort to create a formula more accurate than the MDRD formula. Multiple studies have shown the CKD-EPI formula to perform better and with less bias than the MDRD formula, especially in patients with higher GFR. Many laboratories still

use the MDRD formula; however, some have converted to the CKD-EPI formula.

$$\text{eGFR (mL/min/1.73 m}^2) = 141 \times \min(S_{Cr}/k, 1)^{a} \times \max(S_{Cr}/k, 1)^{-1.209} \times 0.993^{\text{Age}} \times (1.018 \text{ if female}) \times (1.159 \text{ if Black}) \quad \textbf{(Eq. 21.4)}$$

In this formula, k is 0.7 for females and 0.9 for males, a is –0.329 for females and –0.411 for males, min indicates the minimum of S_{Cr}/k or 1, and max indicates the maximum of S_{Cr}/k or 1.

It is important to highlight the recent discussions on race in the use of the MDRD and CKD-EPI formulas. The National Kidney Foundation, alongside the American Society for Nephrology, developed a task force called Reassessing the Inclusion of Race in Diagnosing Kidney Disease. The aim is to create race-neutral formulas for assessing kidney function due to the overestimation the current formulas produce. See Chapter 7, *Nonprotein Nitrogen Compounds*, or more information.

Cystatin C

Cystatin C is a low-molecular-weight protein produced at a steady rate by most body tissues. It is freely filtered by the glomerulus, reabsorbed, and catabolized by the proximal tubule. Levels of cystatin C rise more quickly than creatinine levels in acute kidney injury (AKI). Plasma concentrations appear to be unaffected by diet, gender, race, age, and muscle mass. Studies have shown measurement of cystatin C to be at least as useful as serum creatinine and creatinine clearance in detecting early changes in kidney function. A rise in cystatin C is often detectable before there is a measurable decrease in the GFR or increase in creatinine. Cystatin C can be measured by immunoassay methods.[11] Some findings suggest that an equation that uses both serum creatinine and cystatin C with age, sex, and race would be better than equations that use only one of these serum markers.[12,13]

Biologic Variation

When the same test is performed on various individuals, we find that the mean of each person's results is not the same, showing that individual homeostatic setting points often vary. Biologic variation is defined as the random fluctuation around a homeostatic setting point.[14] This includes the fluctuation around the homeostatic setting point for a single individual, termed *within-subject biologic variation*, and differences between the homeostatic setting points of multiple individuals, termed *between-subject biologic* variation. In the case of creatinine, levels for an individual differ slightly over time, and the mean values of all individuals vary significantly from each other. Therefore, each individual's results span only a small portion of the population-based reference range. This means that for creatinine, within-subject biologic variation is less than between-subject biologic variation. When this is true of a given analyte, the analyte is said to have marked individuality. Interestingly, the between-subject biologic variation is much smaller for cystatin C compared with creatinine, showing that population-based reference values are more useful for cystatin C compared with creatinine. However, the within-subject variation is greater for cystatin C compared with creatinine. As a result, creatinine is more helpful in monitoring renal function over time for a given individual, whereas cystatin C is potentially more useful for detecting minor renal impairment.

β_2-Microglobulin

β_2-Microglobulin (β_2-M) is a small, nonglycosylated peptide (molecular weight, 11.8 kDa) found on the surface of most nucleated cells. The plasma membrane sheds β_2-M at a constant rate, as a relatively intact molecule. β_2-M is easily filtered by the glomerulus, and 99.9% is reabsorbed by the proximal tubules and catabolized. Elevated levels in serum indicate increased cellular turnover as seen in myeloproliferative and lymphoproliferative disorders, inflammation, and renal failure. Both blood and urine β_2-M tests may be ordered to evaluate kidney damage and to distinguish between disorders that affect the glomeruli and the renal tubules. Measurement of serum β_2-M is used clinically to assess renal tubular function in renal transplant patients, with elevated levels indicating organ rejection.

Myoglobin

Myoglobin is a low-molecular-weight protein (16.9 kDa) associated with acute skeletal and cardiac muscle injury. Myoglobin functions to bind and transport oxygen from the plasma membrane to the mitochondria in muscle cells. Blood levels of myoglobin can rise very quickly with severe muscle injury. In **rhabdomyolysis**, myoglobin release from skeletal muscle is sufficient to overload the proximal tubules and cause acute renal failure. Early diagnosis and aggressive treatment of elevated myoglobin may

prevent or lessen the severity of renal failure. Serum and urine myoglobin can be measured easily and rapidly by immunoassays.

Albuminuria

The term **albuminuria**, also referred to as *microalbuminuria*, describes the presence of albumin in the urine. Urine albumin measurement is important in the management of patients with **diabetes mellitus**, who are at serious risk for developing nephropathy over their lifetime. In the early stages of nephropathy, there is renal hypertrophy, hyperfunction, and increased thickness of the glomerular and tubular basement membranes. In this early stage, there are no overt signs of renal dysfunction. In the next 7 to 10 years, there is progression to glomerulosclerosis, with increased glomerular capillary permeability. This permeability allows small (micro) amounts of albumin to pass into the urine. If detected in this early phase, rigid glucose control, along with treatment to prevent hypertension, can be instituted and progression to kidney failure prevented. Quantitative albumin-specific immunoassays, usually using nephelometry or immunoturbidimetry, are widely used. For a 24-hour urine collection, 30 to 300 mg of albumin is diagnostic of albuminuria. A 24-hour urine collection is preferred, but a random urine sample in which the ratio of albumin to creatinine is measured, is common. An albumin to creatinine ratio (ACR) of greater than 30 mg/g is diagnostic of albuminuria.

Neutrophil Gelatinase–Associated Lipocalin

Neutrophil gelatinase–associated lipocalin (NGAL) is a 25-kDa protein expressed by neutrophils and epithelial cells including those of the proximal tubule. The gene encoding NGAL is upregulated in the presence of renal ischemia, tubule injury, and nephrotoxicity.[15] It can be measured in plasma and urine and is elevated within 2 to 6 hours of AKI. NGAL has been shown to be a useful early predictor of AKI and has prognostic value for clinical endpoints, such as initiation of dialysis and mortality. However, urinary NGAL excretion may also arise from systemic stress in the absence of AKI, limiting its specificity. Currently, this test is for research use only.

NephroCheck

NephroCheck is the first FDA-cleared test used to determine if critically ill patients are at risk for developing moderate to severe AKI in the 12 hours following administration of the test. NephroCheck quantifies concentrations of tissue inhibitor of metalloproteinase 2 (TIMP-2) and insulin-like growth factor binding protein 7 (IGFBP-7) in urine specimens and multiplies the results to generate a quantitative AKI risk index (AKI Risk Score). The two biomarkers invoke cell cycle arrest in response to tissue insults that cause AKI.[16] Elevation of these two biomarkers is thought to signal the kidney's attempt to protect itself from harmful insults and signal the increased risk for imminent AKI. The biomarker concentrations have been shown to correspond to severity across all stages of AKI.

Urinalysis

Urinalysis permits a detailed, in-depth assessment of renal status with an easily obtained specimen. Urinalysis also serves as a quick indicator of an individual's glucose status and hepatic–biliary function. Routine urinalysis includes assessment of physical

CASE STUDY 21.1, PART 3

Remember George. In the ED, he was diagnosed with acute renal failure. The doctor ordered intravenous fluids; George's BUN fell to 68 mg/dL, and creatinine fell to 2.2 mg/dL. George's blood culture report was positive for *E. coli*. He was treated with tobramycin and cefepime.

3. What is the cause of his acute renal failure?
4. What is the significance of the patient's large urine hemoglobin?

characteristics, chemical analyses, and a microscopic examination of the sediment from a (random) urine specimen.

Specimen Collection

The importance of a properly collected and stored specimen for urinalysis cannot be overemphasized. Initial morning specimens are preferred, particularly for protein analyses, because they are more concentrated from overnight retention in the bladder. The specimen should be obtained by a clean midstream catch or catheterization. The urine should be freshly collected into a clean, dry container with a tightfitting cover. It must be analyzed within 1 hour of collection if held at room temperature or else refrigerated at 2°C to 8°C for not more than 8 hours before analysis. If not assayed within these time limits, several changes will occur. Bacterial multiplication will cause false-positive nitrite tests, and urease-producing organisms will degrade urea to ammonia and alkalinize the pH. Loss of CO_2 by diffusion into the air adds to this pH elevation, which, in turn, causes cast degeneration and red cell lysis.

Physical Examination

Color. Color intensity of urine correlates with concentration: the darker the color, the more concentrated is the specimen. The various colors observed in urine are a result of different excreted pigments. Yellow and amber are generally due to urochromes (derivatives of urobilin, the end product of bilirubin degradation), whereas a yellowish-brown to green color is a result of bile pigment oxidation. Red and brown after standing are due to porphyrins, whereas reddish brown in fresh specimens comes from hemoglobin or red cells. Brownish black after standing is seen in alkaptonuria (a result of excreted homogentisic acid) and in malignant melanoma (in which the precursor melanogen oxidizes in the air to melanin). Drugs and some foods, such as beets, may also alter urine color.

Odor. Odor has little diagnostic significance. The characteristic pungent odor of fresh urine is due to volatile aromatic acids, in contrast to the typical ammonia odor of urine that has been allowed to stand. Urinary tract infections impart a noxious, fecal smell to urine, whereas the urine of diabetics often smells fruity as a result of ketones. Certain inborn errors of metabolism, such as maple sugar urine disease, are associated with characteristic urine odors.

Clarity. The cloudiness of a urine specimen depends on pH and dissolved solids composition. Turbidity generally may be due to gross bacteriuria, whereas a cloudy appearance is seen in hematuria. Thread-like cloudiness is observed when the specimen is full of mucus. In alkaline urine, suspended precipitates of amorphous phosphates and carbonates may be responsible for turbidity, whereas in acidic urine, amorphous urates may be the cause.[17]

Volume. The volume of urine excreted indicates the balance between fluid ingestion and water lost from the lungs, sweat, and intestine. Most adults produce from 750 to 2000 mL every 24 hours, averaging about 1.5 L per person. Polyuria is observed in diabetes mellitus and diabetes insipidus (in diabetes insipidus, it stems from a lack of AVP), as well as in chronic renal disease, acromegaly (overproduction of the growth hormone somatostatin), and myxedema (hypothyroid edema). Anuria or oliguria (<200 mL/d) is found in nephritis, urinary tract obstruction, AKI, and kidney failure.

Specific Gravity. The specific gravity (SG) of urine is the weight of 1 mL of urine in grams divided by the weight of 1 mL of water. SG gives an indication of the density of a fluid that depends on the concentration of dissolved total solids. SG varies with the solute load to be excreted (consisting primarily of NaCl and urea), as well as with the urine volume. It is used to assess the state of hydration/dehydration of an individual or as an indicator of the concentrating ability of the kidneys.

Specific gravity is measured using the refractive index on most automated urinalysis analyzers. The most common manual method uses a refractometer. This operates on the principle that the refractive index of a urine specimen will vary directly with the total amount of dissolved solids in the sample. This instrument measures the refractive index of the urine as compared with water on a scale that is calibrated directly into the ocular and viewed while held up to a light source. Correct calibration is vital for accuracy. Most recently, an indirect colorimetric reagent strip method for assaying SG has been added to most dipstick screens. Unlike the refractometer, dipsticks measure only ionic solutes and do not take into account glucose or protein.

The normal range for urinary SG is 1.003 to 1.035 g/mL. SG can vary in pathologic states. Low SG can occur in diabetes insipidus, pyelonephritis, and **glomerulonephritis**, in which the renal

concentrating ability has become dysfunctional. High SG can be seen in diabetes mellitus, congestive heart failure, dehydration, adrenal insufficiency, liver disease, and nephrosis. SG will increase about 0.004 units for every 1% change in glucose concentration and about 0.003 units for every 1% change in protein. Fixed SG (isosthenuria) around 1.010 is observed in severe renal damage, in which the kidney excretes urine that is iso-osmotic with the plasma. This generally occurs after an initial period of anuria because the damaged tubules are unable to concentrate or dilute the glomerular filtrate.[16]

pH. Determinations of urinary pH must be performed on fresh specimens because of the significant tendency of urine to alkalinize on standing. Normal urine pH falls within the range of 4.7 to 7.8. Acidity in urine (pH $<$ 7.0) is primarily caused by phosphates, which are excreted as salts conjugated to Na^+, K^+, Ca^{2+}, and NH^+. Acidity also reflects the excretion of the nonvolatile metabolic acids: pyruvate, lactate, and citrate. Owing to the Na^+/H^+ exchange pump mechanism of the renal tubules, pH (H^+ concentration) increases as sodium is retained. Pathologic states, in which increased acidity is observed, include systemic acidosis, as seen in diabetes mellitus, and renal tubular acidosis (RTA). In RTA, the tubules are unable to excrete excess H^+ even though the body is in metabolic acidosis, and urinary pH remains around 6.

Alkaline urine (pH $>$ 7.0) is observed postprandially as a normal reaction to the acidity of gastric HCl dumped into the duodenum and then into the circulation or following ingestion of alkaline food or medications. Urinary tract infections and bacterial contamination also will alkalinize pH. Medications such as potassium citrate and sodium bicarbonate will reduce urine pH. Alkaline urine is also found in Fanconi syndrome, a congenital generalized aminoaciduria resulting from defective proximal tubular function.

Chemical Testing

Routine urine chemical analysis is rapid and easily performed with commercially available reagent strips or dipsticks. These strips are plastic coated with different reagent bands directed toward different analytes. When dipped into urine, a color change signals a deviation from normality. Colors on the dipstick bands are matched against a color chart provided with the reagents. Automated and semiautomated instruments that detect by reflectance photometry provide an alternative to the color chart and offer better precision and standardization. Abnormal results are followed up by specific quantitative or confirmatory urine assays. The analytes routinely tested are glucose, ketones, protein, nitrite, leukocyte esterase, bilirubin/urobilinogen, and hemoglobin/blood.

Glucose. This analyte is normally absent in urine. This test is used to determine whether the reducing substance found in urine is glucose. No other substance excreted in urine is known to give a positive result, including other reducing substances (e.g., galactose, fructose, and lactose). Glucosuria occurs when the renal threshold for glucose is exceeded (typically $>$180 mg/dL). This is most commonly seen in diabetes mellitus.

Ketones. These constituents are normally absent in urine. Their presence is a hallmark of diabetes mellitus and, in the case of ketonuria, signals that the patient may have diabetic ketoacidosis, a potentially life-threatening condition. The clinical significance of these analytes and their testing methods are discussed in Chapter 9, *Carbohydrates*.

Protein. Reagent strips for urinalysis are used as a general qualitative screen for proteinuria. They are primarily specific for albumin, but they may give false-positive results in specimens that are alkaline and highly buffered. Positive dipstick results should be confirmed by more specific chemical assays, as described in Chapter 6, *Amino Acids and Proteins*, or more commonly by microscopic evaluation to detect casts.

Nitrite. This assay semi-quantitates the amount of urinary reduction of nitrate (on the reagent strip pad) to nitrite by the enzymes of certain gram-negative bacteria. A negative result does not mean that no bacteriuria is present. A gram-positive pathogen, such as *Staphylococcus*, *Enterococcus*, or *Streptococcus*, may not produce nitrate-reducing enzymes; alternatively, a urine sample may not have been retained in the bladder long enough to pick up a sufficient number of organisms to register on the reagent strip.[16]

Leukocyte Esterase. White blood cells (WBCs), especially phagocytes, contain esterases. A positive dipstick for esterases indicates possible WBCs in urine.

Bilirubin. Hemoglobin degradation ultimately results in the formation of the waste product bilirubin, which is then converted to urobilinogen in

the gut through bacterial action. Although most of this urobilinogen is excreted as urobilin in the feces, some is excreted in urine as a colorless waste product. This amount is normally too small to be detected as a positive dipstick reaction. In conditions of prehepatic, hepatic, and posthepatic jaundice, however, urine dipstick tests for urobilinogen and bilirubin may be positive or negative, depending on the nature of the patient's jaundice. A more in-depth view of bilirubin metabolism and assay methods is given in Chapter 19, *Liver Function*. Reagent strip tests for bilirubin involve diazotization and formation of a color change.

Urobilinogen. Urobilinogen is normally found in trace amounts in the urine. Therefore, it is reported as normal rather than negative. Methods for urobilinogen differ, but most rely on a modification of the Ehrlich reaction with *p*-dimethylaminobenzaldehyde.[16] Urobilinogen levels increase in liver disease and hemolysis.

Blood. Intact or lysed RBCs will produce a positive dipstick result. The dipstick will be positive in cases of renal trauma/injury, infection, and obstruction that result from calculi or neoplasms.

Microscopic Examination

Centrifuged, decanted urine aliquot leaves behind a sediment of formed elements that is used for microscopic examination.

For cellular elements, evaluation is best accomplished by counting and then taking the average of at least 10 microscopic fields using a standardized urinalysis microscopic system.

Red Blood Cells. Erythrocytes greater in number than 2 per high-power field (HPF) are considered abnormal. Such hematuria may result simply from severe exercise or menstrual blood contamination. However, it may also be indicative of trauma, particularly vascular injury, renal/urinary calculi obstruction, pyelonephritis, or cystitis. Hematuria in conjunction with leukocytes is diagnostic of infection.

White Blood Cells. Leukocytes greater in number than 5 per HPF are considered abnormal. These cells are usually polymorphonuclear phagocytes, commonly known as segmented neutrophils. They are observed when there is acute glomerulonephritis, urinary tract infection, or inflammation of any type. In hypotonic urine (low osmotic concentration), WBCs can become enlarged, exhibiting a sparkling effect in their cytoplasmic granules. These cells possess a noticeable Brownian motion and are called glitter cells, but they have no pathologic significance.

Epithelial Cells. Several types of epithelial cells are frequently encountered in normal urine because they are continuously sloughed off the lining of the nephrons and urinary tract. Large, flat, squamous vaginal epithelia are often seen in urine specimens from female patients, and samples heavily contaminated with vaginal discharge may show clumps or sheets of these cells. Renal epithelial cells are round, uninucleated cells and, if present in numbers greater than 2 per LPF, indicate clinically significant active tubular injury or degeneration. Transitional bladder epithelial cells (urothelial cells) may be flat, cuboidal, or columnar and also can be observed in urine on occasion. Large numbers will be seen only in cases of urinary catheterization, bladder inflammation, or neoplasm.

Miscellaneous Elements. Spermatozoa are often seen in the urine of both males and females. Spermatozoa should be reported in all patients (males and females) younger than age 10 as recommended by the CAP Clinical Microscopy Resource Committee. As a reminder, always follow the laboratory's standard operating procedure for reporting spermatozoa. Yeast cells are also frequently found in urine specimens. Because they are extremely refractile and of a similar size to RBCs, they can easily be mistaken under low magnification. Higher power examination for budding or mycelial forms differentiates these fungal elements from erythrocytes. Parasites found in urine are generally contaminants from fecal or vaginal material. In fecal contaminant category, the most commonly encountered organism is *Enterobius vermicularis* (pinworm) infestation in children. In the vaginal contaminant category, the most common is the intensely motile flagellate, *Trichomonas vaginalis*, a common sexually transmitted infection. A true urinary parasite, sometimes seen in patients from endemic areas of the world, is the ova of the trematode *Schistosoma haematobium* (schistosomiasis). This condition will usually occur in conjunction with a significant hematuria.[16]

Bacteria. Normal urine is usually sterile and contains no bacteria. Small numbers of organisms seen in a fresh urine specimen usually represent skin or contamination. In fresh specimens, however, large numbers of organisms, or small numbers accompanied by WBCs and the symptoms of urinary tract infection,

are highly diagnostic for true infection. Clinically significant bacteriuria is considered to be more than 20 organisms per HPF or, alternatively, 10^5 or greater registered on a microbiologic colony count. Most pathogens seen in urine are gram-negative coliforms (microscopic "rods") such as *Escherichia coli* and *Proteus spp*. Asymptomatic bacteriuria, in which there are significant numbers of bacteria without appreciable clinical symptoms, occurs somewhat commonly in young girls, pregnant women, and patients with diabetes. This condition must be taken seriously because, if left untreated, it may result in pyelonephritis and, subsequently, permanent renal damage.

Casts. Casts are precipitated, cylindrical impressions of the nephrons. They comprise Tamm-Horsfall mucoprotein (uromucoid) from the tubular epithelia in the ascending limb of the loop of Henle. Casts form whenever there is sufficient renal stasis, increased urine salt or protein concentration, and decreased urine pH. In patients with severe renal disease, truly accurate classification of casts may require use of "cytospin" centrifugation and Papanicolaou staining for adequate differentiation. Unlike cells, casts should be examined under low power and are most often located around the edges of the coverslip.

Hyaline. The matrix of these casts is clear and gelatinous, without embedded cellular or particulate matter. They may be difficult to visualize unless a high-intensity lamp is used. Their presence indicates glomerular leakage of protein. This leakage may be temporary (as a result of fever, upright posture, dehydration, or emotional stress) or may be permanent. Their occasional presence is not considered pathologic.

Granular. These casts are descriptively classified as either coarse or finely granular. The type of embedded particulate matter is simply a matter of the amount of degeneration that the epithelial cell inclusions have undergone. Their occasional presence is not pathologic; however, large numbers may be found in chronic lead toxicity and pyelonephritis.

Cellular. Several different types of casts are included in this category. RBC or erythrocytic casts are always considered pathologic because they are diagnostic for glomerular inflammation that results in renal hematuria. They are seen in subacute bacterial endocarditis, kidney infarcts, collagen diseases, and acute glomerulonephritis. WBC or leukocytic casts are also always considered pathologic because they are diagnostic for inflammation of the nephrons. They are observed in pyelonephritis, **nephrotic syndrome**, and acute glomerulonephritis. In asymptomatic pyelonephritis, these casts may be the only clue to detection. Epithelial cell casts are sometimes formed by fusion of renal tubular epithelia after desquamation; their occasional presence is normal. Many, however, are observed in severe desquamative processes and renal stasis that occur in heavy metal poisoning, renal toxicity, eclampsia, nephrotic syndrome, and amyloidosis. Waxy casts are uniformly yellowish, refractile, and brittle appearing, with sharply defined, often broken edges. They are almost always pathologic because they indicate tubular inflammation or deterioration. They are formed by renal stasis in the collecting ducts and are, therefore, found in chronic renal diseases. Fatty casts are abnormal, coarse, granular casts with lipid inclusions that appear as refractile globules of different sizes. Broad casts (renal failure) may be up to two to six times wider than "regular" casts and may be cellular, waxy, or granular in composition. Like waxy casts, they are derived from the collecting ducts in severe renal stasis.

Crystals

Acid pH. Crystals seen in urine with pH values of less than 7.0 include calcium oxalate, which are normal colorless octahedrons or "envelopes"; they may have an almost starlike appearance. Also seen are amorphous urates, which are normal, pink-red masses that look like grains of sand. Uric acid crystals found in this environment are normal, yellow to red-brown crystals that appear in extremely irregular shapes, such as rosettes, prisms, or rhomboids. Cholesterol crystals in acid urine are clear, flat, rectangular plates with notched corners. They may be seen in nephrotic syndrome and in conditions producing chyluria and are always considered abnormal. Cystine crystals are also sometimes observed in acid urine; they are highly pathologic and appear as colorless, refractile, nearly flat hexagons, somewhat similar to uric acid. These are observed in homocystinuria (an aminoaciduria resulting in developmental disability) and cystinuria (an inherited defect of cystine reabsorption resulting in renal calculi).

Alkaline pH. Crystals seen in urine with pH values greater than 7.0 include amorphous phosphates, which are normal crystals that appear as fine, colorless masses, resembling sand. Also seen are calcium carbonate crystals, which are normal forms that appear as small, colorless dumbbells or spheres. Triple phosphate crystals are also observed in alkaline

urines; they are colorless prisms of three to six sides, resembling "coffin lids." Ammonium biurate crystals are normal forms occasionally found in this environment, appearing as spiny, yellow-brown spheres, or "thorn apples."

Pathological Crystals. Pathological crystals include leucine, tyrosine, cystine, sulfonamides, and cholesterol. Sulfonamide crystals are abnormal precipitates shaped like yellow-brown sheaves, clusters, or needles, formed in patients undergoing antimicrobial therapy with sulfa drugs. Leucine crystals found in acid or neutral urine and appear yellowish brown in color with concentric circles having radial striations resembling the dark rings of a tree trunk. Tyrosine crystals are abnormal too. They are found in acidic or neutral pH urines and are shaped like clusters of smooth, yellow needles or spheres. These are sometimes seen in patients with severe liver disease.[16]

Pathophysiology

Glomerular Diseases

Disorders or diseases that directly damage the renal glomeruli may, at least initially, exhibit normal tubular function. With time, however, disease progression involves the renal tubules as well. The following syndromes have discrete symptoms that are recognizable by their patterns of clinical laboratory findings.

Acute Glomerulonephritis

Pathologic lesions in acute glomerulonephritis primarily involve the glomerulus. Histologic examination shows large, inflamed glomeruli with a decreased capillary lumen. Abnormal laboratory findings usually include rapid onset of hematuria and proteinuria (usually albumin and generally <3 g/d). The rapid development of a decreased GFR, anemia, elevated blood urea nitrogen (BUN) and serum creatinine, oliguria, sodium and water retention (with consequent hypertension and some localized edema), and, sometimes, congestive heart failure is typical. Numerous hyaline and granular casts are generally seen on urinalysis. The actual RBC casts are regarded as highly suggestive of this syndrome. Acute glomerulonephritis is often related to recent infection by group A β-hemolytic streptococci. It is theorized that circulating immune complexes trigger a strong inflammatory response in the glomerular basement membrane, resulting in a direct injury to the glomerulus itself. Other possible causes include drug-related exposures, acute kidney infections due to other bacterial (and possibly viral) agents, and other systemic immune complex diseases, such as systemic lupus erythematosus and bacterial endocarditis.

Chronic Glomerulonephritis

Lengthy glomerular inflammation may lead to glomerular scarring and the eventual loss of functioning nephrons. This process often goes undetected for lengthy periods because only minor decreases in renal function occur at first and only slight proteinuria and hematuria are observed. Gradual development of uremia (or azotemia, excess nitrogen compounds in the blood) may be the first sign of this process.

Nephrotic Syndrome

Nephrotic syndrome (**Figure 21.6**) can be caused by several different diseases that result in injury and increased permeability of the glomerular basement membrane. This defect almost always yields several abnormal findings, such as massive proteinuria (>3.5 g/d) and resultant hypoalbuminemia. The subsequent decreased plasma oncotic pressure causes a generalized edema as a result of the movement of body fluids out of vascular and into interstitial spaces. Other hallmarks of this syndrome are hyperlipidemia and lipiduria. Lipiduria takes the form of oval fat bodies in the urine. These bodies are degenerated renal tubular cells containing reabsorbed lipoproteins. Primary causes are associated directly with glomerular disease states.

Tubular Diseases

Tubular defects occur to a certain extent in the progression of all renal diseases as the GFR falls. In some instances, however, this aspect of the overall dysfunction becomes predominant. The result is decreased excretion and reabsorption of certain substances or reduced urinary concentrating capability. Clinically, the most important defect is RTA, the primary tubular disorder affecting acid–base balance. This disease can be classified into two types, depending on the nature of the tubular defect. In distal RTA, the renal tubules are unable to keep up the vital pH gradient between the blood and tubular fluid. In proximal RTA, there is decreased bicarbonate reabsorption, resulting in hyperchloremic acidosis. In general, reduced reabsorption in the proximal tubule is manifested by findings of abnormally low serum values for phosphorus and uric acid and by glucose and amino acids in the

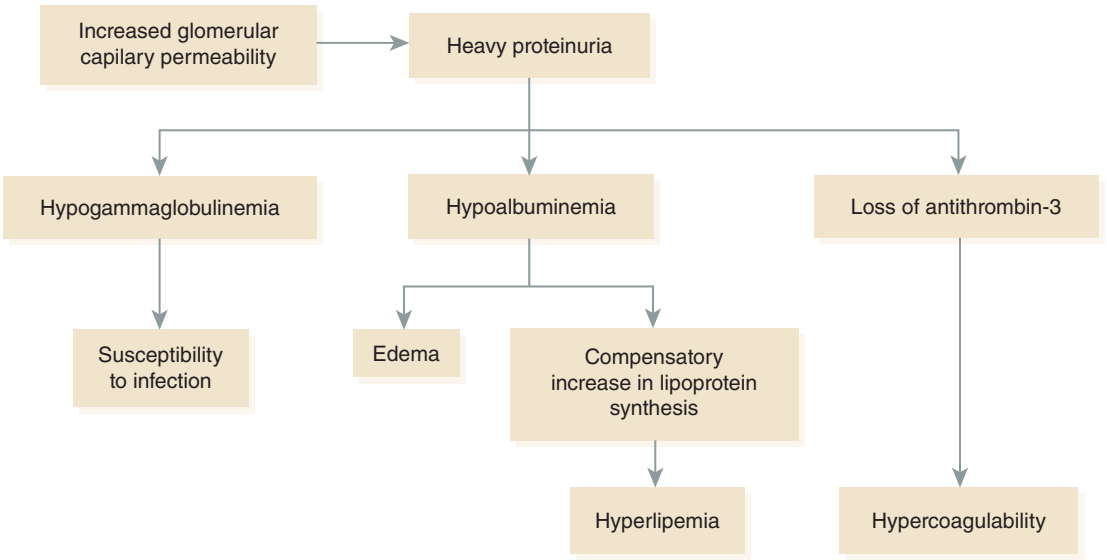


Figure 21.6 Pathophysiology of nephrotic syndrome.

urine. In addition, there may be some proteinuria (usually <2 g/d). Acute or chronic inflammation of the tubules and surrounding interstitium also may occur as a result of radiation toxicity, renal transplant rejection, viral–fungal–bacterial infections, and a reaction to medications. This is known as interstitial nephritis. Characteristic clinical findings in these cases are decreases in GFR, urinary concentrating ability, and metabolic acid excretion; leukocyte casts in the urine; and inappropriate control of sodium balance.[4,7]

Urinary Tract Infection/ Obstruction

Infection

The site of infection may be either in the kidneys (pyelonephritis) or in the urinary bladder (cystitis). In general, a microbiologic colony count of more than 10^5 colonies/mL is considered diagnostic for infection in either locale. Bacteriuria (as evidenced by positive nitrite dipstick findings for some organisms), hematuria, and pyuria (leukocytes in the urine, as shown by positive leukocyte esterase dipstick) are all frequently encountered abnormal laboratory results in these cases. In particular, WBC (leukocyte) casts in the urine is considered diagnostic for pyelonephritis.[4,7,16]

Obstruction

Renal obstructions can cause disease in one of two ways. Either they may gradually raise the intratubular pressure until nephrons necrose and chronic renal failure ensues, or they may predispose the urinary tract to repeated infections. Obstructions may be located in the upper or lower urinary tract. Blockages in the upper tract are characterized by a constricting lesion below a dilated collecting duct. Obstructions of the lower tract are evidenced by the residual urine in the bladder after cessation of micturition (urination); symptoms include slowness of voiding, both initially and throughout urination. Causes of obstructions can include neoplasms (e.g., prostate/bladder carcinoma or lymph node tumors constricting ureters), acquired diseases (e.g., urethral strictures or renal calculi), and congenital deformities of the lower urinary tract. The clinical symptoms of advancing obstructive disease include decreased urinary concentrating capability, diminished metabolic acid excretion, decreased GFR, and reduced renal blood flow. Laboratory tests useful in determining the nature of the blockage are urinalysis, urine culture, BUN, serum creatinine, and complete blood count (CBC). Final diagnosis is usually made by radiologic imaging techniques.[5,8]

Renal Calculi

Renal calculi, or kidney stones, are formed by the combination of various crystallized substances, which are listed in **Table 21.2**. Of these, calcium oxalate stones are by far the most commonly encountered.

It is currently believed that recurrence of calculi in susceptible individuals is a result of several causes but mainly a reduced urine flow rate (related to a decreased fluid intake) and saturation of the urine with large amounts of essentially insoluble

Table 21.2 Types of Kidney Stones

Stone Composition	Cause of Stone Formation
Calcium oxalate	Hyperparathyroidism
	High urine calcium
	Vitamin D toxicity
	Sarcoidosis
	Osteoporosis
Magnesium ammonium phosphate	Infectious processes
Calcium phosphate	Excess alkali consumption
	Infection with urease-producing organisms
Uric acid	Gout
	High levels of uric acid in blood and urine
Cystine	Inherited cystinuria

substances. Chemical analysis of stones is important in determining the cause of the condition. Specialized x-ray diffraction and infrared spectroscopy techniques are widely used for this purpose. Clinical symptoms are, of course, similar to those encountered in other obstructive processes: hematuria, urinary tract infections, and characteristic abdominal pain.[4,7,16]

Renal Failure

Acute Kidney Injury

Acute kidney injury (AKI) is a sudden, sharp decline in renal function because of an acute toxic or hypoxic insult to the kidneys. AKI is a common and serious condition that occurs in approximately 50% of patients admitted to the intensive care unit.[18] The risk, injury, failure, loss of function, end-stage renal disease (RIFLE) and the Acute Kidney Injury Network definitions of AKI are based on changes in both serum creatinine and urinary output; however, both display poor specificity and sensitivity for the early detection of AKI.[19] Novel urine and plasma biomarkers, such as NGAL and NephroCheck (TIMP-2 and IGFBP-7), have emerged as excellent biomarkers for the early prediction and prognosis of AKI. AKI is subdivided into three types, depending on the location of the precipitating defect:

- **Prerenal AKI**: The defect lies in the blood supply before it reaches the kidney. Causes can include cardiovascular system failure and consequent hypovolemia.
- **Intrinsic AKI**: The defect involves the kidney. The most common cause is acute tubular necrosis; other causes include vascular obstructions/inflammations and glomerulonephritis.
- **Postrenal AKI**: The defect lies in the urinary tract after it exits the kidney. Generally, acute renal failure occurs as a consequence of lower urinary tract obstruction or rupture of the urinary bladder.

CASE STUDY 21.1, PART 4

Remember George. His condition continued to deteriorate, and he died 5 days after admission. Cause of death was multiorgan failure secondary to AIDS, sepsis, and alcoholic cirrhosis.

Test	Patient Results	Reference Range
Alkaline phosphatase	443 U/L	45–122
Aspartate aminotransferase	305 U/L	9–45
Alanine aminotransferase	78 U/L	8–63
Gamma glutamyl transpeptidase	724 U/L	11–50
Total bilirubin	2.7 mg/dL	0.2–1.0
Direct bilirubin	2.4 mg/dL	0–0.2

5. How would you interpret this patient's liver function tests, considering his clinical history?

Toxic insults to the kidney that are severe enough to initiate AKI include hemolytic transfusion reactions, myoglobinuria due to rhabdomyolysis, heavy metal/solvent poisonings, antifreeze ingestion, and analgesic and aminoglycoside toxicities. These conditions directly damage the renal tubules. Hypoxic insults include conditions that severely compromise renal blood flow, such as septic or hemorrhagic shock, burns, and cardiac failure. The most commonly observed symptoms of acute renal failure are oliguria and anuria (<400 mL/d). The diminished ability to excrete electrolytes and water results in a significant increase in extracellular fluid volume, leading to peripheral edema, hypertension, and congestive heart failure. Most prominent, however, is the onset of the uremic syndrome or kidney failure, in which increased BUN and serum creatinine values are observed along with the preceding symptoms. The outcome of this disease is either recovery or, in the case of irreversible renal damage, progression to chronic renal failure.[4,7]

Chronic Kidney Disease

CKD is a clinical syndrome that occurs when there is a gradual decline in renal function over time (**Figure 21.7**). It is defined by the presence of kidney damage or decreased kidney function for 3 months

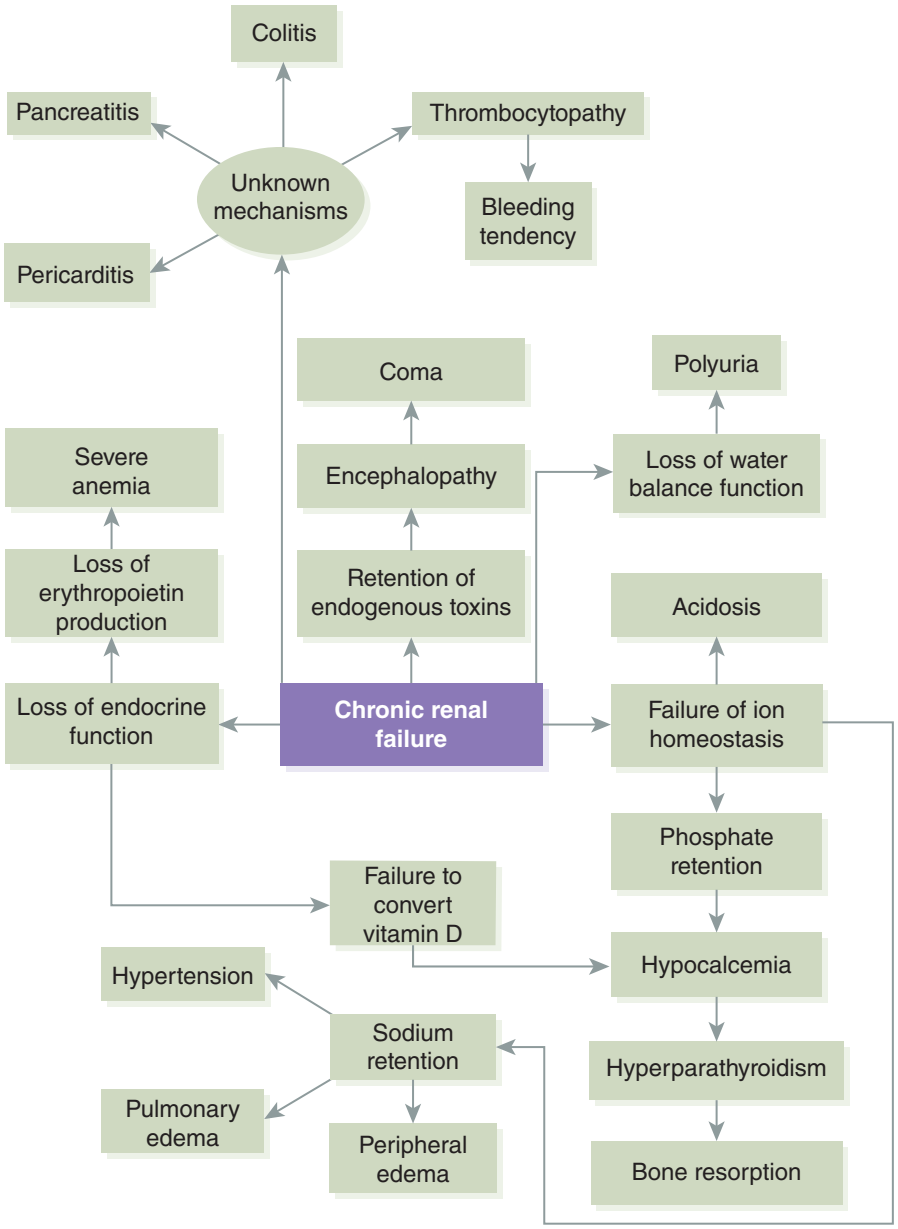


Figure 21.7 Pathophysiology of chronic kidney disease.

Table 21.3 Systematic Classification of CKD Stages

Stage	Description	GFR (mL/min per 1.73 m^2)
1	Kidney damage with normal or ↑GFR	>90
2	Kidney damage with normal or ↓GFR	60–89
3A	Moderate ↓ GFR	45–59
3B	Moderate to severe ↓ GFR	30–44
4	Severe ↓ GFR	15–29
5	Kidney failure	<15
5D	End-stage renal disease; patients undergoing chronic dialysis	
5T	End-stage renal disease; patients who have undergone kidney transplantation	

GFR, glomerular filtration rate; CKD, chronic kidney disease. Stages 1 to 5 illustrate the progression of CKD.

Data from National Kidney Foundation. *K/DOQI Clinical Practice Guidelines for Chronic Kidney Disease: Executive Summary*. New York, NY: National Kidney Foundation; 2002:16.

or more, irrespective of the cause. According to the Centers for Disease Control and Prevention, 1 in 10 U.S. adults has some level of CKD. Early detection and treatment are needed to prevent progression to kidney failure and complications such as coronary vascular disease. The National Kidney Foundation formulated guidelines for earlier diagnosis, treatment, and prevention of further disease progression in 2002. See **Table 21.3** for the five stages of CKD. Since the original classification was published, stage 3 CKD has been subdivided into GFR stages 3A and 3B to reflect the continuous association more accurately between lower GFR and risk for mortality and adverse kidney outcomes. Patients receiving treatment with dialysis are subclassified as GFR stage 5D to highlight the level of care required. Patients who have undergone kidney transplantation are subclassified as stage 5T. GFR and evidence of kidney damage based on measurement of proteinuria or other markers form the basis of the classifications.[20]

In recent years, albuminuria staging has been added to eGFR staging due to the graded increase in risk for mortality, progression of CKD, and end-stage renal disease at higher levels of albuminuria, independent of eGFR. The three stages include A1 (ACR < 30 mg/g) optimal to mild increase with increased risk between being associated with levels between 10 and 29 mg/g; A2 (ACR 30 to 299 mg/g) moderate increase; and A3 (ACR ≥ 300 mg/g) severe increase.

The conditions that can precipitate AKI also may lead to chronic renal failure. Several other causes for this syndrome are listed in **Table 21.4**.

Increasing Incidence of CKD

There is an increasing incidence of CKD in the United States due to the increase in diabetes, the aging population, obesity, and metabolic syndrome. Diabetes mellitus can have profound effects on the renal system. According to the National Kidney Foundation, diabetes is the leading cause of CKD and accounts for 45% of people diagnosed with kidney failure each year (www.kidney.org). The effects are primarily glomerular, but they may affect all kidney structures as well and are theorized to be caused by the abnormally hyperglycemic environment that constantly bathes the vascular system.[4,7]

Typically, diabetes affects the kidneys by causing them to become glucosuric, polyuric, and nocturic. These states are caused by the heavy demands made on the kidneys to diurese hyperosmotic urine. In addition, a mild proteinuria (albuminuria) often develops between 10 and 15 years after the original diagnosis (see the section "Albuminuria"). Hypertension often manifests next, further exacerbating the renal damage. Eventually, chronic renal insufficiency or nephrotic syndrome may evolve, and each may be identified by their characteristic symptoms and laboratory findings. Early treatment of diabetes that focuses on tight control of blood glucose and

Table 21.4 Etiology of Chronic Renal Failure

Etiology	Examples
Renal circulatory diseases	Renal vein thrombosis, malignant hypertension
Primary glomerular diseases	Systemic lupus erythematosus, chronic glomerulonephritis
Renal sequelae to metabolic disease	Gout, diabetes mellitus, amyloidosis
Inflammatory diseases	Tuberculosis, chronic pyelonephritis
Renal obstructions	Prostatic enlargement, calculi
Congenital renal deformity	Polycystic kidneys, renal hypoplasia
Miscellaneous conditions	Radiation nephritis

prevention of high blood pressure may prolong the onset of chronic renal failure.

Aside from hypertension and diabetes, age is the key predictor of CDK. Due to the decline in fertility and increase in the average life span, the percentage of the population aged 65 years or older is projected to increase from 12.4% in 2000 to 19.6% in 2030, according to the U.S. Census Bureau. This continual rise, over the previous decade and those to come, contributes significantly to the increasing incidence of CKD.

Epidemiologic evidence links obesity to CKD and kidney failure. However, diabetes mellitus and hypertension have potential confounding roles because obesity is a risk factor for diabetes and hypertension, the two most common causes of CKD and kidney failure. Recent studies show that obesity itself increases the risk of kidney injury.[21] As an individual gains weight, the nephron number remains the same; however, the GFR increases to meet the higher metabolic demands, which results in damage to the kidney.

The metabolic syndrome, characterized by the presence of at least three of the following risk factors—abdominal obesity, hypertension, low high-density lipoprotein (HDL) cholesterol level, hypertriglyceridemia, or hyperglycemia—is a prevalent disorder in the United States. In a population study of a representative sample of the U.S. general population, the risk of CKD and albuminuria increased progressively with a greater number of components of metabolic syndrome.[22] Individuals with metabolic syndrome had a 2.6-fold increased risk of developing CKD compared with individuals without metabolic syndrome.[22] Interventions that target biochemical components of metabolic syndrome may reduce the risk of CDK.

Renal Hypertension

Renal disease–induced hypertension can be caused by decreased perfusion to all or part of the kidney (ischemia). Lack of perfusion may be caused by traumatic injury or narrowing of an artery or intrarenal arterioles. Chronic ischemia of any kind results in nephron dysfunction and eventual necrosis. The resulting changes in blood and body fluid volumes within the kidney trigger the activation of the renin–angiotensin–aldosterone system, setting off vasoconstriction that is manifested as persistent hypertension.

Renal hypertension can be evaluated by monitoring serum aldosterone, Na^+, and renin levels. As a result of the effect of aldosterone, there will be increased serum Na^+, decreased serum K^+, and increased urine K^+.

Therapy for Acute Kidney Injury

Dialysis. In patients with AKI, uremic symptoms, uncontrolled hyperkalemia, and acidosis have traditionally been indications that the kidneys are unable to excrete the body's waste products and a substitute method in the form of dialysis was necessary. Dialysis is often instituted before this stage, however. Several forms of dialysis are available; however, they

CASE STUDY 21.2, PART 2

Remember Arthur. The ED physician ordered a chemistry panel, arterial blood gas (ABG), prothrombin time with INR, complete urinalysis, and a urine toxicology screen. Calcium oxalate crystals were found in his urine (pH 5). His urine toxicology screen was negative. The panel results are:

© brusinski/iStock/Getty Images Plus/Getty Images.

Analyte	Patient Value	Reference Range	Analyte	Patient Value	Reference Range
Na^+	139 mmol/L	136–145 mmol/L	Osmolality (serum)	416 mOsm/kg	
K^+	4.3 mmol/L	3.5–4.3 mmol/L	Blood gases (arterial)		
Cl^-	94 mmol/L	98–107 mmol/L	pH	7.26	7.35–7.45
CO_2	10 mmol/L	23–29 mmol/L	pCO_2	13 mmHg	35–45 mmHg
BUN	9 mg/dL	8–20 mg/dL	pO_2	129 mmHg	80–100 mmHg
Creatinine	2.2 mg/dL	0.5–1.0 mg/dL	HCO_3^-	18 mmol/L	22–26 mmol/L
Glucose	167 mg/dL	fasting 70–99 mg/dL	Prothrombin time	12.8 secs	10–13 secs
Ca^{2+}	9.9 mg/dL	8.6-10.2 mg/dL	International normalized ratio (INR)	1.1	1.1 or less
Albumin	4.5 g/dL	3.5–5.5 g/dL			

Treatment with fomepizole (4-MP) was initiated to prevent further metabolism of ethylene glycol. Arthur became agitated and was intubated for airway protection and placed on a propofol (sedative–hypnotic) drip. Nephrology was called, and hemodialysis was started. Laboratory test results the morning after the first hemodialysis are listed below. The initial ethylene glycol level from time of presentation to the ED was received at this time and is included.

Analyte	Patient Value	Reference Range	Analyte	Patient Value	Reference Range
Na^+	143 mmol/L	136–145 mmol/L	Albumin	3.1 g/dL	3.5–5.5 g/dL
K^+	4.1 mmol/L	3.5–4.3 mmol/L	Mg^{2+}	1.6 mg/dL	1.5–2.4 mg/dL
Cl^-	108 mmol/L	98–107 mmol/L	CK	405 U/L	24–204 U/L
CO_2	18 mmol/L	23–29 mmol/L	Osmolality	364 mOsm/kg	
BUN	8 mg/dL	8–20 mg/dL	Blood gases (arterial)		
Creatinine	3.1 mg/dL	0.5–1.0 mg/dL	pH	7.40	7.35–7.45
Glucose	194 mg/dL	fasting 70–99 mg/dL	pCO_2	26 mmHg	35–45 mmHg
Ca^{2+}	7.7 mg/dL	8.6–10.2 mg/dL	pO_2	192 mmHg	80–100 mg Hg
Ethylene glycol	499.4 mg/dL	None Detected	HCO_3^-	15 mmol/L	22–26 mmol/L

The patient remained intubated and sedated. He continued to receive hemodialysis two times a day for 4 hours at a time. The patient also continued to receive fomepizole every 12 hours and following each round of hemodialysis. This treatment continued until the osmolar and anion gaps were normalized.

1. What caused the formation of the calcium oxalate crystals? What is the clinical significance of calcium oxalate crystal formation?
2. Calculate the anion and osmolar gaps for both sets of laboratory results. In this case, what is the cause of the anion and osmolar gaps? How are these values clinically significant?
3. What are other common drugs and toxins that can cause acute renal failure?

all use a semipermeable membrane surrounded by a dialysate bath.

In traditional **hemodialysis** (removal of waste from blood), the membrane is synthetic and outside the body. Arterial blood and dialysate are pumped at high rates (150 to 250 and 500 mL/min, respectively) in opposite directions. The blood is returned to the venous circulation and the dialysate discarded. The diffusion of low-molecular-weight solutes (<500 Da) into the dialysate is favored by this process, but mid-molecular-weight solutes (500 to 2 kDa) are inadequately cleared. Creatinine clearance is about 150 to 160 mL/min.

In peritoneal dialysis, the peritoneal wall acts as the dialysate membrane, and gravity is used to introduce and remove the dialysate. Two variations of this form are available: continuous ambulatory peritoneal dialysis (CAPD) and continuous cycling peritoneal dialysis; however, the process is continuous in both, being performed 24 hours a day, 7 days a week. This method is not as rigorous as the traditional method. Small solutes (e.g., potassium) have significantly lower clearance rates compared with the traditional method, but more large solutes are cleared and steady-state levels of blood analytes are maintained.

Continuous arteriovenous **hemofiltration** (ultrafiltration of blood), continuous venovenous hemofiltration, continuous arteriovenous hemodialysis, and continuous venovenous hemodialysis together make up the slow, continuous renal replacement therapies developed to treat AKI in critically ill patients in intensive care settings. In these methods, the semipermeable membrane is again outside the body. Solutes up to 5 kDa (the pore size of the membranes) and water are slowly (10 mL/min) and continuously filtered from the blood in the first two methods, causing minimal changes in plasma osmolality. Volume loss can be replaced in the form of parenteral nutrition and intravenous medications. The final two methods are similar to the filtration methods, but a continuous trickle of dialysis fluid is pumped past the dialysis membrane, resulting in continuous diffusion and a doubling of the urea clearance.

Therapy for Kidney Failure

For patients with irreversible renal failure, dialysis and transplantation are the only two therapeutic options. Initiation of either treatment occurs when the GFR falls to 5 mL/min (10 to 15 mL/min in patients with diabetic nephropathy).

Dialysis. Traditional hemodialysis, its more recent, high-efficiency form, and peritoneal dialysis are the available methods. The clinical laboratory used in conjunction with a hemodialysis facility must be able to adequately monitor procedural efficiency in a wide variety of areas. Renal dialysis has basic goals, and specific laboratory tests should be performed to evaluate the achievement of each goal.

Transplantation. The most efficient hemodialysis techniques provide only 10% to 12% of the small solute removal of two normal kidneys and considerably less removal of larger solutes. Even patients who receive regular dialysis have physical disabilities and decreased quality of life. Kidney transplantation offers the greatest chance for full return to a healthy, productive life. However, this option is limited by the significant shortage of donor organs. For kidney failure patients, waiting for an organ donation can vary from several months to several years.

Renal transplantation removes a kidney from a compatible donor and transplants it into a recipient suffering from irreversible renal failure. The donor can be a cadaver or a live individual (80% and 20%, respectively, of all kidney transplants in the United States). For this procedure to be successful, the body's immune response to the transplanted organ must be suppressed. Therefore, the donor and recipient are carefully screened for ABO blood group, human leukocyte antigen (HLA) compatibility, and preformed HLA antibodies. The HLA system is the major inhibitor to transplantation.

Although kidney transplants have the capacity to function for decades, the mean half-life of a cadaveric transplant is approximately 7 years. The mortality rate is not significantly different from hemodialysis. Three-year graft survival figures vary from 65% to 85%, with live grafts doing better. It has been reported that there is no difference in patient survival among hemodialysis, CAPD, and cadaveric kidney transplantation. Live related donor transplantation is associated with a better patient survival than other therapeutic options for kidney failure.

CASE STUDY 21.3, PART 2

Remember Geraldine. Her daughter told the hospitalist she had been treated 3 weeks before at the hospital for a urinary tract infection. Geraldine was admitted to the hospital for a diagnostic workup and transfusion. Her admitting laboratory results were:

© cobalt/iStock/Getty Images Plus/Getty Images.

Analyte	Patient Value	Reference Range	Analyte	Patient Value	Reference Range
Na^+	129 mmol/L	136–145 mmol/L	Hct	25.6%	36–48 %
K^+	3.7 mmol/L	3.5–4.3 mmol/L	Hgb	8.5 g/dL	12–16 g/dL
Cl^-	97 mmol/L	98–108 mmol/L	WBC	9700	4500–10,000
CO_2	19 mmol/L	23–29 mmol/L			
BUN	52 mg/dL	8–20 mg/dL			
Creatinine	3.2 mg/dL	0.5–1.0 mg/dL			

Urine culture was positive for *Citrobacter*. Urinalysis results are listed:

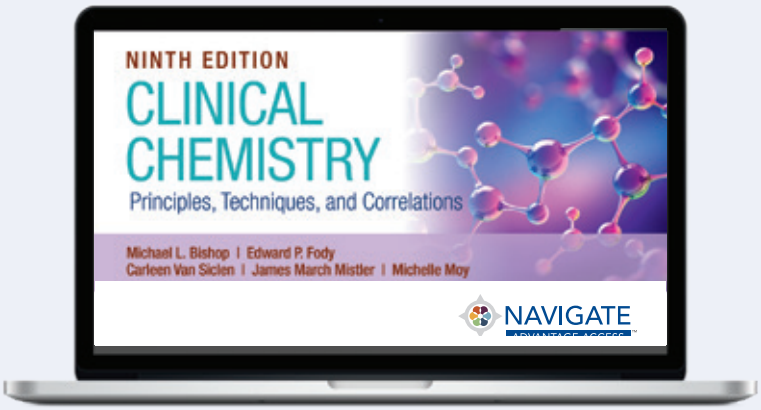
Analyte	Patient Value	Reference Range
Color	Yellow/Cloudy	(Straw/yellow) (Clear)
Specific gravity	1.015	1.003–1.030
pH	5	5–9
Blood	Large	Negative
Protein	2	Negative/Trace
Glucose	Negative	Negative
Ketones	Negative	Negative
Nitrites	Negative	Negative
RBC	>25	0-2/HPF
WBC	1–4	0-5/HPF
Casts	Granular, 1–4	

The patient's renal function continued to decline, and she was put on hemodialysis. A renal biopsy was performed that showed end-stage crescentic glomerulonephritis. Two days later, the patient sustained a perforated duodenal ulcer, which required surgery and blood transfusion. Subsequently, she developed coagulopathy and liver failure. Her condition continued to deteriorate in the next few days, and she died following removal of life support.

1. Looking at the UA, what is the significance of the results of 2+ protein and greater than 25 RBCs?
2. What is the most likely cause of glomerulonephritis?
3. Why was the patient put on hemodialysis?

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Vander A, Sherman JH, Luciano DS. *Human Physiology: The Mechanisms of Body Function*. 7th ed. New York, NY: McGraw-Hill; 1998:503, 508, 519.
2. Kaplan A, Szabo LL, Jack R, et al. *Clinical Chemistry: Interpretation and Techniques*. 4th ed. Baltimore, MD: Williams & Wilkins; 1995:156–157.
3. Davies A, Blakely A, Kidd C. *Human Physiology*. London: Harcourt Publishers; 2001:747.
4. Rock RC, Walker WG, Jennings CD. Nitrogen metabolites and renal function. In: Tietz NW, ed. *Fundamentals of Clinical Chemistry*. 3rd ed. Philadelphia, PA: WB Saunders; 1987:669.
5. Russell PT, Sherwin JE, Obernolte R, et al. Nonprotein nitrogenous compounds. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 2nd ed. St. Louis, MO: CV Mosby; 1989:1005.
6. Fraser D, Jones G, Kooh SW, et al. Calcium and phosphate metabolism. In: Tietz NW, ed. *Fundamentals of Clinical Chemistry*. 3rd ed. Philadelphia, PA: WB Saunders; 1987:705.
7. First MR. Renal function. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 4th ed. St. Louis, MO: CV Mosby; 2003:477–491.
8. Kleinman LI, Lorenz JM. Physiology and pathophysiology of body water and electrolytes. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 4th ed. St. Louis, MO: CV Mosby; 2003:441–461.
9. Sherwin JE, Bruegger BB. Acid–base control and acid–base disorders. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 4th ed. St. Louis, MO: CV Mosby; 2003:462–476.
10. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612.
11. Lab Tests Online. Cystatin C at a glance. Accessed from http://www.labtestsonline.org/understanding/analytes/cystatin_c/glance.html August 27, 2021.
12. Stevens LA, Coresh J, Schmid CH, et al. Estimated GFR using serum cystatin C alone and in combination with serum creatinine: a pooled analysis of 3418 individuals with CKD. *Am J Kidney Dis*. 2008;51:395–406.
13. Schwartz GJ, Munoz A, Schneide MF, et al. New equations to estimate GFR in children with CKD. *J Am Soc Nephrol*. 2009;20:629–637.
14. Fraser CG. *Biological Variation: From Principles to Practice*. Washington, DC: AACC Press; 2001.
15. Haase M, Bellomo R, Haase-Fielitz A. Neutrophil gelatinase associated lipocalin. *Curr Opin Crit Care*. 2010;16:526–532.
16. Kashani K, Al-Khafaji A, Ardiles T, et al. Discovery and validation of cell cycle arrest biomarkers in human acute kidney injury. *Crit Care*. 2013;17:R25.
17. Kaplan LA, Pesce AJ. Examination of urine. In: Kaplan LA, Pesce AJ, eds. *Clinical Chemistry: Theory, Analysis, and Correlation*. 4th ed. St. Louis, MO: CV Mosby; 2003: 1092–1109.
18. Mandelbaum T, Scott DJ, Lee J, et al. Outcome of critically ill patients with acute kidney injury using the Acute Kidney Injury Network criteria. *Crit Care Med*. 2011;39(12):2659–2664.
19. Mehta RL, Kellum JA, Shah SV, et al. Acute Kidney Injury Network: report of an initiative to improve outcomes in acute kidney injury. *Crit Care*. 2007;11:R31.
20. Mitchum C. Implementing the new kidney disease testing guidelines. *Clin Lab News*. 2002 September 14.
21. Kramer H, Luke A. Obesity and kidney disease: a big dilemma. *Curr Opin Nephrol Hypertens*. 2007;16:237–241.
22. Chen J, Muntner P, Hamm LL, et al. The metabolic syndrome and chronic kidney disease in U.S. adults. *Ann Intern Med*. 2004;140:167–174.

CHAPTER 22

Pancreatic Function

Gerald D. Redwine, John E. Lee Sang, and Amanda Rivera-Begeman

CHAPTER OUTLINE

KEY TERMS

Carotenoids
Cholecystokinin (CCK)
D-xylose
Fecal elastase
Gastrin
Islets of Langerhans
Lactose
Pancreatitis
Secretin
Steatorrhea
Zollinger-Ellison syndrome

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

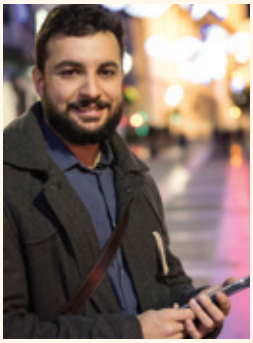
- Discuss the physiologic role of the pancreas in the digestive process.
- List the hormones excreted by the pancreas, together with their physiologic roles.
- Describe the following pancreatic disorders and list the associated laboratory tests that would aid in diagnosis: acute pancreatitis, chronic pancreatitis, pancreatic carcinoma, cystic fibrosis, and pancreatic malabsorption.
- List the tests used to assess intestinal function.
- Evaluate a patient's condition, given clinical data.

The gastrointestinal (GI) system is composed of the mouth, esophagus, stomach, small intestine, and large intestine. Digestion, which is primarily a function of the small intestine, is the process by which starches, proteins, lipids, nucleic acids, and other complex molecules are degraded to simple constituents (molecules) for absorption and use in the body. This chapter discusses the physiology of the pancreas, intestinal physiology, pathologic aspects of intestinal function, and tests of pancreatic and intestinal function.

The pancreas is a large gland that is involved in the digestive process but located outside of the gastrointestinal (GI) system. It is composed of both endocrine and exocrine tissues. The liver is the other major external gland that is involved in the digestive process, and it is covered in Chapter 19, *Liver Function*. The *endocrine* functions of the pancreas

CASE STUDY 22.1, PART 1

Aran, a 38-year-old man, entered the emergency department with the complaint of severe midabdominal pain of 6 hours' duration. A friend, who had driven him to the hospital, stated that Aran fainted as he was being helped into the automobile.



CASE STUDY 22.2, PART 1

Alejandro, a 56-year-old man who has an alcohol use disorder, presents with a 2-week history of midabdominal pain. He also describes clay-colored stools, mild icterus, nausea, vomiting, and a 10-lb weight loss.

CASE STUDY 22.3, PART 1

Helga and Bjorn bring their 7-year-old son Sven to the pediatrician with the complaint of frequent fevers and failure to grow. Sven has had three bouts of pneumonia during the past 2 years and has been plagued by chronic bronchitis, which causes him to cough up copious amounts of thick, yellow, mucoid sputum. Despite a big appetite, he has gained only 2 lbs. in the past 2 years and remains short in stature. Sven especially likes salty foods, and usually has three or four bulky, foul-smelling bowel movements daily. Helga and Bjorn say Sven's 9-year-old sister is in excellent health.

include the production of insulin and glucagon; both hormones are involved in carbohydrate metabolism. The *exocrine* function of the pancreas involves the production of enzymes used in the digestive process.

Physiology of Pancreatic Function

As a digestive gland, the pancreas weighs about 70 to 105 g and is second in size only to the liver. It is located behind the peritoneal cavity across the upper abdomen at about the level of the first and second lumbar vertebrae, about 1 to 2 inches above the umbilicus (**Figure 22.1**). It is located in the curve made by the duodenum (**Figure 22.2**). The liver and pancreas combine bile and enzymes that enter the alimentary tract through the ampulla of Vater, which is formed by the union of the pancreatic duct and the common bile duct at the major duodenal papilla. The combined entry is strategically located near the middle of the relatively short 20- to 30-cm duodenum. After exiting the antrum of the stomach, a bolus of food enters through the pylorus

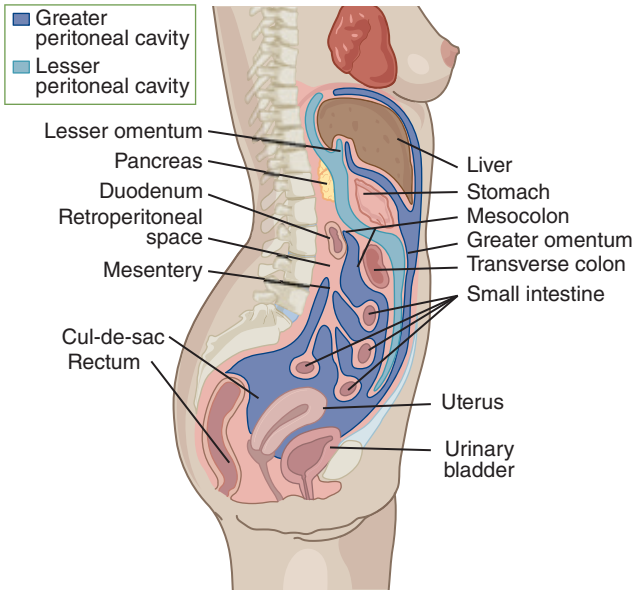


Figure 22.1 Peritoneum and mesenteries. The parietal peritoneum lines the abdominal cavity, and the visceral peritoneum covers abdominal organs. Retroperitoneal organs are covered by the parietal peritoneum. The mesenteries are membranes that connect abdominal organs to each other and to the body wall.

as a fibrous restrictor called the sphincter of Oddi relaxes in response to duodenal release of cholecystokinin (CCK).[1,2]

The pancreas is composed of two morphologically and functionally different tissues: endocrine tissue and exocrine tissue. The *endocrine* (hormone-releasing) component is by far the smaller of the two, comprising only 2% of the total tissue. It consists of the **islets of Langerhans**, which are well-delineated, spherical, or ovoid clusters composed of five known cell types (**Figure 22.3**). The islet cells, most with Greek alphabet cell names, secrete five hormones into the blood: alpha cells secrete glucagon, beta cells secrete insulin, delta cells secrete somatostatin, gamma cells secrete pancreatic polypeptide (PP) hormones, and epsilon cells secrete ghrelin.[3,4] These hormones are secreted into the surrounding blood vessels and carried to target organs or tissues. Researchers have a particular interest in growing the beta cells that produce insulin from stem cells, making future use of the resulting therapeutic insulin production from these cells a possibility. The larger, 98% *exocrine* pancreatic (enzyme-secreting) component secretes about 1.5 to 2 L/d of fluid, which is rich in digestive enzymes, into ducts that ultimately empty into the duodenum.

This digestive fluid is produced by pancreatic acinar cells (grape-like clusters), which line the pancreas and are connected by small ducts. These small ducts empty into progressively larger ducts, eventually forming one major pancreatic duct and a smaller accessory duct. The major pancreatic duct and the common bile duct open into the duodenum at the ampulla of Vater (hepatopancreatic duct) (Figure 22.2). Normal, protein-rich pancreatic fluid is clear, colorless, and watery, with an alkaline pH that can reach up to 8.3. This alkalinity is caused by the high concentration of sodium bicarbonate present in pancreatic fluid, which is used to neutralize the hydrochloric acid in gastric fluid from the stomach as it enters the duodenum. The bicarbonate and chloride concentrations vary reciprocally so that they total about 150 mmol/L. Pancreatic fluid has about the same concentrations of potassium and sodium as serum.

The digestive enzymes, or their proenzymes secreted by the pancreas, are capable of digesting the three major classes of food substances (proteins,

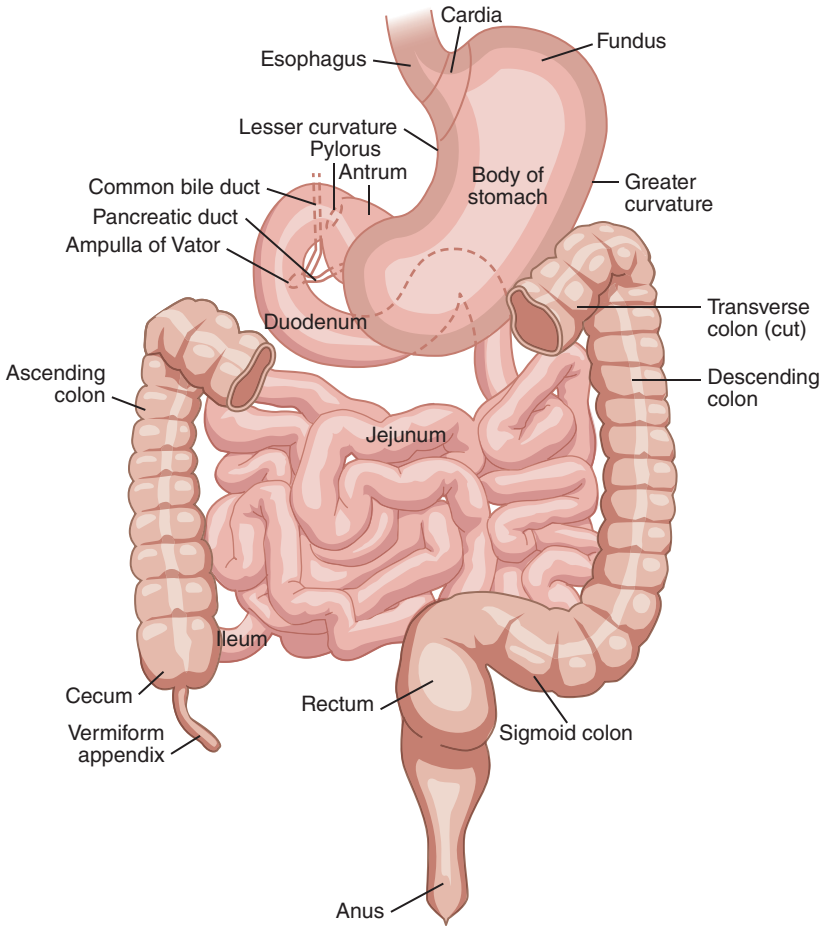


Figure 22.2 The abdominal structures of the alimentary tract.

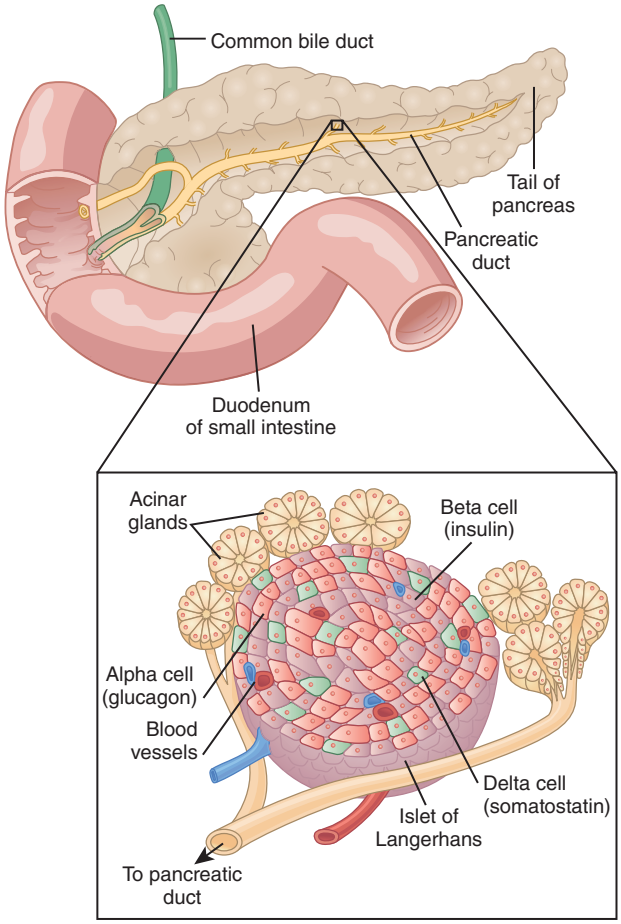


Figure 22.3 Diagram of the pancreas, its relationship to the duodenum, the islets of Langerhans (endocrine gland), and the acinar cells (exocrine glands).

carbohydrates, and fats) and include (1) the proteolytic enzymes trypsin, chymotrypsin, elastase, collagenase, leucine aminopeptidase, and some carboxypeptidases; (2) lipid-digesting enzymes, primarily lipase and lecithinase; (3) carbohydrate-splitting pancreatic amylase; and (4) several nucleases (ribonuclease), which separate the nitrogen-containing bases from their sugar-phosphate strands.

Pancreatic activity is under both nervous and endocrine control. Branches of the vagus nerve can cause a small amount of pancreatic fluid secretion when food is smelled or seen, and these secretions may increase as the bolus of food reaches the stomach. Most of the pancreatic action, however, is under the hormonal control of **secretin** and **cholecystokinin (CCK)**; formerly called *pancreozymin*). Secretin is responsible for the production of bicarbonate-rich and, therefore, alkaline pancreatic fluid, which protects the lining of the intestine from damage. Secretin is synthesized in response to the acidic contents of the stomach reaching the duodenum. It can also affect gastrin activity in the stomach. This pancreatic fluid contains few digestive enzymes. CCK, in the presence of fats or amino acids in the duodenum, is produced by the cells of the intestinal mucosa and is responsible for release of enzymes from the acinar cells by the pancreas into the pancreatic fluid.

Diseases of the Pancreas

Other than trauma, only three disease states cause more than 95% of the medical attention devoted to the pancreas. If they affect the endocrine function of the pancreas, these diseases can result in altered digestion and nutrient metabolism. The role of the pancreas in diabetes mellitus is discussed in Chapter 9, *Carbohydrates*.

Cystic fibrosis (known by various other terms, such as *fibrocystic disease of the pancreas* and *mucoviscidosis*) is an inherited autosomal recessive disorder characterized by dysfunction of mucous and exocrine glands throughout the body. The disease is most common in white people in the United States, occurring in about1 in 6000 live births in the United States.[5] It has various manifestations and can initially present in such widely varying ways as intestinal obstruction of the newborn, excessive pulmonary infections in childhood, or, uncommonly, as pancreatogenous malabsorption in adults. The disease causes the small and large ducts and the acini to dilate and convert into small cysts filled with mucus, eventually resulting in the prevention of pancreatic secretions reaching the duodenum or, depending on the age of the patient, a plug that blocks the lumen of the bowel, leading to obstruction. As the disease progresses, destruction and fibrous scarring of the pancreas increases with a corresponding decrease in function. Cystic fibrosis is transmitted as an autosomal recessive disorder with a high degree of penetrance. It occurs primarily in persons of northern European descent. The cystic fibrosis transmembrane conductance regulator (*CFTR)*gene occurs on chromosome 7, and more than 1500 mutations causing this disorder have been identified; however, some occur more commonly than others.[6] In areas of high frequency, such as Brittany in western France, more than 10% of the population may carry a cystic fibrosis genetic mutation, and 1 in 3000 infants may be affected, making it the most common genetic disorder in these populations, with a predicted increase of 45% in the combined population of infants and adults. Genetic screening is now expanded with international standards for newborn screening.[7-13]

Pancreatic carcinoma is the fourth most frequent form of fatal cancer and causes about 38,000 deaths each year in the United States, which represents about 7% of all deaths from malignant neoplasms. The disease is more common in males than in females and in Blacks/African Americans than in other races. The 5-year survival rate is about 6%, and most patients die within 1 year of diagnosis. Most pancreatic tumors arise as adenocarcinomas of the ductal epithelium. Because the pancreas has a rich supply of nerves, pain is a prominent feature of the disease. If the tumor arises in the body or tail of the pancreas, detection does not often occur until an advanced stage of the disease because of its central location and the associated vague symptoms. Cancer of the head of the pancreas is usually detected earlier because of its proximity to the common bile duct. Signs of these tumors are jaundice, weight loss, anorexia, and nausea. Jaundice is associated with signs of post-hepatic hyperbilirubinemia (intrahepatic cholestasis) and low levels of fecal bilirubin, resulting in clay-colored stools. However, findings are not specific for pancreatic tumors, and other causes of obstruction must be ruled out.

Islet cell tumors of the pancreas affect the endocrine capability of the pancreas. If the tumor occurs in beta cells, hyperinsulinism is often seen, resulting in low blood glucose levels, sometimes followed by hypoglycemic shock. Pancreatic cell tumors that overproduce gastrin are called *gastrinomas*; they cause **Zollinger-Ellison syndrome** and can be duodenal in origin. These tumors are associated with watery diarrhea, recurring peptic ulcer, and significant gastric hypersecretion and hyperacidity. Pancreatic cell tumors that secrete glucagon are rare; the hypersecretion of glucagon is associated with diabetes mellitus. Recommendations now include pancreatic polypeptide (PP) and other biochemicals in addition to glucagon and gastrin levels for gastrointestinal evaluations.[14,15]

Pancreatitis, or inflammation of the pancreas, is ultimately caused by autodigestion of the pancreas as a result of reflux of bile or duodenal contents into the pancreatic duct. Pathologic changes can include acute edema, with large amounts of fluid accumulating in the retroperitoneal space and an associated decrease in effective circulating blood volume; cellular infiltration, leading to necrosis of the acinar cells, with hemorrhage as a possible result of necrotic blood vessels; and intrahepatic and extrahepatic pancreatic fat necrosis. Pancreatitis is generally classified as acute (no permanent damage to the pancreas), chronic (irreversible injury), or relapsing/recurrent, which can also be acute or chronic. It commonly occurs in midlife. Painful episodes can occur intermittently, usually reaching a maximum within minutes or hours, lasting for several days or weeks, and frequently accompanied by nausea and vomiting. Pancreatitis is often associated with alcohol abuse or biliary tract diseases such as gallstones, but patients with hyperlipoproteinemia and those with hyperparathyroidism are also at a significantly increased risk for this disease.

Other etiologic factors associated with acute pancreatitis include mumps, obstruction caused by biliary tract disease, pancreatic tumors, tissue injury, atherosclerotic disease, shock, pregnancy, hypercalcemia, hereditary pancreatitis, immunologic factors associated with postrenal transplantation, and hypersensitivity. Symptoms of acute pancreatitis include severe abdominal pain that is generalized or in the upper quadrants and often radiates toward the back or down the right or left flank. The etiology of chronic pancreatitis is similar to that of acute pancreatitis, but chronic, excessive alcohol consumption appears to be the most common predisposing factor.

Laboratory findings include increased amylase, lipase, triglycerides, and hypercalcemia, which is often associated with underlying hyperparathyroidism. Hypocalcemia may be found and has been attributed to the sudden removal of large amounts of calcium from the extracellular fluid. This occurs due to either impaired mobilization or as a result of calcium fixation by fatty acids liberated by the increased lipase action on triglycerides. Hypoproteinemia is attributable mainly to the notable loss of plasma into the retroperitoneal space. A shift of arterial blood flow from the inflamed pancreatic cells to less affected or normal cells causes oxygen deprivation and tissue hypoxia in the area of damage, including the surrounding organs and tissue. A BUN level should also be monitored as it is an independent predictive marker in the severity of acute pancreatitis.

All three conditions can result in severely diminished pancreatic exocrine function, which can significantly compromise digestion and absorption of ingested nutrients. These conditions are the essence of the general malabsorption syndrome, which embodies abdominal bloating and discomfort; the frequent passage of bulky, malodorous feces; and weight loss. Failure to digest or absorb fats, known as **steatorrhea**, renders a greasy appearance to feces (more than 5 g of fecal fat per 24 hours). The malabsorption syndrome typically involves abnormal

digestion or absorption of proteins, polysaccharides, carbohydrates, and other complex molecules, as well as lipids. Severely deranged absorption and metabolism of electrolytes, water, vitamins (particularly fat-soluble vitamins A, D, E, and K), and minerals can also occur. Malabsorption can involve a single substance, such as vitamin B_{12}, which results in a megaloblastic anemia (pernicious anemia), or lactose caused by a lactase deficiency. In addition to pancreatic exocrine deficiency, the malabsorption syndrome can be caused by biliary obstruction, which deprives the small intestine of the emulsifying effect of bile, and various diseases of the small intestine, which inhibit absorption of digested products.

Tests of Pancreatic Function

Depending on etiology and clinical picture, exocrine pancreatic dysfunction may be suspected when there is evidence of increased amylase and lipase.[16-19] The reader is referred to Chapter 8, *Enzymes*, for an in-depth discussion of these enzymes. Pancreatic exocrine function can be measured directly or indirectly. Direct testing is the most sensitive approach and utilizes endoscopy or the Dreiling tube method after stimulation with secretin or cholecystokinin (CCK). The combined secretin-CCK test, typically performed at specialized centers, simultaneously assesses ductal and acinar secretory capacity. Indirect laboratory tests of exocrine pancreatic function include those used for detection of malabsorption (e.g., qualitative fecal fat analysis, fecal elastase, and fecal chymotrypsin) and the bilirubin test for extrahepatic biliary obstruction. Pancreatic endocrine-related tests (e.g., gastrin, insulin, and glucose) reflect changes in the endocrine cells of the pancreas.

The advantage of direct tests is that the chemical and cytologic examinations are performed on actual pancreatic secretions. Evaluation of pancreatic fluid after stimulation may include measurement of the total volume of pancreatic fluid and bicarbonate levels and enzymes. Morphologic examination of pancreatic fluid (i.e., cytologic evaluation) can often establish the presence, or at least the suspicion, of malignant neoplasms, although the precise localization of the primary organ of involvement (i.e., pancreas, biliary system, ampulla of Vater, or duodenum) is not possible by duodenal aspiration alone.

Due to the advances in imaging techniques such as MRI (magnetic resonance imaging) and MRCP (magnetic resonance cholangiopancreatography), stimulation tests for exocrine pancreatic function are used less often. None of these laboratory tests have proved especially useful in the diagnosis of mild or acute pancreatic disease in which the acute phase has subsided. Rather, most of these tests have found clinical utility in excluding the pancreas from diagnosis. The sweat test, used for screening cystic fibrosis, is not specific for assessing pancreatic involvement but, when used along with the clinical picture at the time of testing, can provide important diagnostic information. The following pancreatic function tests are reviewed briefly: secretin/CCK test, fecal fat analysis, fecal elastase, sweat chloride testing, and amylase and lipase interpretation.

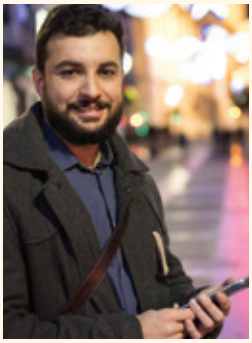
Cholecystokinin (CCK) Test

The hormone cholecystokinin (CCK), which is secreted by the cells in the stomach, stimulates the release of bile into the intestine and the secretion of enzymes by the pancreas. It has an important physiological role in the regulation of pancreatic secretion, gallbladder contraction, and intestinal motility. Another hormone that is important for pancreatic function is secretin, which secretes bicarbonate from the ductal cells of the pancreas. The CCK test is a direct determination of the exocrine secretory capacity of the pancreas. The test involves intubation of the duodenum without contamination by gastric fluid, which would neutralize any bicarbonate. The test is performed after a 6-hour or overnight fast. Pancreatic secretion is stimulated by intravenously administered secretin in a dose varying from 2 to 3 U/kg of body weight, followed by CCK administration.

Several protocols with varying dosing regimens have been established for this test. Pancreatic secretions are collected variously for 30, 60, or 80 minutes after administration of the stimulants, either as 10-minute specimens or as a single, pooled collection. The pH, secretory rate, enzyme activities (e.g., trypsin, amylase, or lipase), and amount of bicarbonate are determined. The average amount of bicarbonate normally excreted per hour is about 15 mmol/L for men and 12 mmol/L for women, with an average flow of 2 mL/kg. Assessment of enzymes must be taken in view of total volume output. Decreased pancreatic flow and increase in enzyme concentrations are associated with pancreatic obstruction. Low concentrations of bicarbonate and enzymes are associated with cystic fibrosis, chronic pancreatitis, pancreatic cysts, calcification, and edema of the pancreas.

CASE STUDY 22.1, PART 2

Remember Aran, who visited the emergency department. His laboratory results are shown below.



© FotoAndalucia/Shutterstock.

Analyte	Patient Value	Reference Range
Serum amylase	640 U/L	(30–220 U/L)
Serum sodium	133 mmol/L	(135–145 mmol/L)
Potassium	3.4 mmol/L	(3.8–5.5 mmol/L)
Calcium	8.8 mg/dL	(9.0–10.1 mg/dL)
Blood urea nitrogen	32 mg/dL	(6–20 mg/dL)

Based on these results, consider the following questions.

1. These laboratory results correlate with what disease?
2. What is the cause for the low serum calcium?
3. What is the cause for the increased blood urea nitrogen?

CASE STUDY 22.2, PART 2

Remember Alejandro, the 56-year-old man with alcohol use disorder. His physician received the following laboratory results.

© Morsa Images/DigitalVision/Getty Images.

Analyte	Patient Value	Reference Range
Bilirubin, Serum	4.2 mg/dL	0.2–1.0 mg/dL
Lactate	6.2 mmol/L	0.3–2.0 mmol/L
Alanine aminotransferase	76 IU/L	7–45 U/L
Alkaline phosphatase	462 IU/L	53–128 U/L
Amylase	80 IU/L	30–220 U/L
Bilirubin, Urine	3+	Negative

Considering the results, address the following questions.

1. What organ system is primarily involved?
2. What are the major diagnostic considerations?
3. What do the laboratory results reveal? What additional laboratory tests would be useful in establishing a diagnosis?
4. What other ancillary studies or procedures might be required?

Fecal Fat Analysis

Fecal lipids are derived from four sources: unabsorbed ingested lipids, lipids excreted into the intestine (predominantly in the bile), cells shed into the intestine, and metabolism of intestinal bacteria. Patients on a lipid-free diet still excrete 1 to 4 g of lipid in the feces in a 24-hour period. Even with a lipid-rich diet, the fecal fat does not normally exceed about 7 g in 24-hour period. Normal fecal lipid is composed of about 60% fatty acids; 30% sterols, higher alcohols, and carotenoids; 10% triglycerides; and small amounts of cholesterol and phospholipids. Although significantly increased fecal fat can be caused by biliary obstruction, severe steatorrhea is usually associated with exocrine pancreatic insufficiency or disease of the small intestine.

Qualitative Screening Test for Fecal Fat

Various screening tests have been devised for detecting steatorrhea. These tests commonly use fat-soluble stains (e.g., Sudan III, Sudan IV, Oil Red O, or Nile

blue sulfate), which dissolve in and color lipid droplets in order to visualize the fat droplets. Of greater importance than the particular technical procedure is the level of experience and dependability of the clinical laboratorian performing the test.

Sudan Staining for Fecal Fat

Neutral fats (triglycerides) and many other lipids stain yellow-orange to red with Sudan III because the dye is much more soluble in lipid than in water or ethanol.[20,21] Free fatty acids do not stain appreciably unless the specimen is heated in the presence of the stain with 36% acetic acid. The slide may be examined warm or cool and the number of fat droplets assessed. As the slide cools, the fatty acids crystallize out in long, colorless, needle-like sheaves. Detection of meat fiber is accomplished by a third aliquot of fecal sample mixed on a slide with 10% alcohol and a solution of eosin, stained for 3 minutes. The meat fiber should stain as rectangular cross-striated fibers. Splitting the sample and detecting neutral fats, fatty acids, and undigested meat fibers can provide diagnostic information. Increases in fats and undigested meat fibers are indicative of patients with steatorrhea of pancreatic origin. A representative fecal specimen is used for analysis.

Normal feces can have up to 40 or 50 small (1 to 5 mm) neutral lipid droplets per high-power microscope field. Steatorrhea is characterized by an increase in the number and size of stainable droplets, often with some fat globules in the 50 to 100-mm range. Fatty acid assessment greater than 100 stained, small droplets, along with the presence of meat fiber, is expected in patients with steatorrhea.

Quantitative Fecal Fat Analysis

The definitive test for steatorrhea is the quantitative fecal fat determination, usually on a 72-hour stool collection, although the collection period may be increased to up to 5 days. Traditional methods for fecal fat determination are the gravimetric and titrimetric methods. Newer methods involve the use of infrared and nuclear magnetic resonance spectroscopy.[22,23] In the gravimetric method, fatty acid soaps (predominantly calcium and magnesium salts of fatty acids) are converted to free fatty acids, followed by extraction of most of the lipids into an organic solvent, which is then evaporated so that the lipid residue can be weighed. In titrimetric methods, lipids are saponified with hydroxide, and the fatty acid salts are converted to free fatty acids using acid. The free fatty acids, along with various unsaponified lipids, are then extracted with an organic solvent, and the fatty acids are titrated with hydroxide after evaporation of the solvent and redissolving of the residue in ethanol. The titration methods measure only saponifiable fatty acids and consequently, render results about 20% lower than those from gravimetric methods. A further objection is that titrimetric methods use an assumed average molecular weight for fatty acids to convert moles of fatty acids to grams of lipid.

At one time, it was common to measure the amount of free fatty acids as a percentage of total lipids on the presumption that a high percentage of free fatty acids indicates adequate pancreatic lipase activity. This method is no longer considered reliable because of spurious results, particularly caused by lipase produced by intestinal bacteria.

It is essential that patients be placed on a lipid-rich diet for at least 2 days before instituting the fecal collection. The diet must contain at least 50 g, and preferably 100 g, of lipids each day. Fecal collections should extend for 3 or more successive days.

There are various ways to express fecal lipid excretion. Expressing lipid excretion as a percentage of wet or dry fecal weight is open to serious challenge because of wide variations in both fecal water content and dry residue as a result of dietary intake. The most widely accepted approach is to report the grams of fecal fat excreted in a 24-hour period.

Gravimetric Method for Fecal Fat Determination

The entire fecal specimen is emulsified with water. An aliquot is acidified to convert all fatty acid soaps to free fatty acids, which are then extracted with other soluble lipids into petroleum ether and ethanol. After evaporation of the organic solvents, the lipid residue is weighed. All feces for a 3-day period are collected in tared containers. The containers *must not* have a wax coating. The specimen must be kept refrigerated.

Total lipid does not change significantly during 5 days' storage of the specimen at refrigerator temperatures. Patients must not ingest castor oil, mineral oil, or other oily laxatives and must not use rectal suppositories containing oil or lipid for 2 days before and during the test. The reference range for fecal lipids in adults is 1 to 7 g per 24 hours.

Fecal Enzymes

More than 90% of exocrine pancreatic function must be lost before a 72-hour fecal fat test becomes positive, but it is still helpful in the context of monitoring

the degree of fat malabsorption. Alternatively, a noninvasive indirect indicator of moderate and severe exocrine pancreatic dysfunction is testing for **fecal elastase**. Fecal elastase-1 is a chymotrypsin-like enzyme that is secreted by the pancreas. It remains stable during transport through the gastrointestinal tract and is deposited into the feces. In exocrine pancreas insufficiency cases, the fecal fat is increased, and the fecal enzymes are decreased. Elastase-1 is secreted by the exocrine pancreas and passes unmodified in the feces. Therefore, there is a direct correlation between the amount of pancreatic elastase-1 and fecal elastase-1.[24]

The reference range for normal fecal elastase is greater than 200 μg/g. A fecal elastase of 100–200 μg/g suggests moderate pancreatic insufficiency, and a fecal elastase less than 100 μg/g suggests severe pancreatic insufficiency. Fecal chymotrypsin is another fecal enzyme that can be measured, but it is less sensitive and specific than fecal elastase at detecting pancreatic insufficiency and can be affected by pancreatic enzyme supplements.

Sweat Electrolyte Testing

In addition to molecular testing for the *CFTR* gene, the measurement of chloride concentration in sweat is still a useful laboratory test for the diagnosis of cystic fibrosis, though it is a sophisticated and difficult test offered in a limited number of medical centers.[25-27] Patients with cystic fibrosis have reduced ability to resorb chloride from sweat. Significantly elevated concentrations of both ions occur in more than 99% of affected patients. The two- to five-fold increases in sweat sodium and chloride are diagnostic of cystic fibrosis in children. Even in adults, no other condition causes an increase in sweat chloride and sodium above 80 mmol/L. Sweat potassium is also increased, but less significantly so, and is not generally relied on for diagnosis. Contrary to some assertions, sweat electrolyte determinations do not distinguish heterozygote carriers of cystic fibrosis from normal homozygotes.

Older methods for acquiring sweat specimens required skilled laboratorians who frequently performed the test. Methods of inducing sweating included applying plastic bags or wrapping the patient in blankets, which was fraught with serious risks of dehydration, electrolyte disturbances, and hyperpyrexia. In 1959, pilocarpine administration by iontophoresis was reported as an efficient method for sweat collection and stimulation. Iontophoresis uses an electric current that causes pilocarpine to migrate into a limited skin area, usually the inside of the forearm, toward the negative electrode from a moistened pad on the positive electrode. A collection vessel is then applied to the skin. The sweat is then analyzed for chloride. For confirmation, the test should be repeated. Commercially available surface electrodes that analyze the sweat chloride are readily available. For details, the reader is referred to Chapter 23, *Body Fluid Analysis*.

It is widely accepted that sweat chloride concentrations greater than 60 mmol/L are diagnostic of cystic fibrosis in children. Sweat sodium and chloride concentrations in female patients undergo fluctuation with the menstrual cycle and reach a peak 5 to 10 days before the onset of menstruation but do not overlap with the ranges associated with cystic fibrosis.

Cystic fibrosis is due to the dysfunction of the chloride transporter (*CFTR* gene product). Most of the time, interpretation of sweat testing is straightforward; however, a spectrum of disease is now being recognized with over 1500 known mutations. From these, the American College of Medical Genetics/American College of Obstetricians and Gynecologists (ACMG/ACOG) recommend a DNA panel of 23 mutations proven effective in screening for carriers of the CFTR gene for genetic counseling.[6,28] In addition, the fecal elastase test is also sensitive to pancreatic insufficiency due to cystic fibrosis.

Serum Enzymes

Lipase is the recommended enzyme for detecting pancreatic disease over amylase of the past or a combination, with cost, sensitivity, and failure to increase the diagnostic value being cited as the reasons.[29,30] Unlike amylase, lipase is almost exclusively a pancreatic function test. Reliable analytic methods make lipase more desirable because it is more sensitive, more specific, and more accurate than amylase. Lipase increases in serum within 24 hours in acute pancreatitis, persists for 8 to 14 days from reabsorption, and is not cleared like amylase. Consequently, some physicians consider lipase more sensitive than amylase as an indicator of acute pancreatitis or other causes of pancreatic necrosis.[18,31]

Amylase, however, remains an alternative, useful in the diagnosis of acute pancreatitis, in which significant increases in serum concentrations occur in about 75% of patients. Typically, amylase in serum increases within 3 to 6 hours of the onset of the disease, reaches a peak in about 24 hours, and

CASE STUDY 22.3, PART 2

Remember Sven, the 7-year-old who had bouts of pneumonia and chronic bronchitis. Evaluate Sven based on the following questions.

1. What disease do the symptoms described above correlate with?
2. What additional laboratory test should be ordered, and what are the expected results?
3. What other clinical laboratory tests would likely be abnormal?

because of its clearance by the kidneys, returns to normal within 3 to 5 days, often making urine amylase a more sensitive indicator of acute pancreatitis. The magnitude of the enzymes' elevation cannot be correlated with the severity of the disease.

Determination of the renal clearance of amylase is useful in detecting minor or intermittent increases in the serum concentration of this enzyme. To correct for diminished glomerular function, the most useful expression is the ratio of amylase clearance to creatinine clearance, as follows:

$$\frac{\%\ \text{Amylase}}{\text{Creatinine clearance ratio}} = 100 \times \frac{\text{UA}}{\text{SA}} \times \frac{\text{SC}}{\text{UC}} \quad \text{(Eq. 22.1)}$$

where UA is urine amylase, SA is serum amylase, SC is serum creatinine, and UC is urine creatinine.

Reference values are less than 3.1%. Significantly increased values, averaging about 8% or 9%, occur in acute pancreatitis but may also occur in other conditions, such as burns, sepsis, and diabetic ketoacidosis.

Both amylase and lipase may be significantly increased in serum in many other conditions (e.g., opiate administration, pancreatic carcinoma, intestinal infarction, obstruction or perforation, and pancreatic trauma). Amylase levels are also frequently increased in mumps, cholecystitis, hepatitis, cirrhosis, ruptured ectopic pregnancy, and macroamylasemia, which is a benign condition in which amylase binds to an immunoglobulin molecule, causing chronic elevation of serum amylase values but normal urine amylase levels. Lipase levels are often significantly increased in bone fractures and in association with fat embolism. Of note, amylase is also measured in saliva and pancreatic cyst fluid.

Tests of Gastric Secretion and Intestinal Function

Gastrin is a peptide hormone that enhances gastric growth, gastric motility, and secretion of hydrochloric acid.[8,32] It is present in G cells of the gastric antrum and duodenum. Gastric analysis begins with a 1-hour basal level following an overnight fast, then stimulation with 6 μg/kg pentagastrin subcutaneous administration, and subsequent a 1-hour test.[33] Alternately, secretin stimulation for serum gastrin can be ordered. The hormone, secretin, stimulates secretion of bicarbonate from the ductal cells of the pancreas. Patients are instructed to fast for 8 hours prior to this test. Serum gastrin measurements are obtained at baseline, as well as 2 minutes, 5 minutes, 10 minutes, 20 minutes, and 30 minutes after secretin stimulation.

Although several conditions may cause increased gastrin (such as atrophic gastritis), plasma gastrin testing is typically performed to diagnose Zollinger-Ellison syndrome (ZES), a gastrinoma, exhibiting increased stomach acid leading to peptic ulcers.[34]

Clinical chemistry testing of intestinal function focuses almost entirely on the evaluation of absorption and its derangements in various disease states. These intestinal diseases and disorders include tropical sprue, celiac disease, Whipple's disease, Crohn's disease, primary intestinal lymphoma, small intestinal resection, intestinal lymphangiectasia, ischemia, amyloidosis, giardiasis, and lactose intolerance. **Lactose** is a disaccharide sugar composed of glucose and galactose present in mammalian milk. Lactase is an enzyme produced in the small intestine that digests lactose. Loss or deficient lactase is a normal adult development; however, when present with abdominal cramps, bloating, and diarrhea after

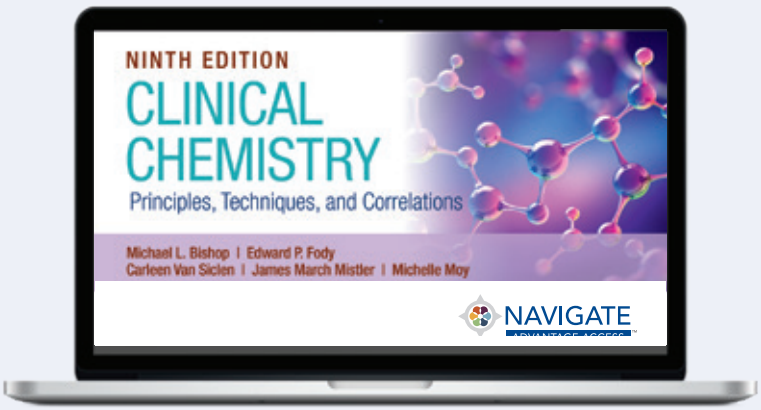
ingestion of mammalian milk or its milk products, it is called lactose intolerance. Lactose intolerance has a high prevalence in persons of African, Latin American, and Asian descent and a low to moderate prevalence in those of European descent, though rates vary widely among populations.[35,36] The lactose intolerance test administered after ingestion of a dose of lactose and subsequent serum glucose testing has been replaced largely by a breath test.[36]

D-xylose is an exogenously administered simple (pentose) monosaccharide sugar. It is not ordinarily present in the blood in significant measurable quantities because it does not require pancreatic lytic enzymes for absorption and can therefore be used to differentiate malabsorption from an intestinal etiology or exocrine pancreatic insufficiency.

Carotenoids are phytochemicals, the chief precursors of vitamin A in humans found in over 600 species of plants and some fungi and algae. The six most common carotenoids in the human body account for 90% of all the carotenoids: α-carotene, β-carotene, β-cryptoxanthin, lutein, lycopene, and zeaxanthin.[37] Malabsorption, starvation, diet, and fever can cause diminished serum carotenoid concentrations.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Shames B. Anatomy and physiology of the duodenum. In: Yeo CB, DeMeester SR, McFadden DW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 1. Philadelphia, PA: Elsevier; 2019:786–803. doi:10.1016/B978-0-323-40232-3.00068-6
2. Darrah R, Nelson R, Damato EG, et al. Growth deficiency in cystic fibrosis is observable at birth and predictive of early pulmonary function. *Biol Res Nursing*. 2016;18(5):498–504.
3. Arabi YM, Dunia J, Al-Dorzi HM, et al. Leptin, ghrelin, and leptin/ghrelin ratio in critically ill patients. *Nutrients*. 2020;12(1):36.
4. Keogh RH, Szczesniak R, Taylor-Robinson D, Bilton D. Up-to-date and projected estimates of survival for people with cystic fibrosis using baseline characteristics: A longitudinal study using UK patient registry data. *J Cystic Fibrosis*. 2018;17(2):218–227.
5. Rutledge A, Knauer M, Bhayana V. Utilization of amylase and lipase testing according to Choosing Wisely recommendations. *Clin Biochem*. 2020;86:71–71.
6. Sontag MK, Yusuf C, Grosse SD, et al. Infants with congenital disorders identified through newborn screening - United States, 2015–2017. *MMWR: Morb Mortal Wkly Rep*. 2020;69(36):1265–1268.
7. Scotet V, Gillet D, Dugueperoux I, et al. Spatial and temporal distribution of cystic fibrosis and of its mutations in Brittany, France: a retrospective study from 1960. *Hum Genet*. 2002;111:247–254.
8. Karczeski BA, Cutting GR. Molecular testing for cystic fibrosis and CFTR-related conditions. In: Grody WW, Nakamura RM, Strom CM, Kiechle FL, eds. *Molecular Diagnostics*. San Diego, CA: Academic Press. 2010:149–165.
9. Burgel P-R, Bellis G, Elborn JS. Modelling future trends in cystic fibrosis demography using the French Cystic Fibrosis Registry: update and sensitivity analysis. *Eur Respir J*. 2017;50(2):1700763. doi: 10.1183/13993003.00763-2017
10. Leeds JS, Oppong K, Sanders DS. The role of fecal elastase-1 in detecting exocrine pancreatic disease. *Nat Rev Gastroenterol Hepatol*. 2011;8(7):405–415.
11. Mais D. *Practical Clinical Pathology*, 2nd ed. The American Society for Clinical Pathology; 2017.
12. Fusco J, El-Gohary Y, Gittes GK. Anatomy, physiology, and embryology of the pancreas. In: Yeo CB, Matthews JB, Fleshman JW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 2. Philadelphia, PA: Elsevier; 2019:1062–1075. doi:10.1016/B978-0-323-40232-3.00090-X

13. Dillhoff ME, Ellison EC. Zollinger-Ellison syndrome. In: Yeo CB, DeMeester SR, McFadden DW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 1. Philadelphia, PA: Elsevier; 2019:702-711. doi:10.1016/B978-0-323-40232-3.00060-1
14. Barry L, McFadden DW. Gastrointestinal carcinoid tumors. In: Yeo CB, DeMeester SR, McFadden DW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 1. Philadelphia, PA: Elsevier; 2019:939–950. doi:10.1016/B978-0-323-40232-3.00080-7
15. Saldinger PF, Bellorin-Marin OE. Anatomy, embryology, anomalies, and physiology of the biliary tract. In: Yeo CB, Matthews JB, Fleshman JW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 2. Philadelphia, PA: Elsevier; 2019:1249–1266. doi:10.1016/B978-0-323-40232-3.00106-0
16. Huang J-h, He D, Chen L, et al. GC-MS based metabolomics strategy to distinguish threetypes of acute pancreatitis. *Pancreatology.* 2019;19(5):630–637.
17. Friedemann E, Trond E, Palwasha J, et al. Fecal fat analyses in chronic pancreatitis importance of fat ingestion before stool collection. *PLoS One*. 2017;12(1):e0169993-e0169993.
18. Lippi G, Valentino M, Cervellin G. Laboratory diagnosis of acute pancreatitis: in search of the Holy Grail. *Crit Rev Clin Lab Sci.* 2012;49(1):18–31.
19. Giannakopoulos A, Katelaris A, Noni M, et al. Hyperthyrotropinemia in newly diagnosed cystic fibrosis patients with pancreatic insufficiency reversed by enzyme therapy. *Eur J Pediatr.* 2018;177(5):775–779.
20. Fine KD, Ogunji F. A new method of quantitative fecal fat microscopy and its correlation with chemically measured fecal fat output. *Am J Clin Pathol*. 2000;113(4):528–534.
21. Berstad A, Erchinger F, Hjartholm AS. Fecal fat determination with a modified titration method. *Scand J Gastroenterol.* 2010;45(5):603–607.
22. Korpi-Steiner NL, Ward JN, Kumar V, McConnell JR. Comparative analysis of fecal fat quantitation via nuclear magnetic resonance spectroscopy (H-1 NMR) and gravimetry. *Clin Chim Acta.* 2009;400(1-2):33–36.
23. Voortman G, Gerrits J, Altavilla M, et al. Quantitative determination of fecal fatty acids and triglycerides by Fourier transform infrared analysis with a sodium chloride transmission flow cell. *Clin Chem Lab Med.* 2002;40:795–798.
24. Strom CM, Crossley B, Buller-Buerkle A, et al. Cystic fibrosis testing 8 years on: lessons learned from carrier screening and sequencing analysis. *Genet Med.* 2011;13(2):166–172.
25. Campbell K. Fecal elastase test, fecal fat, pancreatitis, sweat testing: a guide to the interpretation of pathology tests. In: *Sonic Pathology Handbook: Sonic Healthcare Limited*. 2014:302–303, 327, 663–665, 826–827. Accessed at https://www.southernpath.com.au/media/6932/pathology-handbook.pdf on September 4, 2021.
26. Kinyon K. New cystic fibrosis diagnostic guidelines: What does it mean for your practice? *J Nurse Pract.* 2019;15(4):274–276.
27. Sinclair G, McMahon V, Schellenberg A, et al. Performance of a three-tier (IRT-DNA-IRT) cystic fibrosis screening algorithm in British Columbia. *Int J Neonatal Screen.* 2020;6(46):46.
28. Ismail OZ, Bhayana V. Lipase or amylase for the diagnosis of acute pancreatitis? *Clin Biochem*. 2017;50(18):1275–1280.
29. Barben J, Castellani C, Dankert-Roelse J, et al. The expansion and performance of national newborn screening programmes for cystic fibrosis in Europe. *J Cyst Fibros.* 2017;16(2):207–213.
30. Bouwense SAW, Gooszen HG, van Santvoort HC, Besselink MGH. Acute pancreatitis. In: Yeo CB, Matthews JB, Fleshman JW, eds. *Shackelford's Surgery of the Alimentary Tract*, 8th ed., Vol. 2. Philadelphia, PA: Elsevier; 2019:1076-1084. doi:10.1016/B978-0-323-40232-3.00091-1
31. Holzer H, Reisman A, Marqueen KE, et al. "Lipase only, please": reducing unnecessary amylase testing. *Jt Comm J Qual Patient Saf*. 2019;45(11):742–749.
32. Prosapio JG, Sankar P, Jialal I. Physiology, gastrin. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021.
33. Thakker RV, Newey PJ, Walls GV, et al. Clinical practice guidelines for multiple endocrine neoplasia type 1 (MEN1). *J Clin Endocrinol Metab.* 2012;97(9):2990–3011.
34. Norton JA, Fraker DL, Alexander HR, et al. Surgery to cure the Zollinger-Ellison syndrome. *N Engl J Med.* 1999;341(9):635–644.
35. Chin EL, Huang L, Bouzid YY, et al. Association of lactase persistence genotypes (rs4988235) and ethnicity with dairy intake in a healthy U.S. population. *Nutrients.* 2019;11(8):1860.
36. Catanzaro R, Sciuto M, Marotta F. Lactose intolerance—old and new knowledge on pathophysiological mechanisms, diagnosis, and treatment. *SN Comprehensive Clin Med.* 2021;3(2):499–509.
37. Hammond BR, Jr., Renzi LM. Carotenoids. *Advances in Nutrition.* 2013;4(4):474–476.

PART 5

Specialty Areas of Clinical Chemistry

CHAPTER 23

Body Fluid Analysis

Kyle B. Riding

CHAPTER OUTLINE

KEY TERMS

Amniocentesis
Amniotic fluid (AF)
Ascites
Blood–brain barrier
Cerebrospinal fluid (CSF)
Effusion
Exudate
Hypoglycorrhachia
L/S ratio
Oligoclonal bands
Otorrhea
Pericardial fluid
Peritoneal fluid
Pleural fluid
Respiratory distress syndrome (RDS)
Rhinorrhea
Serous fluid
Subarachnoid hemorrhage
Surfactant
Synovial fluid
Thoracentesis
Transudate
Traumatic tap

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Identify the source of amniotic fluid, cerebrospinal fluid, sweat, synovial fluid, pleural fluid, pericardial fluid, and peritoneal fluid.
- Describe the physiologic purpose of amniotic fluid, cerebrospinal fluid, sweat, synovial fluid, pleural fluid, pericardial fluid, and peritoneal fluid.
- Discuss the clinical utility and methods used to test amniotic fluid, cerebrospinal fluid, sweat, synovial fluid, pleural fluid, pericardial fluid, and peritoneal fluid.
- Correlate patient status when given appropriate laboratory results obtained from amniotic fluid, cerebrospinal fluid, sweat, synovial fluid, pleural fluid, pericardial fluid, and peritoneal fluid.
- Differentiate between a transudate and an exudate, both in terms of their respective causes and laboratory results associated with each.

CASE STUDY 23.1 PART 1

Four generations of the Doe family live in the southeastern United States: 72-year-old grandmother Nancy, 44-year-old father Chad, Chad's oldest daughter, Gina, 23 years old, and Gina's 18-month-old son Joseph. As a whole, the family has had no previous significant medical history, yet have met with their primary care physician and medical specialists throughout their lifespan for various conditions.

This chapter is designed to acquaint the reader with bodily fluids that are frequently analyzed in the clinical chemistry laboratory. The source, method of collection, physiologic purpose, and clinical utility of laboratory measurements for each of these body fluids are emphasized. References to other chapters in this textbook are provided to add clarity to the pathophysiology of a condition, when warranted.

Cerebrospinal Fluid

Cerebrospinal fluid (CSF) is the liquid that surrounds the brain and spinal cord. The brain and spinal cord are covered by the meninges, which consist of three layers: the dura mater, arachnoid, and pia mater (**Figure 23.1**). The CSF flows between the arachnoid mater and the pia mater in an area referred to as the subarachnoid space. The three functions of the CSF are (1) physical support and protection, (2) provision of a controlled chemical environment to supply nutrients to the tissues and removal of wastes, and (3) intracerebral and extracerebral transport (**Figure 23.2**).

The major function of CSF is to serve as a buoyant cushion for the brain. The denser brain floats in the less dense fluid, allowing movement within the skull. The significance of the cushioning function is demonstrated by the result of a blow to the head. The initial shock is distributed across the entire brain, instead of inflicting damage to one localized area. The brain may be bruised at the side opposite the blow, depending on the force imparted.

The second major function of CSF is the maintenance of a constant chemical matrix for the central nervous system (CNS). Serum components of analytes of interest may vary greatly over time, but constituent levels of clinically significant compounds in CSF are maintained within narrow limits.

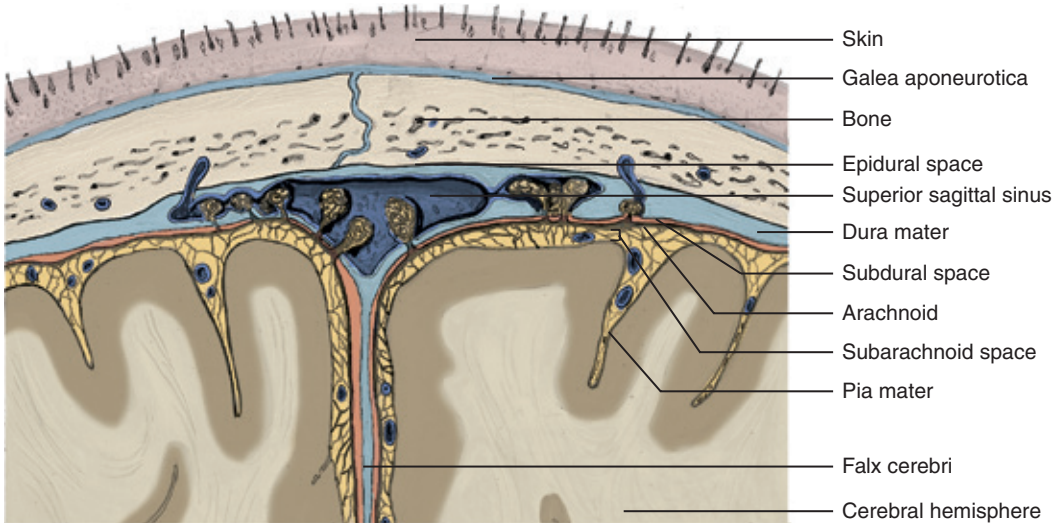


Figure 23.1 Meninges of the brain.

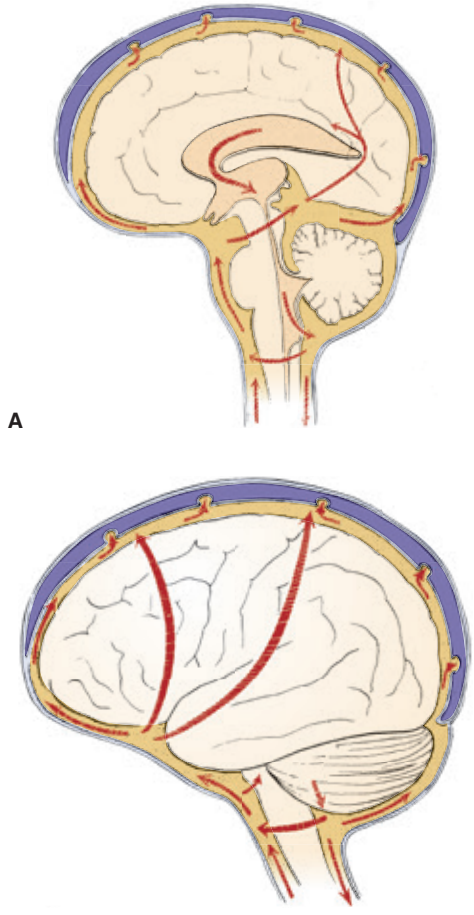


Figure 23.2 Major pathways of CSF. **(A)** Sagittal view. **(B)** Lateral view.

Reprinted from Milhorat TH. *Hydrocephalus and the Cerebrospinal Fluid.* Baltimore, MD: Lippincott Williams & Wilkins; 1972:25, with permission.

The transport function is described as a neuroendocrine role. The CSF is involved in the distribution of hypophyseal hormones within the brain and the clearance of hormones from the brain to the blood.

The total CSF volume is about 150 mL or about 8% of the total CNS cavity volume. The fluid is formed predominantly at the choroid plexus deep within the brain and by the ependymal cells lining the ventricles. The endothelial cells of the choroid plexuses have very tight-fitting junctures to control the passage of substances across their membranes, which is termed the **blood–brain barrier**. Damage to the blood–brain barrier is frequently the reason for abnormal chemistry results in CSF analysis.

CSF is formed at an average rate of about 0.4 mL/min, about 500 mL/day. Formation is a result of selective ultrafiltration of plasma and active secretion by the epithelial membranes. Absorption of CSF occurs at outpouchings in the dura called *arachnoid villi*, also known as granulations, which protrude through the dura to the venous sinuses of the brain and into the bloodstream. The granulations act as one-way valves to maintain an excretion volume equal to the production volume.

Specimens of CSF are obtained by lumbar puncture, usually at the interspace of vertebrae L3 to L4 (lumbar vertebrae) or lower, using aseptic technique (**Figure 23.3**). The fluid obtained is usually separated into three numbered aliquots: (1) for chemistry and serology, (2) for microbiology, and (3) for hematology. In some instances, a fourth tube may be collected for cytology analysis. The order of the tubes reflects the preferred order of collection for minimizing interference from less-than-optimal collection technique, with tube 3 presumably the least contaminated by cells such as red blood cells (RBC). It is paramount to remember that this specimen is of limited volume and should be analyzed immediately. Any remaining sample should be preserved because of its limited availability.

Laboratory investigation of CSF is indicated for cases of suspected CNS infection, demyelinating disease, malignancy, and hemorrhage into the CNS. As with all patient samples entering the laboratory, physical examination of the specimen is the first and often the most important observation made. Normal CSF is clear, colorless, free of clots, and free of blood. Differences from these standards indicate a probable pathology and merit further examination.

Cloudy fluids usually are due to cellular particles or debris present in the sample. Not only do cloudy fluids require microscopic examination, they are often an indication of infectious process within the CNS requiring immediate antimicrobial and corticosteroid therapy.[1] The biochemical workup that follows the identification of cloudy CSF is critical to the treatment of patients with these infectious processes.

In terms of alterations to color, a yellow to brown (i.e., *xanthochromic*) or red color indicates the presence of degraded or fresh blood, respectively. The two most common reasons for blood and hemoglobin pigments to be found in CSF are **traumatic tap** and **subarachnoid hemorrhage**. Traumatic tap is the artifactual presence of blood or derivatives due to the damage of blood vessels during the lumbar puncture. Hemorrhage results from a breakdown of the barrier of the CNS and circulatory system, often from trauma or underlying coagulopathies. Because of the life-threatening complications associated with hemorrhage, it is important to determine the etiology of blood in the CSF sample.

There are several ways to differentiate a traumatic tap from a subarachnoid hemorrhage. The first method involves comparing the amount of blood

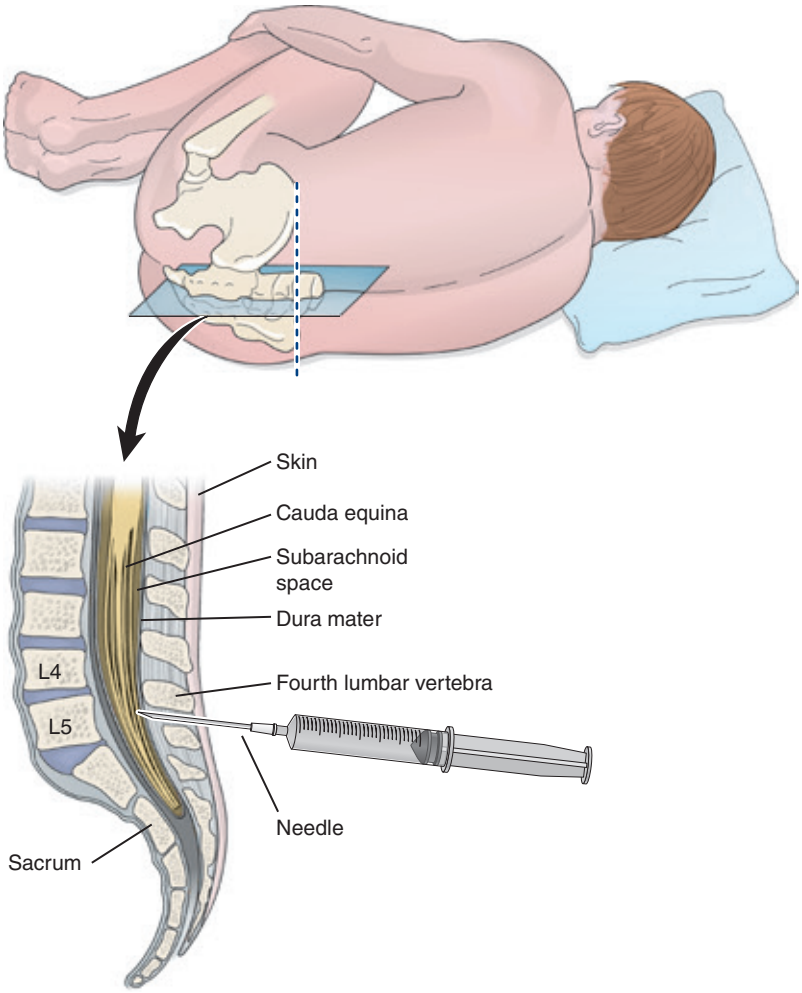


Figure 23.3 Placement of the needle for CSF collection.

present between the collected tubes. If there is a significant difference in the amount of blood present between the first and last tubes collected (i.e., later tubes gradually clearing and less bloody), then the puncture was traumatic. If all tubes collected show the same degree of blood, then a subarachnoid hemorrhage is most likely. **Figure 23.4** demonstrates the difference in appearance of normal clear CSF, red CSF in hemorrhage, xanthochromic CSF from an older hemorrhage demonstrating the build-up of heme derivatives, and CSF from a traumatic tap.[2]

A second method of discriminating between a traumatic tap and subarachnoid hemorrhage involves comparison of the appearance of the CSF precentrifugation and postcentrifugation. If the supernatant of the postcentrifugation sample appears colorless, the presence of blood is likely due to a traumatic tap. This colorless supernatant is due to the fact the red blood cells have not had time to lyse in the solution. If the supernatant appears pink, red, or xanthochromic, hemorrhage should be suspected. This is due to red blood cells lysing easily in CSF due prolonged time in the CSF matrix, which is comprised of lower levels of proteins and lipids in comparison to plasma. The resultant in vivo hemolysis causes the release of free oxyhemoglobin that eventually changes to methemoglobin, and then after about 12 hours, bilirubin is formed (accounting for the xanthochromia). Gradual decrease in CSF color after a hemorrhage occurs over the first 2 days, clearing in about 2 to 4 weeks.[3]

The final method to differentiate a traumatic tap from a hemorrhage may also involve checking for the presence of a clot and/or hemosiderin-containing macrophages. Samples exhibiting signs of a traumatic tap may have clots present within the CSF and will exhibit no hemosiderin. Conversely, samples from a patient experiencing a subarachnoid hemorrhage will show no signs of fibrin clots but may contain hemosiderin-laden macrophages.

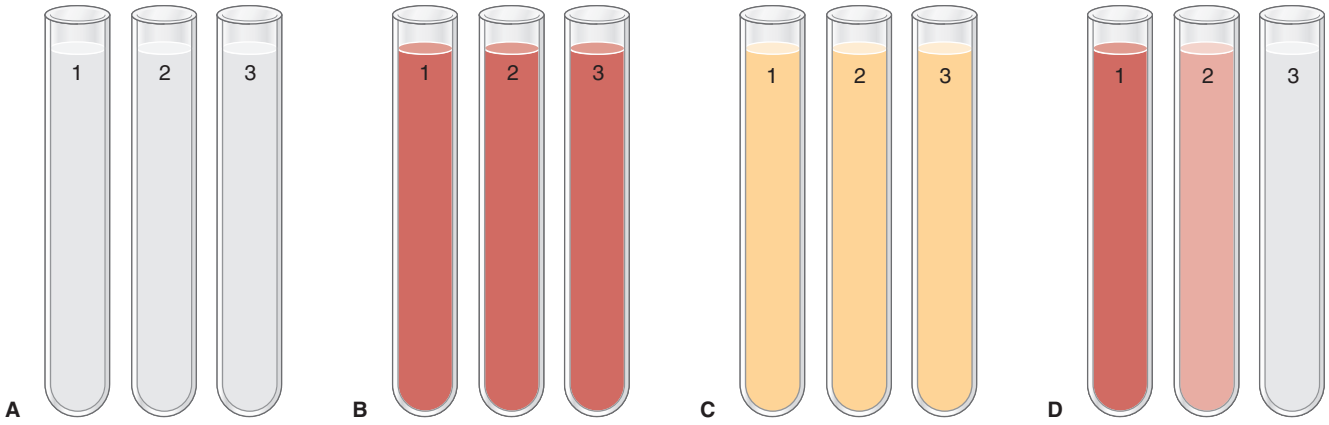


Figure 23.4 Comparison of CSF appearance between **(A)** normal CSF (all three tubes with the same clarity), **(B)** red CSF from fresh hemorrhage (all three tubes with the same *red* color and opaqueness), **(C)** xanthochromic CSF from old hemorrhage (all three tubes with the same yellow color and clear), and **(D)** CSF from a traumatic tap (decreasing amounts of redness in each successive tube).

A wide variety of biochemical markers with CSF samples may be measured. However, in routine practice, the number of useful indicators involves only four analytes. The tests of interest are glucose, protein (total and specific), lactate, and glutamine. Before any chemical analysis, the sample should be centrifuged to avoid contamination by cellular elements.

Glucose enters the spinal fluid predominantly via a facilitative transport compared with a passive (diffusional) or an active (energy-dependent) transport. It is carried across the epithelial membrane by a stereospecific carrier species. The carrier mechanism is responsible for transport of lipid-insoluble materials across the membrane into the CSF. Generally, this is a "downhill" process consistent with a concentration gradient. This mechanism causes the CSF glucose concentration to be about two-thirds that of plasma.

Due to its dependence on plasma glucose concentrations, an isolated CSF glucose concentration may be misleading, and it is recommended that a plasma sample be obtained 2 to 4 hours prior to the tap so that plasma and CSF glucose levels can be compared. However, it is unreasonable to assume that this recommendation is always feasible in situations that involve the CNS. Normal CSF glucose is considered to be 60% to 70% that of the plasma glucose. Increased CSF glucose levels are not clinically significant, usually providing only confirmation of hyperglycemia. It is important to note that as plasma hyperglycemia progresses and exceeds 600 mg/dL, the CSF glucose level does not remain at 60% to 70% of plasma values, most likely due to cellular carriers being saturated and incapable of further transport. This saturation will cause the plasma/CSF glucose to decrease. For this reason, caution should be taken when comparing the plasma and CSF glucose values in a patient with severe hyperglycemia.

Decreased CSF glucose levels (*hypoglycorrhachia*) can be the result of (1) disorder in carrier-mediated transport of glucose into CSF, (2) active metabolism of glucose by cells or organisms, or (3) increased metabolism by the CNS. The pathophysiological mechanism of altered transport remains unclear, but it is speculated to be the cause of **hypoglycorrhachia** in tubercular meningitis and sarcoidosis. Acute bacterial, amebic, fungal, and trichinosis meningitis are examples of glucose consumption by organisms, whereas diffuse meningeal neoplasia and brain tumor are examples of glucose consumption by CNS tissue. Consumption of glucose is usually accompanied by an increased lactate level due to anaerobic glycolysis by organisms or cerebral tissue.

Protein levels in CSF reflect the selective ultrafiltration of the CSF blood–brain barrier. All protein usually found in plasma is found in CSF but at much lower levels. Total protein is about 0.5%, or 1%, that of plasma, but the specific fractional protein concentrations in CSF are not proportional to the plasma levels because of the specificity of the ultrafiltration process. Because of this relationship of CSF and serum proteins, serum analysis should accompany CSF protein analysis.

A decreased level of CSF total protein can arise from (1) decreased dialysis of proteins from the plasma, (2) increased protein loss (i.e., removal of excessive volumes of CSF), or (3) leakage from a tear in the dura, CSF otorrhea, or CSF rhinorrhea. A dural tear can occur as a result of a previous lumbar

puncture or from severe trauma. **Otorrhea** and **rhinorrhea** refer to the leakage of CSF into the ear or the nose, respectively. Identification of CSF leakage is best done by an analysis for β-transferrin, a protein unique to the CSF.

An increased level of CSF total protein is a useful—albeit non-specific—indicator of pathologic states. Increases may be caused by (1) lysis of contaminant blood from traumatic tap, (2) increased permeability of the epithelial membrane, (3) increased production by CNS tissue, or (4) obstruction. The presence of any amount of blood from hemorrhage or traumatic tap can elevate CSF protein levels. Changes in protein permeability often occur due to infectious processes. In particular, the blood–brain barrier becomes more permeable from bacterial or fungal infections. However, current critical care management algorithms do not necessitate that CSF protein be used as a diagnostic marker in infectious meningitis—though it can still provide value in narrowing down the etiological agent.[1] An increase in CNS production occurs in subacute sclerosing panencephalitis (SSPE) or multiple sclerosis (MS). There may also be combinations of permeability and production, as seen with collagen vascular diseases. An obstructive process, such as tumor or abscess, would also cause increased protein.

More sensitive information about non-infectious conditions can be obtained by analysis of the protein fractions present. A comparison to the serum protein electrophoresis pattern is necessary for accurate conclusions. Under normal conditions, prealbumin is present in CSF in higher concentration than in serum. Although the respective proteins can be determined in both serum and CSF, the proteins of greatest interest are albumin and immunoglobulin G (IgG). Because albumin is produced solely in the liver, its presence in CSF must occur by means of blood–brain barrier passage. IgG, however, can arise by local synthesis from plasma cells within the CSF. The measurement of albumin in both serum and CSF is then used to normalize the IgG values from each matrix to determine the source of the IgG.

To determine the integrity of the blood–brain barrier, a CSF/serum albumin index is calculated as follows:

$$\frac{\text{CSF albumin (mg/dL)}}{\text{Serum albumin (mg/dL)}} = \text{CSF serum albumin index}$$

(**Eq. 23.1**)

An index value less than 9 indicates an intact blood–brain barrier.

This index can then be used to calculate the IgG index to determine CNS synthesis of IgG to aid the diagnosis of demyelinating diseases, such as MS and SSPE. MS is the most common inflammatory demyelinating disease of the CNS:

$$\frac{\text{CSF IgG/serum IgG}}{\text{CSF albumin/serum albumin}} = \text{CSF IgG index}$$

(**Eq. 23.2**)

The normal value is less than 0.73.

Increases in serum albumin cause increases in the CSF levels because of membrane permeability. However, increased CSF IgG without concomitant CSF albumin increase suggests local production (MS or SSPE). Increases in both permeability and production are found with bacterial meningitis. Methods to analyze IgG and albumin CSF levels are the same as for serum but are optimized for the lower levels found. Please refer to Chapter 6, *Amino Acids and Proteins*, for further explanation.

In addition to assessment of total protein, CSF albumin, and CSF IgG, less frequent CSF protein markers of inflammation may be of use when clarifying if a disease process is infectious in origin. For example, a recent study on pediatric patients found that the CSF demonstrated consistently elevated interleukin-6 (IL-6), interleukin-8 (IL-8), and monocyte chemoattractant protein-1 (MCP-1) relative to serum concentrations in patients with coronavirus infections of the CNS.[4] While such proteinaceous markers may be of esoteric use in the CSF at this time, their use may become more prevalent as the kinetics of these inflammatory compounds are further studied.

Increased CSF protein levels that do not have a clear etiology usually indicate the need for electrophoretic separation of the respective proteins. In addition, normal CSF protein levels in a patient presenting with a constellation of findings indicative of an inflammatory disease of the CNS also usually require CSF protein electrophoresis. At times, this separation demonstrates multiple banding of the gamma globulin region representing similar yet pathology-related IgG bands. This observation is referred to as **oligoclonal bands** and involve a small number of clones of IgG from the same cell type with nearly identical electrophoretic properties. This occurrence is usually associated with inflammatory diseases and MS or SSPE. These types of disorders would stimulate the immunocompetent cells that produce similar bands across the gamma globulin region. The recognition of an oligoclonal

pattern supersedes the report of normal protein levels and is a cause for concern if the corresponding serum separation does not demonstrate identical banding.

Another protein thought to be specific for MS is myelin basic protein (MBP). Initial reports suggested high specificity, but MBP has also been found in nondemyelinating disorders and does not always occur in demyelinating disorders. MBP levels are used to monitor therapy of MS. Current international guidelines for the diagnosis of MS recognize both an elevated IgG index and the presence of CSF oligoclonal bands that are not found in the serum as supporting evidence. Immunoassay procedures are available for analyzing MBP.[5]

One final group of proteins gaining notoriety are those that are potential markers of prodromal Alzheimer's disease. Traditionally, the use of amyloid precursor proteins in combination with tau proteins was used to identify early stages or identify patients at high risk for disease progression. Novel markers found in the CSF, such as apoprotein-E, appear to be promising and could improve our diagnostic ability on this significant cause of dementia.[6]

Lactate is a useful indicator of anaerobic metabolism within the CSF. While the level of lactate in the CSF can potentially parallel that of the plasma, CSF elevations can occur independent of plasma concentrations under situations where anaerobic metabolism is occurring at an increased rate solely within the CNS. An increased lactate level with a normal to decreased glucose level has been suggested as a readily accessible indicator for bacterial meningitis. This finding is helpful in differentiating bacterial meningitis from other forms of infectious meningitis.[7] Analysis of glucose and lactate in CSF is easily accomplished by techniques used for plasma and serum. It is important that provision for the analysis of glucose or lactate in CSF be immediate or that the specimen be preserved with an antiglycolytic, such as sodium fluoride.

Glutamine is formed by the combination of ammonia and glutamate. It is hypothesized that the synthesis of glutamine is a means to reduce ammonia

CASE STUDY 23.1, PART 2

Remember the Doe family.

© digitalskillet/Shutterstock.

Part 2A

While laboratory analysis often involved the collection of routine blood and urine specimens, the cases in this chapter will highlight times when they needed other body fluids collected for diagnostic purposes.

1. Besides blood and urine, what are other body fluids that can be collected for diagnostic purposes?
2. Do you believe that there are body fluids that are only clinically relevant at certain time points in life? Or, instead, is the diagnostic utility of these fluids equal across the lifespan?

Part 2B

The youngest member of the Doe family, 18-month-old Joseph, is rushed to the emergency department with a fever of 103.2°F and lethargy. As part of the diagnostic assessment, a CSF sample was collected. Macroscopic, cellular, and biochemical results for this sample are as follows:

CSF Color: Colorless
CSF Clarity: Slightly Cloudy

CSF WBC Count: 95 cells/μL—(80% lymphocytes, 15% monocytes, 5% neutrophils)
CSF RBC Count: 2 cells/μL

CSF Total Protein: 72 mg/dL
Plasma Total Protein: 6.8 g/dL
CSF Glucose: 55 mg/dL
Plasma Glucose: 90 mg/dL

3. Based upon the above data, what is the most likely cause of Joseph's medical emergency?
4. What is the clinical significance of the slightly cloudy CSF?
5. Calculate the CSF Glucose : Plasma Glucose ratio. Is the result within expected limits?

levels within the central nervous system. While the level of ammonia within CSF is difficult to accurately determine due to its relative instability, the level of glutamine within CSF samples is reflective of the amount of ammonia present within the central nervous system. Glutamine elevations are frequently seen in patients with hepatic encephalopathy along with various other pathologies.

Recently, the use of neuron-specific enolase (NSE) has been used to test the cerebral spinal fluid of patients with physical symptoms of neurological conditions similar to that of Creutzfeldt-Jakob disease as well as in patients with small cell and non-small cell lung cancer. Further discussion on the usefulness of NSE can be found in Chapter 28, *Tumor Markers*.

Serous Fluids

The lungs, heart, and abdominal cavities are surrounded by two serous membranes: the parietal membrane lining the cavity wall and the visceral membrane lining the organs. Serous fluids, an ultrafiltrate of plasma, are located between the membranes. When serum dialyzes across these membranes, the fluid formed is called **serous fluid**—specifically, pleural (lung), pericardial (heart), and peritoneal (abdominal) fluid.

The formation of serous fluid is a continuous process driven by the hydrostatic pressure of the systemic circulation and maintenance of oncotic pressure due to protein. The potential space is usually fluid-filled with no gases present. The fluid reduces or eliminates friction caused by expansion and contraction of the encased organs.

Serous fluids, along with other extracellular fluids, account for 45% of the body's water volume and share common characteristics in regard to their chemical composition.[8] Various forces control the movement of fluids within the different extracellular compartments of the body. The tissue's colloidal osmotic pressure (interstitial fluid pressure), along with the capillary's hydrostatic pressure (filtration pressure), regulates the outward flow of fluid from the capillary. The colloidal osmotic pressure of the capillary and the tissue's hydrostatic pressure regulate the inward flow of fluid into the capillary from the tissue. **Figure 23.5** illustrates the direction of these forces. Furthermore, the lymphatic system removes the fluids entering into the interstitial space. **Figure 23.6** shows the normal flow of fluids among the bloodstream, tissues, and lymphatic vessels.

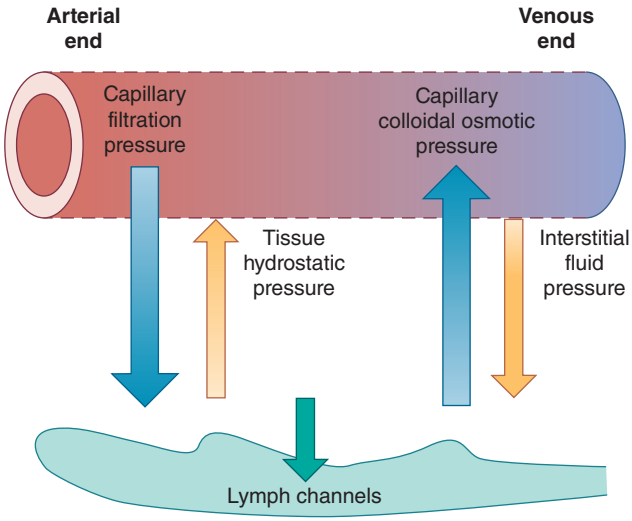


Figure 23.5 Forces governing the exchange of fluid at the capillary level.

Examining these biochemical and cellular alterations that occur to serous fluids can assist in diagnosing and monitoring the patient's condition.[9] Such alterations occur due to an imbalance in pressures that lead to excessive egress of fluid into tissue spaces that cause an accumulation of fluid that is cause for analysis of the serous fluid. This accumulation is called an **effusion**. Effusions are further classified as **transudate** or **exudate**. Proper classification of effusions is critical to patient care.

A transudate may occur during various systemic disorders that disrupt fluid filtration, fluid reabsorption, or both. These include conditions such as congestive heart failure, hepatic cirrhosis, and nephrotic syndrome. Exudative effusions occur during inflammatory processes that result in damage to blood vessel walls, body cavity membrane damage, or decreased reabsorption by the lymphatic system. Infections, inflammations, hemorrhages, and malignancies are examples of pathologic processes that can cause the formation of exudates. The formation of transudates and exudates can damage tissues and body cavity membranes and alter lymphatic function.

Various laboratory tests are used to differentiate between transudates and exudates, including fluid appearance, specific gravity, amylase, glucose, lactate dehydrogenase (LD), and proteins. Additional tests such as ammonia, lipids, and pH may be useful in confirming the cause of an effusion for specific body sites.

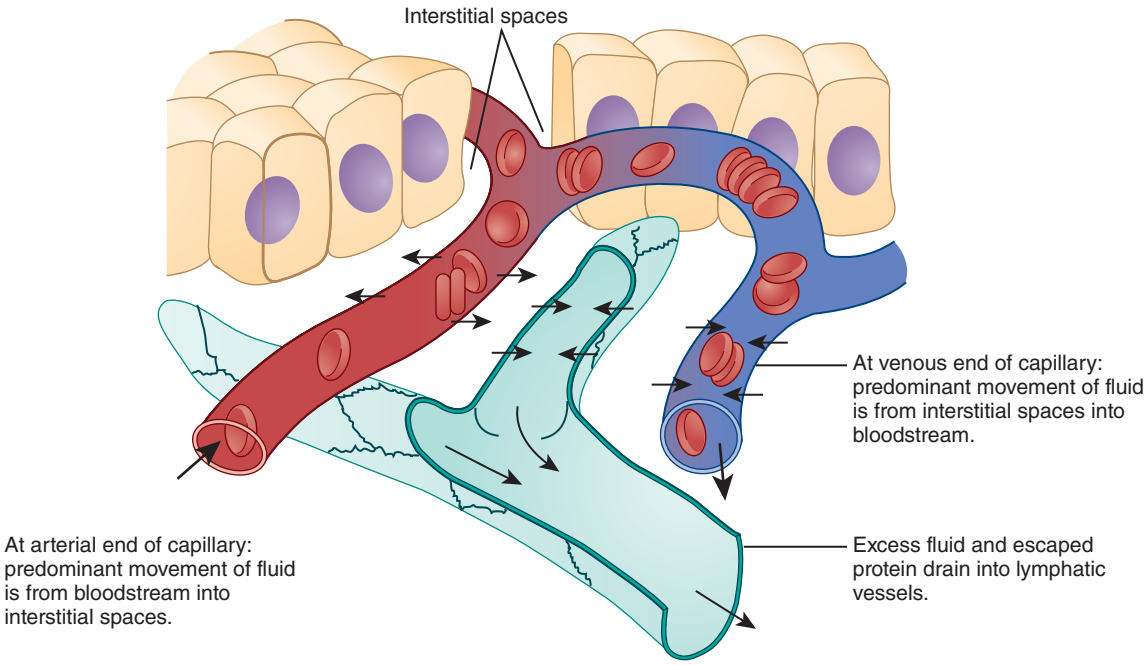


Figure 23.6 Exchanges through capillary membranes in the formation and removal of interstitial fluid.

Pleural Fluid

The outer layer of the pleural space, the parietal layer, is served by the systemic circulation and the inner layer, the visceral layer, by the bronchial circulation. As with all serous fluids, **pleural fluid** is formed by the ultrafiltration of plasma through the parietal layer and is essentially interstitial fluid of the systemic circulation. With normal conditions, there is 3 to 20 mL of pleural fluid in the pleural space. The fluid exits by drainage into the lymphatics of the visceral pleura and the visceral circulation. Any alteration in the rate of formation or removal of the pleural fluid affects the volume, causing an effusion.

Once an effusion is identified, it is then necessary to classify the nature of the effusion by analysis of the pleural fluid. The fluid is removed from the pleural space by needle and syringe during visualization by radiology. The procedure is called **thoracentesis**, and the fluid is called *thoracentesis fluid* or *pleural fluid* (**Figure 23.7**). Samples are collected as follows: a heparinized tube for culture, an ethylenediaminetetraacetic acid (EDTA) tube for microscopy, a sodium fluorescein (NaF) tube for glucose and lactate, and a nonanticoagulated tube for further biochemical testing.

Transudates are secondary to remote (nonpleural) systemic pathologies and exhibit biochemical and cellular abnormalities consistent with noninflammatory changes in fluid dynamics within the pleural space. An example would be hypoproteinemia due to malnutrition leading to decreased oncotic pressure and subsequent decrease in fluid reabsorption into the capillaries. An exudate indicates primary involvement of the pleura and lung, such as infection, and demands immediate attention. An example of an exudative process would be infection-mediated damage to the membranes,

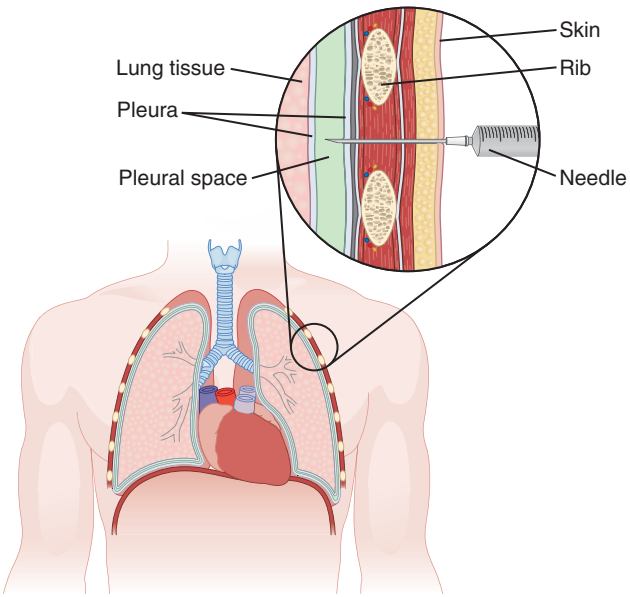


Figure 23.7 Thoracentesis. A needle is inserted into the pleural space to withdraw fluid.

From Cohen BJ. *Medical Terminology.* 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.

Table 23.1 Causes of Pleural Effusions

Transudative	Exudative
Congestive heart failure[a]	Bacterial pneumonia[a]
Nephrotic syndrome	Tuberculosis
Hypoproteinemia	Pulmonary abscess
Hepatic cirrhosis	Malignancy (lymphatic obstruction)
Chronic renal failure	Viral/fungal infection
	Pulmonary infarction
	Pleurisy
	Pulmonary malignancy
	Lymphoma
	Pleural mesothelioma

[a]Most common cause.

allowing increased fluid entry into the pleural space (**Table 23.1**).

The assignment of fluid to either the transudate or exudate category had previously been based on the protein concentration of the fluid. This criterion has been replaced by the use of a series of fluid/plasma (F/P) ratios known as Light's criteria. These criteria specifically state that if the F/P ratio for total protein is greater than 0.5 or the ratio for LD is greater than 0.6 (or the pleural fluid LD/upper limit serum LD ratio reference range is >0.67), the fluid is an exudate. It is important for the clinical laboratorian to note that recent studies have demonstrated the classification of transudate and exudate based upon Light's criteria can differ based upon variability in analytical systems—mostly due to variance in LD determinations.[10]

For pleural fluid, an additional characterization can be made by determining the pleural fluid cholesterol, the fluid-to-serum cholesterol ratio, and the fluid-to-serum bilirubin ratio. Exudates are indicated by fluid cholesterol greater than 60 mg/dL, fluid-to-serum cholesterol ratio greater than 0.3, and fluid-to-bilirubin ratio of 0.6 or higher.

Further characterization of the exudate by the chemistry laboratory may involve analysis for glucose, lactate, amylase, triglyceride, pH, or uric acid. A decrease in glucose or increase in lactate would suggest infection or inflammation. An increase in amylase compared with that of serum suggests pancreatitis. Grossly elevated triglyceride levels (2 to 10 times serum levels) could indicate thoracic duct leakage. The use of pH measurements, performed as one would perform a blood–gas level determination, has gained favor. Succinctly, pH less than 7.2 suggests infection, and pH close to 6.0 indicates esophageal rupture. Finally, uric acid levels have been shown to be significantly lower in exudates when compared to transudates.[11] The methodologies for these analyses are the same as those employed for the serum and blood constituents and, therefore, are feasible in the clinical laboratory.

One additional diagnostic challenge in differentiating transudates and exudates is the impact of diuretic therapy on fluid biomarkers. One study found that 20% to 30% of transudative effusions with cardiac or hepatic etiology are misclassified under Light's criteria as exudates.[12] This is most likely due to diuretic therapies used in these conditions causing elevations in many commonly measured pleural fluid analytes—particularly total protein and LD.[13] However, in the case of those patients on diuretics due to underlying cardiac abnormalities it has been found that the addition of pleural fluid testing of NT-proBNP helps correctly identify transudative effusions more accurately than the Light's criteria alone.[14] Pleural fluid BNP levels are also more accurate than Light's criteria but have weaker performance than the NT-pro-BNP. Unfortunately, no clear candidate for better identifying pleural transudates in patients on diuretics due to hepatic dysfunction or failure has been found to date.

Pericardial Fluid

The relationship of the pericardium, **pericardial fluid**, and the heart is similar to that with the lungs and pleural fluid. Mechanisms of formation and drainage are the same. Pericardial effusions are an accumulation of fluid around the heart caused by damage to the mesothelium and are usually always exudates. Figure 23.8 illustrates the pericardium surrounding the heart. Normally, the pericardium contains less than 50 mL of fluid. The procedure for removing excess pericardial fluid, pericardiocentesis, is dangerous and therefore rarely performed (**Figure 23.8**). However, this procedure is necessary to obtain a sample if cultures are needed to investigate an infection or if cytology is needed for suspected malignancy.[8] Adenosine deaminase testing may also be requested in suspected cases of tubercular effusions.

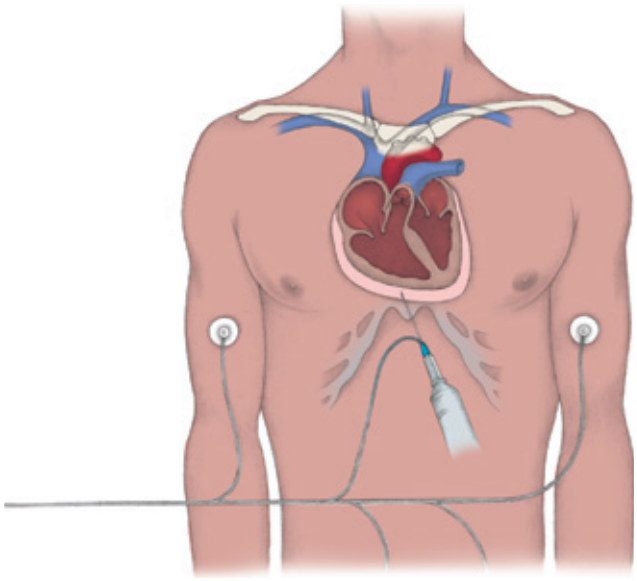


Figure 23.8 Aspirating pericardial fluid. In pericardiocentesis, a needle and a syringe are inserted through the chest wall into the pericardial sac. Electrocardiographic monitoring, with a lead wire attached to the needle and electrodes placed on the limbs (right arm, left arm, and left leg), helps ensure proper needle placement and avoids damage to the heart.

From *Nursing Procedures*. 4th ed. Ambler, PA: Lippincott Williams & Wilkins; 2004.

Peritoneal Fluid

Excess fluid (>50 mL) in the peritoneal cavity indicates disease. The presence of excess **peritoneal fluid** is called **ascites**, and the fluid is called *ascitic fluid*. The process of obtaining samples of this fluid by needle aspiration is paracentesis (**Figure 23.9**). Usually, the fluid is visualized by ultrasound to confirm its presence and volume before paracentesis is attempted.

The same mechanisms that cause serous effusions in other body cavities are the same for the peritoneal cavity. Specifically, a disturbance in the rate of dialysis secondary to a remote pathology is a transudate, compared with a primary pathology of the peritoneal membrane causing an exudate. The multiple factors that apply to this large space, such as renal and hepatic function, tend to cloud the distinction. The most common cause of transudative ascites is portal hypertension. Obstructions to hepatic flow, such as cirrhosis, congestive heart failure, and hypoalbuminemia for any reason also demonstrate high incidence. The exudative causes of ascites are predominantly metastatic ovarian, prostate, and colon cancer and infective peritonitis.

The recommended method for differentiating causes of peritoneal effusions is the serum–ascites albumin gradient (SAAG).[15] The SAAG is calculated by subtracting the fluid albumin level from the serum albumin level. A difference of 1.1 g/dL or more is used to indicate a transudative process, while a difference of less than 1.1 g/dL indicates an exudative process. A neutrophil count greater than 250 cells/μm indicates peritonitis. Measurement of the tumor markers CEA and CA125 is indicated in suspected cases of malignancy.

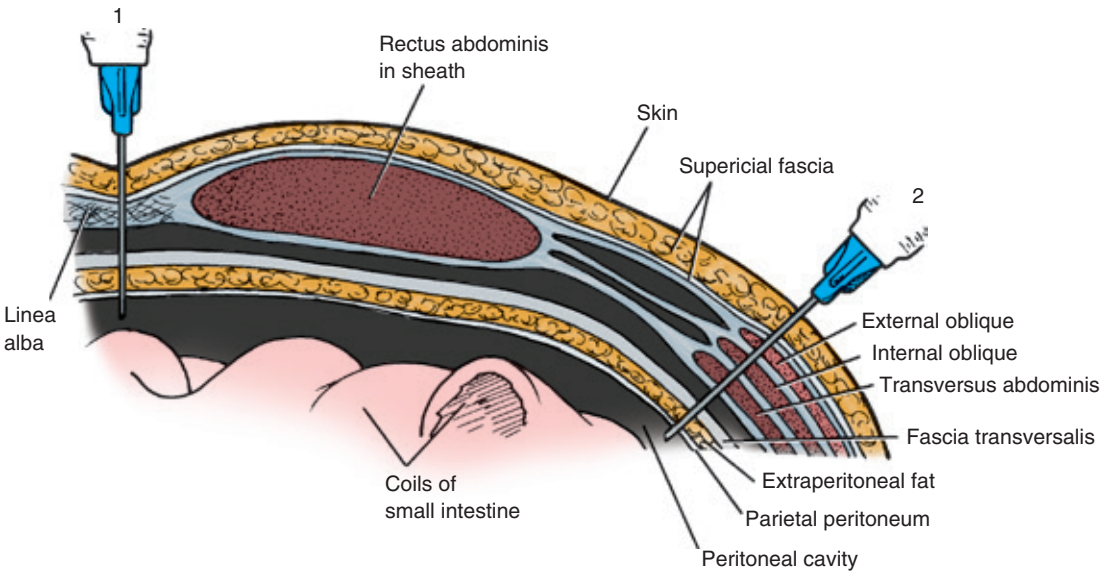


Figure 23.9 Paracentesis of the abdominal cavity in midline.

From Snell RS. *Clinical Anatomy*. 7th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2003.

CASE STUDY 23.1, PART 3

Remember Nancy. She is the senior member of the Doe family, is 72 years old. Nancy has been on a thiazide diuretic for early-stage congestive heart failure for the past five years. She is admitted to the hospital for suspicions of newly developed pleural effusion. The pleural effusion is confirmed, and a thoracentesis is performed. Ultimately, the thoracentesis was determined to be a transudative effusion due to her congestive heart failure.

6. What is the difference between a transudate and exudate? Does a transudate make sense in Nancy's case?
7. Explain the pathogenesis of a pleural effusion secondary to congestive heart failure.
8. What impact could her diuretic therapy have on the biochemical analysis of her pleural fluid? Does this have clinical significance?

Amniotic Fluid

The amniotic sac provides an enclosed environment for fetal development. This sac is bilayered as the result of a fusion of the amniotic (inner) and chorionic (outer) membranes at an early stage of fetal development. The fetus is suspended in **amniotic fluid (AF)** within the sac. The AF provides a cushioning medium for the fetus, regulates the temperature of the fetal environment, allows fetal movement, and serves as a matrix for influx and efflux of constituents such as glucose, sodium, and potassium.

Depending on the interval of the gestational period, the fluid may be derived from different sources, but ultimately, the mother is the source of the AF. At the initiation of pregnancy, some maternal secretions cross the amnion and contribute to the volume. Shortly after formation of the placenta and embryo and fusion of membranes, AF is largely derived by transudation across the fetal skin. In the last half of pregnancy, the skin becomes substantially less permeable, and fetal micturition, or urination, becomes the major volume source.

The fate of the fluid also varies with the period of gestation. A bidirectional exchange is presumed to occur across the membranes and at the placenta. Similarly, during early pregnancy, the fetal skin is involved in exchange of AF. In the last half of pregnancy, the mechanism of fetal swallowing is the major fate of AF.

There is a dynamic balance established between production and clearance; fetal urination and swallowing maintain this balance. The continual swallowing maintains intimate contact of the AF with the fetal gastrointestinal tract, buccal cavity, and bronchotracheal tree. This contact is evidenced by the sloughed cellular material from the fetus that provides us with an indication of various fetal developmental milestones and functional stages.

A sample of fluid is obtained to analyze the cellular material and the biochemical constituents of the AF. This fluid is collected via transabdominal **amniocentesis** (amniotic sac puncture), which is performed under aseptic conditions. Before an attempt is made to obtain fluid, the positions of the placenta, fetus, and fluid pockets are visualized using ultrasonography. Aspiration of anything except fluid could lead to erroneous conclusions, as well as possible harm to the fetus (**Figure 23.10**).

Amniocentesis and subsequent AF analysis are performed to test for congenital diseases, neural tube defects (NTDs), hemolytic disease of the fetus, and fetal pulmonary development. Emphasis is placed on those conditions diagnosed or monitored using the

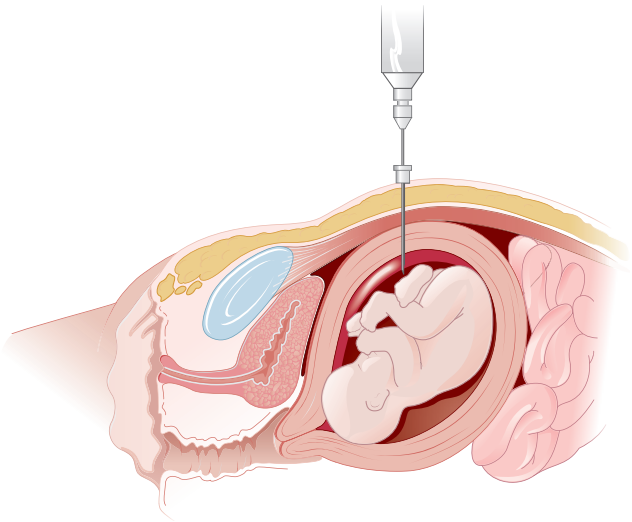


Figure 23.10 Amniocentesis. A sample is removed from the amniotic sac for fetal abnormality testing.

clinical biochemistry laboratory. For a more detailed explanation of amniotic fluid testing, see *Chapter 24: Pregnancy and Prenatal Testing*.

Hemolytic Disease of the Fetus and Newborn

The analysis of AF to screen for hemolytic disease of the fetus and newborn (erythroblastosis fetalis) was the first recognized laboratory procedure performed on AF. Hemolytic disease of the newborn is a syndrome of the fetus resulting from incompatibility between maternal and fetal blood, because of differences in either the ABO or Rh blood group systems. Maternal antibodies to fetal erythrocytes cause a hemolytic reaction that can vary in severity. The resultant hemoglobin breakdown products, predominantly bilirubin, appear in the AF and provide a measure of the severity of the incompatibility reaction.

The method used most commonly is a direct spectrophotometric scan of undiluted AF and subsequent calculation of the relative bilirubin amount. Classically, absorbance due to the presence of bilirubin is reported instead of a concentration of bilirubin. The method consisted of spectrophotometric scanning the AF sample between the wavelengths of 550 to 350 nm against a water blank. The common method, the method of Liley,[16] then plotted the readings of absorbances on the y-axis at 5-nm intervals against wavelength along the x-axis, using semilogarithmic paper. The baseline is constructed by creating a line between two absorbance readings, the absorbance points at 550 and 350 nm. These points are connected using a straight line. An absorbance peak is seen at 450 nm and is a result of bilirubin measured within the AF. The larger the difference between baseline absorbance and peak absorbance at 450 nm is an indication of the level of bilirubin present within the AF. This direct observation is referred to as the ΔA_{450}.

To avoid interference in the spectrophotometric scan, specimens should be immediately centrifuged and the fluid separated from the sediment. This will prevent not only particulate interference but also the possibility of increased lysis of red blood cells in the specimen producing hemoglobin in the AF. As with all specimens for bilirubin analysis, AF specimens for bilirubin scans must be protected from light. Specimens are routinely collected in amber-colored tubes. Exposure to light results in the photo-oxidation of bilirubin to biliverdin that will not be detected at 450 nm, resulting in underestimation of the hemolytic disease severity.

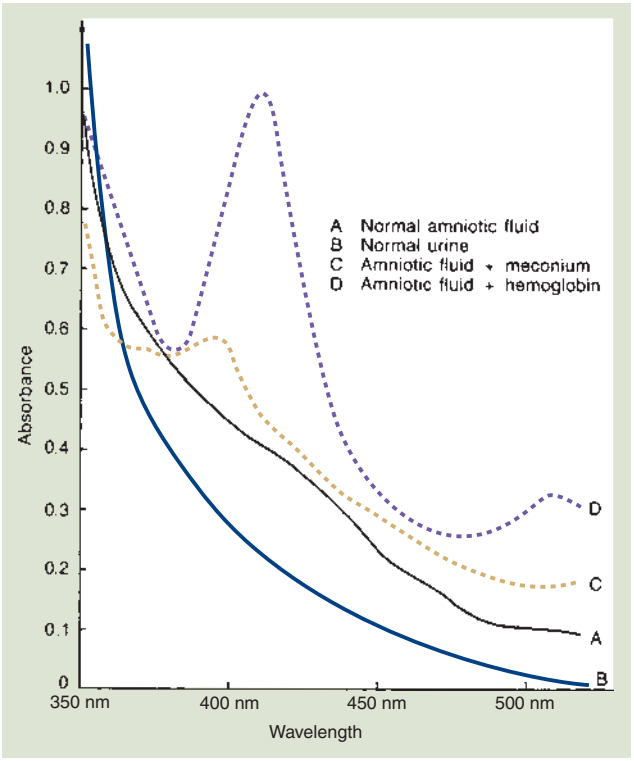


Figure 23.11 Amniotic fluid absorbance scans.

Examples of interferences, compared with a normal specimen, are given in **Figure 23.11**. Each laboratory should compile its own catalog of real examples for spectrophotometric analysis. The presence of hemoglobin is identified by its peak absorbance at 410 to 415 nm; the presence of urine is identified by the broad curve and confirmed by creatinine and urea analyses; and the presence of meconium is identified by the distinctly greenish color of the AF and flat absorbance curve.

Care must be used in the interpretation of the spectra. A decision for treatment can be made based on the degree of hemolysis and gestational age. The rather limited treatment options are immediate delivery, intrauterine transfusion, and observation. The transfusion can be accomplished by means of the umbilical artery and titrated to the desired hematocrit.

In order to determine the rate of fetal prognosis, bilirubin absorbance results are plotted on a Liley chart. The chart is based on weeks of gestation/gestational age across the x-axis and the optical density (OD) of bilirubin measured at 450 nm. Based on the three relatively equal divisions of the Liley Zone chart, if bilirubin results are in the upper third of the chart, noted usually as Zone 1,

Table 23.2 Risk Interpretation of Second Trimester Maternal Serum Quad Testing Analyte Patterns

Increased risk for	hCG	Unconjugated Estriol	AFP	Inhibin A
Open spina bifida (NTD)	Normal	Normal	Increased	Not applicable
Anencephaly (NTD with baby born without parts of skull or brain)	Decreased	Decreased	Increased	Not applicable
Down syndrome (trisomy 21)	Increased	Increased	Decreased	Not applicable
Edwards syndrome (trisomy 18)	Decreased	Decreased	Decreased	Not applicable

the fetus's condition is considered desperate, and the best course of action is immediate delivery or transfusion. In the area designated at Zone 2, the fetus would have a delivery timeframe of 36 to 39 weeks. In Zone 3 the fetus is considered not anemic, and delivery will be at term.

Neural Tube Defects

Screening for NTDs is initially performed using maternal serum. The presence of elevated levels of α-fetoprotein (AFP) is primarily associated with NTDs such as spina bifida and anencephaly. Elevated maternal serum AFP is also found with abdominal hernias into the umbilical cord, cystic hygroma, and poor pregnancy outcome. Low maternal serum AFP is associated with an increased incidence of Down's syndrome and other aneuploidy states.

AFP is a product of first the fetal yolk sac and then the fetal liver. It is released into the fetal circulation and presumably enters the AF by transudation. Entry into the maternal circulation could be by placenta crossover or from the AF. An open NTD (e.g., spina bifida) that causes an increase in amniotic fluid α-fetoprotein (AFAFP) is accompanied by an increase in maternal serum AFP. Under normal conditions, AFAFP would be cleared by fetal swallowing and metabolism. An increased presence overloads this mechanism, causing AFAFP elevation. Both serum and AF are routinely analyzed using immunologic methods. Results are reported as multiples of the median (MoM). Because of the variety of demographics involved in determining normal values, each laboratory should establish its own median values for gestational weeks (usually weeks 15 to 21). Results can then be calculated using the formula:

$$\text{MoM} = \frac{\text{Specimen AFP concentration}}{\text{Median AFP concentration for gestational week}} \quad \text{(Eq. 23.3)}$$

The protocol for AFP testing to determine the presence of an NTD is generally the following (**Table 23.2**)[17]:

1. Maternal serum AFP, usually in conjunction with hCG, unconjugated estriol, and inhibin A assays, known as the prenatal quad test
2. Repeat the AFP, if elevated
3. Diagnostic ultrasound
4. Amniocentesis for confirmation

Interpretation of maternal serum AFP testing is complex, being a function of age, race, weight, gestational age, and level of nutrition. Therefore, the material serum AFP can only be used as a screening tool and should not be considered diagnostics for any of the conditions it is associated with and listed above. Testing of AFAFP is the confirmatory procedure to detect NTDs.

Concern over the difficulty in interpreting AFP tests generated the need for a second test to confirm NTDs and abdominal wall defects. The secondary method used is the assay for a central nervous system (CNS)-specific acetylcholinesterase (AChE). The NTD allows direct or, at least less difficult, passage of AChE into the AF. Analysis for CNS-specific AChE in the AF then offers a degree of confirmation for AFAFP. The methods used for CNS AChE include enzymatic, immunologic, and electrophoretic with inhibition. The latter includes the use of acetylthiocholine as substrate and BW284C51, a specific CNS inhibitor, to differentiate the serum pseudocholinesterase from the CNS-specific AChE.

Fetal Lung Maturity

The primary reason for most AF testing is the need to assess fetal pulmonary maturity. All the organ systems are at jeopardy from prematurity, but the state of the fetal lungs is a priority from the clinical perspective. Consequently, the laboratory is

asked whether sufficient specific phospholipids are reflected in the AF to prevent atelectasis (alveolar collapse) if the fetus were to be delivered. This question is important when preterm delivery is contemplated because of other risk factors in pregnancy, such as preeclampsia and premature rupture of membranes. Risk factors to fetus or mother can be weighed against interventions, such as delay of delivery with steroid administration to the mother to enhance fetal **surfactant** production, or against at-risk postdelivery therapies, such as exogenous surfactant therapy, high-frequency ventilation, and extracorporeal membrane oxygenation.

Alveolar collapse in the neonatal lung may occur during the changeover from placental oxygen to room air as an oxygen source at birth if the proper quantity and type of phospholipid (surfactant) is not present. The ensuing condition, which may vary in the degree of severity, is a type of **respiratory distress syndrome (RDS)**. It has also been referred to as *hyaline membrane disease* because of the hyaline membrane found in affected lungs. Lung maturation is a function of differentiation of alveolar epithelial cells (pneumocytes) into type I and type II cells beginning near the 24th week of pregnancy. The type I cells form the alveolar–capillary membrane for exchange of gases. The type II cells produce and store the surfactants needed for alveolar stability in the form of lamellar bodies.

As the lungs mature, increases occur in phospholipid concentration—particularly the compounds phosphatidylglycerol (PG) and lecithin.[18] These two compounds, comprising 10% and 70% of total phospholipid concentration, respectively, are most important as surfactants. Their presence in high enough levels acts in concert to allow contraction and re-expansion of the neonatal alveoli. Insufficient surfactant allows alveoli to collapse, requiring a great deal of energy to re-expand the alveoli upon inspiration. This not only creates an extreme energy demand on a newborn but probably also causes physical damage to the alveoli with each collapse. The damage may lead to "hyaline" deposition, or the newborn may not have the strength to continue inspiration at the energy cost. The result of either can be fatal.

Tests for assessing fetal lung maturity (FLM) include functional assays and quantitative assays. Functional assays provide a direct physical measure of AF in an attempt to assess surfactant ability to decrease surface tension. These include the "bubble or shake" test and the foam stability index (FSI).[19] Quantitative tests include the lecithin–sphingomyelin ratio (**L/S ratio**), PG, and lamellar body counts.

The FSI, a variant of Clements and colleagues' original "bubble test"[20] that was performed at the bedside, appears acceptable as a rapid, inexpensive, and informative assay. This qualitative, technique-dependent test requires only common equipment. The assay is based on the ability of surfactant to generate a surface tension lower than that of a 0.47 mol fraction ethanol–water solution. If sufficient surfactant is present, a stable ring of foam bubbles remains at the air–liquid interface. As surfactant increases (FLM probability increases), a larger mole fraction of ethanol is required to overcome the surfactant-controlled surface tension. The highest mole fraction used while still maintaining a stable ring of bubbles at the air–liquid interface is reported as the FSI (**Table 23.3**). The test is dependent on technique and can also be skewed by contamination of any kind in the AF (e.g., blood or meconium contamination). Interpretation of the FSI bubble patterns is difficult and technique dependent. Results can vary among clinical laboratorians. Most laboratories have found that an FSI of 0.47 correlates well with an L/S ratio of 2.0.

The quantitative tests were given emphasis primarily by the work of Gluck et al.[21] The phospholipids of importance in these tests are PG, lecithin, and sphingomyelin (SP). Relative amounts of PG and lecithin increase dramatically with pulmonary maturity, whereas SP concentration is relatively constant and provides a baseline for the L/S ratio. Increases in PG and lecithin correspond to the larger amounts of surfactant being produced by the type II pneumocytes as fetal lungs mature.

The classic technique for separation and evaluation of the lipids involves thin-layer chromatography (TLC) of an extract of the AF. The extraction procedure removes most interfering substances and results in a concentrated lipid solution. Current practices use either one- or two-dimensional TLC for identification. Laboratory needs determine if a one-dimensional or a two-dimensional method is performed. An example of the phospholipids

Table 23.3 FSI Determination

	Tube 1	Tube 2	Tube 3
Vol amniotic fluid	0.50	0.50	0.50
Vol 95% EtOH	0.51	0.53	0.55
FSI	0.47	0.48	0.49

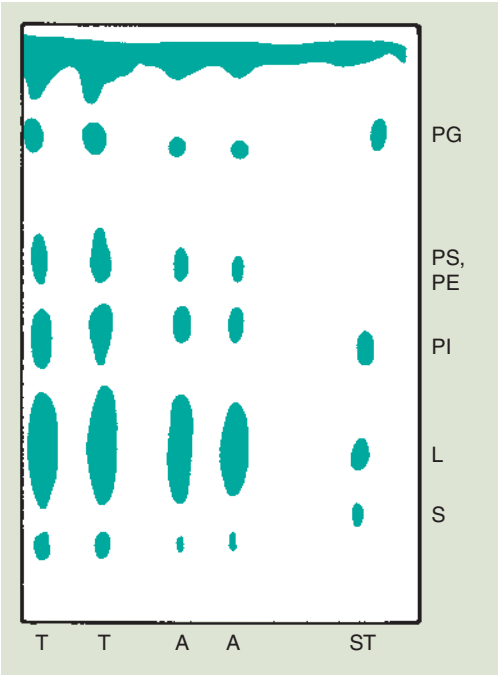


Figure 23.12 Thin-layer chromatogram of AF phospholipids. Standard phospholipids (*ST*), total extract (*T*), and acetone precipitable compounds (*A*) in AF are shown. The phospholipid standards contained, per liter, 2 g each of lecithin and P1, 1 g each of PG and SP, and 0.3 g each of PS and PE; 10 mL of the standard was spotted.

Reprinted from Tsai MY, Marshall JG. Phosphatidylglycerol in 261 samples of amniotic fluid. *Clin Chem.* 1979;25(5):683, with permission. Copyright 1979 American Association for Clinical Chemistry.

separation by one-dimensional TLC is shown in **Figure 23.12**.[22]

The classic breakpoint for judgment of maturity has been an L/S ratio of 2.0. As a result of the time-consuming requirements of performing the L/S ratio, several additional tests have been developed to allow faster determination of FLM. The L/S ratio, however, remains the gold standard by which all methods are compared.

Phosphatidylglycerol

As mentioned previously, an additional phospholipid essential for FLM is PG that increases in proportion to lecithin. In the case of diabetic mothers, however, development of PG is delayed. Therefore, using an L/S ratio of 2.0 as an indicator of FLM cannot be relied upon to ensure that RDS will not occur unless PG is also included in interpretation of the L/S ratio. With the current trend toward less labor-intensive techniques, an immunologic assay using antibody specific for PG can be used to determine FLM.

The AmnioStat-FLM (Irvine Scientific, Santa Ana, CA) immunologic test is designed to measure the adequate presence of PG in AF. Because lecithin production is not affected in diabetic mothers and the levels of PG and lecithin rise at the same rate in unaffected pregnancies, the AmnioStat-FLM can be used to determine whether adequate FLM is present. Good correlation has been shown compared with the L/S ratio. The antibody-specific immunologic assay offers the additional advantage, not present in other assays, of being unaffected by specimen debris such as meconium and blood.

Lamellar Body Counts

The phospholipids produced and secreted by the type II alveolar cells are released in the form of lamellar bodies. As FLM increases, these lamellated packets of surfactant also exhibit an increased presence in the AF. The fact that lamellar bodies are approximately the same size as platelets provides a convenient method to determine

CASE STUDY 23.1, PART 4

Remember Gina. She is the oldest daughter of the Doe family, became pregnant with her first child at 21 years of age. She had a low-risk pregnancy and was prescribed folate supplements to take. During a routine OBGYN office visit, a maternal serum α-fetoprotein level was collected and found to be elevated. This made Gina and her family quite nervous for the health of her child. Ultrasound confirmed the gestational age and that this was a single gestational pregnancy.

9. What condition is maternal serum α-fetoprotein used to screen for?
10. What follow-up test should be completed next? What should be done if it is normal? Abnormal?

© digitalskillet/Shutterstock.

their concentration using the platelet channel on automated hematology analyzers.[23] Amniotic fluid lamellar body counts greater than 50,000/μL are predictive of fetal lung maturity, whereas lamellar body counts less than 15,000/μL are suggestive of fetal lung immaturity and increased risk of neonatal RDS. Acceptable counts must be correlated with the specific instrumentation.[24] A standardized protocol has been developed in an effort to make the assay transferable between laboratories and has allowed a greater incorporation of this assay into practice.[25]

Sweat

The common exocrine sweat glands function in the regulation of body temperature. They are innervated by cholinergic nerve fibers and are a type of exocrine gland. Sweat has been analyzed for its multiple inorganic and organic contents but, with one notable exception, has not been proved as a clinically useful model. That exception is the analysis of sweat for chloride levels in the diagnosis of cystic fibrosis (CF). This test has recently been given the misnomer "sweat electrolytes" in some locations, even though the only electrolyte measurement taken is that of chloride.[26] The sweat test is the single most accepted common diagnostic tool for the clinical identification of this disease. Normally, the coiled lower part of the sweat gland secretes a "presweat" upon cholinergic stimulation. As the presweat traverses the ductal part of the gland going through the dermis, various constituents are resorbed. In CF, the electrolytes, most notably chloride and sodium ions, are improperly resorbed owing to a mutation in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene, which controls a cyclic AMP–regulated chloride channel.

CF (mucoviscidosis) is an autosomal recessive inherited disease that affects the exocrine glands and causes electrolyte and mucous secretion abnormalities. This exocrinopathy is present only in the homozygous state. The frequency of the carrier (heterozygous) state is estimated at 1 of 20 in the United States predominantly affects those of European decent. The observed rate of expression ranks CF as the most common lethal hereditary disease in the United States, with death usually occurring by the third decade. The primary cause of death is pneumonia, secondary to the heavy, abnormally viscous secretions in the lungs. These heavy secretions cause obstruction of the micro airways, predisposing the CF patient to repeated episodes of pneumonia. Patients also experience pancreatic insufficiency due to abnormally viscous secretions obstructing pancreatic ducts. This obstruction ostensibly causes pooling and autoactivation of the pancreatic enzymes. The enzymes then cause destruction of the exocrine pancreatic tissue.

Diagnostic algorithms for CF continue to rely on abnormal sweat chloride, pancreatic or bronchial abnormalities, and family history. The use of blood immunoreactive trypsin, a pancreatic product, is now prevalent in newborn screening programs. The rapidly developing area of molecular genetics provides the definitive methodology. The gene defect causing CF has been localized on chromosome 7, and the most frequent mutations have been characterized. However, approximately 1500 mutations of the *CFTR* gene have been cataloged, with the majority of patients found to be homozygous or heterozygous for the most common mutation.

The sweat glands remain structurally unaffected by CF. Analysis of sweat for both sodium and chloride is valid, but, historically, chloride was and is the major element, leading to use of the sweat chloride test. Because of its importance, a standard method has been suggested by the Cystic Fibrosis Foundation. This method is based on the pilocarpine nitrate iontophoresis method of Gibson and Cooke.[27] Pilocarpine is a cholinergic-like drug used to stimulate the sweat glands. The sweat is absorbed on a gauze pad during the procedure. After collecting sweat by iontophoresis, chloride analysis is performed. Many methods have been suggested, and all are dependent on laboratory requirements. Generally, the sweat is leached into a known volume of distilled water and analyzed for chloride (chloridometer). In general, values greater than 60 mmol/L are considered positive and are usually repeated at another date for confirmation. A repeat sweat test >60 mmol/L with a positive genetic test is diagnostic for CF.

Other tests including osmolarity, conductivity, and chloride electrodes or patches placed on the skin are available but are considered screening tests with abnormal results followed by the Gibson-Cooke reference method. A variety of instrumentation is available for these screening tests.

Although a value of 60 mmol/L is generally recognized for the quantitative pilocarpine iontophoretic test, it is important to consider several factors in interpretation. Not only will there be analytic variation around the cutoff, an epidemiologic borderline area will also occur. Considering this, values over 30 mmol/L for chloride are more appropriate in determining the need for repetition. Some

patients with *CFTR* mutations have been found to have values below 60 mmol/L. Other variables must be considered. Age generally increases the limit—so much so that it is increasingly difficult to classify adults. The patient's state of hydration also affects sweat levels. Because the complete procedure is technically demanding, expertise should be developed before the test is clinically available. A complete description of sweat collection and analysis, including procedural justifications, is available for review from CLSI document C34-A2.[28]

Synovial Fluid

Joints are classified as movable or immovable. The movable joint contains a cavity that is enclosed by a capsule; the inner lining of the capsule is the synovial membrane (**Figure 23.13**). This cavity contains *synovial fluid*, which is formed by ultrafiltration of plasma across the synovial membrane. Along with ultrafiltration of the plasma, the membrane also secretes a mucoprotein rich in hyaluronic acid. This mucoprotein is what causes the **synovial fluid** to be viscous. Synovial fluid functions as a lubricant for the joints and as a transport medium for delivery of nutrients and removal of cell wastes. The volume of fluid found in a large joint, such as the knee, rarely exceeds 3 mL. Normal fluid is clear, colorless to pale yellow, viscous, and nonclotting. Variations are indicative of pathologic conditions and are summarized in **Table 23.4**. Collection of a synovial fluid sample is accomplished by arthrocentesis of the joint under aseptic conditions (**Figure 23.14**). The sample should be collected in a heparin tube for culture, a heparin or liquid EDTA for microscopic analysis, and a fluoride tube for glucose analysis. The viscosity of the fluid often requires the use of hyaluronidase to break down the mucoprotein matrix and allow for appropriate manual or automated aspiration and delivery into reaction vessels.

Chemical analysis of synovial fluid includes the testing of several different analytes. These analytes include total protein, glucose, uric acid, and LD. Total protein within the synovial fluid can be measured using the most common methods performed on serum samples. The normal range for synovial fluid protein is 1 to 3 g/dL. Increased synovial fluid protein levels are seen in ankylosing spondylitis, arthritis, arthropathies that accompany Crohn's disease, gout, psoriasis, Reiter's syndrome, and ulcerative colitis.

Synovial fluid glucose levels are interpreted using serum glucose levels after a 6 to 8-hour fast. Normally, synovial fluid glucose levels are less than 10 mg/dL lower than serum levels. Infectious disorders of the joint demonstrate large decreases in synovial fluid glucose and can be as much as 20 to 100 mg/dL less than serum levels. Other groups of joint disorders typically demonstrate a less degree of decrease in synovial fluid glucose. The ratio of synovial fluid to plasma glucose (normally 0.9:1) remains the most useful mechanism

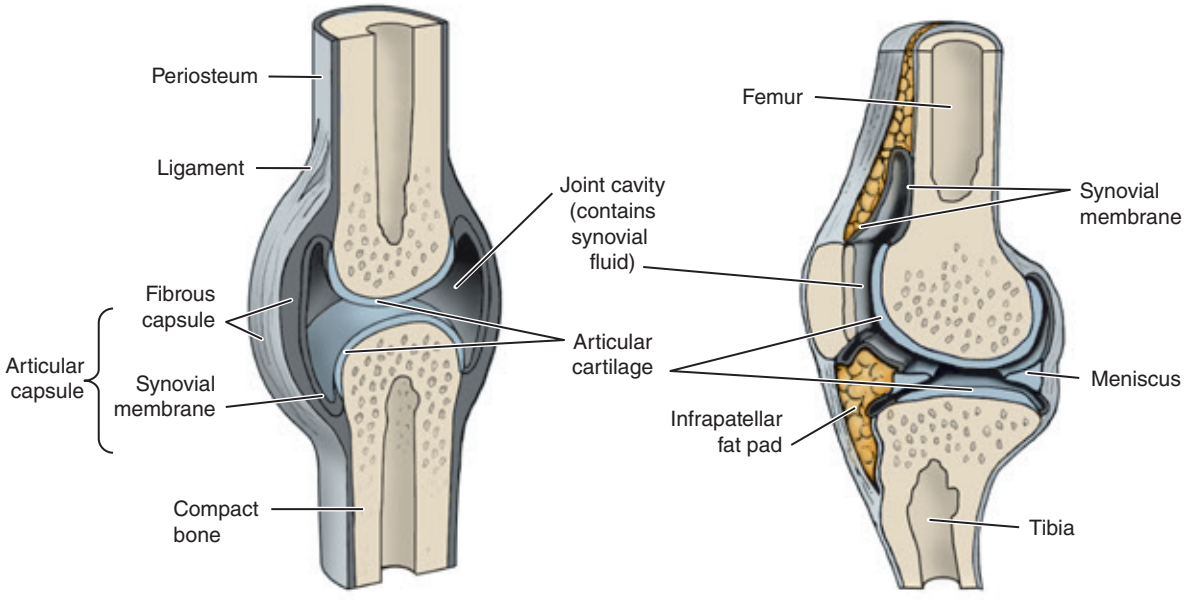


Figure 23.13 Synovial joint and model.

Table 23.4 Classification of Synovial Fluids

Group	Category	Visual	Viscosity	Mucin Clot	Cell Count (Cells/HPF)	Glucose Blood: SF	Other
	Normal	Colorless—straw Good	High	Good	<150 WBCs <25% neutrophils	0–10	
I	Noninflammatory	Yellow Slightly cloudy	Decreased	Fair	≤ 1,000 WBCs <30% neutrophils	0–10	
II	Inflammatory	White, gray, yellow Cloudy, turbid	Absent	Poor	<100,000 WBCs >50% neutrophils	0–4	
III	Septic	White, gray, yellow, or green Cloudy, purulent	Absent	Poor	50,000–200,000 WBCs >90% neutrophils	20–100	Positive cultures
IV	Crystal induced	White Cloudy, turbid, opaque, milky	Absent	Poor	500–200,000 WBCs <90% neutrophils	0–80	Crystals present
V	Hemorrhagic	Sanguinous, xanthochromic, red, or brown Cloudy	Absent	Poor	50–10,000 WBCs <50% neutrophils	0–20	RBCs present

Mundt L, Shanahan K. *Graff's Textbook of Routine Urinalysis and Body Fluids.* 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2011:268.

to evaluate glucose levels within the synovial fluid. Decreased ratios are found in inflammatory (e.g., gout, rheumatoid arthritis [RA], and systemic lupus erythematosus) and septic (e.g., bacterial and viral arthritis) conditions. Standard methods for glucose analysis can be used for synovial fluid analysis.

Synovial fluid uric acid levels should be determined in patients with a monoarthritic episode to confirm or rule out the presence of gouty arthritis since plasma uric acid levels do not correlate with synovial fluid levels. While identifying the presence of uric acid crystals via microscopic methods is diagnostic of gouty arthritis, the determination of synovial fluid uric acid levels can be used in laboratories where crystal identification is not a routine procedure. Synovial fluid uric acid levels normally range from 6 to 8 mg/dL.

Although rarely performed, lactic acid measurements can be helpful in diagnosing septic arthritis. Normally, synovial fluid lactate is less than 25 mg/dL but can be as high as 1000 mg/dL in septic arthritis. LD can be elevated in synovial fluid, while serum levels remain normal. Synovial fluid LD levels are usually increased in RA, infectious arthritis, and gout.

Rheumatoid factor (RF) is an antibody to immunoglobulins. Between 60% and 80% of patients with rheumatoid arthritis have RF in their serum,[29] whereas more than half of these patients will demonstrate RF in synovial fluid. As such, some providers may request the qualitative or quantitative detection in synovial fluid.

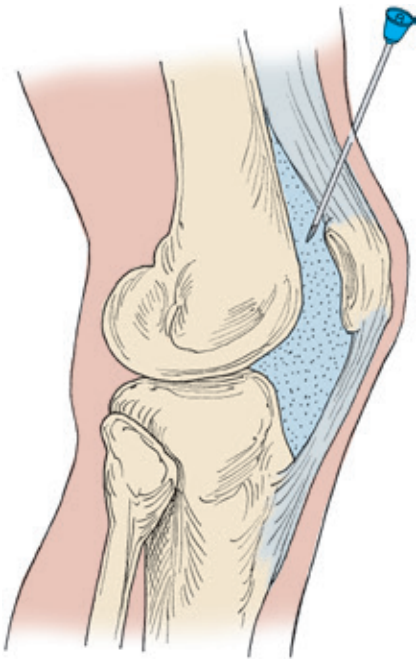


Figure 23.14 Placement of needle in arthrocentesis of knee joint.

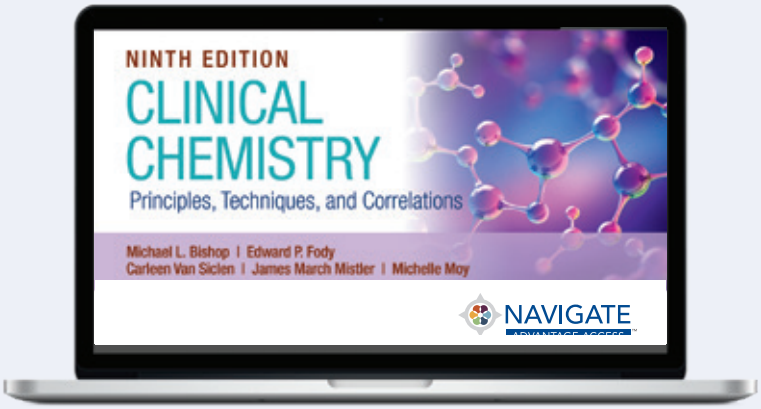
CASE STUDY 23.1, PART 5

Remember Chad. Chad, the father of the Doe family, develops severe, sudden onset pain in the big toe of his left foot. The toe appears swollen and red. Chad reports no physical injury has occurred that he can think of. The physician at the local hospital collects synovial fluid from the swollen joint and sends it to the laboratory for crystal analysis.

11. Based upon the physician's order, predict what condition they are suspecting is the cause.
12. What is the main crystal of interest in synovial fluid for a case such as this?
13. The synovial fluid is found to be extremely viscous upon receipt in the lab. What is causing this, and how can it be addressed to adequately analyze the specimen?

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Meyfroidt G, Kurtz P, Sonneville R. Critical care management of infectious meningitis and encephalitis. *Intens Care Med*. 2020;46:92–201.
2. Brunzel NA. *Fundamentals of Urine and Body Fluid Analysis*. 2nd ed. Philadelphia, PA: Saunders; 2004:327–328.
3. McBride LJ. *Textbook of Urinalysis and Body Fluids: A Clinical Approach*. New York, NY: Lippincott; 1998:200–201.
4. Li Y, Li H, Fan R, et al. Coronavirus infections in the central nervous system and respiratory tract show distinct features in hospitalized children. *Intervirology*. 2016;59: 163–169.
5. Okta M, Okta K, Ma J, et al. Clinical and analytical evaluation of an enzyme immunoassay for myelin basic protein in cerebrospinal fluid. *Clin Chem*. 2000;46:1336–1330.
6. Llano DA, Bundela S, Mudar RA, Devanarayan V. A multivariate predictive modeling approach reveals a novel CSF peptide signature for both Alzheimer's Disease state classification and for predicting future disease progression. *PLoS One*. 2017;12(8):e0182098.
7. Bailey EM, Domenico P, Cunha BA. Bacterial or viral meningitis. *Postgrad Med*. 1990;88:217.
8. Ross DL, Neeley AE. *Textbook of Urinalysis and Body Fluids*. New York, NY: Appleton-Century-Crofts; 1983:275.
9. Calbreath DF. *Clinical Chemistry: A Fundamental Textbook*. Philadelphia, PA: Saunders; 1992:370.
10. Cornes MP, Chadburn AJ, Thomas C, et al. The impact of between analytical platform variability on the classification of pleural effusions into exudate or transudate using Light's criteria. *J Clin Pathol*. 2017;70:607–609.
11. Kumar IV, Vani N, Jothula KY. A study of uric acid levels in serum and pleural fluid as to differentiate between transudate and exudate in pleural effusion. *Int J Clin Biochem Res*. 2020;7(2):216–220.
12. Bielsa S, Porcel JM, Castellote J, et al. Solving the Light's criteria misclassification rate of cardiac and hepatic transudates. *Respirology*. 2012;17(4):721–726.
13. Romero-Candeira S, Fernandez C, Martín C, et al. Influence of diuretics on the concentration of proteins and other components of pleural transudates in patients with heart failure. *Am J Med*. 2001;110(9):681–686.
14. Porcel JM, Martinez-Alonso M, Cao G, et al. Biomarkers of heart failure in pleural fluid. *Chest*. 2009;136(3):671–677.

15. Runyon BA, Montano AA, Akriviadis EA. The serum-ascites albumin gradient is superior to the exudate-transudate concept in the differential diagnosis of ascites. *Ann Intern Med.* 1992;117(3):215–220.
16. Liley AW. Liquor amnil analysis in the management of the pregnancy complicated by rhesus sensitization. *Am J Obstet Gynecol.* 1961;82:1359.
17. Rose NC, Mennuti MT. Maternal serum screening for neural tube defects and fetal chromosome abnormalities. *West J Med.* 1993 Sep;159(3):312–317.
18. Kulovich MV, Hallman MB, Gluck L. The lung profile I. The normal pregnancy. *Am J Obstet Gynecol.* 1979;135(1):57–63.
19. Statland BE, Freer DE. Evaluation of two assays of functional surfactant in amniotic fluid: surface-tension lowering ability and the foam stability index test. *Clin Chem.* 1979;25:1770.
20. Clements JA, Plataker ACG, Tierney DF, et al. Assessment of the risk of the respiratory distress syndrome by a rapid test for surfactant in amniotic fluid. *N Engl J Med.* 1972;286: 1077–1081.
21. Gluck L, Kulovich MV, Borer RC Jr, et al. Diagnosis of the respiratory distress syndrome by amniocentesis. *Am J Obstet Gynecol.* 1971;109:440–445.
22. Tsai MY, Marshall JG. Phosphatidylglycerol in 261 samples of amniotic fluid from normal and diabetic pregnancies, as measured by one-dimensional thin layer chromatography. *Clin Chem.* 1979;25:682–685.
23. Khazardoost H, Yakazadeh S, Borna F, et al. Amniotic fluid lamellar body count and its sensitivity and specificity evaluating fetal lung maturity. *J Obstet Gynecol.* 2005;25: 257–259.
24. Szallasi A, Gronowski AM, Eby CS. Lamellar body count in amniotic fluid: a comparison of four different hematology analyzers. *Clin Chem.* 2003;49:994–997.
25. Neerhof ME, Dohnal JC, Ashwood ER, et al. Lamellar body counts: a consensus on protocol. *Obstet Gynecol.* 2001;97:318–320.
26. Sweat Chloride Test. Lab Tests Online. https://labtestsonline.org/tests/sweat-chloride-test. Accessed on: Mar 21, 2021.
27. Gibson LE, Cooke RE. A test for concentration of electrolytes in cystic fibrosis of the pancreas utilizing pilocarpine by iontophoresis. *Pediatrics.* 1959;23:545–549.
28. Clinical Laboratory Standards Institute. *Sweat Testing: Sample Collection and Quantitative Analysis Approved Guidelines.* Villanova, PA: Clinical Laboratory Standards Institute, reaffirmed 2005. (CLSI document C34-A2.)
29. Johnson PM, Faulk WP. Rheumatoid factor: its nature, specificity, and production in rheumatoid arthritis. *Clin Immunol Immunopathol.* 1976;6(3):414–430.

CHAPTER 24

Pregnancy and Prenatal Testing

Heather L. Phillips and Michelle R. Campbell

CHAPTER OUTLINE

KEY TERMS

Blastocyst
Conception
Embryo
Fetus
Gestation
Ovum
Ovulation
Placenta
Zygote

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Describe human development from conception to birth.
- Discuss the biological changes (e.g., endocrine, renal, immune) that occur within the mother's body during pregnancy.
- State the importance of testing the maternal serum α-fetoprotein in the second trimester of pregnancy and correlate levels with normal and disease states.
- Specify the methods of analysis for α-fetoprotein.
- Explain the purpose of testing maternal unconjugated estriol concentrations.
- Correlate expected laboratory results and clinical symptoms to the following fetal complications: neural tube defects, Down syndrome, trisomy 18, isoimmunization, and preterm delivery.
- Interpret laboratory results and clinical symptoms to the following maternal and pregnancy complications: preeclampsia, hyperemesis gravidarum, and ectopic pregnancy.

The clinical laboratory plays a critical role in the diagnostic testing of the pregnant female and her unborn **fetus**. The health of the fetus and mother are intertwined during pregnancy, making it critical to have a firm understanding of both the pregnancy and fetal development. Laboratory tests are used to determine the presence of pregnancy and monitor the health and wellbeing of the fetus and the mother. It is important to have a fundamental understanding of the development of a pregnancy, the physiology of pregnancy, and maternal adaptions in order to understand the role of in clinical testing for pregnancy health-care.

CASE STUDY 24.1, PART 1

Lorna, a 38-year-old female, is 23 weeks pregnant with her third child. Although the pregnancy was unplanned, she describes it as "typical" and "uneventful" in comparison to her previous two. Both she and her husband have no family history of genetic disorders.

© Oliver Rossi/DigitalVision/Getty Images.

CASE STUDY 24.2, PART 1

Caroline, a 25-year-old female, is at 16 weeks' gestation in her first pregnancy. To date, her pregnancy has been unremarkable.

© Robin Skjoldborg/DigitalVision/Getty Images.

CASE STUDY 24.3, PART 1

Nina, a 32-year-old female, is in the 9th week of pregnancy. Other than morning sickness, her pregnancy has been unremarkable. A transvaginal ultrasound at 8 weeks gestation was normal with an active embryo.

© Leland Bobbe/DigitalVision/Getty Images.

Pregnancy

A normal human pregnancy is approximately 40 weeks when **gestation** is measured from the first day of the last menstrual period. A pregnancy is divided into three sections. The first section is referred to as the first trimester. The first trimester begins at the day of **conception** and lasts for 12 weeks. The 13th week starts the second trimester, during which fetal growth is exponential. The second trimester ends at the 26th week of gestation. The third trimester begins at 26 weeks and ends with the delivery of a newborn. The term *gestation* is used as an indication of fetal development from the last period, while *conception* is used as an indication of development from the date the sperm fertilizes the ovum.

Conception and Embryonic Development

Before conception can occur in the female reproductive system, a mature and healthy **ovum** (egg) must be released from the ovary into the fallopian tube. This is referred to as **ovulation**. Under normal circumstances, ovulation occurs on the 14th day of the typical normal menstrual cycle. Once the ovum has been fertilized in the fallopian tube by a sperm cell, it is then referred to as the **zygote**. Rapid cellular division (meiosis) begins as the zygote travels down the fallopian tube toward the uterus. Once meiosis has created between 50 to 60 cells, the zygote will then form a cavity that becomes the primitive yolk sac. With this development, the cluster of cells becomes a blastocyst, with the immediate goal of beginning implantation into the uterine wall. At this point, the cluster of cells is roughly five days past fertilization and needs a **placenta** to sustain the pregnancy; the placenta is the functional interface between the mother and fetus that secretes hormones and growth factors into the mother with physiological effects. The cells on the outside of the **blastocyst** become trophoblasts, which invade the uterine endometrium to create the placenta.

Once implantation occurs, the cluster of cells is now called an **embryo** (**Figure 24.1**). The cells will continue to rapidly divide and differentiate into various cellular types. The embryo, now nourished by the placenta and protected in a layer of amniotic fluid, will continue to develop and create organs through a process known as *organogenesis*. This process takes up to 10 weeks.

The placenta is critical to maintain the pregnancy and development of the embryo and fetus. The placenta not only provides the nutritional and oxygen supply for the embryo, but it is also responsible for the secretion of hormones necessary to maintain the pregnancy. The placenta produces the hormones and steroids necessary for maternal adaptation, the transformation of the female body to support the pregnancy. The placenta is attached to the uterine wall at 4 weeks' gestation, or 2 weeks after conception. When the placenta begins secreting hormones, the pregnancy can be detected using clinical laboratory testing methods.

Fetal Development

At 10 weeks past conception, the fetus has developed its organs, has a heartbeat, and moves its arms and legs regularly. Under normal conditions, the fetus is roughly 8 cm long and weighs nearly 13 g. During the

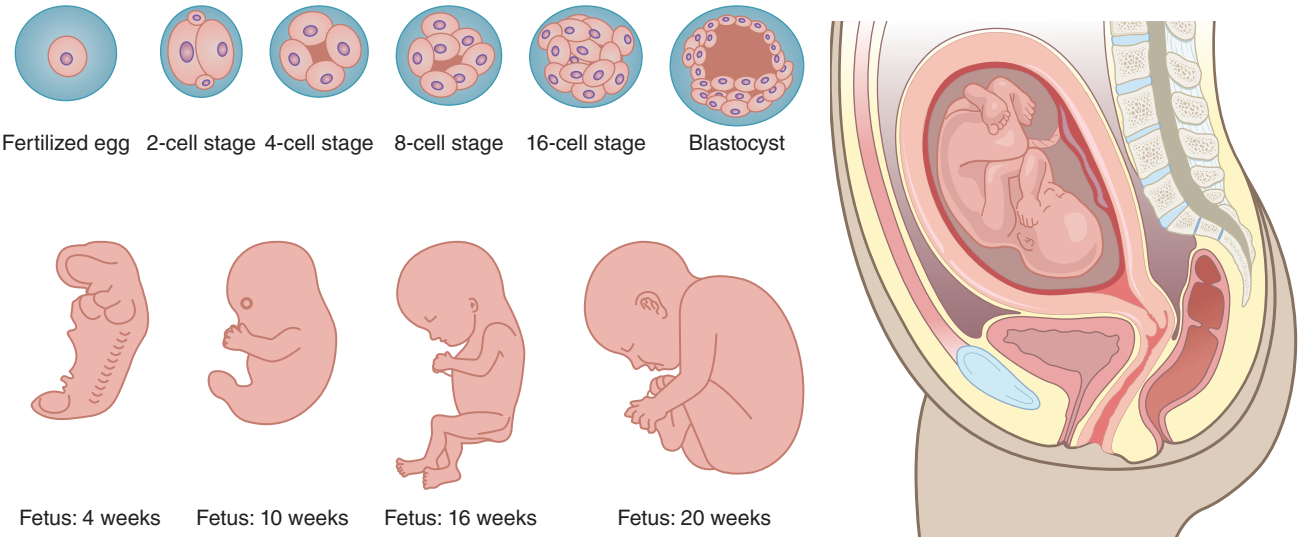


Figure 24.1 A developing human embryo.

second trimester, the fetus will begin to grow rapidly. The fetus will begin to freely move about, can develop hiccups, and can yawn. The genitalia become visible and functionally release sex hormones. The ovaries of a female fetus will be full of millions of immature eggs, and the testes of a male fetus will begin releasing testosterone. The fetus's organs will begin maturing, allowing it to swallow amniotic fluid, urinate, suck its thumb, and respond to sounds outside the womb. When the gestational 26th week has been reached, most fetuses born prematurely can survive if sufficient neonatal care is available.

In the third trimester of development (**Figure 24.2**), the fetus gains at least 200 g of weight per week (about 7 ounces). The eyes are now open and blink at regular intervals, and instead of occasional expansions of the diaphragm, the fetus now practices breathing. The sucking reflex that begins at 34 weeks will be used to gain suckling muscle strength for delivery. The last 4 weeks of pregnancy are used mainly for muscle development, building fat storage, and practicing movements required for life after delivery.

Delivery

A pregnancy that has come to "term" refers to a level of development at which a fetus can be safely delivered via vaginal birth and survive without significant intervention. A healthy fetus is usually delivered between 38 and 42 weeks' gestation. It is unknown exactly what triggers delivery. However, researchers believe labor is induced by hormones released by the fetus influencing the mother's uterus.

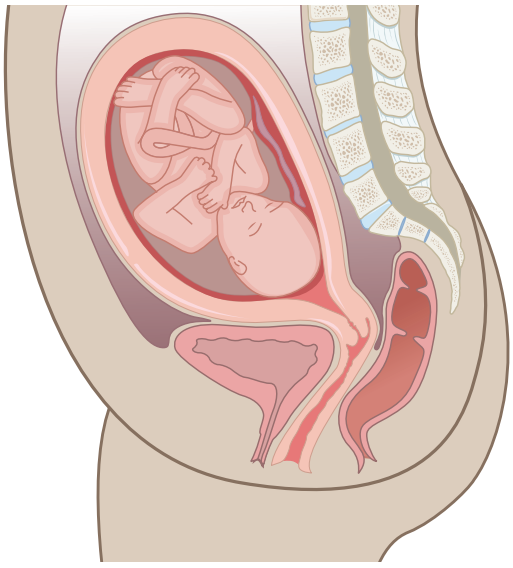


Figure 24.2 A human fetus at the last stage of development.

The hormone progesterone increases steadily throughout pregnancy. Closer to term, the progesterone hormone levels rapidly decrease before labor begins. It is thought commencement of labor begins when progesterone levels decrease to non-pregnant levels. Progesterone prevents the production of prostaglandins and decreases the expression of genes that would cause the formation of receptors on smooth muscle cells whose main purpose is contractability.[1] Ultimately, increased levels of progesterone decrease the contractility of uterine smooth muscle cells. On the other end of the spectrum, estrogen promotes the expression of receptors on muscle tissue cells that aid in contraction. When the progesterone levels decrease, estrogen levels begin to rapidly increase, allowing for smooth muscle contractions.

It is widely known that the hormone oxytocin is responsible for triggering the maternal nursing behavior, as well as the induction of labor, due to its effects on the human brain and mammary glands. When a pregnancy comes to term, the posterior pituitary secretes high concentrations of oxytocin into the mother's peripheral circulation, which triggers uterine contractions and the initiation of parturition (childbirth). High levels of oxytocin also enable the mother to bond better and protect her newborn when it is most vulnerable.[1] Research demonstrates that when low levels of oxytocin are produced, strong bonding between mothers and newborns does not occur, and there is insufficient mammary gland stimulation.[1] The synthetic oxytocin drug Pitocin is used to induce labor or strengthen contractions during labor. Following delivery of the newborn, the mother's body will begin transformation back to its pre-pregnancy state under normal circumstances. A major contributor to this is the expulsion of the placenta, as it provides many of the hormones necessary to support pregnancy and maternal transformation.

Transformation During Pregnancy

During pregnancy, the mother will undergo significant anatomical and physiological changes to accommodate a growing fetus. These changes can mimic both early and advanced stages of disease. For this reason, it is important to understand the normal physiological changes that occur during pregnancy. These changes in the mother's body begin shortly after conception, and many maternal adaptions are largely mediated by the placenta. Under normal circumstances, after delivery of the fetus and expulsion of the placenta,

most physiological changes that occurred in the mother's body during pregnancy will return to the pre-pregnancy state.

Plasma Volume Increase

One of the first changes begins with the plasma increase that occurs in the mother's peripheral circulation. The plasma volume increases steadily over the course of pregnancy and is usually directly proportional to the weight of the growing fetus in the third trimester. Plasma volumes increase by 50% to as much as 85% of the values seen before pregnancy as extracellular volumes increase by 30% to 50%.[2] Plasma osmolality will fall by 10 mOsm/kg by the third trimester of pregnancy. The increase in plasma volume is necessary for the pregnancy to maintain a homeostatic blood pressure for the mother and developing fetus and for uteroplacental perfusion.[3] The increase in plasma dilutes the peripheral whole blood, causing what appears to be a decrease in erythrocyte concentrations. However, it is important to keep in mind this is due to fluid displacement and not the development of a disease process (i.e., a dilutional effect).

Renin–Angiotensin–Aldosterone System

The renin–angiotensin–aldosterone system (RAAS) is activated during pregnancy by the underfilling of arteries and the release of relaxin, a hormone produced by the ovary and the placenta. This causes the hypothalamus to trigger the release of vasopressin (also known as arginine vasopressin, AVP, and previously called anti-diuretic hormone or ADH) from the posterior pituitary. Vasopressin is a steroid hormone that causes the kidneys to retain water by altering the level of porin receptors in the tubular lumen membranes. In addition to the secretion of vasopressin, the activation of the RAAS causes the secretion of aldosterone by the adrenal glands, which allows for the retention of sodium by the distal convoluted tubules, and therefore water, since water will follow sodium. Aldosterone is the primary hormone responsible for the increase of plasma volume seen during pregnancy.[3] Aldosterone exhibits a threefold increase by the end of the first trimester and a 10-fold increase by the third trimester.[2]

The placental production of estrogens stimulates the liver to produce angiotensin. This causes a marginal increase of aldosterone compared to the volume of circulating renin. Progesterone, secreted during pregnancy to prevent uterine contractions, is an aldosterone antagonist and allows for the secretion of sodium even when high levels of aldosterone are present. The increase in the glomerular filtrate rate (GFR) during pregnancy also allows for the excretion of excess sodium in the urine. However, progesterone is a natural antikaluretic and causes the retention of potassium.[4] Therefore, serum potassium levels may be slightly increased during pregnancy but are often within the upper limit of the established reference range.

Maintaining this hypervolemic state would not be possible without pregnancy unless a disease state were present. The threshold for thirst and vasopressin is reset in pregnancy to a lower plasma osmolality level. These changes are mediated by the pregnancy hormones relaxin and human chorionic gonadotropin (hCG).[5]

Natural Decreases in Analytes

Pregnancy also causes fluctuations in many other analytes, such as iron. Additional iron is required for the increased demand of hemoglobin synthesis. A developing fetus will increase the oxygen carrying capacity needs of the mother, requiring more circulating erythrocytes. During pregnancy, a 20% increase in oxygen consumption is required.[2] This will stimulate the secretion of erythropoietin, a hormone produced by the kidneys. Erythropoietin stimulates the development of more erythrocytes to increase the oxygen load, creating an oxygen-rich environment for the developing fetus. The developing fetus also requires high volumes of iron for hemoglobin synthesis to produce the necessary enzymes for proper development.[6] The developing fetus also causes a 10- to 20-fold increase in folate, as well as a twofold increase in vitamin B_{12} requirements.[2]

Pregnancy is often associated with an iodine deficiency due to the active transport of peripheral iodine from the mother to the fetus through the placenta. Iodine is needed for fetal brain and nervous system development. The increased GFR during pregnancy causes an increased iodine excretion, further decreasing available iodine for maternal thyroid hormone production. The World Health Organization (WHO) recommends mothers increase their consumption of iodine by 100 to 200 mg per day.[2]

The dietary intake of protein is important during pregnancy. These proteins are broken down during digestion, and the amino acids are transported across the placenta to the developing fetus. During pregnancy, protein catabolism is decreased due to the increased synthesis of LDL cholesterol by the liver. The mother's body will use the circulating

triglycerides as an energy source, as opposed to performing protein catabolism as the body would do in non-pregnant or fasting states.

The fetus has a high demand for calcium, especially in the third trimester. Due to the increased demand for calcium, the mother's body doubles its calcium absorption and increases its availability of vitamin D. The developing fetus needs 30–50 g/day of free (ionized) calcium to maintain its development and physiological requirements, with most of the calcium needs being in the third trimester. The fetus's calcium requirement drops in the second trimester, and additional calcium in the mother's circulation is stored in the mother's skeleton. If sufficient calcium is not available for the fetus, calcium is released from the mother's bones. When testing for calcium in the mother's serum during pregnancy, the total calcium level will appear to be decreased. This is due to a natural decrease in circulating serum albumin levels causing a resultant decrease in the albumin-bound fraction of calcium. However, the serum ionized calcium levels will remain unchanged.[2] Because of this, it is more clinically relevant to measure ionized calcium levels during pregnancy as opposed to total circulating calcium.

Biochemically, many of the analytes affected are part of a greater analytic cascade. Each laboratory should take care to establish reference ranges for each analyte and, if necessary, establish a specific reference range for pregnant females. **Table 24.1** describes mean maternal serum laboratory values expected during normal pregnancies.

Upregulation of Hormones and Steroids

The upregulation of estradiol, prostaglandins, and endothelium-dependent factors cause peripheral dilation. Cardiac output will increase 20% by the eighth week of pregnancy,[2] and the maternal heart

Table 24.1 Reference Ranges for Maternal Serum Laboratory Values Expected During Normal Pregnancies by Trimester

	First	Second	Third
ALT	3–30 U/L	2–33 U/L	2–25 U/L
ALB	3.1–5.1 g/dL	2.6–4.5 g/dL	2.3–4.2 g/dL
ALP	17–88 U/L	25–126 U/L	38–229 U/L
AMY	24–83 U/L	16–73 U/L	15–87 U/L
AST	3–23 U/L	3–33 U/L	4–32 U/L
Bicarbonate	20–24 mmol/L	20–24 mmol/L	20–24 mmol/L
BUN	7–12 mg/dL	3–13 mg/dL	3–11 mg/dL
Ca	8.8–10.6 mg/dL	8.2–9.0 mg/dL	8.2–9.7 mg/dL
CL	101–105 mmol/L	97–109 mmol/L	97–109 mmol/L
Creat	0.4–0.7 mg/dL	0.4–0.8 mg/dL	0.4–0.9 mg/dL
Ion Ca	4.5–5.1 mg/dL	4.4–5.0 mg/dL	4.4–5.3 mg/dL
OSMO	275–280 mOsm/kg	276–289 mOsm/kg	278–280 mOsm/kg
K	3.6–5.0 mmol/L	3.3–5.0 mmol/L	3.3–5.1 mmol/L
Na	133–148 mmol/L	129–148 mmol/L	130–148 mmol/L
Pre-Alb	15–27 mg/dL	20–27 mg/dL	14–23 mg/dL
TBIL	0.1–0.4 mg/dL	0.1–0.8 mg/dL	0.1–1.1 mg/dL
TP	6.2–7.6 g/dL	5.7–6.9 g/dL	5.6–6.7 g/dL
TSH	0.60–3.40 uIU/mL	0.37–3.60 uIU/mL	0.38–4.04 uIU/mL

rate increases by 10 to 20 beats per minute. Blood pressure decreases in the first and second trimesters but usually returns to non-pregnant levels by the third trimester.[2] During delivery, the cardiac output increases by as much as 50% and then drops to pre-pregnancy levels between the first hour and two weeks after delivery, as long as no pregnancy-related hypertension exists.

Hormones such as relaxin begin rising at conception, peak at the completion of the first trimester, and remain elevated until delivery of the fetus. Relaxin is produced by the corpus luteum and the placenta. Relaxin is important as it plays a role in water balance during pregnancy and stimulates the formation of endothelin, which mediates the vasodilation of the renal arteries.[2]

Plasma estradiol levels secreted by the placenta cause the prolactin-producing cells to begin dividing rapidly in the anterior pituitary after conception. This causes enlargement of the maternal pituitary gland during pregnancy. Plasma concentrations of prolactin begin to increase dramatically during the first trimester and are 10 times higher when the pregnancy comes to term.[2] Estradiol, progesterone, and inhibin cause negative feedback for the hormones follicle-stimulating hormone (FSH) and luteinizing hormone (LH).

During pregnancy, secreted hormones such as human placental lactogen, growth hormone (GH), progesterone, cortisol, and prolactin affect insulin sensitivity in the mother. Insulin receptor signaling is affected on peripheral tissue and skeletal muscle during the second and third trimesters.[7] During the first trimester, insulin sensitivity is increased. During the second and third trimesters, the pregnant mother's body develops an insulin resistance. Insulin is secreted by the pancreatic beta cells in response to elevated circulating glucose concentrations caused by increased GH and cortisol levels. Under normal conditions, the concentrations of circulating insulin and glucose are inversely proportional. Glucose is important for the growing fetus's development, and insulin resistance is important for the maternal increased demands for glucose. Increased circulating insulin and relative hypoglycemia allows the mother's body to undergo lipolysis and use fat metabolism for energy. The circulating glucose and amino acids can then be saved for fetal development. When testing a fasting plasma glucose level during pregnancy, it is expected the plasma glucose levels will be slightly decreased. This is due to the rapid utilization of the available glucose in the mother's peripheral circulation.

Increased Glomerular Filtration Rate

As mentioned previously, the glomerular filtration rate (GFR) increases exponentially during pregnancy due to the prolific vasodilation. This increase in GFR can also be attributed to the increase in plasma volume, which causes a decrease in the oncotic pressure of the glomeruli. The GFR is expected to increase by 50% to 85% during pregnancy, up to 170 mL/min.[2] Glomerular hypertension is avoided during pregnancy when the body decreases the vascular resistance in the renal afferent and efferent arterioles. This occurs even though there is increased renal flow, as the glomerular hydrostatic pressure will remain stable.[2] As the GFR increases, filtration of creatinine, uric acid, and blood urea nitrogen increases, causing a decrease in plasma concentrations. Serum creatinine levels have a mean value of 0.5 mg/dL during pregnancy, and blood urea nitrogen mean values are 8.9 mg/dL.

Tubular reabsorption in the kidneys during pregnancy is also affected. The reabsorption of glucose in the proximal convoluted tubule is less effective. Under healthy conditions, it is estimated 90% of pregnant women spill 1–10g of glucose into their urine daily. The glomerulus also becomes more permeable to albumin during pregnancy, and the excretion of albumin into the urine increases. Proteins may increase in the urine by as much as 300 mg per day.[2] Excreting this volume of proteins should increase the protein concentrations seen in urine above pre-pregnancy values. However, the concentration remains within the upper limit of the established reference range.

Hyperventilation

It is normal for the mother to experience hyperventilation during pregnancy. Hyperventilation is a result of the increased demand for oxygen consumption. During pregnancy, arterial pO_2 increases and arterial pCO_2 decreases. To compensate, serum bicarbonate falls to 18–22 mmol/L (reference range 22–30 mmol/L). Interestingly, during pregnancy it is normal to have a mild, fully compensated respiratory alkalosis.[2] This is due to greater carbon dioxide production and is compensated for by increasing the renal excretion of bicarbonate.[8]

Endocrine

Endocrine changes are also observed during pregnancy, with the most notable changes occurring with the thyroid. The liver increases production of

thyroid-binding globulin (TBG), causing increases in both thyroxine (T_4) and triiodothyronine (T_3). Free T_4 and free T_3 levels are altered slightly but without any notable clinical significance when tested. hCG is structurally similar to thyroid stimulating hormone (TSH) and can bind to TSH receptors. When hCG is increased in the peripheral circulation during pregnancy, TSH levels decrease due to the negative feedback nature of TSH and thyroid hormones.

During pregnancy, adrenocorticotropic hormone (ACTH), cortisol, and free cortisol plasma concentrations increase. ACTH is released from the anterior posterior gland and controls the release of cortisol from the adrenal glands. Cortisol is a powerful glucocorticoid that is responsible for maintaining glucose levels as well as hemodynamics. At the end of the first trimester, cortisol levels increase. By the end of the third trimester, cortisol levels are three times higher than in non-pregnant females.[2] The diurnal variation of cortisol seen in non-pregnant females continues to be exhibited during pregnancy.

Cholesterols

Triglyceride levels remain increased throughout pregnancy. This occurs due to the liver's ability to increase synthesis. Maintaining increased triglyceride levels prevents the mother from using all of her stored adipose tissue, provides energy for the mother, and saves the circulating glucose for the developing fetus. Low-density lipoprotein (LDL) levels are also increased during pregnancy. These circulating cholesterols are used by the placenta to create steroid hormones such as sex steroids and corticosteroids for the developing fetus. Circulating LDL will be 50% higher at term than in a non-pregnant female.[2]

Immune System

The immune system of the mother is also altered to prevent an immune response against the paternal antigens present in the developing fetus. The Th1 pro-inflammatory response is decreased and shifted to a Th2 anti-inflammatory response (T-helper cell).[1] Circulating monocytes, which are capable of expressing major histocompatibility complex II (MHC II), also decrease in the mother. This decreases the ability of antigen presentation and ultimately the stimulation of T cells during pregnancy. It has also been noted that the total number of circulating natural killer cells and pro-inflammatory cytokines, such as interferon-γ (IFN-γ), is also reduced.[1] Progesterone is responsible for reducing the expression of pro-inflammatory cytokines by the uterus and trophoblast, including IL-1β and IL-6, and it may be responsible for the regulation of monocytes.[1]

Neuroactive Hormones

Neuroactive hormones are also secreted by the placenta during pregnancy. Kisspeptin is a hormone secreted by the hypothalamus; however, its function is not completely understood. Kisspeptin both stimulates and inhibits glucose-stimulated insulin secretion depending on the concentration of circulating isoforms that effect pancreatic islets. Kisspeptin rises throughout pregnancy to concentrations that are 10,000 times higher than the concentrations seen in a non-pregnant female.[1] Kisspeptin also affects the maternal cardiovascular system by causing vasoconstrictive effects on the smooth muscle cells.

Thyrotropin-releasing hormone (TRH) is another neuroactive hormone that is important during pregnancy. TRH stimulates the anterior pituitary to release TSH in non-pregnancy. The placenta releases additional TRH to maintain elevated circulating concentrations in the mother's body. This causes elevated secretions of TSH. Thyroid hormones are necessary for euthyroid function, as well as optimal brain development, respirations, and metabolism in the developing fetus.[9]

Analytes for Maternal and Fetal Health Assessment

Analytes such as progesterone, α-fetoprotein, and unconjugated estriol can be measured to determine both the fetal and maternal pregnancy health. Concentrations of each analyte vary depending on the stage of development of the fetus, the stage of pregnancy, and size of the placenta. It is important to have a fundamental understanding of the physiology of pregnancy and maternal health to develop a firm understanding in the testing of the analytes for both maternal and fetal health assessment.

Human Chorionic Gonadotropin

Human chorionic gonadotropin (hCG) is an important hormone for the early detection of pregnancy. Its serial measurement can be used to ensure the development of a healthy pregnancy or to monitor complete expulsion of pregnancy after termination. Under normal pregnancy conditions, the concentration of hCG in the maternal serum rises exponentially

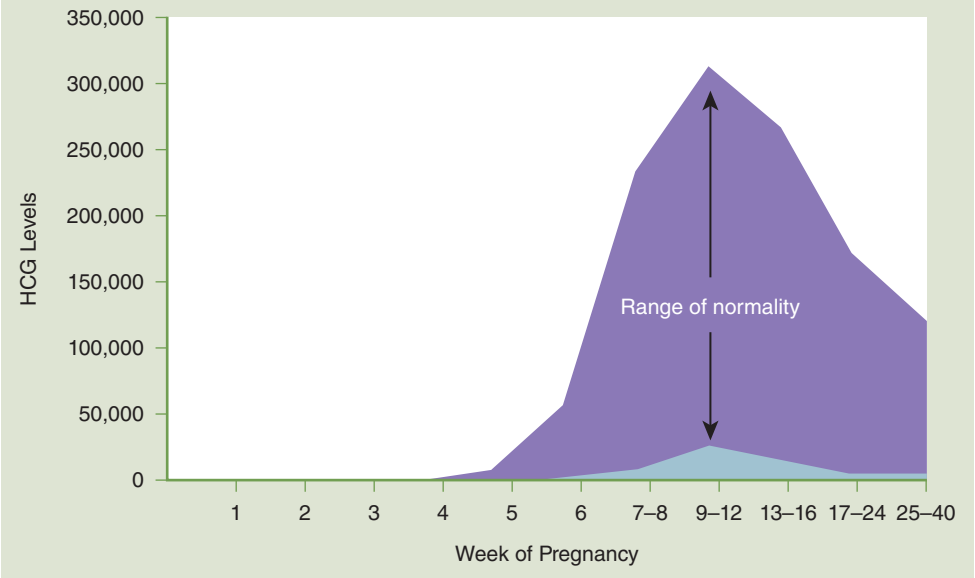


Figure 24.3 Graphical representation of the β-hCG maternal serum concentration compared to gestational age.

in the first trimester. This hormone nearly doubles every 24 hours until the 8th week of pregnancy. At 10 weeks of gestation, the serum hormone level peaks and then decreases until the 16th week when the serum hCG then plateaus and remains constant until the pregnancy comes to full term.

Figure 24.3 provides a graphical representation of the β-hCG maternal serum concentration compared to gestational age.

hCG is a heterodimeric glycoprotein placental hormone. This hormone is noncovalently bound by α and β subunits. It is synthesized by the trophoblastic tissue, which is the membrane that forms the wall of the blastocyst in the early development of the fetus and will eventually become the placenta. hCG is responsible for stimulating the corpus luteum to produce progesterone in maintaining the pregnancy.[10] TSH and hCG have similar biomolecular structures; therefore, hCG can cross-react with TSH, stimulating the thyroid gland. During pregnancy, elevated hCGlevels can be found in the maternal serum and urine.

hCG is composed of four parts: the nicked CG (CGn), the CG β-subunit (CGβ), nicked CG β-subunit (CGβn), and CG β-core fragment (CGβcf).[11] The hCG β-subunit is the only polypeptide known that is responsible for multiple independent molecules such as pituitary hCG, free β-subunit, fetal hCG, and hyperglycosylated hCG.[12] The intact hCG hormone can be tested, as well as the free subunits. Before the development of the placenta, the primary hCG present is hyperglycosylated

CASE STUDY 24.3, PART 2

Remember Nina. She has felt great since her 8-weeks visit to the obstetrician. However, this morning, she woke up to abdominal cramping and light bloody discharge from her vaginal canal. She called her doctor and arranged for an appointment at 8:00 a.m. When she arrived, she had been experiencing severe abdominal cramping for the past 2 hours.

1. Based on Nina's symptoms, what laboratory test should be completed?
2. Should the laboratory test be performed once or serially?

hCG. Hyperglycosylated hCG free β-subunit is a sugar variant made by the placental cytotrophoblast cells of the embryo. It controls the invasion and implantation of pregnancy. This hormone is an autocrine that antagonizes the transforming growth factor (TGF)-β receptor. During pregnancy, the maternal plasma has higher levels of total hCG. This molecule is catabolized by the liver, and the β subunit is degraded by the renal system, creating the CG β-core fragment (CGβcf). The CGβcf is secreted in the urine during pregnancy. It is this key component that is detected by the urine hCG point-of-care (POC) testing cassettes.

The serial monitoring of hCG can also indicate the presence of abnormal pregnancies. Pregnancies whose hCG levels plateau before eight weeks of gestation or that fail to double every 24 hours are assumed to be nonviable. Following the termination of a pregnancy or delivery, it can take 7 to 60 days for the hCG to return to non-pregnant levels.[13] When fetal Down syndrome is present, the hCG will be 2 times higher in the maternal serum. When hCG is paired with other analyte testing and ultrasound, the quality of the evaluation of the pregnancy increases exponentially.

Point-of-Care Testing

β-hCG testing is often first detected qualitatively during a pregnancy with a CLIA-waived POC testing cassette (**Figure 24.4**), also known as a home pregnancy test. CLIA stands for Clinical Laboratory Improvement Amendments, which regulate laboratory testing; a CLIA waiver is needed for any test intended for home use. Result interpretation involves a qualitative negative or a positive result. This form of POC testing is rapid and simple to perform with a low risk of error to detect the onset of pregnancy. In early pregnancy, the primary hCG is hyperglycosylated hCG, which is not detected by the home pregnancy test. While there are over 100 manufacturers of home pregnancy tests, the lower limit of detection for this type of testing is uniformly 20 mIU/mL. Urine home pregnancy testing is inexpensive, does not require a trained medical professional to interpret, does not require an invasive procedure, and can provide results in less than 10 minutes.

False negatives can occur for a variety of reasons when using this method for urine hCG testing. Since the primary hCG secreted in early pregnancy is hyperglycosylated hCG, it is not detected by the POC hCG cassette. Testing conducted too early in a pregnancy, before secretion of the CG β-core fragment (CGβcf) in the urine by the renal system would cause a false negative result. Individuals with renal diseases that prevent the concentration of renal filtrate will likely cause a false negative due to the likelihood of producing a dilute urine. Situations in early pregnancy involving polyuria or other hyperhydration states could prevent the detection of hCG in the urine due to dilution of analytes in the sample. False negatives can occur due to the *hook effect*, which describes the saturation of antibodies in the test due to overwhelming volume of hCG molecules (**Figure 24.5**).[14] Both sets of antibodies in the test bind to the hCG epitope separately, but neither antibody binds to the same hCG molecule. This prevents the sandwiching of the antibodies in the assay, creating a false negative result on the POC test cassette.

False positives and negatives can also occur with POC hCG tests when testing urine. This test does not require trained medical professionals to perform the test, but incorrect result interpretation can occur when results are viewed with an untrained eye. Consumed drugs such as aspirin, carbamazepine, and methadone can contribute to false-positive results.[15] Contamination of the female genital secretions that

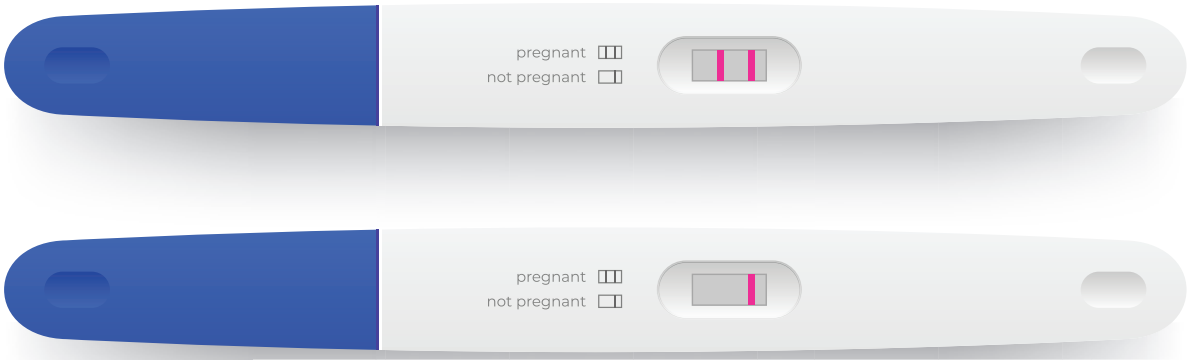


Figure 24.4 A home pregnancy test cassette.

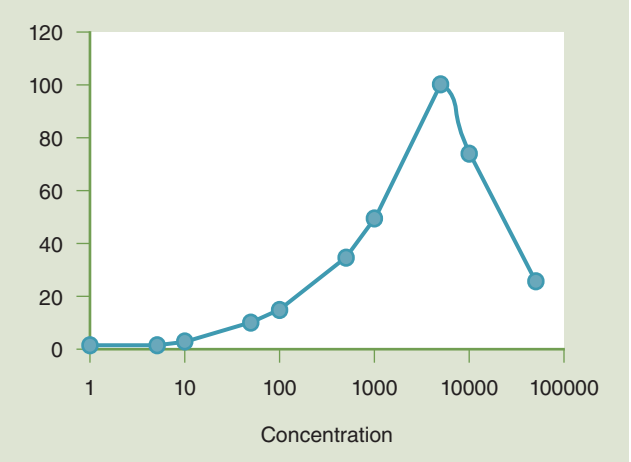


Figure 24.5 Hook effect seen while testing for β-hCG.

Courtesy of Molecular Devices, LLC.

include seminal fluid and strongly alkaline urines can cause false negatives.[15]

Serum Immunoassay Methodologies

A wide variety of automated immunoassays are currently available. Currently, most assays involve monoclonal capture and use tracer or radiographic antibodies targeted toward epitopes found in the intact hCG and associated subunits. The serum immunoassay is a quantitative interpretation of the concentration of hCG in a patient's serum. hCG levels are usually reported in milli-international units of hCG hormone per milliliter of blood (mIU/mL) but can also be reported in international units per liter (IU/L). It is important to be aware of the antibody cross-reactivity of the assay used to detect the various hCG isoforms.

There is no reference range for pregnancy-related quantitative maternal serum β-hCG. There are established ranges, however, that indicate an estimated gestational age based on the quantitative β-hCG result. β-hCG levels vary greatly between pregnant females and between pregnancies in the same female. Non-pregnant females β-hCG can range from 0 mIU/mL to 9 mIU/mL. As non-pregnant females age, hCG naturally increases in the serum. For this reason, it is suggested a cutoff value of 14 mIU/mL is used for non-pregnant females greater than 55 years of age.[16] The lower limit of detection for this assay is 0 mIU/mL.

Interfering factors that contribute to false positives with serum immunoassay begin with the development of heterophilic antibodies. Heterophilic antibodies contribute to false positives because their presence in the serum occurs due to a previous exposure to an animal product that interacts with antibodies present in the assay.[17] Rheumatoid factors can also bind to epitopes of manufactured tracer or radiographic antibodies found in immunoassays to detect hCG, causing a false positive. Individuals with IgA deficiency, chronic renal failure, or end-stage renal disease can also see false positives.[18,19] Individuals who have recently been transfused with erythrocytes or plasma containing hCG could contribute to a false-positive hCG result by immunoassay methods. Exogenous hCG consumed by individuals for hCG doping, weight loss, or assisted reproduction can also cause a false-positive result. False negatives usually occur due to the hook effect described previously.[20]

α-Fetoprotein (AFP)

The embryonic yolk sac produces α-fetoprotein, followed by the parenchymal cells of the fetal liver.[21] Secretion of this glycoprotein are encoded on chromosome 4q25. AFP binds the hormone estradiol and is thought to protect the fetus from immunological attack from the mother, but the physiologic function is not completely understood. During fetal development, this glycoprotein can pass through the placenta into the bloodstream of the mother. For this reason, AFP levels are often tested in the mother's serum and known as maternal serum α-fetoprotein (MS-AFP). At the end of the first trimester, AFP levels begin to rise and then dramatically fall after the 32nd week of pregnancy. After birth, the AFP levels slowly decline to adult levels. AFP has no known function in normal adults.

Maternal Serum α-Fetoprotein

AFP is used as a screening marker for fetal abnormalities and fetal distress and can be an indicator of pregnancy health. Elevated MS-AFP levels can indicate a risk of a number of possible issues in the fetus, including neural tube defects (NTDs) such as spin abifida, anencephaly, omphalocele, gastroschisis, low birth weight, intrauterine growth restriction (IUGR), premature delivery, and intrauterine fetal death. MS-AFP can predict spina bifida in the fetus 80% of the time and anencephaly 90% of the time.[21] For the mother, it can be an indicator of preeclampsia and increased risk of perinatal death. When MS-AFP is decreased, it can be an indicator that the fetus has Down syndrome or is at a higher risk for trisomy 18. The American College of Obstetricians and Gynecologists (ACOG) recommends

AFP screening for all pregnant women between 15 to 20 weeks of pregnancy.[22]

Under non-pregnant conditions, MS-AFP ranges from 0–40 ng/mL and varies with age, body mass, and ethnicity. Black and Asian women usually have higher AFP levels, while White women and those of Hispanic descent often have lower AFP levels. During pregnancy, AFP begins to rise at 14 weeks gestation. The AFP levels increase until the 32nd week of gestation. While MS-AFP concentrations vary, between the 15th and 20th weeks of gestation, levels should be between 10–150 ng/mL when ethnicity, body mass, maternal age, and the gestational age of the fetus are considered.[21]

Many laboratories interpret the AFP results using the multiples of median (MoM). This compares the patient's AFP serum concentration to the median of the reference range. This helps to overcome multiple variables that can influence the results. The MoM is calculated by dividing the mother's AFP serum concentration by the median of the reference range. The MoM MS-AFP reference range is 0.5–2.0 MoM.

Several components influence testing and contribute to false-positive results. If the pregnancy is multiparous, this increases the AFP concentration in the mother's serum. It is recommended that screening procedures do not assume a pregnancy is one fetus. The incorrect calculation of fetal gestation can cause the AFP to be higher than expected. Individuals who have gestational diabetes or who are exposed to tobacco smoke have a higher risk of false-positive results. Laboratory technique highly influences AFP results. Therefore, screenings should consistently be performed at the same laboratory for effective interpretation of absolute results.

Fetal Amniotic Fluid α-Fetoprotein

When MS-AFP levels are higher than expected and ultrasounds do not provide a cause for the increased levels, the amniotic fluid must be tested. The amniotic fluid AFP concentration exhibits an intricate pattern before the 14th week of gestation; the MoM for open spina bifida is 7 and the MoM for anencephaly is 20.

When the amniotic fluid is contaminated with fetal or maternal blood, it interferes with testing. It is recommended to test the amniotic fluid for the presence of fetal hemoglobin when testing amniotic fluid for the AFP concentration.

Testing for AFP

AFP is stable for 7 days at room temperature. However, AFP forms multimers rapidly during storage when the serum is dilute, and antibodies in immunoassays may not be able to bind to the AFP epitopes. AFPs can be tested in the serum, amniotic fluid, and urine.

In the laboratory, AFP concentrations were originally determined by radioimmunoassay (RIA). Today, enzyme-labelled immunoassays (EIAs), monoclonal bead assay, microparticle immunoassay, and polyclonal bead assay are used, which have lower detection capabilities. These methods use a sandwich design to capture the AFP molecule. If electrophoresis is used, AFP is 70 kD in size, and the electrophoresis mobility is between albumin and α_1-globulin.

Unconjugated Estriol

Estriol (E3), a steroid hormone, is the predominate estrogen in pregnancy.[23] Estriol is primarily synthesized in the placenta, with a small amount

CASE STUDY 24.2, PART 2

Remember Caroline. She visits her obstetrician for her 16-week check. Her maternal history only indicated she was an avid runner before her pregnancy and had a recent appendectomy (appendix removal). Fraternal history did not indicate any anomalies. A routine transvaginal ultrasound at 8 weeks' gestation did not note any remarkable findings. Her routine second trimester maternal serum α-fetoprotein was unexpectedly and significantly elevated.

1. Based on the results of the α-fetoprotein, what is the most likely diagnosis?
2. Which testing should be performed following these results of the maternal α-fetoprotein?

© Robin Skjoldborg/DigitalVision/Getty Images.

converted from precursors from the maternal ovaries.[24,25] The unconjugated form of estriol (uE3) is present in maternal serum in low concentrations and is measured as an indicator of fetal health, particularly during the second and third trimesters of pregnancy. Maternal serum uE3 levels increase with gestational age, with increases of approximately 20% to 25% per week in the second trimester, and reach maximum levels when the pregnancy reaches term.[26]

Decreased levels of uE3 during the second trimester have been associated with a variety of disorders of the fetus and adverse pregnancy outcomes including preeclampsia, fetal growth restriction, and low birth weight.[27,28] The most common use of uE3 measurement in maternal serum is for prenatal screening for trisomy 21 (Down syndrome) and trisomy 18 syndrome (also known as Edwards syndrome).[28] Decreased levels of uE3 have also been associated with placental sulfatase deficiency (such as X-linked ichthyosis), chromosomal abnormalities, Smith-Lemli-Opitz syndrome, and fetal demise. Depending on the methodology and/or assay employed for measurement, serum concentrations of uE3 during the second trimester range from 0.70 to 2.50 µg/mL (2.43 to 8.68 nmol/L).[24]

Preanalytical handling of the specimen is of particular importance when uE3 is to be measured. At ambient and refrigerated temperatures, uE3 concentrations increase due to the hydrolysis of conjugated estriol, which could cause falsely elevated results and therefore false-negative results to patients.[29] It is recommended that serum be separated promptly and testing performed within 7 days of collection.[24] Immunoassays performed on automated platforms are most commonly used for the measurement uE3, with many vendors offering nonisotopic assays in place of radioimmunoassays (RIAs).[24,30]

Inhibin A

Inhibin is a glycoprotein with a primary function of suppressing FSH. Inhibin is composed of two subunits, one larger (α) and one smaller (β).[31] Inhibin A is a dimer composed of 2 βA subunits and is commonly measured in maternal serum as part of assessment for Down syndrome.[24] Inhibin A originates from the fetoplacental unit, in addition to the ovarian granulosa cells in females and testicular Sertoli cells in males.[31,32] Concentrations of inhibin A in maternal serum vary during pregnancy, rising at weeks 8–10, reaching the lowest levels at approximately 17 weeks, and then rising again throughout the third trimester.[26,33]

Trisomy 21 (Down syndrome) is associated with increased maternal serum levels of inhibin A during the second trimester of pregnancy[34] and is therefore included in prenatal screening tests.[33] Studies have also suggested that the measurement of second-trimester levels of inhibin A may be useful for early detection or to predict the development of preeclampsia.[32,35] A typical inhibin A value measured at 17 weeks gestation is approximately 175 ng/L (175 pg/mL). Unlike uE3, inhibin A is stable in maternal serum. Inhibin can be measured by enzyme-linked immunosorbent assays (ELISAs) and RIAs. Of particular importance for prenatal screening is to employ a methodology that is specific for dimeric inhibin A. More recently developed assays have employed monoclonal antibodies to achieve this specificity and ensure the detection of dimeric inhibin A alone.[24,36]

CASE STUDY 24.1, PART 3

Remember Lorna. She visits her obstetrician for a routine check and to review the results of the triple marker screen and ultrasound performed earlier in her second trimester at approximately 18 weeks' gestation. Levels of all three markers (AFP, uE3, and hCG) were low. An ultrasound performed around the same time shows that the fetus is smaller than average, suggesting intrauterine growth restriction (IUGR).

1. Based on the results of the triple screen, what do you suspect as a likely diagnosis for the fetus?
2. What other marker is commonly performed for screening purposes during the second trimester?
3. What would be the benefit of adding this other marker?
4. What marker could have been run in her first trimester of pregnancy? Would you expect this marker to present in low or high concentrations, had it been measured?

The Triple and Quadruple Screen Test

In addition to measuring individual markers discussed previously, second trimester prenatal screening commonly combines the measurement of multiple markers in maternal serum with a variety of other characteristics (maternal age, gestation, number of fetuses, etc.) to report risk estimates for aneuploidy and birth defects.[37] Routinely performed between 16 and 20 weeks' gestation, the triple screen measures AFP, hCG, and uE3, while the quadruple (quad) screen also includes the measurement of inhibin A.[38,39] A penta screen incorporating hyperglycosylated hCG is performed in some laboratories, although less commonly than triple and quad testing. After incorporating information on maternal age, gestational age, ethnicity, weight, number of fetuses, and diabetes status, risk estimates are reported for trisomy 18, trisomy 21, and open NTDs.[36,37] AFP results are reported as multiples of median (MoM).[40]

It is critical to note that reported results are *risk estimates only*. Triple and quad tests are intended for screening purposes and are not diagnostic. The triple screen has an approximate detection rate of 69% for trisomy 21, while the addition of inhibin A in the quad screen raises the detection rate to about 81%, depending on the gestational age.[37] Positive screening results should be confirmed with diagnostic tests and consultation with a genetic counselor, as recommended by the patient's physician.[41]

Acetylcholinesterase

Acetylcholinesterase (AChE) is a cholinergic enzyme found in neuromuscular junctions in the muscle and nerves. This enzyme immediately breaks down acetylcholineby hydrolyzation, which is a naturally occurring neurotransmitter. This analyte is important for the detection of NTDs of the fetus during pregnancy. When this disease state is suspected, the presence of acetylcholinesterase can be measured in the amniotic fluid. A polyacrylamide gel electrophoresis is performed. A 99.76% specificity is exhibited for this type of electrophoresis. Cholinesterase will be present in the sample under normal pregnancy and fetal conditions, but the presence of AChE can be diagnostic for an open neural tube defect, abdominal wall defects, or anencephaly. The detection rate of disease is high when testing this analyte: 97% for anencephaly, 99% for open spina bifida, and 94% for abdominal wall defects. On polyacrylamide gel electrophoresis, AChE presents as a slow migrating band.

Pregnancy Associated Plasma Protein A (PAPP-A)

Pregnancy-associated plasma protein A (PAPP-A) is a large glycoprotein produced in the placenta and commonly measured in maternal serum. PAPP-A in maternal circulation is in a tetrameric form bound to the proform of eosinophil major basic protein (proMBP).[42] PAPP-A functions in angiogenesis, matrix mineralization, and to prevent a maternal immune response to fetus.[43] Maternal serum concentrations increase over the course of pregnancy and reach maximum levels when the pregnancy reaches full term.[24]

Unlike other maternal serum markers commonly measured in prenatal screening, the measurement of PAPP-A is most valuable for screening purposes during the first trimester of pregnancy. Low PAPP-A concentrations in the first trimester are seen in trisomy 13 (Patau syndrome), trisomy 18 (Edwards syndrome), trisomy 21 (Down syndrome), and Turner syndrome.[42,44] Complications such as intrauterine growth restriction (IUGR), premature delivery, and preeclampsia have also been correlated with low levels of PAPP-A during early pregnancy.[24] PAPP-A is included in first trimester screening partnered with hCG and nuchal translucency screening by ultrasound.[37] PAPP-A concentrations are typically expressed as MoM, similar to AFP. Early assays to measure PAPP-A were in the form of ELISAs, while today there are several commercial immunoassays available on automated platforms.[24]

Progesterone

Progesterone is a steroid hormone produced by the placenta during pregnancy. This placental hormone plays an important role in the establishment and maintenance of a healthy pregnancy. It reduces the contractability of smooth muscle, regulates local and systemic inflammation, and blocks natural killer (NK) cell degranulation. The maternal progesterone serum concentration increases as soon as trophoblastic cells imbed in the endometrium of the uterus. The concentration continues to rise until term of the pregnancy, when the progesterone levels begin to rapidly fall. This is required for the smooth muscle contraction requirements of the uterus during delivery of the fetus. Maternal serum progesterone levels can be monitored when spontaneous abortion is suspected, or if the mother is at a high risk for natural fetal abortion. Degradation of placental cells or removal of the placenta will cause a decrease in maternal serum progesterone due to the placenta production of the hormone.

Glucose

An oral glucose tolerance test (OGTT) should be conducted during the 24th to 28th weeks of pregnancy in all patients in order to screen for gestational diabetes mellitus (GDM). Patients may be tested earlier if they are at a higher risk of developing diabetes, such as those older than 35 years or who have obesity. GDM is one of the most common pregnancy complications and affects about 10% of all pregnancies. If patients develop GDM, their overall likelihood of developing type 2 diabetes later in life increases. Maternal GDM not only is problematic for the mother, but it also causes severe complications with the fetus. Fetuses can become macrosomic, experience low blood sugar and breathing difficulties after birth, and can develop several congenital issues, including central nervous system (CNS) malformations and cardiac anomalies. Fetuses also can die in utero if GDM is not managed properly.

There are three glucose tests commonly ordered during pregnancy to access the overall ability of the mother to maintain regular glucose levels and to determine maternal and fetal health. The glucose challenge or the 1-hour oral glucose tolerance test (OGTT) is often performed first as a screening tool for GDM. The mother is asked to continue with normal activities but should fast for at least 8 hours prior to testing. If the mother cannot tolerate fasting, she may continue to consume food normally, thought this may alter interpretation of the test. She is then asked to consume a sugary solution with 50 g of glucose. The mother's blood glucose concentration is tested 1 hour after the glucose was consumed to determine if the levels fall to within the desired blood glucose range. OGTT results lower than 140 mg/dL (7.8 mmol/L) are considered normal. Maternal blood glucose levels above 140 mg/dL indicate that another, longer OGTT test should be performed to confirm a diagnosis of GDM, while levels higher than 190 mg/dL (10.6 mmol/L) after the 1-hour test are diagnostic for GDM.

If the 1-hour OGTT screen results are elevated but not diagnostic, either a 2-hour or 3-hour OGTT must be completed to get a better assessment of the mother's glucose and insulin resistance status. The mother must fast for 8 to 12 hours before beginning either test. The 2-hour and 3-hour glucose tests are performed by obtaining a fasting maternal serum glucose level before the mother consumes a drink containing 100 g of glucose. The maternal serum glucose concentration is determined every hour for either 2 or 3 hours, depending on which test was administered. Result interpretation is shown in **Box 24.1**.

BOX 24.1 OGTT Results for Diagnosing Gestational Diabetes Mellitus (GDM)

A diagnosis of GDM is made if results meet or exceed:

Fasting glucose:	95 mg/dL
1 hour:	180 mg/dL
2 hours:	155 mg/dL
3 hours:	140 mg/dL

Fetal Fibronectin

Fetal fibronectin is a protein present in high concentrations in the mucosal lining of the uterus and is present in high concentrations in amniotic fluid. Normally, low concentrations of this protein are found in cervicovaginal secretions. It appears at 39 weeks within 7 to 14 days of delivery due to proteolysis of the amniotic membranes. It is useful for assessing the risk that preterm delivery will occur in the next 7 to 14 days in women with signs and symptoms of preterm labor. A positive result indicates that preterm delivery is likely within the next 7 to 14 days. On the other hand, a low or negative value is very predictive of those patients between 24 and 32 weeks who will not go into labor. The absence of fetal fibronectin at 39 weeks suggests the patient will not deliver within the next 2 weeks.

Common Diseases Associated with the Developing Fetus

Not all fetal diseases or congenital anomalies can be detected with laboratory screening, even when lab work is compared with results from other ancillary services. The common anomalies are NTDs, Down syndrome, trisomy 18, and isoimmunization. Details regarding each disease state and laboratory detection are discussed below.

Neural Tube Defects

Neural tube defects are one of the most common birth defects seen in the developing fetus. These NTDs often occur long before a pregnancy has even been detected. In the early days of development, the embryo begins the formation of the CNS. Nineteen days after conception, cells that will become the brain and spinal cord are strategically aligned on a flat plate that rolls up into a long, hollow tube. This hollow tube drops into the back of the embryo, and the back skin covers

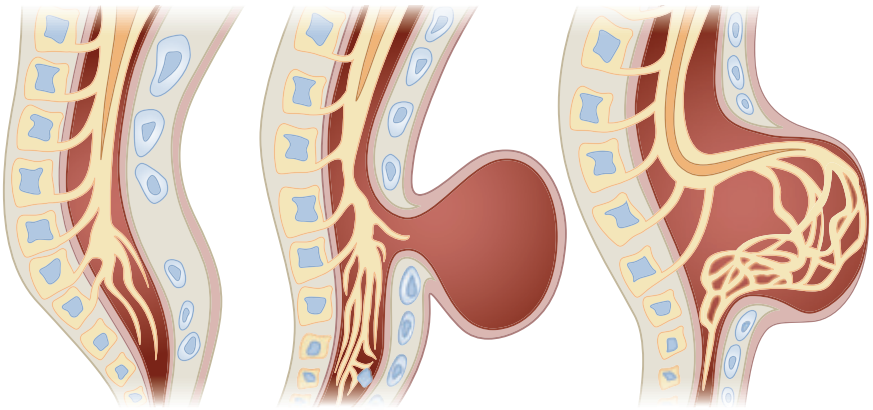


Figure 24.6 Incomplete neural tube development leading to spina bifida.

the tube. The completed formation of the neural tube occurs at 4 weeks' gestation.

An NTD occurs when the neural tube does not close completely. Failure of this fusion leads to permanent defects. During embryonic development, components that should be inside the neural tube will protrude to the outside of the embryo's body or will be left open to expose the CNS to damage. The extent of the defect depends on the location where fusion did not occur on the neural tube. Neural tube defects can be open or closed.

Myelomeningocele, often referred to as spina bifida, occurs when fusion is not complete along the spinal cord of the developing fetus. This defect causes death or impairments to the nerves where the cells did not fuse on the neural tube. These fetuses often have limited lower extremity mobility and sensation, neuropathies of the bowel and bladder, and possible renal involvement. Anencephaly is the absence of a major portion of the brain, skull, and scalp. This occurs when the proximal end of the neural tube fails to fuse together. Encephalocele is an opening or sac that contains part of the brain and is often protruding from the opening of the neural tube that did not close during embryonic development.

Currently, there is no known direct cause for NTDs. There are over 200 gene mutations known to contribute to the development of NTDs. For unknown reasons, if no known risk factors are present, a pre-conception diet rich in folic acid and supplements of 0.4 mg of folic acid daily appear to decrease the risk of having a fetus with a NTD. When both maternal and fraternal risk factors are present, pre-conception supplementation with 4.0–5.0 mg of folic acid and vitamin B_{12} are needed to reduce the risk of developing a fetus with a NTD. Despite this, a deficiency in folic acid has not been shown cause NTDs.

Laboratory testing for NTDs begin with MS-AFP levels. NTDs that are not covered with skin are referred to as open NTDs. In these situations, neural fluids mix with amniotic fluids and can enter the mother's circulation. If high volumes of AFP are present in the mother's peripheral circulation, concerns over NTDs arise, and further testing is warranted.

Genetic testing must be conducted on fetal cells. These are obtained by a diagnostic amniocentesis. A chromosomal microarray, a fetal amniotic fluid AFP, and AChE are tested to determine if genetic anomalies are present and if the associated analytes are elevated.[45]

Down Syndrome

Down syndrome is the most common serious genetic anomaly. An extra copy of the long arm region q22.1 to q22.3 of chromosome 21 results in mental deficiency, hypotonia, congenital heart defects, and a flat facial profile. Usually, a fetus with Down syndrome has three copies of chromosome 21, but 5% of fetuses have Down syndrome caused by translocations. Maternal age severely increases the risk of having a fetus with Down syndrome. At the maternal age of 20, the risk of having a fetus with Down syndrome is 1 in 1200.[24] At the maternal age of 40, the risk of having a fetus with down syndrome is 1 in 75.[24] It is often recommended that mothers over the age of 35 years have an amniocentesis for genetic screening. However, Down syndrome is only detected 20% of the time with this method.[24] When results from triple test results are considered with amniocentesis

results, over 60% of fetuses with Down syndrome are detected.

Trisomy 18

Trisomy 18 occurs when the fetus receives an extra copy of chromosome 18. The rate of live births with trisomy 18 appears to be much lower than how often this abnormality actually occurs. Although trisomy 18 is seen in 1 in 8000 births,[24] it is estimated that 80% of fetuses with trisomy 18 do not live past eight weeks' gestation. Of those fetuses that survive, 70% will not make it past the second or third trimester.[24] After birth, 50% of the newborns die within 5 days and 90% of affected infants will die within 100 days.

The triple test is used to detect 60% of trisomy 18 cases.[24] AFP, hCG, and uE3 levels are lower than the expected values. While the triple test is beneficial in detecting a substantial number of trisomy 18 cases, false-positive test results have been documented. The quad test is more accurate and has fewer false-positive test results. If the screen is positive, then an amniocentesis is performed for diagnostic testing.

Isoimmunization

The isoimmunization of a fetus involves a fetal hemolytic disorder that occurs when maternal antibodies are directed against fetal erythrocytic antigens. Maternal antibodies can be developed against any antigens found on the surface of the fetal erythrocytes. This disease can be severe and may be fatal to the fetus or the newborn (hemolytic disease of the newborn). The extent of the hemolysis can be determined by the amount of total bilirubin present in the amniotic fluid.

The most common cause of isoimmunization occurs when the mother is Rh negative and is carrying a Rh-positive fetus. During pregnancy small amounts of fetal erythrocytes pass through the placental barrier into maternal circulation. The mother's plasma cells react by creating IgG antibodies against the glycoproteins found on the surface of the fetal erythrocytes, known as RhD antibodies. These IgG antibodies can cross the placenta and attach to the fetal erythrocytes, eventually causing hemolysis of the erythrocyte.

The greatest concern of isoimmunization is the anemia that ensues from the lysing of fetal erythrocytes. The fetal anemia that results creates a heavy burden on the fetal heart to supply adequate oxygen levels to the fetal tissue.[24] Erythropoiesis begins to replace the lost cells immediately. This occurs in the fetal bone marrow and extramedullary of liver and spleen. Extramedullary erythropoiesis of the liver destroys the fetal hepatocytes.[24] This causes decreased volumes of fetal serum albumin, leading to a cascade of problematic events. When these symptoms are severe enough, fetal congestive heart failure, edema, and pleural and pericardial effusions are imminent. At this point, the disease condition is referred to as *hydrops fetalis*, and intrauterine demise shortly follows.

Isoimmunization of fetal cells can be prevented with early laboratory testing. Maternal blood typing and antibody screening should be conducted by the second trimester of pregnancy. Mothers who are Rh-negative can be given a preventative immunoglobulin against RhD in both the second trimester and after fetal delivery.

Preterm Delivery

A pregnancy is considered *at term* between the 38th and 42nd week of pregnancy. When a fetus is delivered earlier than the 37th week of gestation, the pregnancy is labeled as *pre-term*. Preterm delivery can occur for many reasons, whether spontaneously or through emergency cesarian section due to the decline of fetal or maternal health.

The most common critical problem encountered with preterm deliveries is Respiratory Distress Syndrome (RDS).[24] The disease is caused by a deficiency in fetal pulmonary surfactant. Surfactant keeps the alveoli from collapsing during inspiration. Fetal surfactant production begins slowly at 20 weeks' gestation, but not enough surfactant is present to sustain life until 36 weeks' gestation. Fetal lungs contain 100 times the surfactant of adult lungs.[24] This is largely due to the fact that at delivery, fetal lungs must be able to overcome the transition from breathing amniotic fluid to breathing room air without the aid of their mother's supplemental oxygen. If preterm labor is suspected, the mother can be injected with steroids. This causes the fetal lungs to produce more surfactant and can prevent the development of RDS.

Common Diseases Associated with the Mother and Pregnancy

Inappropriate adaptation of maternal physiology may lead to complications of pregnancy. Maternal adaptations to pregnancy are largely mediated by the placenta.

Preeclampsia

Preeclampsia, characterized by hypertension, proteinuria, and edema in the second and third trimester, occurs in 5% of pregnancies and is not completely understood. If left untreated, preeclampsia can progress to eclampsia, which is characterized by the development of convulsions. Eclampsia causes abnormal endothelial reactivity that causes fibrin to be deposited in organs throughout the body. Deaths that occur from eclampsia are often caused by CNS complications. Often, the liver is injured due to ischemia thatcauses hemorrhage, infarctions, or hepatic failure.

In preeclampsia, there is greater than a fourfold increase of vasopressinase, an enzyme produced by the placenta during pregnancy that metabolizes vasopressin, often leading to a transient diabetes insipidus.[46] As a consequence of maternal plasma volume expansion, the secretion of atrial natriuretic peptides increases by 40% in the third trimester and rises further during the first week postpartum. The levels of natriuretic peptides are higher in pregnant women with chronic hypertension and preeclampsia.[47] Also, steroid hormones are implicated in pregnancy complications such as preeclampsia. Moreover, placental estrogen and progesterone levels are reduced in preeclamptic patients compared with healthy pregnant women.[1] Research highlights the importance of regulating kisspeptin production during gestation; increased placental kisspeptin is associated with preeclampsia.[1]

Hyperemesis Gravidarum

Up to 70% of pregnancies are complicated with hyperemesis gravidarum (morning sickness) that starts between the fourth and eighth weeks of pregnancy and can last for 15 weeks.[24] Under some circumstances, severe morning sickness results in malnutrition of both the mother and fetus. Low-birth-weight babies are common for those mothers who suffer with hyperemesis gravidarum.[24] Increased nausea and excessive vomiting often result in dehydration, electrolyte imbalance, ketonuria, weight loss, and vitamin or mineral deficiencies.[1] In these cases, intravenous fluid and vitamin substitution is commonly required. Thiamine supplementation is important to avoid the development of Wernicke's encephalopathy.[2]

The levels of hCG peak at the end of the first trimester when the trophoblast is most actively producing hCG, correlating with the nausea symptoms.[2] Nausea is also more frequent in pregnancies with high levels of hCG, such as in twin pregnancies.

Thyroid hormones may also be involved in the development of nausea symptoms, as a strong association with nausea and abnormal thyroid function tests has been found.[1] TSH and hCG have similar biomolecular structures, and therefore hCG cross-reacts with TSH, stimulating the thyroid gland.

The nausea symptoms usually resolve by week 20, but about 10% to 20% of the patients experience symptoms beyond week 20; some have nausea until the end of the pregnancy.[24] In most cases, minor dietary modification and observation of electrolyte balance is sufficient.

Ectopic Pregnancy

In the first 5 days after conception, the egg is migrating through the fallopian tubes. On rare occasions, the trophoblastic cells will prematurely attach to the lining of the fallopian tube instead of attaching to uterine endometrium. When this occurs, the condition is called *ectopic pregnancy*. While most of the time ectopic pregnancies occur in the fallopian tube (98%), they can occur in other locations such as the ovary, cesarean section scars, cervix, and the peritoneal cavity. It is estimated that 2% of all pregnancies are ectopic.[48] Twenty-five percent of women with ectopic pregnancies experience the following symptoms: vaginal bleeding and lower quadrant pain and have an adnexal mass.[48] The major concern for an ectopic pregnancy is the rupture of the fallopian tube, causing severe internal hemorrhage.

Risk factors for ectopic pregnancies are a previous history of pelvis inflammatory disease (PID), often caused when females contract sexually transmitted infections (STIs), previous ectopic pregnancies, and tubal surgeries. Low to moderate risks for ectopic pregnancy are seen with use of intrauterine contraceptive devices (IUDs), smoking, and assisted reproductive technology, including *in vitro* fertilization (IVF).

When a female presents with symptoms that are concerning for an ectopic pregnancy, a β-hCG pregnancy test is immediately ordered to determine if the female is pregnant or not. If the β-hCG is positive, a transvaginal ultrasound will be ordered to look for the presence of an embryo in the fallopian tube. Currently, the treatment for an ectopic pregnancy is based on the gestational age and includes medications to dissolve the cells or surgical removal of the fallopian tube to prevent a life-threatening hemorrhage from occurring.

Analyte Deficiency

Analyte deficiency can happen during pregnancy. Some analytes are more subjected to decreasing than other. As discussed, pregnancy is associated with a relative iodine deficiency. The causes for this are active transport of iodine from the mother to the fetal-placental unit and increased iodine excretion in the urine.[2] The World Health Organization recommends an increase in iodine intake in pregnancy from 100–200 mg/d.[2] If iodine intake is maintained in pregnancy, the size of the thyroid gland remains unchanged, and therefore the presence of goiter should always be investigated. The thyroid gland is 25% larger in patients who are iodine deficient.[2]

Calcium levels are often deficient in the mother during the second and third trimesters of pregnancy. The fetus's skeletal system is growing exponentially and requires large volumes of calcium. If ionized calcium is not readily available for the developing fetus, calcium stored by the mother's skeletal system is removed and used for fetal development. It is important the pregnant mother consumes foods high in calcium or a nutritional supplement formulated for pregnant women to prevent the unnecessary use of maternal skeletal calcium. Research does not demonstrate an increase in osteoporosis in mothers with low ionized calcium during pregnancy.

Due to the high demand of maternal resources by the fetus during pregnancy, it is highly recommended the mother consume additional nutritional supplements often referred to as pre-natal vitamins. The supplements should include broad spectrum vitamins and minerals such as iron and iodine and should contain other micronutrients when possible.

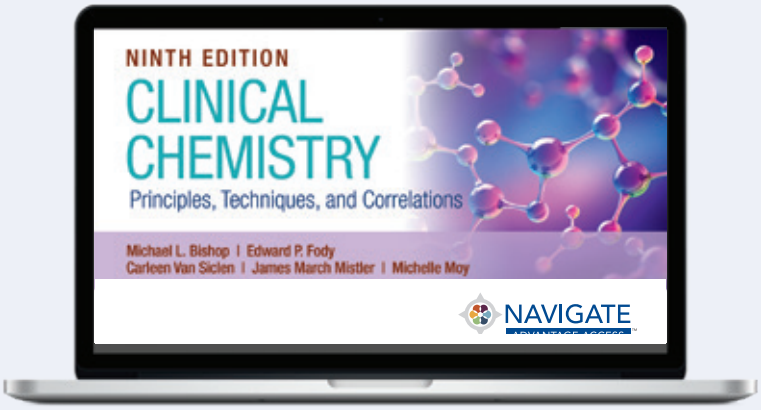
Other Clinical Impairments

During pregnancy there is also an increase in serum levels of deoxycorticosterone, corticosteroid-binding globulin (CBG), ACTH, cortisol and free cortisol. These changes cause a state of physiological hypercortisolism and may be clinically manifested by striae, facial plethora, rising blood pressure, and/or impaired glucose tolerance.[31]

Increased calcium absorption is associated with an increase in calcium excretion in the urine, and these changes begin from 12 weeks.[2] During periods of fasting, urinary calcium values are low or normal, confirming that hypercalciuria is the consequence of increased absorption. Pregnancy is therefore a risk factor for kidney stones.[2]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Duval C, Dilworth MR, Tunster SJ, et al. PTHrP is essential for normal morphogenetic and functional development of the murine placenta. *Dev Biol.* 2017;430(2):325–336.
2. Gordon MC. Maternal physiology. In Gabbe SG, Niebyl JR, Simpson JL et al., eds. *Obstetrics. Normal and Problem Pregnancies.* 6th ed. Philadelphia, PA: Saunders Elsevier; 2012: 42–65.
3. Lumbers ER, Pringle KG. Roles of the circulating renin-angiotensin-aldosterone system in human pregnancy. *Am J PhysiolReg IntegrCompar Physiol.* 2014;306(2):R91–R101.
4. Gonzalez-Campoy JM, Romero CJ, Knox FG. Escape from the sodium-retaining effects of mineralocorticoids: role of ANF and intrarenal hormone systems. *Kidney Int.* 1989;35(3): 767–777. doi:10.1038/ki.1989.51

5. Davison JM, Gilmore EA, Dürr J, et al. Altered osmotic thresholds for vasopressin secretion and thirst in human pregnancy. *Am J Physiol.* 1984:246(1 Pt 2):F105–F109. doi:10.1152/ajprenal.1984.246.1.F105
6. Soma-Pillay P, Nelson-Piercy C, Tolppanen H, Mebazaa A. Physiological changes in pregnancy. *Cardiovasc J Afr.* 2016;27(2):89–94. doi:10.5830/CVJA-2016-021
7. Newbern D, Freemark M. Placental hormones and the control of maternal metabolism and fetal growth. *Curr Opin Endocrinol Diabet Obes.* 2011;18(6):409–416. https://doi.org/10.1097/MED.0b013e32834c800d
8. Napso T, Yong HEJ, Lopez-Tello J, Sferruzzi-Perri AN. The role of placental hormones in mediating maternal adaptations to support pregnancy and lactation. *Front Physiol*. 2018;9:1091.
9. Mirabito Colafella, KM, Samuel CS, Denton KM. Relaxin contributes to the regulation of arterial pressure in adult female mice. *Clin Sci.* 2017;131(23): 2795–2805.
10. Ashwood ER, Knight GJ. Clinical chemistry of pregnancy. In Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics.* St. Louis, MO: Elsevier Saunders; 2006: 2153–2206.
11. Sigel CS, Grenache DG. Detection of unexpected isoforms of human chorionic gonadotropin by qualitative tests. *Clin Chem*. 2007;53(5):989–990. https://doi.org/10.1373/clinchem.2007.085399
12. Cole LA. Summary: hCG, a remarkable molecule. In Cole LA, Butler SA, eds., *Human Chorionic Conadotropin (hCG).* 2nd ed. Waltham, MA: Elsevier; 2015: 411–412.
13. Butts SF, Guo W, Cary MS, et al. Predicting the decline in human chorionic gonadotropin in a resolving pregnancy of unknown location. *Obstet Gynecol*. 2013 Aug;122(2 Pt 1):337–343.
14. Griffey RT, Trent CJ, Bavolek RA, et al. Hook-like effect causes false-negative point-of-care urine pregnancy testing in emergency patients. *J Emerg Med.* 2013 Jan;44(1):155–160.
15. Chard T. Pregnancy tests: a review. *Hum Reprod*. 1992 May;7(5):701–710.
16. Betz D, Fane K. Human chorionic gonadotropin. [Updated 2020 Aug 30]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan. Available from: https://www.ncbi.nlm.nih.gov/books/NBK532950/
17. Check JH, Nowroozi K, Chase JS, et al. False-positive human chorionic gonadotropin levels caused by a heterophile antibody with the immunoradiometric assay. *Am J Obstet Gynecol*. 1988 Jan;158(1):99–100.
18. Knight AK, Bingemann T, Cole L, Cunningham-Rundles C. Frequent false positive beta human chorionic gonadotropin tests in immunoglobulin A deficiency. *Clin Exp Immunol.* 2005 Aug;141(2):333–337.
19. Fahy BG, Gouzd VA, Atallah JN. Pregnancy tests with end-stage renal disease. *J Clin Anesth.* 2008 Dec;20(8):609–613.
20. Delbeke FT, Van Eenoo P, De Backer P. Detection of human chorionic gonadotrophin misuse in sports. *Int J Sports Med*. 1998 May;19(4):287–290.
21. Adigun OO, Yarrarapu SNS, Khetarpal S. Alpha fetoprotein. [Updated 2020 Sep 11]. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021 Jan. Available from: https://www.ncbi.nlm.nih.gov/books/NBK430750/
22. Başbug D, Başbug A, Gülerman C. Is unexplained elevated maternal serum alpha-fetoprotein still important predictor for adverse pregnancy outcome? *Ginekol Pol*. 2017;88(6): 325–330.
23. Ali ES, Mangold C, Peiris AN. Estriol: emerging clinical benefits. *Menopause*. 2017;24(9):1081–1085.
24. Yarbrough ML, Stout M, Gronowski AM. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier, Inc. 2018:1655–1696.
25. Newby D, Aitken DA, Howatson AG, Connor JM. Placental synthesis of estriol in Down's syndrome pregnancies. *Placenta*. 2000;21(2–3):263–267.
26. Palomaki GE, Bradley LA, McDowell GA, et al. Technical standards and guidelines: Prenatal screening for Down syndrome: This new section on Prenatal Screening for Down Syndrome, together with the new section on "Prenatal Screening for Open Neural Tube Defects," replaces the previous Section H of the American College of Medical Genetics Standards and Guidelines for Clinical Genetics Laboratories. *Genet Med.* 2005;7:344–354.
27. Settiyanan T, Wanapirak C, Sirichotiyakul S, et al. Association between isolated abnormal levels of maternal serum unconjugated estriol in the second trimester and adverse pregnancy outcomes. *J Matern Fetal Neonatal Med.* 2016;29(13):2093–2097.
28. Long W, Zhou Q, Wang H, et al. Second-trimester maternal serum screening biomarkers inthe risk assessment for preeclampsia. *Ann Clin Lab Sci*. 2018;48(3):308–313.
29. Lambert-Messerlian GM, Eklund EE, Malone FD, et al. Stability of first- and second-trimester serum markers after storage and shipment. *Prenat Diagn.* 20016;26:17–21.
30. Maharjan AS, Wyness SP, Ray JA, et al. Detection and characterization of estradiol (E2) and unconjugated estriol (uE3) immunoassay interference due to anti-bovine alkaline phosphatase (ALP) antibodies. *Pract Lab Med.* 2019;17:e00131.
31. De Kretser DM, Robertson DM. The isolation and physiology of inhibin and related proteins. *Biol Reprod.* 1989;40(1): 33–47.
32. Broumand F, Lak SS, Nemati F, Mazidi A. A study of the diagnostic value of Inhibin A Tests for occurrence of preeclampsia in pregnant women. *Electron Physician*. 2018;10(1):6186–6192.
33. Yue CY, Zhang CY, Ni YH, Ying CM. Spradley FT, Ed. Are serum levels of inhibin A in second trimester predictors of adverse pregnancy outcome? *PLoS One*. 2020;15(5):e0232634.
34. Noble PL, Wallace EM, Snijders RJ, Groome NP, Nicolaides KH. Maternal serum inhibin-A and free beta-hCG concentrations in trisomy 21 pregnancies at 10 to 14 weeks of gestation. *Br J Obstet Gynaecol*. 1997;104(3):367–371.
35. Ree PH, Hahn WB, Chang SW, et al. Early detection of preeclampsia using inhibin A and other second-trimester serum markers. *Fetal Diagn Ther*. 2011;29:280–286.
36. Wilson G, Liitti P, Pölönen T, Sairanen M, Spencer K. A technical and clinical evaluation of a new assay for inhibin A and its use in second trimester Down syndrome screening. *Clin Chem Lab Med.* 2016;54(9):1473–1479.
37. Carlson LM, Vora NL. Prenatal diagnosis: screening and diagnostic tools. *Obstet Gynecol Clin North Am*. 2017;44(2): 245–256.
38. Dey M, Sharma S, Aggarwal S. Prenatal screening methods for aneuploidies. *North Am J Med Sci*. 2013;5(3):182–190.

39. Wilson KL, Czerwinski JL, Hoskovec JM, et al. NSGC Practice Guideline: prenatal screening and diagnostic testing options for chromosome aneuploidy. *J Genet Couns*. 2013; 22(1):4–15.
40. Test ID: QUAD1 Quad Screen (Second Trimester) Maternal, Serum. QUAD1 - Clinical: Quad Screen (Second Trimester) Maternal, Serum. https://www.mayocliniclabs.com/test-catalog/Clinical+and+Interpretive/113145. Accessed May 5, 2021.
41. Langaker GS. Prenatal Genetic Screening. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing; May 21, 2020.
42. Kirkegaard I, Uldbjerg N, Oxvig C. Biology of pregnancy-associated plasma protein-A in relation to prenatal diagnostics: an overview. *Acta ObstetGynecol Scand*. 2010; 89(9):1118–1125.
43. Patil M, Panchanadikar TM, Wagh G. Variation of Papp-A level in the first trimester of pregnancy and its clinical outcome. *J Obstet Gynaecol India*. 2014;64(2):116–119.
44. Santorum M, Wright D, Syngelaki A, et al. Accuracy of first-trimester combined test in screening for trisomies 21, 18 and 13. *Ultrasound Obstet Gynecol*. 2017;49(6): 714–720.
45. Wilson RD. Guide No. 410: Prevention screening, diagnosis, and pregnancy management for fetal neural tube defects. *J Obstet Gynecol Can*. 2021;23(2):124–139.
46. Marques P, Gunawardana K, Grossman A. Transient diabetes insipidus in pregnancy. *Endocrinol Diabetes Metab Case Rep*. 2015;2015:150078. doi:10.1530/EDM-15-0078.
47. Castro LC, Hobel CJ, Gornbein J. Plasma levels of atrial natriuretic peptide in normal and hypertensive pregnancies: a meta-analysis. *Am J Obstet Gynecol*. 1994;171(6): 1642–1651.
48. Lee R, Dupuis C, Chen B, et al. Diagnosing ectopic pregnancy in the emergency setting. *Ultrasonography*. 2018;37(1): 78–87. doi:10.14366/usg.17044

CHAPTER 25

Therapeutic Drug Monitoring

Jie Gao and Floyd Josephat

CHAPTER OUTLINE

KEY TERMS

Bioavailability
Distribution
Drug half-life *(T½)*
Peak drug concentration
Pharmacodynamics
Pharmacogenomics
Pharmacokinetics
Therapeutic drug monitoring (TDM)
Therapeutic range
Trough drug concentration

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to:

- Discuss drug characteristics that make therapeutic drug monitoring essential.
- Identify factors that influence the absorption of an orally administered drug.
- List factors that influence the rate of drug elimination.
- Define drug distribution and discuss factors that influence it.
- Calculate volume of distribution, elimination constant, and drug half-life.
- Correlate drug concentrations to pharmacokinetic and pharmacodynamic parameters.
- State specimen collection and handling requirements for therapeutic drug monitoring.
- Specify the therapeutic category or use of each drug presented in this chapter.
- Describe potential toxic side effects/toxicity of the therapeutic drugs discussed.
- Apply knowledge of therapeutic drug monitoring to interpret laboratory results.

Therapeutic drug monitoring (TDM) is the measurement of drugs (medications that are prescribed by medical providers) and/or their metabolites in body fluids, usually in blood, to maintain therapeutic benefits. The **therapeutic range** of a drug is the dose (or concentration) range of a drug within which the drug produces the desired therapeutic effect. A drug therapy may lead to inefficacy or toxicity if its blood concentration is outside of the therapeutic range. Laboratory personnel play an essential role in TDM because appropriate timing of specimen collection, accurate measurement of drug concentrations, and

CASE STUDY 25.1, PART 1

John, a 56-year-old man, was brought to the emergency department (ED) by his wife with signs of renal failure. She reports that John has congestive heart failure that has been successfully treated with digoxin for several years, but he recently developed the renal failure. Since the couple has been traveling internationally, his wife is concerned regarding his adherence with taking his prescribed drugs as directed by his cardiologist.

© Thinkstock/Stockbyte/Getty Images.

CASE STUDY 25.2, PART 1

Thelma is a 72-year-old woman who lives alone. Her daughter brought Thelma to the emergency department because Thelma says she has not felt well for the last 2 days. She was admitted through the ED 2 hours ago to the Cardiac Care Unit with a cardiac arrhythmia. Thelma states she has been taking her medications, but she does not remember when or if she took them today. Her daughter indicates Thelma is taking oral procainamide.

© Octavio Burguette/Shutterstock.

CASE STUDY 25.3, PART 1

Adriene is a patient who has been on successful oral phenytoin therapy for seizure disorders for the last 6 months. She tells her physician that she has experienced intermittent severe diarrhea for the past 2 weeks. After the prolonged diarrhea, Adriene had a seizure for the first time in "a long time."

© Rido/Shutterstock.

timely reporting of results are critical to achieve safe and effective pharmaceutical therapies for patients.

Most drugs are approved for use with recommended dosage regimens for patients who will achieve therapeutic benefits without the need of frequent adjustments of dosages. However, TDM is usually employed when a drug has a narrow therapeutic range or marked pharmacokinetic variabilities or critical adverse effects. The main purposes of TDM are to (1) ensure correct drug dosages for the therapeutic range, and (2) identify drug–drug interactions if multiple drugs are taken together.

Patient age, gender, genetics, diet, co-administered drugs, and even naturopathic agents can influence drug concentrations and efficacy. Therefore, TDM is often used to establish a dosing regimen that fits each individual situation and need.

TDM may also be used to identify patients who are nonadherent or to reoptimize a dosing regimen because of drug–drug interactions or patient's physiologic changes that may affect drug concentrations at sites of action.

TDM is performed based on studies of pharmacokinetics and pharmacodynamics. This chapter introduces these concepts and their influence on therapeutic effects. Some drugs commonly monitored by TDM testing in the clinical laboratory are also briefly discussed.

Pharmacokinetics

Pharmacokinetics is a study of the movement of drugs in the body. It provides a time course of drug concentrations in the body as a function of absorption, **distribution**, metabolism, and excretion. There are many factors that influence drug absorption, distribution, metabolism, and elimination (**Figure 25.1**);

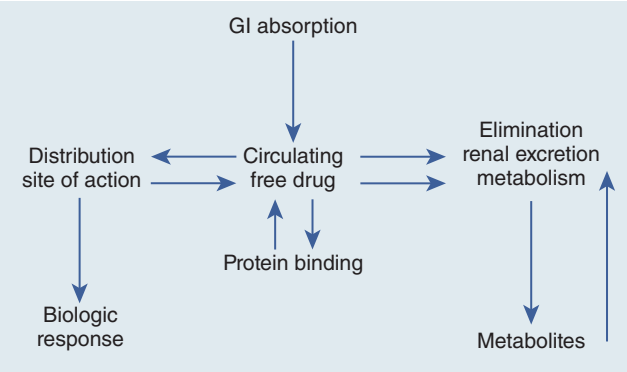


Figure 25.1 Overview of factors that influence the circulating concentration of an orally administered drug. GI, gastrointestinal.

therefore, the process of achieving therapeutic drug concentrations is by no means straightforward.

Routes of Administration

To achieve optimal therapeutic benefit, a drug must be at an appropriate concentration at its site of action. For instance, a cardioactive drug might need to reach cardiac myocytes at a dosage concentration that is effectively maintained in the therapeutic range for days to weeks. Oral administration is the most common route of delivery, as it is the least invasive for patients. Intravenous (IV) administration via an IV drip or an implanted venous access port offers the most direct and effective delivery to their sites of action. The fraction of the administered dose that eventually reaches its site of action is defined as its **bioavailability**.

Some medications can also be injected directly into muscle tissue through intramuscular (IM) injections or just under the skin with a subcutaneous (SC) injection. Others can be inhaled as an aerosol via a nebulizer or can be absorbed through the skin (transcutaneous) through use of a transdermal patch. Rectal delivery through administration of a suppository is commonly used in infants.

Each method of administration presents with different characteristics that affect circulating drug concentrations. Our discussions of drug absorption, distribution, metabolism, and elimination will focus on the most commonly administered therapeutic drugs for hospitalized patients, which laboratorians are responsible for quantitating in clinical chemistry.

Drug Absorption

When oral administration of drugs is utilized, the efficiency of absorption from the gastrointestinal tract depends on several factors, including (1) dissociation from its administered form, (2) solubility in gastrointestinal fluids, and (3) diffusion across gastrointestinal membranes. Tablets and capsules require dissolution before being absorbed, whereas liquid solutions tend to be more rapidly absorbed. Some drugs are subject to uptake by active transport mechanisms intended for dietary constituents; however, most are absorbed by passive diffusion from the gastrointestinal tract into the bloodstream. This process requires that the drug be in a hydrophobic, or nonionized, state. Because of gastric acidity, weak acids are efficiently absorbed in the stomach, but weak bases are preferentially absorbed in the

intestine where the pH is more alkaline. For most drugs, absorption from the gastrointestinal tract occurs in a predictable manner in healthy people; however, changes in intestinal motility, pH, inflammation, as well as the presence of food or other drugs, may dramatically alter absorption rates. For instance, a patient with inflammatory bowel syndrome may have a compromised gastrointestinal tract, which may affect normal absorption of some drugs. Absorption can also be affected by coadministration of drugs that affect gastrointestinal function such as antacids, fiber, kaolin, sucralfate, cholestyramine, and antiulcer medications. Morphine may also slow gastrointestinal motility, thereby influencing the rate of drug absorption.

Additionally, drug absorption rates may change with age, pregnancy, or pathologic conditions. In these instances, predicting the final circulating concentration in blood from a standard oral dose can be difficult. However, with the use of TDM, effective oral dosage regimens can be determined.

Drug Distribution

The free fraction of circulating drug is subject to diffusion out of the vasculature and into interstitial and intracellular spaces. The movement of a drug between blood circulation and tissues/organs and the relative proportion of the drug in the tissues define the drug's distribution. The ability to leave circulation is largely dependent on the lipid solubility of the drug. Drugs that are highly hydrophobic can easily cross cell membranes to enter cells or partition into lipid compartments, such as adipose tissue and nerve cells. Drugs that are polar and not ionized can also cross cell membranes but do not sequester into lipid compartments. Ionized species diffuse out of the vasculature but at a slow rate. The volume of distribution relates the dose of the drug in the body to the concentration of the drug that is measured in the specimen and is expressed mathematically as follows:

$$V_d = D/C \qquad \text{(Eq. 25.1)}$$

where V_d is the volume of distribution in liters, D is the dose of the drug in milligrams (mg) or grams (g), and C is the drug concentration in plasma (mg/L or g/L). Drugs that are hydrophobic can have large V_d values due to partitioning into hydrophilic compartments. Substances that are ionized or are primarily bound to proteins in plasma have small V_d values due to sequestration in the vasculature.

Free Versus Bound Drugs

Most drugs in circulation are subject to binding with serum protein constituents, and most form drug–protein complexes. An important aspect regarding drug dynamics is that typically only the *free* or *unbound* fraction can interact with its site of action and result in a biologic response. For this reason, the free drug fraction is also termed *active* fraction. At a standard dose, the total plasma drug concentration may be within the therapeutic range, but patients may experience toxic adverse effects due to a high free fraction or no therapeutic benefit due to a low free fraction. These alterations in active fraction may occur secondary to changes in plasma protein content during inflammation, malignancies, pregnancy, hepatic disease, nephrotic syndrome, malnutrition, and acid–base disturbances. Albumin represents the majority of protein constituents in plasma and the major transporter of drugs, and changes in its concentration can affect the free versus bound status of many drugs. Additionally, increases in plasma α_1-acid glycoprotein during inflammation will lead to increased binding of drugs such as propranolol, quinidine, chlorpromazine, cocaine, and benzodiazepines. The fraction of free drugs may also be influenced by other drugs or endogenous substances, such as urea, bilirubin, or hormones that compete for binding sites of proteins. Measurement of the free drug fraction should be considered for drugs that are highly protein-bound or when clinical signs are inconsistent with the total drug measurement.

Drug Metabolism

All substances absorbed from the intestine (except from the rectum) enter the hepatic portal system. In this system, circulating blood from the gastrointestinal tract is routed through the liver before it enters into general circulation. *First-pass effect* is a phenomenon in which a drug is metabolized, resulting in a reduced concentration of the drug before reaching the circulatory system. Certain drugs are subject to significant hepatic metabolism during passage through the liver, as the liver is a major site of drug metabolism. Liver metabolism varies in every individual, as it is influenced by genetics. These variations in drug metabolism related to genetics are examined in the discipline of **pharmacogenomics**. In addition to genetic variations, a patient with impaired liver function may have a reduced capacity to metabolize certain drugs. This concern is particularly important if the efficacy of a drug depends on a therapeutically

active metabolite that results from the action of liver enzymes on the original drug. This enzymatic process is referred to as *biotransformation*. Patients with liver disorders may require dose adjustments of the drug as the rate of metabolism and the subsequent elimination processes are altered.

Most drugs are *xenobiotics*, which are exogenous substances that are capable of entering biochemical pathways intended for endogenous substances. There are many potential biochemical pathways in which drugs can be acted on or biotransformed. The biochemical pathway responsible for a large portion of drug metabolism is the hepatic mixed-function oxidase (MFO) system. The basic function of this system involves taking hydrophobic substances and, through a series of enzymatic reactions, converting them into water-soluble products. These products then can be either transported into the bile or released into general circulation for elimination by renal filtration.

There are many enzymes involved in the MFO system and are commonly divided into two functional groups or phases. Phase I reactions produce reactive intermediates. Phase II reactions conjugate functional groups, such as glutathione, glycine, phosphate, and sulfate, to reactive sites on the intermediates, resulting in water-soluble products. The MFO system is nonspecific and allows many different endogenous and exogenous substances to go through this series of reactions. Although there are many potential substrates for this pathway, the products formed from an individual substance are specific. For example, acetaminophen is metabolized in the MFO pathway, ultimately leading to the formation of a glutathione conjugate following phase II reactions. In the presence of too much acetaminophen, as in the case of an overdose, the MFO system may be overwhelmed and cannot effectively metabolize it to a safe, water-soluble end product for elimination by the kidneys. In this case, the conjugating group for a given drug can become depleted in phase II reactions, and an accumulation of phase I products occurs. Excessive phase I products may result in toxic effects, and in the case of acetaminophen, irreversible damage to hepatocytes may occur.

It is also worth noting that the MFO system can be induced. This is seen as an increase in the synthesis and activity of the rate-limiting enzymes within this pathway. Induction of the MFO system typically results in accelerated clearance and a corresponding shorter **drug half-life (T½)** (explained in the next section). Hepatic diseases characterized by a loss of functions may result in slower rates of clearance and a corresponding longer drug half-life. For example, cirrhosis results in irreversible damage and fibrosis of the liver, rendering nonfunctional hepatocytes. Consequently, xenobiotics may not be effectively metabolized by the MFO system, thereby reducing the rates of metabolism and elimination while increasing the opportunity for toxicity. Thus, certain drugs may affect their own rates of elimination. Due to the variability of induction, TDM is commonly used to establish the appropriate dosage regimens for these drugs.

Because many substrates may enter the MFO system, competitive and noncompetitive drug–drug interactions can result in altered rates of elimination of the involved drugs. Interactions are not limited to between drugs but may also occur between a drug and food (e.g., grapefruit) or beverages (e.g., alcohol or caffeine). For example, metabolism of acetaminophen by the MFO system is altered in the presence of alcohol, which renders it more toxic. In most instances, the degree of alteration is unpredictable.

For some drugs, there is a considerable amount of variance in the rate of hepatic and nonhepatic drug metabolism within a normal population. This results in a highly variable rate of clearance, even in the absence of disease. Establishing dosage regimens for these drugs is, in many instances, aided by TDM. With increasing use of molecular genetic testing, the identification of common genetic variants involved in drug-metabolizing pathways, and an individual genome is now assisting in the establishment of an individualized dosage regimen. TDM is foundational to personalized medicine.

Drug Elimination

Drugs can be cleared from the body by various mechanisms. The free fraction of a drug or its metabolites are subject to glomerular filtration, renal secretion, or both. For those drugs that are not secreted nor reabsorbed, the elimination rate of the free drug fraction directly relates to the glomerular filtration rate (GFR). Decreases in GFR result in increased drug half-lives and elevated plasma concentrations. Aminoglycoside antibiotics and cyclosporine, an immunosuppressant drug, are examples of drugs that are not secreted or reabsorbed by the renal tubules.

Independent of the clearance mechanism, decreases in the plasma drug concentration most often occur as a first-order process, indicating an exponential rate of loss. A first-order elimination follows the general equation:

$$\Delta C/\Delta T = -kC \qquad \textbf{(Eq. 25.2)}$$

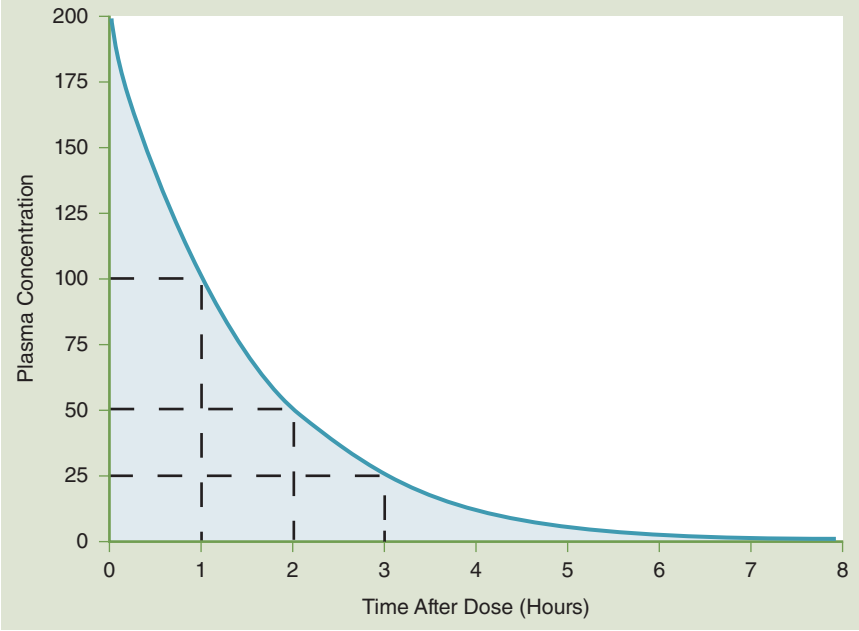


Figure 25.2 First-order drug elimination. This graph demonstrates exponential rate of loss on a linear scale. Hash-marked lines are representative of drug half-life.

This equation defines the changing rate of drug concentration ($\Delta C/\Delta T$) and is directly related to the concentration of drug (C) and the constant (k). The k value is a proportionality factor that describes the percent change per unit time. It is commonly referred to as the *elimination constant* or the *rate of elimination* and is negative because it is a decreasing value. The graphic solution to this equation is an exponential function that declines in the predicted curvilinear manner, asymptotically approaching zero (**Figure 25.2**). The graph shown in Figure 25.2 illustrates a large change at high drug concentrations and a smaller change at low drug concentrations; however, the rate, or percent lost, remains the same. Plotting it in semilogarithmic dimensions (**Figure 25.3**) can linearize this function.

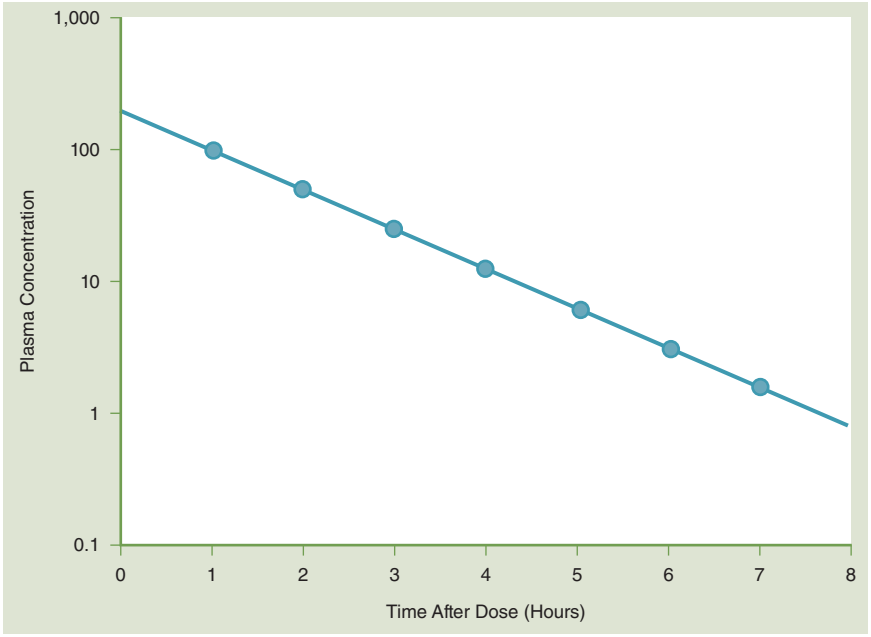


Figure 25.3 Semilogarithmic plot of exponential rate of drug elimination. The slope of this plot is equal to the rate of elimination (k).

Drugs are also eliminated through hepatic metabolism, renal filtration, or a combination of the two. Many drugs metabolized by the liver are then excreted in bile for elimination. For some drugs, elimination by these routes is highly variable and may be altered by functional changes in either organ system. In these situations, the elimination rate and the estimated blood concentration of a drug after a given time period are important factors in establishing an effective and safe dosage regimen. Equation 25.3 and Figures 25.2 and 25.3 are useful in determining the rate of elimination and concentration of a drug after a given time period. The following equation illustrates how the drug concentration can be estimated based on an initial concentration, the elimination constant, and time since the initial dose is administered:

$$C_T = C_0 e^{-kt} \qquad \text{(Eq. 25.3)}$$

where C_0 is the initial concentration of drug, C_T is the concentration of drug after the time period (T), k is the elimination constant, and T is the time period evaluated. This is the most useful form of the elimination equation. By measuring the initial drug concentration and the concentration after a period of time (T), the elimination constant can be determined. Once k is known, the amount of drug that will be present after a certain time period can be determined.

Elimination Example

The concentration of gentamicin is 10 µg/mL at 12:00. At 16:00, the gentamicin concentration is 6 µg/mL. What is the elimination constant (k) for gentamicin in this patient? Using Equation 25.3:

$$C_T = C_0 e^{-kt}$$
$$C_0 = 10 \text{ µg/mL}, C_T = 6 \text{ µg/mL}, T = 4 \text{ hours}, k = ?$$

Substituting these values into the elimination equation (Eq. 25.3) yields:

$$6 \text{ µg/mL} = (10 \text{ µg/mL})\, e^{-k[4 \text{ h}]}$$

Divide both sides by 10 µg/mL and note that the concentration units can be canceled:

$$0.6 = e^{-k[4 \text{ h}]}$$

To eliminate the exponent, take the natural logarithm of both sides:

$$\ln (0.6) = \ln (e^{-k[4 \text{ h}]})$$

Solve the natural log:

$$-0.51 = -k(4 \text{ h})$$

Multiply both sides by –1:

$$0.51 = k(4 \text{ h})$$

Divide both sides by 4 hours:

$$0.13/\text{h} = k$$

The calculated value for k indicates the patient is eliminating gentamicin at a rate of 13% per hour.

In this same patient on the same day, what would the predicted blood concentration of gentamicin be at midnight (0000)?

For C_0, either the 12:00 or 16:00 value can be used as long as the correct corresponding time duration is used. In this example, the 16:00 value of 6 µg/mL will be used.

$$C_0 = 6 \text{ µg/mL}, \text{T} = 8 \text{ hours}, \text{k} = 0.13/\text{h}, C_T = ?$$

Substituting these values into the elimination equation (Eq. 25.3) yields:

$$C_T = (6 \text{ µg/mL})\, e^{-0.13/\text{h}\,(8 \text{ h})}$$

Solve for the exponent and note that the time unit can be cancelled:

$$C = (6 \text{ µg/mL})\, e^{-1.04}$$
$$C_T = (6 \text{ µg/mL})\, (0.35)$$
$$C_T = 2.1 \text{ µg/mL}$$

The calculated C_T value indicates the patient would have a serum gentamicin concentration of 2.1 µg/mL at midnight (0000).

Although the elimination constant (k) is a useful value, it is not a common nomenclature in the clinical setting. Instead, the term drug half-life (T½) is used. One drug half-life is the time needed for the blood concentration of the drug to decrease by one-half. It can be determined graphically (**Figure 25.2**) or by conversion of the elimination constant (k) to drug half-life (T½) using the formula given in Equation 25.4. Of these two methods, the calculation provides the easiest and most accurate way to determine the drug half-life. In this calculation, 0.693 is the logarithm of 2 and represents the exponential rate of elimination (assuming elimination is by first-order kinetics). Referring to the example above, the half-life of gentamicin in this patient would be calculated as follows:

$$T\,½ = 0.693/k$$
$$T\,½ = 0.693/(0.13/\text{h}) \qquad \text{(Eq. 25.4)}$$
$$T\,½ = 5.33 \text{ hours}$$

The gentamicin half-life value for this patient indicates that after 5.33 hours, the concentration of

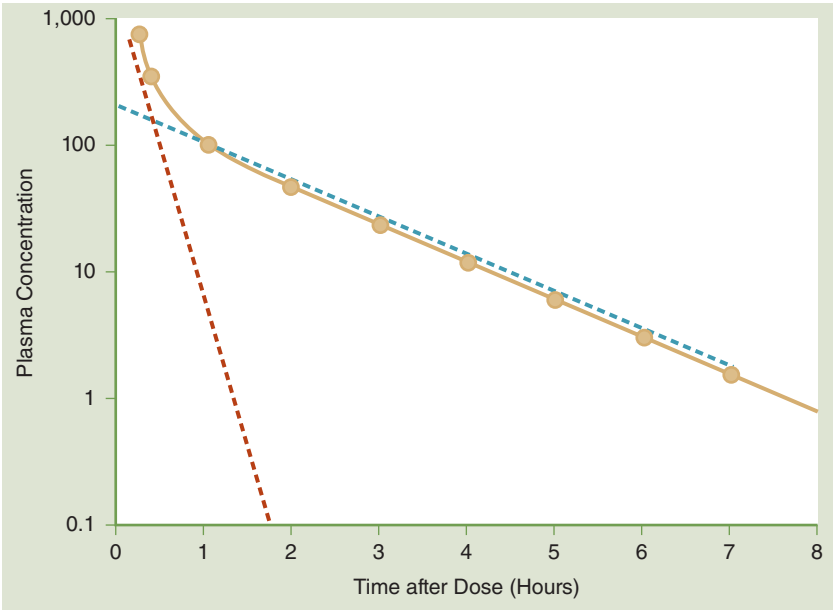


Figure 25.4 Semilogarithmic elimination plot of a drug subject to distribution. The initial rate of elimination is influenced by the distribution (*red line*) and terminal elimination rates (*blue line*). After distribution is complete (1.5 hours), elimination is first-order.

drug in the blood would be one-half of the initial concentration.

Figure 25.3 is an ideal plot of drug elimination after an intravenous (IV) bolus assuming there is no previous distribution of the drug. A drug that distributes outside of the vascular space would produce an elimination graph such as in **Figure 25.4**. The rapid rate of change seen immediately after the initial IV bolus is a result of distribution and elimination. The elimination constant (k) can be determined only after distribution is complete. **Figure 25.5** is a plot of drug concentration in blood vs. time after oral

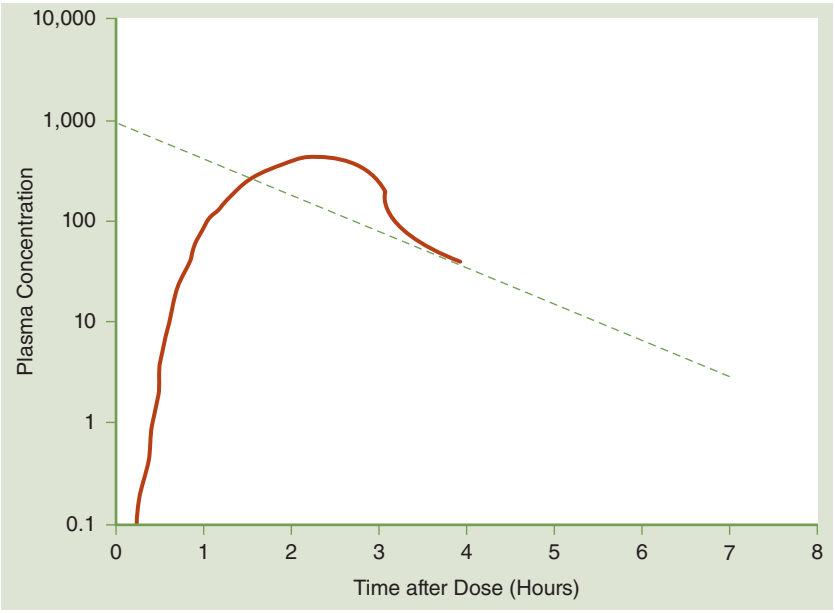


Figure 25.5 Blood concentrations of a drug after oral administration. After oral administration at time 0, blood concentration increases (*solid line*) after a brief lag period. Blood concentrations peak when the rates of elimination and distribution exceed the rate of absorption. First-order elimination (*dotted line*) occurs when absorption and distribution are complete.

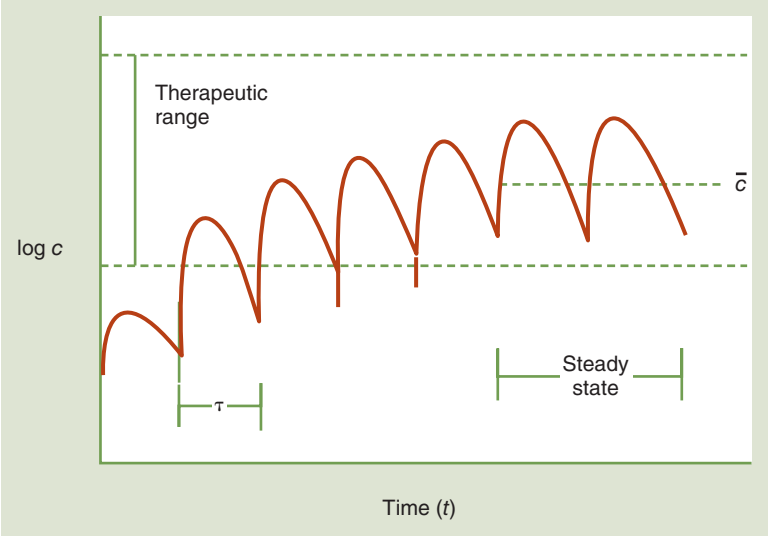


Figure 25.6 Steady-state kinetics in a multiple-dosage regimen. The character τ indicates the dosage interval. Equal doses at this interval reach steady state after six or seven dosage intervals. $\bar{c}$ means drug concentration.

administration. As absorbed drug enters the circulation, it is subject to simultaneous distribution and elimination. Blood concentrations rise when the rate of absorption exceeds the rate of distribution and elimination and declines as elimination and distribution exceed absorption. The rate of elimination can only be determined after absorption and distribution are complete.

Most drugs are not administered as a single bolus but are delivered on a scheduled basis (e.g., a prescribed dose of 100 mg given once every 8 hours). With this type of administration, the blood drug concentration oscillates between a maximum level and a minimum level, which are referred to as the **peak drug concentration** and the **trough drug concentration**, respectively. The goal of a multiple-dosage regimen is to achieve trough and peak concentrations within the therapeutic range. Evaluation of this oscillating function cannot be done immediately after initiation of a scheduled dosage regimen. Approximately five to seven doses are required before a *steady-state* oscillation is achieved, as demonstrated in **Figure 25.6**.

After the first oral dose, absorption and distribution occur, followed only by elimination. Before the concentration of the drug drops significantly, the second dose is given. The peak of the second dose is added to what remained from the first dose. Because elimination is first-order, the higher concentration results in a larger amount of elimination. The third through the seventh scheduled doses all have the same effect of increasing the blood concentration and the amount of elimination. By the end of the seventh dose, the amount of drug administered in a single dose is equal to the amount eliminated during the dosage period. At this point, a steady state is established, and the peak and trough concentrations can be evaluated.

Pharmacodynamics

Pharmacodynamics is the study of the biochemical and physiological effects of drugs and their mechanisms of action. It describes the relationship between a drug's concentrations at its site of action and its pharmacological responses, including both therapeutic and adverse effects. The most common mechanism of action is that a drug activates its receptors at the site of action. The extent of pharmacological responses is related to the concentrations of the drug at the site. This relationship is described by the *dose–response curve*, which plots the drug dose (or concentration) against its pharmacological responses (**Figure 25.7A**). This relationship can be influenced by several factors such as demographic factors (e.g., age, gender), the density of receptors at the site of action, the presence of other drugs that compete for binding at the same receptor and so on. If a drug concentration is high enough, a maximum effect (E_{max}) can be achieved. When the logarithm of drug concentration is plotted versus pharmacological responses (**Figure 25.7B**), the concentration at which 50% of the maximum effect is seen is referred to as the *50% effective concentration* or $EC_{50,}$ which is an important determinant of drug activities. Constant exposure of receptors to drugs may lead to a reduced response, which is referred to as *tolerance*.

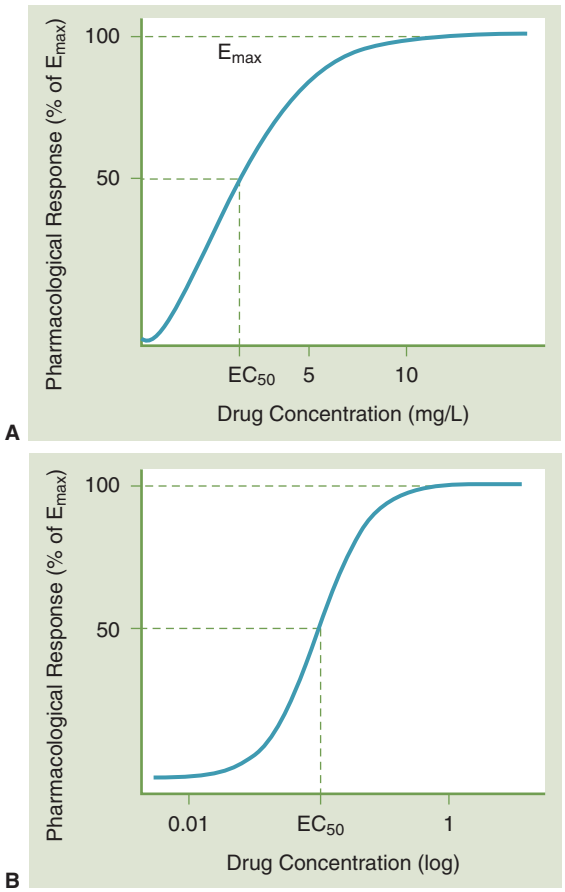


Figure 25.7 Typical dose (concentration)–response curve for the effect of a drug on a target site of action (as a percentage of the maximal effect). **(A)** Drug concentration against the responses. **(B)** The logarithm of drug concentration against the responses.

Tolerance can be described in terms of the dose–response curve. As shown in **Figure 25.8**, an increasing amount of EC_{50} is needed to achieve the same pharmacological response.

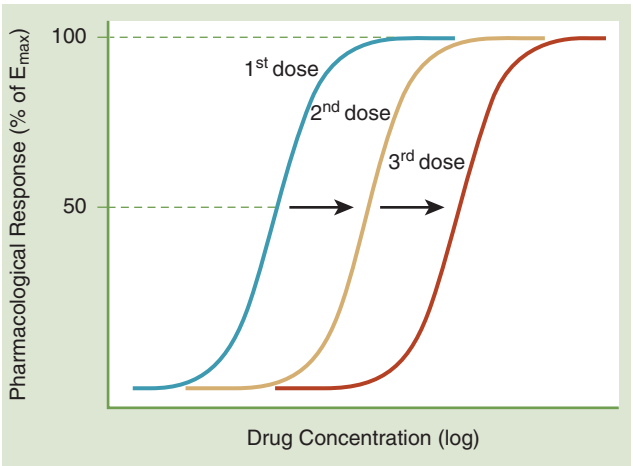


Figure 25.8 Dose (concentration)–response curves for drug tolerance with repeated dosing.

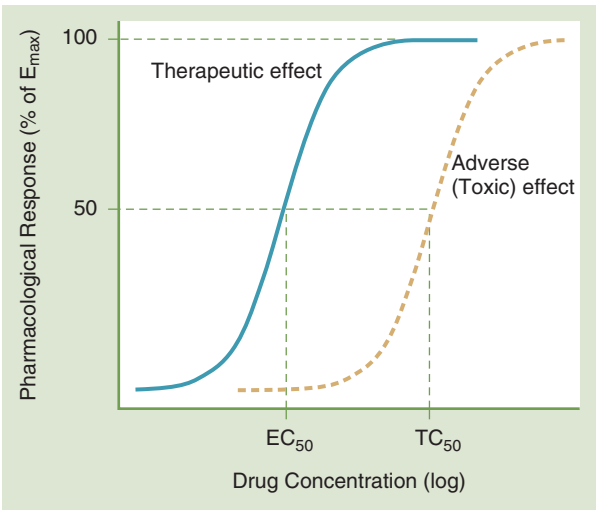


Figure 25.9 Dose (concentration)–response curves for the therapeutic and adverse effects.

As shown in **Figure 25.9**, the adverse (or toxic) effects of drugs are often dose-related in the same way to the therapeutic effects. The ratio between the EC_{50} and the 50% toxic concentration (TC_{50}) as shown in **Eq. 25.5** is known as the *therapeutic index* (TI), which indicates the relative safety of a drug. The TIs of most drugs are greater than 100, but some drugs, such as digoxin, warfarin, insulin, phenytoin, and opioids have therapeutic indices less than 10. TDM is often employed to titrate the doses of such drugs with low therapeutic indices in order to maximize therapeutic benefits but avoid adverse effects.

$$\textit{Therapeutic Index} = TC_{50}/EC_{50} \qquad \textbf{(Eq. 25.5)}$$

Specimen Collection

Accurate timing of specimen collection is the single most important factor in TDM. In general, trough concentrations are drawn right before the next dose, and peak concentrations are drawn 1 hour after an orally administered dose. The peak level for intravenously administered aminoglycosides is 90 minutes after completion of the infusion. This rule of thumb must always be used within the clinical context of the situation. Drugs that are absorbed at a slow rate may require several hours before the peak drug concentrations can be evaluated. In all situations, the determination of drug peak concentrations should be performed only after a steady state has been achieved.

Serum is the specimen of choice for measuring circulating concentrations of most drugs. Care must be taken that the appropriate container is used when collecting these specimens, as some drugs may be

absorbed into the gel of serum separator collection tubes. It is necessary to follow manufacturer recommendations to prevent preanalytic error, and failure to do so may result in falsely low values. Calcium-binding anticoagulants add cations that may interfere with analysis or cause a drug to distribute differently between red blood cells and plasma. As a result, specimen tubes that contain ethylenediaminetetraacetic acid (EDTA), citrate, or oxalate are generally considered unacceptable specimen types for TDM analysis. However, there is one notable exception: EDTA whole blood is the specimen of choice for measuring immunosuppressive drugs.

Pharmacogenomics

The effectiveness of a drug within a population depends in part on how patients respond. For any medication, there are patients who are *responders* and patients who are *nonresponders*. Responders are patients benefiting from the therapeutic and desired effects of a drug, while nonresponders do not demonstrate a beneficial effect from the initiation of a given drug regimen. The therapeutic effectiveness of drugs in responders and nonresponders has recently been attributed to the interindividual variation in genetic polymorphisms of the patients' drug metabolism pathways. As previously mentioned, pharmacogenomics is the science of studying these variations and developing drug therapies to compensate for the genetic differences impacting therapy regimens.

One of the most prominent genes that affect drug metabolism is *CYP450*, the gene that encodes cytochrome P450 (CYP450), which is a family of enzymes within the MFO system previously described. The differences in drug metabolism in a population are attributable to the variations in the enzymes as a result of genetic polymorphism. The three variations most often linked to differences in drug metabolism are CYP2D6, CYP2C9, and CYP3A4. This information can be used to personalize drug doses to the degree that is appropriate for the CYP450 profile of the patient. For example, if the patient's CYP450 profile indicates they have genes known to metabolize a drug more slowly, they would be given lower doses of the drug to avoid toxic concentrations. Alternatively, if a patient's CYP450 profile indicates the presence of genes predisposing the patient to an increased rate of metabolism, an increased dose would be needed to maintain therapeutic drug concentrations. Pharmacogenetic profiling can also be used to predict drug–drug interactions or as an indicator if the drug will provide any therapeutic benefit at all.[1]

Cardioactive Drugs

Cardioactive drugs are the most widely used medicines. Many types of cardioactive drugs have been developed to treat cardiac complications, and only the inotropic agents (cardiac glycosides) and antiarrhythmics frequently involve TDM practice to avoid their serious adverse effects.

Digoxin

Digoxin (Lanoxin) is a cardiac glycoside used in the treatment of cardiac arrythmias as well as congestive heart failure.[2] It functions by inhibiting membrane Na^+, K^+-ATPase. This causes a decrease in intracellular potassium, which results in increased intracellular calcium in cardiac myocytes. The increased intracellular calcium improves cardiac contractility. This effect is seen in the plasma concentration range of 0.8 to 2.0 ng/mL.[3] Although levels >2.0 ng/mL can be used to control ventricular tachycardia, toxic serum concentrations in this range often mimic the cardiac arrhythmias for which the drug was originally prescribed.[2] Other adverse effects include premature ventricular contractions (PVCs), atrioventricular node blockage, nausea, vomiting, and visual disturbances.

The absorption of orally administered digoxin is influenced by dietary factors, gastrointestinal motility, and formulation of the drug. In circulation, its plasma protein binding rate is about 25%. The unbound or free form of digoxin is sequestered into muscle cells, and at equilibrium, the tissue concentration is 15 to 30 times greater than that of blood. Elimination of digoxin occurs primarily by renal filtration of unbound digoxin. The remainder is metabolized by the liver. The half-life of plasma digoxin is 38 hours in an average adult. The major contributing factor to the extended half-life is the slow release of tissue digoxin back into circulation.

Because of the narrow therapeutic range, establishing a dosage regimen of digoxin usually requires assessment of blood concentrations after initial dosing to ensure that effective and nontoxic plasma concentrations are achieved. In addition, the GFR can have dramatic effects on blood concentrations of digoxin. Frequent dose adjustments, in conjunction with measurement of blood digoxin concentrations, should be performed in patients with renal diseases. Low potassium and magnesium potentiate digoxin actions,[3] and adjustment of blood concentrations below the therapeutic range may be necessary to avoid toxicities. Thyroid function may also influence the actions of digoxin. Hyperthyroid patients display a resistance

to digoxin actions, and hypothyroid patients become more sensitive to digoxin.[4]

To evaluate the peak blood concentration of digoxin, the timing of specimen collection is crucial. In an average adult, the peak blood concentrations of digoxin occur between 2 and 3 hours after oral administration; however, its uptake into the tissues is a relatively slow process. As a result, its peak blood concentrations do not correlate well with therapeutic effects on tissues. It has been established that the significant effects of digoxin occur 8 to 10 hours after an oral dose, and specimens should be drawn within this window. Specimens collected before this time are misleading and should not be considered valid representations of the therapy to the patient.

Immunoassays are commercially available to measure total digoxin (digoxin and its active metabolites) concentrations in serum.[3] It should be noted that this assay measures both bound and free digoxin, so increased values will be seen in patients undergoing treatment with Digibind, an antidote for digoxin overdose. Newborns, pregnant women, and patients with renal and hepatic failure produce endogenous digoxin-like immunoreactive factors (DLIF) that cross-react with the antibodies used to measure digoxin and may cause falsely elevated concentrations. Biotin is another interfering substance.[5]

Quinidine

Quinidine (Quinidex Extentabs, Cardioquin, or Quinora) is a natural product extracted from the bark of the "fever tree" (*Cinchona* spp.) native to Central and South America. It is commonly used to treat various cardiac arrhythmias. The two most common formulations are quinidine sulfate and quinidine gluconate. Quinidine is generally administered orally, as its gastrointestinal absorption is complete and rapid, and its peak blood concentrations occur 2 hours after an oral administration. The therapeutic range of quinidine is 2 to 5 µg/mL. Approximately 70% to 80% of circulating quinidine is bound to plasma proteins.[3] Quinidine has a half-life of 6 to 8 hours and is primarily eliminated by hepatic metabolism. Induction

CASE STUDY 25.1, PART 2

Remember John. Laboratory records indicate semiannual peak digoxin concentrations that have all been within the therapeutic range. A serum specimen was collected upon admission, and the results received by the ED physician are shown in the table below. The physician reviews the results with John and his wife and indicates that John is not exhibiting signs or symptoms of digoxin toxicity.

© Thinkstock/Stockbyte/Getty Images.

Laboratory Results

Analyte	Patient Value	Reference Range
Sodium	129	135–145 mmol/L
Potassium	5.5	3.5–5.3 mmol/L
Chloride	113	97–107 mmol/L
tCO_2	16	21–31 mmol/L
Urea nitrogen	180	5–20 mg/dL
Creatinine	4.5	0.6–1.0 mg/dL
Digoxin	2.5	0.9–2.0 ng/mL

1. After reviewing the results, should the physician be concerned about any of these results, and why?
2. If the specimen was collected without consideration of the time of the patient's last digoxin dose, how might this affect the interpretation of the test result?
3. Other than the time relative to dose administration, what additional laboratory results should be taken into consideration when interpreting the digoxin result?
4. How can the test methodology affect the interpretation of the digoxin result?

of hepatic metabolism, such as by barbiturates, increases the clearance rate of quinidine, whereas impairment of this system, as seen in late-stage liver disease, may extend the half-life of quinidine in circulation. The most common adverse effects of quinidine toxicity are nausea, vomiting, and abdominal discomfort. However, more serious side effects such as thrombocytopenia and tinnitus can occur, which may need medical intervention. Signs of cardiovascular toxicity, such as PVCs, may be seen when blood concentrations are twice the upper limit of the therapeutic range.

In most instances, monitoring of quinidine therapy only involves measurement of trough concentration to ensure blood concentrations are within the therapeutic range. Assessment of peak concentrations is performed only when symptoms of toxicity are present. Serum quinidine concentrations are routinely determined by immunoassays.

Procainamide and *N*-acetylprocainamide

Like quinidine, procainamide (Procanbid, Procan SR, Pronestyl) is also an antiarrhythmic drug. The oral route is the most common administration method as its absorption is rapid and complete, with peak blood concentrations occurring about 1 hour after administration.[3] Approximately 20% of absorbed procainamide is bound to plasma proteins. Procainamide has a half-life of approximately 4 hours and is eliminated by a combination of both renal filtration and hepatic metabolism. *N*-acetylprocainamide (NAPA), a hepatic metabolite of procainamide, demonstrates similar antiarrhythmic potential to the parent drug. The therapeutic range for procainamide is 4–10 µg/mL and for NAPA is 12–18 µg/mL. Alteration in either renal or hepatic function can lead to increased plasma concentrations, which results in myocardial depression and arrhythmia. Both procainamide and its active metabolite are measured by immunoassay to determine the total antiarrhythmic potential of procainamide therapies. Toxicity may occur when the sum of procainamide and NAPA exceeds 40 µg/mL.

Disopyramide

Disopyramide (Norpace) is another antiarrhythmic drug used in the treatment of cardiac abnormalities. Disopyramide may be administered as a quinidine substitute when the adverse effects of quinidine are unacceptable. It is most commonly administered orally, as disopyramide is absorbed completely and rapidly, with plasma concentrations occurring about 1 to 2 hours after administration.[3] Disopyramide binds to several plasma proteins. Its plasma protein binding is highly variable between individuals and is concentration dependent. Therefore, the pharmacodynamics of disopyramide have not been fully determined. In most patients, the therapeutic range of disopyramide is 3.0–7.5 µg/mL. However, interpretation of disopyramide results should take the clinical perspective into consideration. The adverse effects of disopyramide are dose dependent.[3] Anticholinergic effects, such as dry mouth and constipation, may be seen at concentrations greater than 4.5 µg/mL. Cardiac effects of drug toxicity, such as bradycardia and atrioventricular node blockage, are usually seen at concentrations greater than 10 µg/mL. Disopyramide has a half-life of approximately 7 hours and is primarily eliminated by renal filtration and, to a lesser extent, by hepatic metabolism. In conditions with a decreased GFR, the drug half-life is prolonged, and the blood concentrations

CASE STUDY 25.2, PART 2

Remember Thelma. Initial laboratory results collected in the ED are not indicative of renal or hepatic disease. After receiving procainamide results below the therapeutic range, the cardiologist orders an IV loading dose resulting in a serum concentration of 6 µg/mL. The therapeutic range for procainamide is 4 to 8 µg/mL, and its half-life is 4 hours. Four hours after the initial loading dose, another equivalent dose is given as an IV bolus, resulting in a serum concentration of 7.5 µg/mL.

1. What results did the cardiologist receive from the laboratory to rule out renal and hepatic disease?
2. Does the serum concentration after the second dose seem appropriate? If not, what would be the predicted serum concentration at this time?
3. What factors would influence the rate of elimination of this drug?

are elevated. The concentrations of disopyramide are commonly determined by chromatographic methods or immunoassay in serum or plasma.

Antibiotics

Aminoglycosides

Aminoglycosides are a group of antibiotics used to treat gram-negative and some gram-positive bacterial infections. There are many individual antibiotics within this classification, but the most frequently prescribed aminoglycosides are gentamicin, tobramycin, amikacin, and kanamycin. These antibiotics share a common pharmacological mechanism of inhibiting bacterial protein synthesis but vary in effectiveness against different strains of bacteria. The blood concentrations of aminoglycosides above the therapeutic ranges may cause serious toxicities, especially nephrotoxicity and ototoxicity. Aminoglycosides can disrupt inner ear cochlear and vestibular membranes, which results in hearing and balance impairment. The ototoxic effects are irreversible and may be seen with repeated high-level exposures.[6] Nephrotoxicity is also a major concern for patients on this type of antibiotic therapy as aminoglycosides damage the proximal renal tubules, which results in electrolyte imbalance and proteinuria. These effects are usually reversible; however, extended high-level exposures may result in necrosis of renal tubular cells and subsequent renal failure.[7] Toxic concentrations for this type of drug differ for each antibiotic and peak concentrations.

Aminoglycosides are poorly absorbed in the gastrointestinal tract, and their administration is limited to intravenous infusions or intramuscular injections. Therefore, these drugs are not used in an outpatient setting. The peak concentrations of aminoglycosides are achieved in 1 to 2 hours after administration. Aminoglycosides have a half-life of approximately 2 to 3 hours. They are eliminated by renal filtration. Appropriate dosing adjustments must be made in patients with compromised renal function. Immunoassay is the primary method to quantify aminoglycoside levels in serum.

Gentamicin

Gentamicin is an antibiotic used to treat life-threatening blood infections (bacteremia and septicemia) caused by gram-negative bacilli, particularly Enterobacteriaceae, *Acinetobacter* and *Pseudomonas* species. It is also used in combination with *S. aureus*. Conventional dosing of gentamicin is usually given 2 to 3 times per day by intravenous or intramuscular injections in doses to achieve peak blood concentration between 3.0 to 12.0 mcg/mL depending on the type of infection. Gentamicin also may be administered at higher doses (usually 5–7 mg/kg) once per day to patients with good renal function known as pulse dosing. Dosing amount or interval must be decreased to accommodate for reduced renal function. Ototoxicity and nephrotoxicity are the primary toxicities associated with gentamicin. Monitoring of serum levels and symptoms consistent with ototoxicity is important. The therapeutic range for the trough level is <2.0 µg/mL, and peak targets are generally between 3.0 and 12.0 mcg/mL for conventional dosing.

Tobramycin

Tobramycin is an antibiotic used to treat bacteremia and septicemia by gram-negative bacilli including Enterobacteriaceae and *Pseudomonas aeruginosa*. Toxicities include ototoxicity and nephrotoxicity. Monitoring of serum levels, renal function, and symptoms consistent with ototoxicity is important. Audiology testing should be considered at baseline prior to initiation of therapy. Therapeutic goal trough levels should be below 2.0 µg/mL for conventional dosing. Prolonged exposure to trough levels exceeding 2.0 µg/mL may lead to toxicity. The therapeutic goal for peak levels is 3.0–12.0 µg/mL.

Amikacin

Amikacin is an aminoglycoside used to treat severe blood infections by susceptible strains of gram-negative bacteria. Aminoglycosides induce bacterial death by irreversibly binding bacterial ribosomes to inhibit protein synthesis. Amikacin can be administered orally to reduce intestinal flora. Peak serum concentrations are seen 30 minutes after intravenous infusion or 60 minutes after intramuscular administration. Serum half-lives in patients with normal renal function are 2 to 3 hours. Amikacin is principally excreted by the kidneys. It may accumulate in the kidney at 50 to 100 times the serum concentration. Nephrotoxicity can present as dizziness, vertigo, or, if severe, ataxia and a Ménière disease–like syndrome. Auditory manifestations include simple tinnitus or any degree of hearing loss, which may be temporary or permanent. Nephrotoxicity is most frequently manifested by transient proteinuria or azotemia. The therapeutic range for trough levels are <8.0 µg/mL, and the peak levels are 20.0–35.0 µg/mL.

Vancomycin

Vancomycin is a glycopeptide antibiotic against gram-positive cocci and bacilli infections. Because of its poor gastrointestinal absorption, vancomycin is administered by intravenous infusion and reaches peak concentrations in 1 hour after dosing. The plasma protein binding level of vancomycin is approximately 55%. The pharmacodynamic properties of vancomycin have not been firmly established. Many toxic effects occur when the concentration exceeds the therapeutic range (10 to 20 µg/mL).[7] The main adverse effects of vancomycin include "red man syndrome," nephrotoxicity, and ototoxicity. Red man syndrome is characterized by an erythemic flushing of the extremities.[8] The nephrotoxic effects occur more frequently at trough concentrations greater than 20 µg/mL, and the ototoxic effects occur more frequently when its peak concentrations exceed 40 µg/mL. Because vancomycin has a long distribution phase, generally only trough concentrations are monitored to ensure the drug concentration is within the therapeutic range. Vancomycin has a half-life of 4 to 6 hours and is primarily eliminated by renal filtration and secretion. The serum concentration of vancomycin is routinely determined by immunoassay methods.

Antiepileptic Drugs

Antiepileptic drugs (AEDs) are used to treat seizure disorders and to suppress seizures but are only effective while the drug metabolites are in the body. The therapeutic ranges of AEDs are determined as the concentrations that provide therapeutic benefit with no or minimal acceptable adverse effects. AEDs have complicated pharmacokinetics, which cause significant interindividual variability in therapeutic ranges. In addition, the first generationof AEDs have narrow therapeutic ranges, and the therapeutic ranges of the second generation of AEDs have not been clearly identified. Therefore, TDM has been a critical part of AED treatments in clinical practice. TDM has been used to establish the therapeutic baseline, to reassess a modified medication regimen, and to assess patient adherence.[9] Most AEDs in serum are analyzed by immunoassay or chromatography. In a normal physiologic state, the total AED concentration (both free and protein-bound AEDs) may be sufficient for therapeutic monitoring purposes. However, a free drug measurement may be necessary when plasma protein concentrations change. This may include pregnancy, late-stage renal or hepatic disease, malnutrition, or when a known drug–drug interaction may occur.[10] The TDM of AEDs generally prefers a trough concentration measured at the end of the dosing interval.[10]

Primidone

Primidone is used for control of grand mal seizures that are refractory (resistant) to other antiepileptic drugs. Primidone is rapidly absorbed after oral administration, achieves peak serum concentrations after 2 hours, and has a half-life of 8 hours. It is not highly protein-bound, as approximately only 10% is bound to protein in the blood. Primidone is converted by the liver to its metabolites phenobarbital and phenylethylmalonamide. Primidone is preferred over phenobarbital when a steady state needs to be established quickly. At steady state, which is achieved approximately 2 weeks after therapy is initiated, blood levels of primidone range from 5.0 to 12.0 µg/mL for adults. When monitoring primidone and phenobarbital levels simultaneously, the specimen should be drawn just before the next dose is administered. Toxicity associated with primidone is primarily due to the accumulation of phenobarbital. Both primidone and phenobarbital are measured to assess the total AED potential, and primidone is measured by homogenous enzyme immunoassay.

Phenobarbital

Phenobarbital (Luminal or Solfoton) is a slow-acting barbiturate that effectively controls several types of seizures. Absorption of oral phenobarbital is slow but complete. For most patients, its peak plasma concentrations are reached about 10 hours after an oral dose, and approximately 50% of circulating phenobarbital is bound to plasma proteins. The half-life of phenobarbital is 70 to 100 hours. It is eliminated primarily by hepatic metabolism, and renal filtration is also significant. Compromised renal or hepatic function decrease its rate of elimination. Because of the slow rate of absorption and long half-life, blood concentrations of phenobarbital do not change dramatically within a dosing interval, and only trough concentrations are typically evaluated unless toxicity is suspected. Toxic adverse effects of phenobarbital include drowsiness, fatigue, depression, and reduced mental capacity.

Phenobarbital clearance occurs by the hepatic MFO system. It is also important to note that phenobarbital is a potent inducer of the MFO system. After initiation of therapy, dose adjustment is usually required after the induction period is complete. For most individuals, this occurs within 10 to 15 days

after the initial dose. The therapeutic range for phenobarbital is 20.0 to 40.0 µg/mL for adults, and it is measured by immunoassay techniques.

Phenytoin and Free Phenytoin

Phenytoin (Dilantin) is a common therapeutic agent used in the treatment of seizure disorders. It is also used as a short-term prophylactic agent in brain injury to prevent loss of functional tissue. Phenytoin is primarily administered as an oral preparation, and its gastrointestinal absorption is variable and sometimes incomplete. Peak concentrations are achieved 3 to 12 hours post dose. Circulating phenytoin has a high but variable degree of protein binding (87% to 97%) and can be easily displaced by other highly protein-bound drugs.[11] Like most drugs, the unbound, or free, fraction is the biologically active form of the drug. Reduced protein binding may occur with anemia, hypoalbuminemia, and when coadministered with other drugs with similar binding properties. In these situations, symptoms of toxicity may be observed even though the total drug concentration is within the therapeutic range. The most significant adverse effect of phenytoin toxicity is initiation of seizures. Thus, seizures in a patient receiving phenytoin therapy may be a result of subtherapeutic or toxic concentrations. Additional adverse effects of phenytoin include hirsutism, gingival hyperplasia, vitamin D deficiency, and folate deficiency. Phenytoin has a half-life of 6 to 24 hours and is eliminated by hepatic metabolism. At therapeutic concentrations, the elimination pathway of phenytoin may become saturated (zero-order kinetics); therefore, relatively small changes in dosage or elimination can have dramatic effects on blood drug concentrations. Phenytoin is also an inducer of the hepatic MFO pathway, which reduces the half-life of concurrently administered drugs that are eliminated by this pathway.[9] For most patients, serum phenytoin concentrations of 10 to 20 µg/mL are considered effective. In some situations, however, the therapeutic range must be individualized to suit the clinical situation.

As discussed, phenytoin is highly protein-bound (90%), mostly to albumin, and the remaining 10% circulates in the free, unbound form. Free phenytoin is the active form of the drug. Increased free phenytoin produces an enhanced pharmacologic effect. Additionally, the free fraction is readily available to the liver to be metabolized, and therefore it is cleared more quickly. The therapeutic range for free phenytoin in serum is 1 to 2 µg/mL. This has been well correlated with the pharmacologic actions of this drug. In patients with altered plasma protein binding, determination of the free fraction aids in dosage adjustment. Free phenytoin levels are the best indicator of adequate therapy in renal failure. Like most other TDMs, it is measured by immunoassay.

Valproic Acid

Valproic acid, or valproate (Depakote), is used in the treatment of petit mal and absence seizures.[11] It is administered as an oral preparation, as gastrointestinal absorption is rapid and complete. Circulating valproic acid is highly protein-bound (93%), and peak concentrations are reached in 1 to 4 hours after dosing. The percentage of protein-bound valproic acid decreases in renal failure, in late liver disease, and with coadministration of drugs that compete for its binding sites. Valproic acid is eliminated by hepatic metabolism, which may be induced by coadministration of other AEDs but is inhibited by administration of felbamate,

CASE STUDY 25.3, PART 2

Remember Adriene. Her doctor orders phenytoin, and the results are extremely low. Her dose is increased until her serum concentration returns to the therapeutic range. Adriene reports the diarrhea is resolved. Several days after the dose adjustment, she has another seizure.

1. What is the most probable cause of the initial low serum phenytoin concentration?
2. Would determination of free serum phenytoin aids in resolving the cause of the initial seizure?
3. What assays other than the determination of serum phenytoin would aid in this situation?
4. What is the most probable cause of the seizure after the diarrhea had been resolved?

which is another AED.[11] Without coadministration of these drugs, valproic acid has a half-life of 11 to 17 hours. The therapeutic range for valproic acid is relatively wide (50 to 120 µg/mL), and determination of blood concentrations is primarily performed to ensure that toxic levels (>120 µg/mL) are not present. Nausea, lethargy, and weight gain are the most common adverse effects of valproic acid toxicity; however, pancreatitis, hyperammonemia, and hallucinations have been associated with significantly elevated concentrations (>200 µg/mL). Hepatic dysfunction occasionally occurs in some patients even at therapeutic concentrations; therefore, hepatic markers should be checked frequently during the first 6 months of therapy. Many factors influence the portion of valproic acid bound to plasma proteins, so measurement of the free fraction provides a more reliable index of therapeutic and toxic concentrations.

Carbamazepine

Carbamazepine (Tegretol) is an effective treatment for various seizure disorders. However, because of its serious adverse effects including agranulocytosis and aplastic anemia, it is only used when patients do not respond well to other AEDs. Orally administered carbamazepine is absorbed with a high degree of variability, and approximately 70% to 80% of circulating carbamazepine is bound to plasma proteins. Peak concentrations are achieved in 4 to 8 hours after oral dosing, and the half-life for carbamazepine is 10 to 20 hours. It is eliminated primarily by hepatic metabolism, so liver dysfunction can result in accumulation of carbamazepine. Carbamazepine is an inducer of hepatic metabolism, so its blood concentrations must be frequently monitored. When indicated, dosage adjustments are guided by monitoring serum levels.

Carbamazepine toxicity is diverse and variable.[12] Certain effects occur in a dose-dependent manner, while others do not. There are several idiosyncratic effects of carbamazepine, which affect a portion of the population at therapeutic concentrations including rashes, leukopenia, nausea, vertigo, and febrile reactions.[12] Of these, leukopenia is most serious. Leukocyte counts are commonly assessed at baseline before treatment is initiated to detect this possible toxic effect. Liver function tests should also be evaluated during this time period as mild, transient liver dysfunction is commonly seen during initiation of therapy. Persistent increases in liver markers or significant leukopenia commonly result in discontinuation of carbamazepine therapy. The therapeutic range for carbamazepine is 4 to 12 µg/mL, and serum concentrations greater than 15 µg/mL are associated with toxicity, possibly resulting in aplastic anemia.[12]

Ethosuximide

Ethosuximide (Zarontin) is used for controlling petit mal seizures. It is administered as an oral preparation, with peak concentrations achieved within 2 to 4 hours after dosing. The therapeutic range of ethosuximide is 40 to 100 µg/mL. Toxicities associated with high plasma concentrations are rare, tolerable, and generally self-limiting. Common adverse effects include nausea, vomiting, anorexia, dizziness, and lethargy. Less than 5% of circulating drug is bound to proteins in the plasma. Ethosuximide is metabolized in the liver; however, approximately 20% is excreted through renal filtration. The half-life for ethosuximide is 40 to 60 hours. TDM for ethosuximide is performed to ensure that blood concentrations are in the therapeutic range.

Felbamate

Felbamate (Felbatol) is primarily indicated for use in severe epilepsies such as in children with the mixed seizure disorder, Lennox-Gastaut syndrome, and in adults with refractory epilepsy.[9–11] Felbamate is most commonly administered orally as it is nearly completely absorbed by the gastrointestinal tract, and peak blood concentrations can be achieved within 1 to 4 hours.[13] In circulation, approximately 30% of felbamate is bound to plasma proteins and has a half-life of 14 to 22 hours in adults. Felbamate is eliminated by renal and hepatic metabolism, so impairments of hepatic or renal function can significantly increase the half-life of the drug in circulation. Hepatic metabolism is enhanced by enzyme inducers such as phenobarbital, primidone, phenytoin, and carbamazepine and results in a decreased half-life.

TDM may be indicated due to a narrow therapeutic range as 25 to 60 µg/mL[13] but should only be considered after steady state has been reached. Documented adverse side effects of felbamate include fatal aplastic anemia and hepatic failure.[9,11]

Gabapentin

Gabapentin (Neurontin) is administered orally with a maximum bioavailability of 60%, which is reduced when antacids are administered concurrently. This drug may be administered as a monotherapy or in conjunction with other AEDs for patients suffering

from complex partial seizures with or without generalized seizures and for pain management in some scenarios.[9,11] Peak concentrations for gabapentin are achieved 2 to 3 hours following dosage. Gabapentin does not bind to plasma proteins and is not metabolized in the liver. It is eliminated unchanged by the kidneys and has a half-life of approximately 5 to 9 hours in patients with normal renal function.[13] Children require a higher dose than adults to maintain a comparable half-life as they eliminate the drug faster than adults. Due to exclusive renal clearance of this drug, impaired kidney function increases the half-life of the circulating drug in a linear manner.[13] Therapeutic concentrations are reported to be between 12 and 20 µg/mL although a wide range of serum concentrations have been reported in association with seizure control. Multiple daily doses may be the preferred dosage regimen to avoid adverse drug effects at abnormally high blood concentrations and breakthrough seizures at the trough concentrations.[9] Adverse effects associated with gabapentin toxicity are generally mild and may include fatigue, ataxia, dizziness, and weight gain. Gabapentin is the most frequently used AED in patients with liver disease and in treating partial-onset seizures in patients with acute intermittent porphyria.[9]

Lamotrigine

Lamotrigine (Lamictal) is used to treat patients with partial and generalized seizures.[9] It is orally administered and is rapidly and completely absorbed from the gastrointestinal tract, reaching peak concentrations 3 hours after administration. Once in circulation, approximately 55% of lamotrigine is protein-bound and biologically inactive.[13] Hepatic metabolism accounts for the majority of elimination, and in patients undergoing monotherapy, its half-life is 15 to 30 hours.[11,13] The rate of elimination for lamotrigine is highly dependent on patient age and physiologic condition. Younger infants tend to metabolize this drug slower than older infants, and children metabolize lamotrigine twice as quickly as adults. Marked increases in clearance occur during pregnancy, peaking at 32 weeks of gestation.[9]

Lamotrigine clearance is increased by enzyme-inducing AEDs such as phenobarbital, primidone, phenytoin, and carbamazepine; however, valproic acid is an inhibitor of lamotrigine metabolism and may increase its half-life to 60 hours. Because of these drug–drug interactions, TDM is essential to maintain therapeutic drug concentrations. Although individual therapeutic ranges may vary, a concentration range of 2.5 to 15 µg/mL has been noted as efficacious, and an increasing concentration seems to correlate well with increased risk of toxicity.[13] A small percentage of patients taking lamotrigine have developed a rash. Other adverse effects associated with toxicity include neurological effects, such as dizziness, and gastrointestinal disturbances.

Levetiracetam

Levetiracetam (Keppra) is an orally administered AED that does not bind to plasma proteins, so it is almost completely bioavailable and reaches peak concentration in 1 hour after dosing.[9,11,13] Levetiracetam is used in partial and generalized seizures. Sixty-five percent of levetiracetam is excreted unchanged by the kidneys, and it has a half-life of 6 to 8 hours, although the rate of elimination is increased in children and pregnant females and decreased in the elderly. The rate of clearance for this drug correlates well with the GFR, which may be of use in patients with renal impairment. The need for TDM of levetiracetam is not as pronounced as for other AEDs due to its lack of pharmacokinetic variability but may be useful in monitoring adherence and fluctuating concentrations during pregnancy.[9] Therapeutic concentrations are 12.0 to 35.0 µg/mL. Adverse effects are minimal but include dizziness and weakness.

Oxcarbazepine

Oxcarbazepine (Trileptal) is an orally administered prodrug that is almost immediately metabolized to licarbazepine.[11,13] It is indicated for treatment of partial seizures and secondarily in generalized tonic–clonic seizures. In circulation, almost 40% is bound to plasma proteins, and peak concentrations are achieved about 8 hours after dosing. It is metabolized by the liver into two active enantiomeric forms: (R)-licarbazepine and (S)-licarbazepine, which are metabolized further by glucuronide conjugation.[11,13] In adults, the half-life of this drug is 8 to 10 hours. Children have a higher clearance rate of oxcarbazepine and need a higher dosing regimen to obtain the optimal blood concentration compared with adults. In the elderly population, the drug clearance is reduced by 30%, so a lower dosage regimen is needed to maintain therapeutic concentrations.[13] Both metabolism and clearance of the drug are reduced in patients with marked renal dysfunction, and appropriate dosage adjustments must be made to avoid toxicity. The metabolism of licarbazepine is sensitive to enzyme inducers such as phenytoin and phenobarbital, which may decrease the blood concentration by

20% to 40%.[13] TDM may be indicated when therapeutic benefits are not being met, when drug–drug interactions are suspected, or during pregnancy. Although not well defined, therapeutic effects of licarbazepine have been reported at serum concentrations of 12 to 35 µg/mL. Its adverse effects are similar to those of carbamazepine.

Tiagabine

Tiagabine (Gabitril) is used in the treatment of partial seizures.[11] Gastrointestinal absorption of tiagabine is rapid and nearly complete. Its peak concentrations are in 1 to 2 hours after dosing. Approximately 96% of circulating tiagabine is protein-bound, and its half-life is variable but in the range of 4 to 13 hours. Due to its significant protein binding, the ratio of free to bound drug is affected by other protein-binding drugs such as valproic acid, naproxen, and salicylates and by pregnancy.[13] It is highly metabolized by the hepatic MFO pathway,[9,13] so hepatic dysfunction can increase the half-life of the drug. TDM may be indicated due to intraindividual and interindividual variations. Therapeutic benefits of the drug have been observed at concentrations of 20 to 100 ng/mL.[13] Adverse side effects on central nervous system (CNS) have been noted including confusion, difficulty in speaking clearly (stuttering), mild sedation, and a tingling sensation in the body's extremities, or paresthesia, especially in the hands and fingers.[9,14]

Topiramate

Topiramate (Topamax) is used in the treatment of partial and generalized seizures.[9] Only 15% of topiramate is bound to plasma proteins and is almost completely bioavailable after oral administration. Peak concentrations are achieved within 1 to 4 hours after dosing.[11,13] The half-life of topiramate is 20 to 30 hours, and the majority of this drug is eliminated by renal filtration, although some is eliminated by hepatic metabolism. The dose-to-blood concentration ratio in children is less than that of adults such that children require a higher dose to maintain plasma topiramate concentrations compared to adults.[13] Blood concentrations are increased secondary to renal insufficiency but may be decreased when used with other enzyme-inducing AEDs. The therapeutic range for topiramate is based on collection of trough serum specimens and is reported to be less than 25 mg/L. Adverse CNS side effects of topiramate include change of taste with particular foods (e.g., diet soda and beer) and a sensation of "pins and needles" in the extremities.[9,15] TDM may be indicated when a steady state is reached to provide an effective individual baseline concentration and may also be employed when therapeutic benefits are not achieved or drug–drug interactions need to be monitored.[13]

Zonisamide

Zonisamide (Zonegran) is an anticonvulsant used in adjunctive therapy for partial and generalized seizures. This drug is administered orally and absorbed from the gastrointestinal tract on the order of 65% or higher. Peak blood concentrations are reached 4 to 7 hours after dosing, and approximately 60% of the drug is bound to plasma proteins and accumulates extensively in erthyrocytes.[13] The majority of zonisamide is metabolized by the liver via acetylation, and oxidation followed by glucuronide conjugation and then renal excretion. The half-life of zonisamide is 50 to 70 hours in patients receiving monotherapy and may be reduced to 25 to 35 hours when other enzyme-inducing AEDs are administered concomitantly.[13] Children require higher doses to achieve therapeutic blood concentrations compared to those of adults.[13] Clinicians treating patients with liver or kidney disease should exercise caution as blood zonisamide concentrations may increase proportionally to the level and type of organ impairments. The therapeutic range of zonisamide in serum has been reported as 10 to 40 µg/mL but may vary among patients.[13] Symptoms of zonisamide toxicity include difficulty breathing, low blood pressure, slow heart rate, and possible loss of consciousness. TDM may be indicated to establish a baseline level after a steady state has been achieved, to detect drug–drug interactions, or at therapeutic failure.

Psychoactive Drugs

Lithium

Lithium is a mood-altering drug primarily used in the treatment of bipolar disorder, recurrent depression, and aggressive or self-mutilating behavior, though it may also be used as a preventative treatment for migraines and cluster headaches. Its gastrointestinal absorption is complete and rapid, so lithium is administered orally, with peak blood concentrations reached 2 to 4 hours after a dose is administered. Lithium is a cationic metal that does not bind to proteins. Lithium has a half-life of 10 to 35 hours and is eliminated predominately by renal filtration. It is subject to reabsorption in the renal tubules. Compromises in renal function

usually result in lithium accumulation. Correlations between serum lithium concentrations and therapeutic responses have not been well established. Its therapeutic range is 0.5 to 1.2 mmol/L in a large portion of the patient population.[16] The purpose of TDM for lithium is to avoid toxic effects associated with high blood concentrations. Lithium at 1.5 to 2 mmol/L in serum may cause apathy, lethargy, speech difficulties, and muscle weakness.[16] The serum concentrations of lithium greater than 2 mmol/L are associated with renal impairment, hypothyroidism, and CNS disturbances such as muscle rigidity, seizures, and possible coma. Determination of serum lithium concentrations is commonly measured by colorimetric methods, which have replaced older methods such as flame photometry, ion-selective electrodes, and atomic absorption photometry. Test tubes that contain lithium anticoagulants must be avoided to prevent falsely increased specimen results.

Tricyclic Antidepressants

Tricyclic antidepressants (TCAs) are a class of drugs used to treat depression, insomnia, extreme apathy, and loss of libido. Among all TCAs, imipramine, amitriptyline, and doxepin are most commonly monitored in clinical laboratories.[16] Desipramine and nortriptyline are active metabolites of imipramine and amitriptyline, respectively. The TCAs are orally administered and demonstrate a varying degree of absorption. In many patients, they slow gastric emptying and intestinal motility, which significantly slows the rate of absorption. As a result, peak concentrations are reached in the range of 2 to 12 hours. Approximately 85% to 95% of TCAs are protein-bound. For most TCAs, therapeutic effects are not seen for the first 2 to 4 weeks after initiation of therapy. TCAs are eliminated by hepatic metabolism, and many of their metabolic products have therapeutic actions. Their rates of metabolism are influenced by a wide variety of factors. As a result, the half-life of TCAs varies considerably among patients (17 to 40 hours). The rate of elimination can also be influenced by coadministration of other drugs that are eliminated by hepatic metabolism. The toxicities of TCAs are dose dependent; serum concentrations twice the upper limit of the therapeutic range can lead to drowsiness, constipation, blurred vision, and memory loss. Even higher levels may cause seizure, cardiac arrhythmia, and unconsciousness.

Because of the high variability in half-life and absorption, blood concentrations of the TCAs should not be evaluated until a steady state has been achieved. TDM is indicated to determine the therapeutic efficacy and potential toxicity. Many of the immunoassays for TCAs use polyclonal antibodies, which cross-react among the different TCAs and their metabolites. These immunoassays are used for TCA screening rather than TDM, and the results are reported out as "total tricyclics."[16] Some immunoassays employ an extraction step to separate the parent drugs from their metabolites prior to analysis, and interpretation of these results requires an in-depth understanding of these assays. Chromatographic methods provide simultaneous evaluation of both the parent drugs and their metabolites, which provides a basis for unambiguous interpretation of results.[16]

Clozapine

Clozapine (Clozaril, FazaClo) is an antipsychotic used in the treatment of otherwise treatment-refractory schizophrenia. Absorption is rapid and almost complete, and approximately 97% of circulating drug is bound to plasma proteins. Peak concentrations are achieved within 2 hours of administration. Clozapine is metabolized in the liver and has a half-life of 8 to 16 hours. Research has shown that beneficial effects of clozapine have been demonstrated at 350 to 420 ng/mL. TDM may be indicated to check for adherence and in patients with altered pharmacokinetics. TDM may also be used to avoid toxicity, which can result in seizures.[17]

Olanzapine

Olanzapine (Zyprexa) is a thienobenzodiazepine derivative that effectively treats schizophrenia, acute manic episodes, and the recurrence of bipolar disorders.[17] It can be administered as a fast-acting intramuscular (IM) injection, however, it is more commonly administered orally. The drug is absorbed well in the gastrointestinal tract; however, an estimated 40% is inactivated by first-pass metabolism. Peak concentrations are reached 5 to 8 hours after dosing, and approximately 93% is bound to plasma proteins in circulation. Olanzapine is metabolized in the liver and has a variable half-life of 21 to 54 hours. Women and nonsmokers tend to have lower clearance and thus higher blood concentrations of olanzapine compared with men and smokers.[17] There is indication that the blood concentrations of olanzapine correlate well with clinical outcomes and that TDM may help to optimize clinical responses while balancing the occurrence of adverse effects in a therapeutic range of 20 to 50 ng/mL.[17] Adverse effects of olanzapine include tachycardia, decreased consciousness, and possible coma.

Immunosuppressive Drugs

Transplantation medicine is a rapidly emerging discipline within clinical medicine. The clinical laboratory plays many important roles in transplantation programs. Among these responsibilities, monitoring of immunosuppressive drugs used to prevent organ rejection is a key concern. Most immunosuppressive drugs require establishment of individual dosage regimens to optimize therapeutic outcomes and minimize toxicity.

Cyclosporine

Cyclosporine (Gengraf, Neoral, Sandimmune) is a cyclic polypeptide that has a potent immunosup pressive activity. Its primary clinical use is suppression of host-versus-graft rejection of heterotopic transplanted organs. It is administered as an oral preparation with an absorption rate of 5% to 50% and peak concentrations within 1 to 6 hours. Because of this high variability, the relationship between oral dosage and blood concentration is poor; therefore, TDM is an important part of establishing a dosage regimen. More than 98% of circulating cyclosporine is protein-bound, and cyclosporine appears to sequester in cells, including erythrocytes.[18] Erythrocyte content is highly temperature dependent; therefore, evaluation of blood cyclosporine concentrations requires rigorous control of specimen temperatures. Therefore, whole blood specimens are used to avoid this preanalytical variable. The pharmacodynamics of cyclosporine have been well established. Cyclosporine is eliminated by hepatic metabolism and has a half-life of approximately 12 hours. Whole blood trough concentrations range from 100 to 400 ng/mL. Target steady-state trough concentrations vary depending on the type of organ transplant, level of immunosuppression, and time after transplantation. Results should be interpreted in conjunction with clinical information and any signs or symptoms of toxicity or organ rejection. Cyclosporine results above the upper limit have been associated with cyclosporine toxicity. The adverse effects of cyclosporine are primarily renal tubular and glomerular dysfunction, which may result in hypertension.[19] Several immunoassays are available for the determination of whole blood cyclosporine concentrations, although many show cross-reactivity with inactive metabolites. Chromatographic methods are available and provide separation and quantitation of the parent drug from its metabolites.[20]

Tacrolimus

Tacrolimus FK-506 (Astagraf, Envarsus, Hecoria, Prograf) is an orally administered immunosuppressive drug that is 100 times more potent than cyclosporine.[21] Early use of tacrolimus suggested a low degree of toxicity compared to cyclosporine at therapeutic concentrations; however, after extensive uses in clinical practice, both drugs appear to have comparable degrees of nephrotoxicity. Tacrolimus has been associated with thrombus formation at concentrations above its therapeutic range (5–15 ng/mL).

Many aspects of tacrolimus pharmacokinetics are similar to those of cyclosporine. Gastrointestinal uptake is highly variable, with peak blood concentrations achieved in 1 to 3 hours. More than 98% of circulating tacrolimus is bound to proteins in the plasma. Tacrolimus has a half-life of 10 to 12 hours; it is eliminated almost exclusively by hepatic metabolism, and its metabolites are primarily secreted into bile for excretion. Increases in immunoreactive tacrolimus may be seen in cholestasis as a result of cross-reactivity with these metabolites. Because of the high potency of tacrolimus, circulating therapeutic concentrations are low. This limits the methodologies capable of measuring whole blood concentrations. Currently, the most common method is high-performance liquid chromatography-tandem mass spectrometry; however, several immunoassays are also available. As with cyclosporine, therapeutic ranges for tacrolimus are dependent on the transplant organ and time from transplantation. The blood concentrations correlate well with therapeutic and toxic effects, and whole blood is the preferred specimen for tacrolimus TDM. Some of the adverse effects associated with tacrolimus toxicity include anemia, leukopenia, thrombocytopenia, and hyperlipidemia.

Sirolimus

Sirolimus (Rapamune) is an antifungal agent with immunosuppressive activity that is used to prevent graft rejection in patients receiving a kidney transplant. Sirolimus is rapidly absorbed after oral administration, and peak blood concentrations are achieved 1 to 2 hours after dosing. To increase the therapeutic efficacy, sirolimus is commonly coadministered with cyclosporine or tacrolimus as the bioavailability of sirolimus is 15% when taken in conjunction with cyclosporine.[20,22,23] Sirolimus has a long half-life of 62 hours and is predominantly metabolized in the liver. Blood concentrations are extensively affected by the first-pass metabolism and also by individual

differences in absorption, distribution, metabolism, and excretion, demonstrating the need for TDM. This drug is also extremely potent and requires TDM due to its inherent toxicity. Adverse effects associated with toxicity include thrombocytopenia, anemia, leukopenia, infections, and hyperlipidemia.[22,23] Sirolimus binds more readily to lipoproteins than plasma proteins, making whole blood the ideal specimen for analysis.[23] Approximately 92% of circulating sirolimus is bound. Initial TDM is performed using a trough specimen drawn during its steady state. Subsequent monitoring is performed by collecting trough specimens on a weekly basis for the first month, followed by a biweekly sampling pattern in the second month. These specimens are analyzed and used to establish a safe and effective therapeutic range. A therapeutic range of 4 to 12 µg/L is used when sirolimus is administered in conjunction with cyclosporine, and a range of 12 to 20 µg/L is used if cyclosporine therapy is not used or discontinued.[23] Sirolimus concentrations can be measured using chromatographic methods or by high-performance liquid chromatography-tandem mass spectrometry.[20,22]

Everolimus

Everolimus is an immunosuppressive agent derived from sirolimus (rapamycin). Everolimus has a shorter half-life than sirolimus, which allows for more rapid achievement of steady-state pharmacokinetics. Everolimus is metabolized by CYP3A4; therefore, inducers or inhibitors of that enzyme may require dose adjustments. The most common adverse side effects include hyperlipidemia, thrombocytopenia, and nephrotoxicity. Everolimus is FDA approved for prophylaxis of graft rejection in solid organ transplant. Measuring its blood drug concentrations is common practice for its use in organ transplant programs. Therapeutic targets vary depending on the organ transplant; heart and kidney transplant guidelines target trough blood concentrations between 3 and 8 ng/mL. Whole blood samples are measured by liquid chromatography-tandem mass spectrometry.

Mycophenolic Acid

Mycophenolate mofetil is a prodrug that is rapidly converted in the liver to its active form, mycophenolic acid (MPA).[19] MPA is a lymphocyte proliferation inhibitor that is used most commonly as supplemental therapy with cyclosporine and tacrolimus in renal transplant patients.[13] As with the other anti-rejection drugs, low trough concentrations of MPA increase the risk of acute rejection, while high concentrations imply toxicity. MPA is administered orally and absorbed under neutral pH conditions in the intestine.[24] Interindividual variation of gastrointestinal tract physiology influences the absorption of MPA; however, peak concentrations are generally achieved 1 to 2 hours after dosing. Once in circulation, MPA is 95% protein-bound. The degree to which MPA is protein-bound varies both intra- and interindividually and is dependent on circulating albumin concentrations, renal function, and the concentration of other drugs that may competitively bind to plasma albumin.[24] MPA is primarily eliminated by renal excretion (>90%) and has a half-life of approximately 17 hours. The therapeutic range for MPA is reported to be 1 to 3.5 µg/mL; increased levels that lead to toxicity may cause nausea, vomiting, diarrhea, and abdominal pain. Serum concentrations of MPA and its metabolites can be assayed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) or, more commonly, immunoassay, though immunoassays are generally considered less specific.[24] As with most immunoassay methods, cross-reactivity between MPA and its active metabolite (AcMPAG) should be taken into account along with the clinical picture when evaluating a dosage regimen.

Antineoplastics

Assessment of the therapeutic benefit and toxicity of most antineoplastic drugs is not aided by TDM because their pharmacodynamics are hard to establish. Many of these agents are rapidly metabolized or incorporated into cellular macromolecular structures within seconds to minutes of their administration. In addition, the therapeutic ranges for many of these drugs include concentrations associated with toxic effects. Considering that most antineoplastic agents are administered intravenously as a single bolus, the actual delivered dose is more relevant than circulating concentrations.

Methotrexate

Methotrexate (Otrexup, Rasuvo) is one of the few antineoplastic drugs in which TDM may offer some benefits to determine a therapeutic regimen. High-dose methotrexate followed by leucovorin rescue has been shown to be an effective therapy for various neoplastic conditions.[25] This therapy involves the relative rate of mitosis of normal versus neoplastic

cells. In general, neoplastic cells divide more rapidly than normal cells, and methotrexate inhibits DNA synthesis in all cells. Neoplastic cells, which rapidly divide, have a higher requirement for DNA and are susceptible to depravation of this essential constituent before normal cells. The efficacy of methotrexate therapy is dependent on a controlled period of inhibition, one that is selectively detrimental to neoplastic cells. This is accomplished by administration of leucovorin, which reverses the action of methotrexate at a specific time after methotrexate infusion. This is referred to as *leucovorin rescue*. Failure to stop the action of methotrexate results in cytotoxic effects to most cells. Evaluation of serum methotrexate concentrations, after the inhibitory time period has passed, is used to determine how much leucovorin is needed to counteract many of the toxic effects of methotrexate.[25] Serum concentrations of methotrexate are commonly monitored during high-dose therapy to identify the time at which active intervention by leucovorin rescue should be initiated. Criteria for serum concentrations indicative of a potential for toxicity after single-bolus, high-dose therapy are as follows: Methotrexate >10 µmol/L 24 hours after dose; >1 µmol/L 48 hours after dose; >0.1 µmol/L 72 hours after dose.

Methotrexate is administered orally with peak blood concentrations 1 hour after dosing. Approximately 50% of methotrexate is bound to plasma proteins in circulation. Methotrexate has a half-life of 5 to 9 hours and is predominantly excreted through the renal system. Trough serum specimens are preferred for determination of methotrexate concentrations. The therapeutic range for trough specimens is less than 1 µmol/L 48 hours after dose.

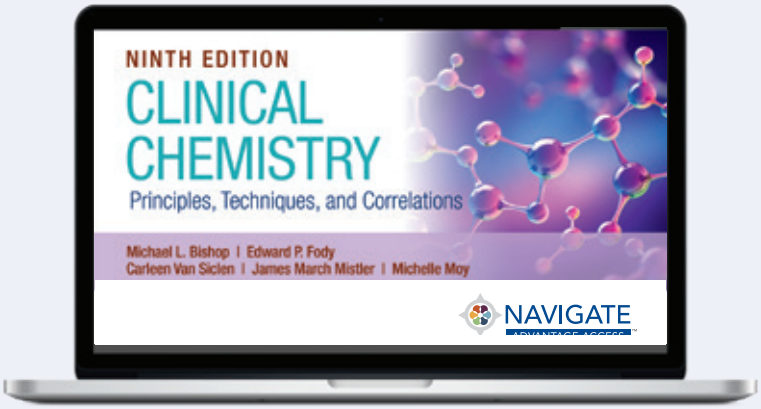
Bronchodilators

Theophylline

Theophylline (Theo-Dur, Theo-24, Uniphyl) is used in the treatment of respiratory disorders, such as asthma and stable chronic obstructive pulmonary disease, for patients that have difficulty using an inhaler or those with nocturnal symptoms. Absorption can be variable, with peak blood concentrations achieved 1 to 2 hours after dosing when a rapid-release formation is administered or within 4 to 8 hours for a modified-release preparation. Approximately 50% to 65% of circulating drug is bound to plasma proteins, primarily albumin. Theophylline has a half-life of 3 to 8 hours. It is predominantly metabolized in the liver; however, about 20% is eliminated through the renal system. Beneficial effects have been demonstrated at 10 to 20 µg/L. Though infrequent, concentrations above 20 µg/L may lead to serious adverse effects including insomnia, tachycardia, seizures, arrhythmias, and possible cardiorespiratory arrest. There is a poor correlation between dosage and blood concentrations; however, TDM may initially be useful in optimizing the dosage or the confirmation of toxicity when suspected.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Lynch T, Price A. The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am Fam Physician.* 2007;76(3)391–396.
2. Goldberger AL. Cardiac arrhythmias due to digoxin toxicity. In Traub SJ, Parikh N, (eds). *UpToDate*. Waltham, MA: UpToDate; 2021.
3. Tuzimski T, Petruczynik A. Review of chromatographic methods coupled with modern detection techniques applied in the therapeutic drugs monitoring (TDM). *Molecules*. 2020;25(17):4026. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7504794
4. Jurgens G, Graudal NA, Kampmann JP. Therapeutic drug monitoring of antiarrhythmic drugs. *Clin Pharmacokinet.* 2003;42:647–663.
5. Datta P, Hinz V, Klee G. Comparison of four digoxin immunoassays with respect to interference from digoxin-like immunoreactive factors. *Clin Biochem.* 1996;29(6):541–547.
6. McDermott JH, Wolf J, Hoshitsuki K, et al. Clinical Pharmacogenetics Implementation Consortium Guideline for the use of aminoglycosides based on MT-RNR1 genotype. *Clin Pharmacol Ther*. 2021 May 25. doi: 10.1002/cpt.2309. Online ahead of print.
7. Winston L, Benowitz N. Once-daily dosing of aminoglycosides: how much monitoring is truly required? *Am J Med.* 2003;114:239–240.
8. Martel TJ, Jamil RT, King KC. Red man syndrome. *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021.
9. Perucca E. An introduction to antiepileptic drugs. *Epilepsia.* 2005;46:31–37.
10. Eadie MJ. Therapeutic drug monitoring—antiepileptic drugs. *Br J Clin Pharmacol.* 2001;52:11S–19S.
11. Israni RK, Kasbekar N, Haynes K, et al. Use of antiepileptic drugs in patients with kidney disease. *Semin Dial.* 2006;19:408–416.
12. Al Khalili Y, Sekhon S, Jain S. Carbamazepine toxicity. In *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2021.
13. Johannessen SI, Tomson T. Pharmacokinetic variability of newer antiepileptic drugs. *Clin Pharmacokinet.* 2006;45: 1061–1075.
14. Bauer J, Cooper-Mahkorn C. Tiagabine: Efficacy and safety in partial seizures – current status. *Neuropsychiatr Dis Treat*. 2008;4(4):731–736.
15. Marmura MJ. Safety of topiramate for treating migraines. *Expert Opin Drug Saf*. 2014;13(9):1241–1247.
16. Eap CB, Grunder G, Baumann P, et al. Tools for optimising pharmacotherapy in psychiatry (therapeutic drug monitoring, molecular brain imaging and pharmacogenetic tests): focus on antidepressants. *World J Biol Psychiatr*. 2021 May 12; 1–68. doi: 10.1080/15622975.2021.1878427
17. Mauri MC, Volonteri LS, Colasanti A, et al. Clinical pharmacokinetics of atypical antipsychotics. *Clin Pharmacokinet.* 2007;46:359–388.
18. Masuda S, Inui K. An update review on individualized dosage of calcineurin inhibitors in organ transplant recipients. *Pharmacol Ther.* 2006;112:184–198.
19. Taylor AL, Watson CJE, Bradley JA. Immunosuppressive agents in solid organ transplantation: mechanisms of action and therapeutic efficacy. *Crit Rev Oncol Hematol.* 2005;56: 23–46.
20. Christians U, Schmitz V, Schoning W, et al. Toxicodynamic therapeutic drug monitoring of immunosuppressants: promises, reality, and challenges. *Ther Drug Monit*. 2008 Apr;30(2):151–158.
21. Baraldo M, Furlanut M. Chronopharmacokinetics of cyclosporine and tacrolimus. *Clin Pharmacokinet.* 2006;45: 775–788.
22. Antunes NJ, Kipper K, Couchman L, et al. Simultaneous quantification of cyclosporin, tacrolimus, sirolimus and everolimus in whole blood by UHPLC-MS/MS for therapeutic drug monitoring. *Biomed Chromatogr*. 2021 Jun; 35(6):e5071.
23. Stenton SB, Partovi N, Ensom MHH. Sirolimus: the evidence for clinical pharmacokinetic monitoring. *Clin Pharmacokinet.* 2005;44:769–786.
24. Elbarbry FA, Shoker AS. Therapeutic drug measurement of mycophenolic acid derivatives in transplant patients. *Clin Biochem.* 2007;40:752–764.
25. Baldo P, Fornasier G, Ciolfi L, et al. Pharmacovigilance in oncology. *Int J Clin Pharm*. 2018 Aug;40(4):832–841.

CHAPTER 26

Toxicology

Christopher R. Swartz

CHAPTER OUTLINE

KEY TERMS

Bioaccumulation
Body burden
Dose–response relationship
Drugs of abuse (DOA)
ED_{50}
LD_{50}
Poison
Quantal dose–response relationship
TD_{50}
Therapeutic index
Toxicant
Toxicokinetics
Toxicology
Toxin
Xenobiotic

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Identify the routes for toxin exposure and factors influencing absorption including dose.
- Distinguish between acute and chronic toxicity.
- List the major toxicants discussed.
- Explain the pathologic mechanisms of the major toxicants.
- Compare and contrast specimen types, including advantages and disadvantages of each.
- Discuss specimen collection, handling, and processing for toxicology testing.
- Differentiate between quantitative and qualitative tests in toxicology.

- Select common qualitative and quantitative toxicology test methods.
- Evaluate toxicity in the clinical laboratory, given patient test results.
- Calculate osmolal gap and correlate with the presence of osmotically active substances.
- Specify the role of the laboratorian in the evaluation of exposure to toxins and poisons.
- Correlate patient results to suspected poisoning cases and recommend additional testing.

In 2019, there were over 2 million **poison** exposures reported in the United States.[1] Many of these exposures were caused by different poisons or toxic substances found in various household items, assorted prescription or illicit drugs, heavy metals, and other miscellaneous materials. The highest incidence (number of poison exposures reported per 100,000 population) in the United States occurred in young children (5 years old or less), followed by teenagers and adults.[1] The ability of the laboratorian to assist health care providers through identification and quantification of a specific toxic substance is of critical importance to patient care, especially in the case of acute poison exposures.

This chapter seeks to provide a broad overview of toxicology, with a particular emphasis on applications in clinical toxicology. The laboratorian will associate common toxins or poisons with physiological effects and pathological mechanisms caused by acute or chronic exposure, along with appropriate medical treatments. Suitable specimens for toxicological testing will be described, and proper collection and handling will be discussed. The laboratorian will be introduced to screening and confirmatory testing methodologies for common toxicological analytes.

Toxicology is the study of the adverse effects of xenobiotics in humans. **Xenobiotics** are chemicals and drugs that are not normally found in or produced by the body. The scope of toxicology is very broad and includes three major disciplines: mechanistic, descriptive, and regulatory toxicology. *Mechanistic toxicology* elucidates the cellular, molecular, and biochemical effects of xenobiotics within the context of a dose–response relationship between the xenobiotic and its adverse effect(s). Mechanistic studies provide a basis for rational therapy design and the development of laboratory tests to assess the degree of exposure in individuals. *Descriptive toxicology* uses the results from animal experiments to predict what level of exposure will cause harm in humans. This process is known as *risk assessment*. In *regulatory toxicology*, combined data from mechanistic and descriptive studies are used to establish standards that define the level of exposure that will not pose a risk to public health or safety. Typically, regulatory toxicologists work for, or in conjunction with, government agencies.

There are also a number of specialties within toxicology, including forensic, clinical, and environmental toxicology. *Forensic toxicology* is primarily concerned with the medical and legal consequences of exposure to chemicals or drugs. A major focus of forensic toxicology is establishing and validating the analytic performance of test methods used to generate evidence in legal situations, including cause of death. *Clinical toxicology* focuses on the relationships between xenobiotics and disease states. This area emphasizes not only diagnostic testing but also therapeutic intervention. *Environmental toxicology* includes the evaluation of

CASE STUDY 26.1, PART 1

Emily, a 45-year-old female, has become socially withdrawn. Emily has also been apathetic and suffers from insomnia. Emily's mother is worried about her and made an appointment for Emily to see her primary care physician. After reviewing Emily's recent history, her physician suspects that Emily has major depressive disorder and refers Emily to a psychiatrist. Her physician also ordered laboratory tests after her appointment.

environmental chemical pollutants and their impact on human health.

Within the organizational structure of a typical clinical laboratory, toxicology is usually considered a specialty of clinical chemistry because the qualitative and quantitative methodologies used to measure xenobiotics overlap with this discipline. However, appropriate diagnosis and management of patients with acute poisoning or chronic exposure to xenobiotics requires an integrated approach from all sections of the clinical laboratory.[2,3]

Xenobiotics, Poisons, and Toxins

The terms *xenobiotic, poison*, and *toxin* are often used interchangeably; however, there are some important distinctions that should be made between them. Xenobiotics are defined as exogenous agents that can have an adverse effect on a living organism. This term is more often used to describe environmental exposure to chemicals or drugs. Examples of environmental drug exposures include antibiotics and antidepressants; chemical exposures might include perfluorinated and brominated compounds. Similarly, poisons are also exogenous agents that have an adverse effect on a biological system; however, this term is more often used when describing substances from an animal, plant, mineral, or gas. Examples include venoms from poisonous snakes or spiders, poison hemlock, arsenic, lead, and carbon monoxide. **Toxins**, however, are endogenous substances biologically synthesized either in living cells or in microorganisms. Examples include botulinum toxin produced from the microorganism *Clostridium botulinum*, hemotoxins produced from venomous snakes, and mycotoxins produced from fungus. The terms **toxicant** and *toxic* refer to substances that are not produced within a living cell or microorganism and are more commonly used to describe environmental chemicals.

From a clinical standpoint, almost 50% of poisoning cases are due to intentional suicide attempts; accidental exposure accounts for about 30%; and the remaining cases are a result of homicide or occupational exposure. Of these, suicide has the highest mortality rate. Accidental exposure occurs most frequently in children; however, accidental drug overdose of either therapeutic or illicit drugs is relatively common in adolescents and adults. Occupational exposure primarily occurs in industrial and agricultural settings but is an expanding area of concern as we learn more about the role of various chemical agents and their contribution to disease.

Routes of Exposure

Toxins can enter the body via several routes; however, ingestion, inhalation, and transdermal absorption are the most common. Ingestion is most often observed in the clinical setting. For most toxins to exert a systemic effect, they must be absorbed into circulation. Absorption of toxins from the gastrointestinal tract occurs via several mechanisms. Some toxins are taken up by processes intended for dietary nutrients; however, most are passively absorbed through diffusion. Diffusion requires that the substance be able to cross the cellular barriers of the gastrointestinal tract. Hydrophobic substances do have the ability to diffuse across cell membranes and therefore can be absorbed anywhere along the gastrointestinal tract. Ionized substances, however, cannot passively diffuse across the membranes. Weak acids can become protonated in gastric acid, resulting in a nonionized species, which can then be absorbed in the stomach. In a similar manner, weak bases can be absorbed in the intestine where the pH is largely neutral or slightly alkaline. Other factors that influence the absorbance of toxins from the gastrointestinal tract include the rate of dissolution, gastrointestinal motility, resistance to degradation in the gastrointestinal tract, and interaction with other substances. Toxins that are not absorbed from the gastrointestinal tract do not produce systemic effects but may produce local effects, such as diarrhea, bleeding, and malabsorption, which may cause systemic effects secondary to toxin exposure.

Dose–Response Relationship

The concept that all substances have the potential to cause harm, even water, is a central theme in toxicology. Paracelsus (1493–1591) pioneered the use of chemicals in medicine and coined the term "the dose makes the poison." Understanding this **dose–response relationship** is fundamental and essential to modern toxicology. To enable assessment of substances' potential to cause pathologic effects, it is necessary to establish an index of the relative toxicities of the substances. Several systems are available, but most correlate the dose of a xenobiotic that will result in harmful effects. One such system correlates a single, acute oral dose range with the probability of a lethal outcome in an average 70-kg male (154 lbs.) (**Table 26.1**). This is a useful system to compare the relative toxicities of substances,

Table 26.1 Toxicity Rating

Toxicity Rating	Lethal Oral Dose in Average Adult
Super toxic	<5 mg/kg
Extremely toxic	5–50 mg/kg
Very toxic	50–500 mg/kg
Moderately toxic	0.5–5 g/kg
Slightly toxic	5–15 g/kg
Practically nontoxic	>15 g/kg

Data from Klaassen CD. Principles of toxicology. In: Klaassen CD, Amdur MO, Doull J, eds. *Toxicology: the Basic Science of Poisons.* 3rd ed. New York, NY: Macmillan; 1986:13.

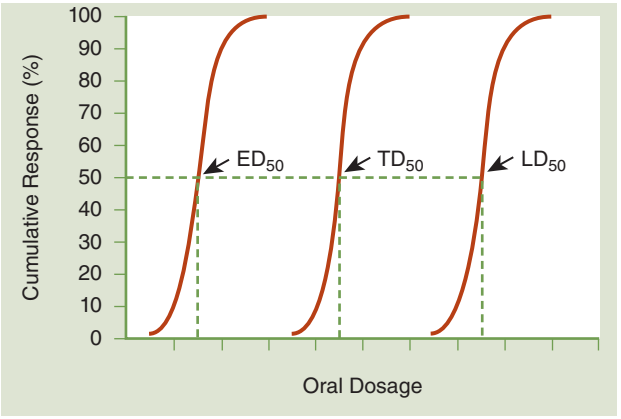


Figure 26.1 Dose–response relationship. Comparison of responses of a therapeutic drug over a range of doses. The ED_{50} is the dose of drug in which 50% of treated individuals will experience benefit. The TD_{50} is the dose of drug in which 50% of individuals will experience toxic adverse effects. The LD_{50} is the dose of drug in which 50% of individuals will result in mortality.

as the predicted response is death, which is a valid endpoint. However, most xenobiotics can express pathologic effects other than death at lower degrees of exposure; therefore, other indices have been developed.

A more in-depth characterization can be acquired by evaluating data from a cumulative frequency histogram of toxic responses over a range of doses. This experimental approach is typically used to evaluate responses over a wide range of concentrations. One response monitored is the toxic response or the response associated with an early pathologic effect at lower than lethal doses. This response has been determined to be an indicator of the toxic effects specific for that toxin. For a substance that exerts early toxic effects by damaging liver cells, the response monitored may be increases in serum alanine aminotransferase (ALT) or γ-glutamyltransferase (GGT) activity. The dose–response relationship implies that there will be an increase in the toxic response as the dose is increased. It should be noted that not all individuals display a toxic response at the same dose. The population variance can be seen in a cumulative frequency histogram of the percentage of people producing a toxic response over a range of concentrations (**Figure 26.1**). The **TD_{50}** (toxic dose) is the predicted dose that would produce a toxic response in 50% of the population. If the monitored response is death, the **LD_{50}** (lethal dose) is the predicted dose that would result in death in 50% of the population. Similar experiments can be used to evaluate the doses of therapeutic drugs. The **ED_{50}** (effective dose) is the dose that would be predicted to be effective or have a therapeutic benefit in 50% of the population. The **therapeutic index** is the ratio of the TD_{50} (or LD_{50}) to the ED_{50}. Drugs with a large therapeutic index demonstrate fewer toxic adverse effects when the dose of the drug is in the therapeutic range.

Dose–response relationships may apply to an individual or a population. The individual dose–response relationship relates to the individual's health status as well as changes in xenobiotic exposure levels. A **quantal dose–response relationship** describes the change in health effects of a defined population based on changes in the exposure to the xenobiotic.

Acute and Chronic Toxicity

Acute toxicity and *chronic toxicity* are terms used to relate the duration and frequency of exposure to observed toxic effects. Acute toxicity is usually associated with a single, short-term exposure to a substance in which the dose is sufficient to cause immediate toxic effects. Chronic toxicity is generally associated with repeated and frequent exposure for extended periods of time (months to years) at doses that are insufficient to cause an immediate acute response. In many instances, chronic exposure is related to an accumulation of the toxicant or the toxic effects within the individual. Chronic toxicity may affect different systems from those associated with acute toxicity; therefore, dose–response relationships may differ for acute and chronic exposures for the same xenobiotic. The initial management of patients thought to suffer from acute exposure to potentially toxic substances includes the ABCDE approach: Airway, Breathing, Circulation, Disability, and Exposure.[4]

Analysis of Toxic Agents

Toxicology testing may be performed to screen for the presence of a number of agents that might be present (e.g., drug screens, heavy metal panels) or as targeted testing. Targeted testing might be performed when an environmental risk of exposure is known (e.g., industrial workers, chemical plants), to support the investigation of an exposure (e.g., chemical spill, suicide attempt), to comply with occupational regulations or guidelines (e.g., Occupational Safety and Health Administration), or to confirm clinical suspicions of poisoning (e.g., arsenic, cyanide). Due to nonspecific signs and symptoms of toxicity and the fact that frequently the duration and extent of exposure are unknown, diagnosis of most toxic element exposures depends heavily on laboratory testing.

In general, clinical toxicology testing is performed on urine or blood specimens. Other suitable specimens for forensic toxicological analysis include plasma, serum, nails, hair, and oral fluid.[5] In selecting the best specimen for a specific test, it is important to recognize that toxic agents exhibit unique absorption, distribution, metabolism, and elimination kinetics, also known as **toxicokinetics**. As such, the predicted toxicokinetics of the individual element(s) being tested must be coordinated with the selection of specimen type and timing of collection relative to the time of exposure. An exposure could be missed entirely if testing is performed on an inappropriate specimen. For example, exposure to methylmercury could be missed if testing is performed on urine, as methylmercury is primarily excreted in fecal material. An exposure to arsenic could be missed if testing is performed with blood collected a few days after the exposure due to the short half-life of arsenic in blood.

Preanalytical variables such as elimination patterns, analyte stability, and specimen collection procedures must be considered. For urine testing, 24-hour collections are preferred in order to compensate for variable elimination patterns throughout the day. Reporting results per gram of creatinine is also common to account for variable excretion and renal function. Random urine collections may not provide the most accurate profile of exposure when compared to a 24-hour collection, but they can be useful for screening and qualitative detection of exposure to several potentially toxic agents. Any elevated result that is inconsistent with clinical expectations should be confirmed by testing a second specimen collection or a second specimen type.

One challenge of specimen collection for toxicological studies is that several aspects of the collection, handling, and storage can introduce external contamination into the sample. Common sources of external contamination include patient clothing, skin, hair, collection environment (e.g., dust, aerosols, antiseptic wipes), and specimen handling variables (e.g., container, lid, preservatives). Concentrated acids are commonly used as urine preservatives; however, contaminants may also be introduced in either the acid itself or in the process used to add the acid to the urine (e.g., pipette tips). Specimen containers and lids should also be devoid of contaminating organic and inorganic agents that may interfere with analytical testing. For example, certified "trace element–free" blood collection tubes are available; these tubes commonly have a royal blue top and can be used for most trace elements testing, although a tan-top tube is manufactured specifically for lead determinations. Another consideration when handling biological specimens for metals testing is the use of acid-washed pipette tips, containers, and other supplies to prevent contamination. Laboratories may also need to exercise precautions to prevent loss of toxic agents due to in vitro volatilization and metabolism. For instance, mercury and arsenic are particularly vulnerable to loss and metabolism, respectively, during sample processing and storage. These scenarios represent only a fraction of the many specimen handling considerations necessary to reduce preanalytical error in toxicology testing.

Analysis of toxic agents in a clinical setting is typically a two-step process. The first step is a *screening test*, which is a rapid, simple, qualitative procedure intended to detect the presence of specific substances or classes of toxicants. In general, these procedures have good analytic sensitivity but lack specificity.[6] A negative result can rule out the presence of a drug or toxicant; however, a positive result should be considered a presumptive positive until confirmed by a second, more specific method, which is the second step of the test process; this is also called a *confirmatory test*. Confirmatory tests are generally quantitative and report the concentration of the substance in the specimen, in contrast to qualitative screening tests that provide a result of positive (drug is present) or negative (drug is absent). A variety of analytical methods can be used for screening and confirmatory testing, although immunoassays are commonly used for most drug screens. In some instances, these assays are specific for a single specific drug (e.g., tetrahydrocannabinol [THC]), but in most cases, the assay is used to detect drugs within a general class (e.g., barbiturates and opioids). The reference method for quantitative identification of most organic compounds is gas

CASE STUDY 26.1, PART 2

Recall Emily. Her complete blood cell count was unremarkable except for an elevated erythrocyte mean cell volume. Results of the urinalysis were unremarkable. The serum chemistry testing revealed slightly increased aspartate aminotransferase (AST), elevated alanine aminotransferase (ALT), total bilirubin, and high-density lipoproteins (HDL). All other chemistry results, including glucose, urea, creatinine, cholesterol, sodium, potassium, and arterial blood gases were within the reference ranges.

1. Why did Emily's physician suspect that she had major depressive disorder?
2. What laboratory results are abnormal?
3. Besides major depressive disorder, what other medical conditions could Emily have, based on her laboratory results?

chromatography (GC) coupled with a mass spectrometer (GC-MS) as the detector. Liquid chromatography, in combination with tandem mass spectrometry (LC-MS/MS), has emerged as an important analytical technique for the quantification of various analytes in biological matrices and has the potential to replace immunoassays in screening procedures.[7] Inorganic compounds, including speciation, may be quantitated using inductively coupled plasma-mass spectrometry (ICP-MS) or atomic absorption (AA) methods.[8] Nuclear magnetic resonance (NMR) spectroscopy is widely used for the structural elucidation of organic and inorganic compounds and has found utility in the search for new biomarkers in the clinical laboratory.[9] Potential applications of NMR spectroscopy in the toxicology laboratory include recording NMR spectra of clinical or forensic specimens and comparing the spectra to known libraries to determine if a particular drug or metabolite may be present.[10] Please refer back to Chapter 4, *Analytic Techniques*, for more on these methodologies.

Toxicology of Specific Agents

Many chemical agents encountered on a regular basis have potential for toxicity. The focus of this section is to discuss some of the most commonly encountered nondrug toxins seen in a clinical setting, as well as those that present as medical emergencies with acute exposure.

Alcohols

The toxic effects of alcohol are both general and specific. Exposure to alcohol, like exposure to most volatile organic solvents, initially causes disorientation, confusion, and euphoria but can progress to unconsciousness, paralysis, and possibly death with high-level exposure. Most alcohols display these effects at about equivalent molar concentrations. This similarity suggests a common depressant effect on the central nervous system (CNS) that appears to be mediated by changes in membrane properties. In most cases, recovery from CNS effects is rapid and complete after cessation of exposure.

Distinct from the general CNS effects are the specific toxicities of each type of alcohol, which are usually mediated by biotransformation of alcohols to toxic products. There are several pathways by which short-chain aliphatic alcohols can be metabolized. Of these, hepatic conversion to an aldehyde, by alcohol dehydrogenase (ADH), and further conversion to an acid, by hepatic aldehyde dehydrogenase (ALDH), is the most significant.

$$\text{Alcohol} \xrightarrow{\text{ADH}} \text{Aldehyde} \xrightarrow{\text{ALDH}} \text{Acid} \quad (\textbf{Eq. 26.1})$$

Ethanol

Ethanol (ETOH) exposure is common, and excessive consumption, with its associated consequences, is a leading cause of economic, social, and medical problems throughout the world. Excessive drinking was estimated to cost the United States $249 billion in 2010.[11] Many social and family problems are associated with excessive ethanol consumption, and the burden to the health care system is significant. Ethanol-related disorders are consistently one of the top 10 causes of hospital admissions, and approximately 20% of all hospital admissions have some degree of alcohol-related problems. It is estimated that 80,000 Americans die each year, either directly or indirectly, as a result of excessive alcohol consumption. This correlates to about a fivefold increase in premature mortality. In addition, consumption of ethanol during pregnancy may lead to fetal alcohol

Table 26.2 Common Indicators of Ethanol Abuse

Test	Comments
GGT	Increases can be seen before the onset of pathologic consequences. Increases in serum activity can occur in many conditions unrelated to ethanol use.
AST	Increases in serum activity can occur in many conditions unrelated to ethanol use.
AST/ALT ratio	A ratio of >2.0 is highly specific for ethanol-related liver disease.
HDL	High serum HDL is specific for ethanol consumption.
MCV	Increased erythrocyte MCV is commonly seen with excessive ethanol consumption. Increases are not related to folate or vitamin B_{12} deficiency.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, γ-glutamyltransferase; HDL, high-density lipoprotein; MCV, mean cell volume.

syndrome or fetal alcohol effects, both of which are associated with delayed motor and mental development in children.

Correlations have been established between blood alcohol concentration and the clinical signs and symptoms of acute intoxication. A blood alcohol concentration of 80 mg/dL has been established as the statutory limit for operation of a motor vehicle in the United States, as this concentration is associated with a diminution of judgment and motor function. Determinations of blood ethanol concentration by the laboratory may be used in litigation of drunken driving cases and require appropriate chain-of-custody procedures for specimen collection and documentation of acceptable quality control performance, instrument maintenance procedures, and proficiency testing records. Approximately 50% of the 40,000 to 50,000 annual automobile-related fatalities in the United States involve alcohol as a factor.

Besides the short-term effects of ethanol, most pathophysiologic consequences of ethanol abuse are associated with chronic consumption. Chronic consumption has been associated with compromised function of various organs, tissues, and cell types; however, the liver appears to be affected the most. The pathologic sequence starts with the accumulation of lipids in hepatocytes. With continued consumption, this may progress to alcoholic hepatitis. In about 20% of individuals with long-term, high-level alcohol intake, this develops into a toxic form of hepatitis. Of those who do not progress to toxic hepatitis, progression to liver cirrhosis is common. Cirrhosis of the liver can be characterized as fibrosis leading to functional loss of the hepatocytes. Progress through this sequence is associated with changes in many laboratory tests related to hepatic function, including liver enzymes. Several laboratory indicators have the required diagnostic sensitivity and specificity to identify excessive ethanol consumption, and most correlate well to the progression of ethanol-induced liver disease. **Table 26.2** lists common laboratory indicators of prolonged ethanol consumption.

Several mechanisms have been proposed to mediate the pathologic effects of long-term ethanol consumption. Of these, adduct formation with acetaldehyde appears to play a key role. Hepatic metabolism of ethanol is a two-step enzymatic reaction with acetaldehyde as a reactive intermediate. Most ethanol is converted to acetate, or acetic acid, in this pathway; however, a significant portion of the acetaldehyde intermediate is released in the free state.

$$\text{Ethanol} \rightarrow \text{Acetaldehyde} \rightarrow \text{Acetate} \\ \rightarrow \text{Acetaldehyde adducts} \qquad \textbf{(Eq. 26.2)}$$

The enzymatic method involving alcohol dehydrogenase for the measurement of ethanol mimics this hepatic metabolic pathway and is described later in this chapter.

Extracellular acetaldehyde is a transient species as a result of rapid adduct formation with amine groups of proteins. Many of the pathologic effects of ethanol have been correlated with the formation of these adducts, and formation of acetaldehyde adducts has also been shown to change the structure and function of various proteins.

Methanol

Methanol is a common laboratory solvent that is also found in many household cleaners. It may be ingested accidentally as a component of many commercial products or as a contaminant of homemade liquors. Methanol is initially metabolized by hepatic ADH to the intermediate formaldehyde. Formaldehyde is then

CASE STUDY 26.1, PART 3

Recall Emily. Her physician suspects ethanol abuse. However, she denies alcohol consumption. Subsequent testing revealed a serum γ-glutamyltransferase (GGT) three times the upper limit of the reference range. No ethanol was detected in Emily's serum, and screening tests for infectious forms of hepatitis were negative.

4. Are these results consistent with a patient who is consuming toxic quantities of ethanol?
5. What is the most common method for determination of serum alcohol levels?
6. What additional testing would you recommend to rule in/out ethanol abuse?

rapidly converted to formic acid by hepatic ALDH. The formation of formic acid causes severe metabolic acidosis, which can lead to tissue injury and possible death. Formic acid is also responsible for optic neuropathy that can lead to blindness.

Isopropanol

Isopropanol, also known as rubbing alcohol, is also commercially available. It is metabolized by hepatic ADH to acetone, which is its primary metabolic end product. Both isopropanol and acetone have CNS depressant effects similar to ethanol; however, acetone has a long half-life, and intoxication with isopropanol can therefore result in severe, acute-phase, ethanol-like symptoms that persist for an extended period.

Ethylene Glycol

Ethylene glycol (1,2-ethanediol) is a common component of hydraulic fluid and antifreeze. Ingestion is mainly seen in individuals with alcohol use and dependency issues, as well as by children because of its sweet taste. The immediate effects of ethylene glycol ingestion are similar to those of ethanol; however, metabolism by hepatic ADH and ALDH results in the formation of several toxic species including oxalic acid and glycolic acid, which result in severe metabolic acidosis. This is complicated by the rapid formation and deposition of calcium oxalate crystals in the renal tubules. With high levels of consumption, calcium oxalate crystal formation in the kidneys may result in renal tubular damage.

Determination of Alcohols

From a medicolegal perspective, determinations of blood ethanol concentrations must be accurate and precise. Serum, plasma, and whole blood are acceptable specimens, and correlations have been established between ethanol concentration in these specimens and impairment of psychomotor function. Because ethanol uniformly distributes in total body water, serum, which has greater water content than whole blood, has a higher concentration per unit volume. Because of this difference in distribution, most states have standardized the acceptable specimen types admissible as evidence, and some jurisdictions even mandate a specific method (often GC) be used for legal ethanol determination. When acquiring a specimen for ethanol determination, several preanalytical issues must be considered to ensure the integrity of the sample. One of these requirements is that the venipuncture site must only be cleaned with an alcohol-free disinfectant such as chlorhexidine gluconate or iodine antiseptics. Also, because of the volatile nature of short-chain aliphatic alcohols, specimens must be capped at all times to avoid evaporation. Sealed specimens can be refrigerated or stored at room temperature for up to 14 days without loss of ethanol. Nonsterile specimens or those intended to be stored for longer periods of time should be preserved with sodium fluoride to avoid increases in ethanol content resulting from contamination due to bacterial fermentation.

Several analytic methods can be used to determine the concentration of ethanol in serum. Among these, osmometric, chromatographic, and enzymatic methods are the most commonly used. When osmolality is measured by freezing point depression, increases in serum osmolality correlate well with increases in serum ethanol concentration. The degree of increase in osmolality due to the presence of ethanol is expressed as the difference between the measured and the calculated osmolality, otherwise referred to as the *osmolal gap*. It has been established that serum osmolality

increases by approximately 10 mOsm/kg for each 60 mg/dL increase in serum ethanol; therefore, the osmolal gap is useful for estimating the amount of ethanol present in the serum. The osmolal gap can be calculated as follows:

$$\text{Osmolal gap} = \text{measured osmolality} - \text{calculated osmolality} \quad \text{(Eq. 26.3)}$$

This relationship, however, is not specific to ethanol, and increases in the osmolal gap also occur with other metabolic imbalances. Therefore, use of the osmolal gap for determination of serum or blood ethanol concentration lacks analytic specificity; however, it is a useful screening test. Headspace GC coupled with flame ionization is the established reference method for ethanol determinations and is quite useful as it can simultaneously quantitate other alcohols as well, such as methanol and isopropanol. Analysis begins with dilution of the serum or blood sample with a saturated solution of sodium chloride in a closed container. Volatiles within the liquid specimen partition into the airspace (headspace) of the closed container. Sampling of the headspace provides a clean specimen with little or no matrix effect. Quantitation can be performed by constructing a standard curve or calculating the concentration based on relative changes to an internal standard (*n*-propanol), as shown in **Figure 26.2**.

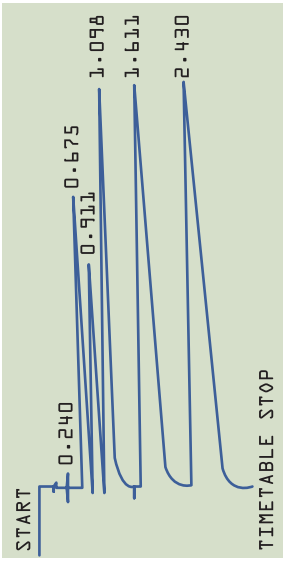
Enzymatic methods for ethanol determination use a nonhuman form of ADH to oxidize ethanol in the specimen to acetaldehyde with simultaneous reduction of NAD^+ to NADH.

$$\text{Ethanol} + NAD^+ \xrightarrow{\text{ADH}} \text{Acetaldehyde} + NADH \quad \text{(Eq. 26.4)}$$

The reduced NADH produced can be monitored directly by absorbance at 340 nm or can be coupled to an indicator reaction. This form of ADH is relatively specific for ethanol, and intoxication with methanol or isopropanol produces a negative or low result; therefore, a negative result by this method does not rule out ingestion of other alcohols. There is good agreement between the enzymatic reactions of ethanol and GC. The enzymatic reactions can be fully automated and do not require specialized instrumentation. Ethanol metabolites, including ethyl glucuronide, ethyl sulfate, and phosphatidylethanol, can also be tested for in clinical and forensic toxicology.[12] Methanol, isopropanol, and ethylene glycol can also can be determined using GC. Ethylene glycol is precipitated, and the supernatant analyzed with GC-flame ionization. Though not commonly available, a commercial veterinary enzymatic assay for the detection of ethylene glycol has been adapted by the clinical laboratory at the University of Iowa Hospitals and Clinics.[13]

Carbon Monoxide

Carbon monoxide is produced by incomplete combustion of carbon-containing substances. The primary environmental sources of carbon monoxide include



Retention Time (Minutes)	Analyte
0.675	Methanol
0.911	Acetone
1.098	Ethanol
1.611	Isopropanol
2.430	*n*-Propanol

Figure 26.2 Headspace gas chromatography of alcohol. The concentration of each alcohol can be determined by comparison to the response from the internal standard *n*-propanol.

gasoline engines, improperly ventilated furnaces, and wood or plastic fires. Carbon monoxide is a colorless, odorless, and tasteless gas that is rapidly absorbed into circulation from inspired air.

When carbon monoxide binds to hemoglobin, it is called *carboxyhemoglobin (COHb)*. The affinity of carbon monoxide for hemoglobin is 200 to 225 times greater than for oxygen. Air is approximately 20% oxygen by volume. If inspired air contained 0.1% carbon monoxide by volume, this would result in 50% carboxyhemoglobinemia at equilibrium. Because both carbon monoxide and oxygen compete for the same binding site, exposure to carbon monoxide results in a decrease in the concentration of oxyhemoglobin, and for this reason, carbon monoxide is considered a very toxic substance. Carbon monoxide expresses its toxic effects by causing a leftward shift in the oxygen–hemoglobin dissociation curve, resulting in a decrease in the amount of oxygen delivered to the tissue. The net decrease in the amount of oxygen delivered to the tissue results in hypoxia. The major toxic effects of carbon monoxide exposure are seen in organs with high oxygen demand, such as the brain and heart. The concentration of COHb, expressed as the percentage of COHb, presents relative to the capacity of the specimen to form COHb, and the corresponding symptoms are detailed in **Table 26.3**. The only treatment for carbon monoxide poisoning is 100% oxygen therapy. In severe cases, hyperbaric oxygen may be used to promote distribution of oxygen to the tissues. The half-life of COHb is roughly 60 to 90 minutes in a patient with normal respiratory function breathing 100% oxygen.

Table 26.3 Symptoms of Carboxyhemoglobinemia

COHb (%)	Symptoms and Comments
0.5	Typical in nonsmokers
5–15	Range of values seen in smokers
10	Shortness of breath with vigorous exercise
20	Shortness of breath with moderate exercise
30	Severe headaches, fatigue, impairment of judgment
40–50	Confusion, fainting on exertion
60–70	Unconsciousness, respiratory failure, death with continuous exposure
80	Immediately fatal

COHb, carboxyhemoglobin.

Several methods are available for the evaluation of carbon monoxide poisoning. Pulse (finger) CO-oximetry has been examined for screening purposes.[14,15] There are two primary quantitative assays for COHb: differential spectrophotometry and GC. Spectrophotometric methods work on the principle that different forms of hemoglobin present with different spectral absorbance curves. By measuring the absorbance at four to six different wavelengths, the concentration of the different species of hemoglobin, including COHb, can be determined by calculation. This is the most common method used for COHb and is the basis for several automated systems. GC methods, however, are considered the reference method for determination of COHb due to their high accuracy and precision. Carbon monoxide is released from hemoglobin after treatment with potassium ferricyanide. After analytic separation, carbon monoxide is detected by changes in thermal conductivity, and the COHb concentration can be determined.

Caustic Agents

Caustic agents are found in many household products and occupational settings. Even though any exposure to a strong acid or alkaline substance is associated with injury, aspiration and ingestion present the greatest hazards. Aspiration is usually associated with pulmonary edema and shock, which can rapidly progress to death. Ingestion produces lesions in the esophagus and gastrointestinal tract, which may produce perforations. This results in hematemesis, abdominal pain, and possibly shock. The onset of metabolic acidosis or alkalosis occurs rapidly after ingestion of caustic agents, and corrective therapy for ingestion is usually by dilution.

Cyanide

Cyanide is classified as a super-toxic substance that can exist as a gas, solid, or in solution. Because of the various forms, cyanide exposure can occur by inhalation, ingestion, or transdermal absorption. Cyanide is used in many industrial processes and is a component of some insecticides and rodenticides. It is also produced as a pyrolysis product from the burning of some plastics, including urea foams that are used as insulation in homes. Thus, carbon monoxide and cyanide exposure may account for a significant portion of the toxicities associated with fires and smoke

inhalation. Ingestion of cyanide is also a common suicide agent.

Cyanide expresses toxicity by binding to heme iron. Binding to mitochondrial cytochrome oxidase causes an uncoupling of oxidative phosphorylation. This leads to rapid depletion of cellular adenosine triphosphate as a result of the inability of oxygen to accept electrons. Increases in cellular oxygen tension and venous pO_2 occur as a result of lack of oxygen utilization. At low levels of exposure, patients experience headaches, dizziness, and respiratory depression, which can rapidly progress to seizure, coma, and death at greater doses. Cyanide clearance is primarily mediated by rapid enzymatic conversion to thiocyanate, a nontoxic product rapidly cleared by renal filtration. Cyanide toxicity is associated with acute exposure at concentrations sufficient to exceed the rate of clearance by this enzymatic process.

Evaluation of cyanide exposure requires a rapid turnaround time, and several test methods are available. Ion-selective electrode (ISE) methods and photometric analysis following two-well microdiffusion separation are the most common methods. Chronic low-level exposure to cyanide is generally evaluated by determining a urinary thiocyanate concentration. High blood lactic acid levels are also commonly found in cyanide poisoning.[16]

Metals and Metalloids

Arsenic

Arsenic is a metalloid that may exist bound to or as a primary constituent of many different organic and inorganic compounds. It exists in both naturally occurring and manmade substances; therefore, exposure to arsenic occurs in various settings. Environmental exposure through air and water is prevalent in many industrialized areas, and occupational exposure occurs in agriculture and smelting industries. Arsenic is also a common homicide and suicide agent. Ingestion of less harmful organic forms of arsenic, such as arsenobetaine and arsenocholine, can occur with foods such as clams, oysters, scallops, mussels, crustaceans (crabs and lobsters), and some bottom-feeding finfish.

Arsenic toxicity is largely dependent on the valence state, solubility, and rate of absorption and elimination. The three major groups for arsenic include arsine gas (arsine trioxide), inorganic forms (trivalent and pentavalent), and organic forms (arsenobetaine and arsenocholine). The rate of absorption largely depends on the form of arsenic. Inhalation of arsine gas demonstrates the most acute toxicity. Organic arsenic-containing compounds, such as those found in seafood, are rapidly absorbed by passive diffusion in the gastrointestinal tract. Other forms are absorbed at a slower rate. Clearance of arsenic is primarily by renal filtration of the free, ionized state. Arsenic is rapidly cleared from the blood, such that blood levels may be normal even when urine levels remain markedly elevated. The "fish arsenic" such as arsenobetaine and arsenocholine is cleared in urine within 48 hours; however, the initial half-life of inorganic arsenic is close to 10 hours. Approximately 70% of inorganic arsenic is secreted in urine, of which 50% of excreted inorganic arsenic has been transformed to the organic form. However, these patterns may vary with the dose and clinical status of the patient. Chronic toxicity of arsenic can be due to low level, persistent exposure that may lead to **bioaccumulation** or increased **body burden** of this metal. Arsenic expresses toxic effects by high-affinity binding to the thiol groups in proteins; therefore, this can reduce the portion available for renal filtration and elimination. Because many proteins are capable of binding arsenic, the toxic symptoms of arsenic poisoning are nonspecific, and arsenic binding to proteins often results in a change in protein structure and function. Many cellular and organ systems are affected with arsenic toxicity; fever, anorexia, and gastrointestinal distress are often seen with chronic or acute arsenic ingestion at low levels. Peripheral and central damage to the nervous system, renal effects, hemopoietic effects, and vascular disease leading to death are associated with high levels of exposure.

Analysis of arsenic is often performed by atomic absorption spectrophotometry (AAS) or, more commonly, ICP-MS. Most forms of arsenic are only detectable in the blood for a few hours. More than 90% of an arsenic exposure is recovered in the urine within 6 days, making urine the specimen of choice for an exposure occurring in the previous week.

Some toxins can bind to sulfhydryl groups in keratin found in hair and fingernails. For this reason, long-term exposure to some toxins may also be assessed in these tissues. Typically, toxic element deposition in hair and fingernails is demonstrated 2 weeks after an exposure, and in arsenic poisoning cases, distinct white lines can be observed in the fingernails, which are referred to as Mees' lines.

Cadmium

Cadmium is a metal found in many industrial processes, with its main use being electroplating and galvanizing, although it is also commonly encountered

during the mining and processing of many other metals. Cadmium is a pigment found in paints and plastics and is the cathodal material of nickel–cadmium batteries. Due to its widespread industrial applications, this element has become a significant environmental pollutant. In the environment, cadmium binds strongly to organic matter where it is immobilized in soil and can be taken up by agricultural crops. Since tobacco leaves accumulate cadmium from the soil, regular use of tobacco-containing products is a common route of human exposure. Smoking is estimated to at least double the lifetime body burden of cadmium. For nonsmokers, human exposure to cadmium is largely through the consumption of shellfish, organ meats, lettuce, spinach, potatoes, grains, peanuts, soybeans, and sunflower seeds.

Cadmium expresses its toxicity primarily by binding to proteins; however, it can also bind to other cellular constituents. Cadmium distributes throughout the body but has a tendency to accumulate in the kidney, where most of its toxic effects are expressed. An early finding of cadmium toxicity is manifested by renal tubular dysfunction in which tubular proteinuria, glucosuria, and aminoaciduria are typically seen. In addition to renal dysfunction, concomitant parathyroid dysfunction and vitamin D deficiency may also occur. *Itai-itai* disease is characterized by severe osteomalacia and osteoporosis from the long-term consumption of cadmium-contaminated rice. Elimination of cadmium is very slow, as the biological half-life of cadmium is 10 to 30 years. Evaluation of excessive cadmium is most commonly accomplished by the determination of whole blood or urinary content using AAS.

Lead

Lead is a byproduct or component of many industrial processes, which has contributed to its widespread presence in the environment. It was a common constituent of household paints before 1972 and is still found in commercial paints and art supplies. Plumbing constructed of lead pipes or joined with leaded connectors has contributed to the lead concentration of water. Gasoline contained tetraethyl lead until 1978. The long-term utilization of leaded gasoline, lead-based paint, and lead-based construction materials has resulted in airborne lead, contaminated soil, and leaded dust.

The lead content of foods is highly variable. In the United States, the average daily intake for an adult is between 75 and 120 µg. This level of intake is not associated with overt toxicity. Because lead is present in all biological systems and because no physiologic or biochemical function has been found, the key issue is identifying the threshold dose that causes toxic effects. Gastrointestinal absorption is influenced by various factors, though the exact factors controlling the rate of absorption are unclear. However, susceptibility to lead toxicity appears to be primarily dependent on age. Adults absorb 5% to 15% of ingested lead, whereas children have a greater degree of absorption, and infants absorb nearly 30% to 40%. Absorbed lead binds with high affinity to many macromolecular structures and is widely distributed throughout the body. Lead distributes into two theoretical compartments: bone and soft tissue, with bone being the largest pool. Lead combines with the matrix of bone and can persist in this compartment for a long period due to its long half-life of almost 20 years. The half-life of lead in soft tissue is somewhat variable, though the reported average half-life is 120 days.

Elimination of lead occurs primarily by renal filtration, but because only a small fraction of total body lead is present in circulation, the elimination rate is slow. Considering the relatively constant rate of exposure and the slow elimination rate, total body lead accumulates over a lifetime. As mentioned, the largest accumulation occurs in bone, but there is also significant accumulation in the kidneys, bone marrow, circulating erythrocytes, and peripheral and central nerves.

Lead toxicity is multifaceted and occurs in a dose-dependent manner (**Figure 26.3**). Abdominal or neurological symptoms manifest after acute exposure only. The neurologic effects of lead are of particular importance. Lead exposure causes encephalopathy characterized by cerebral edema and ischemia. Severe lead poisoning can result in stupor, convulsions, and coma. Lower levels of exposure may not present with these symptoms; however, low-level exposure may result in subclinical effects characterized by behavioral changes, hyperactivity, attention deficit disorder, and a decrease in intelligence quotient (IQ) scores. Higher levels of exposure have also been associated with demyelinization of peripheral nerves resulting in a decrease in nerve conduction velocity.

Since children are more sensitive to these effects than adults, childhood lead testing is mandated in 11 states and for all children enrolled in Medicaid programs. The normal threshold for blood lead levels (BLL) set by the Centers for Disease Control and Prevention (CDC) is 3.5 µg/dL. Growth deficits are seen in children with BLL greater than 3.5 µg/dL, and anemia may occur at a BLL of 20 µg/dL. Children with

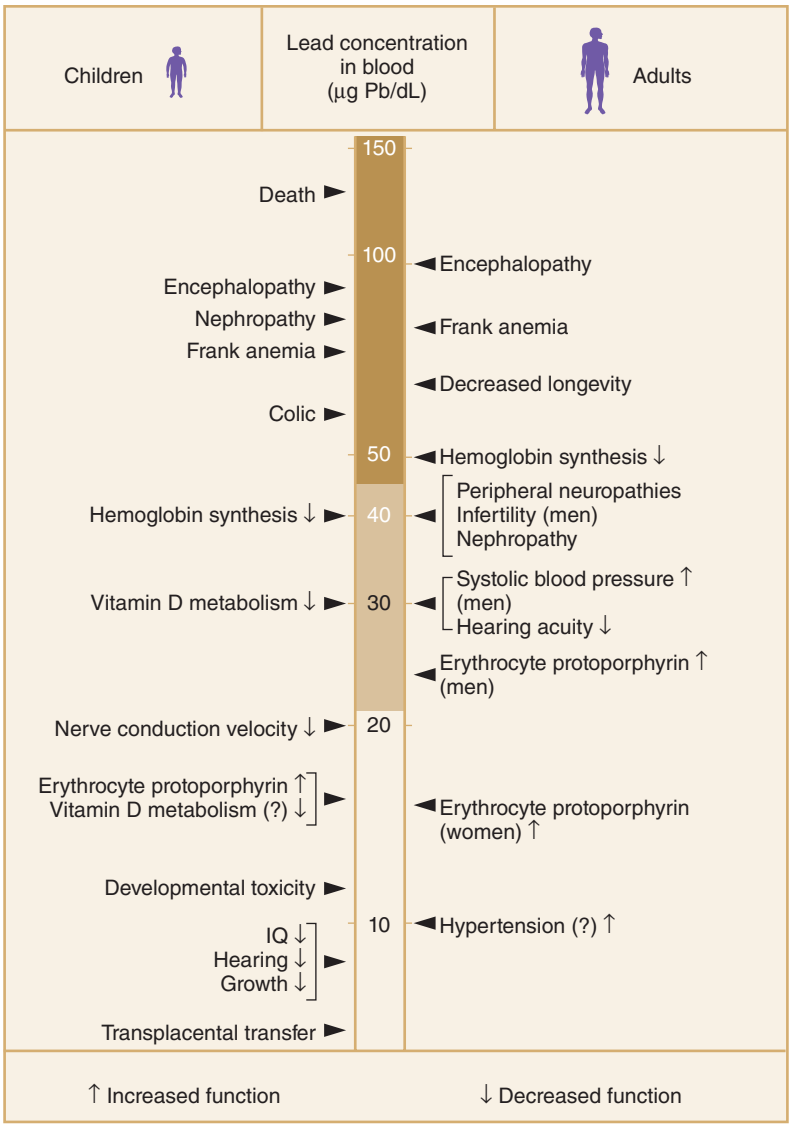


Figure 26.3 Comparison of effects of lead on children and adults.

From Centers for Disease Control and Prevention. (1992). *Case Studies in Environmental Medicine: Lead Toxicity.* U.S. Department of Human Services, Public Health Service, Agency for Toxic Substance and Disease Registry.

a BLL less than 3.5 µg/dL have been reported to suffer from permanent IQ and hearing deficits. As a result, many states have lowered the upper limit to 5 µg/dL. A threshold to identify permanent effects of lead poisoning is not currently known.

Lead is also a potent inhibitor of many enzymes, which leads to many of the toxic effects of lead exposure. The most noteworthy effects are on vitamin D metabolism and the heme synthetic pathway. Decreased serum concentrations of both 25-hydroxy and 1,25-dihydroxy vitamin D are seen in excessive lead exposure, resulting in changes in bone and calcium metabolism. Anemia is the result of the inhibition of the heme synthetic pathway, which results in increases in the concentration of several intermediates in this pathway, including aminolevulinic acid and protoporphyrin. Increases in protoporphyrin result in high concentrations of zinc protoporphyrin in circulating erythrocytes. Zinc protoporphyrin is a highly fluorescent compound, and measurement of this fluorescence has been used to screen for lead toxicity in the clinical laboratory. Increased urinary aminolevulinic acid is a highly sensitive and specific indicator of lead toxicity that correlates well with blood levels. Another hematologic finding of lead poisoning is the presence of basophilic stippling in erythrocytes as a result of inhibition of erythrocytic pyrimidine nucleotidase. This enzyme is responsible for the removal of residual DNA after extrusion of the nucleus. When identified, basophilic stippling is a key indicator of lead toxicity.

Excessive lead exposure has also been associated with hypertension, carcinogenesis, birth defects,

compromised immunity, and several renal effects. Early stages of toxicity are associated with renal tubular dysfunction, resulting in glycosuria, aminoaciduria, and hyperphosphaturia. Late stages are associated with tubular atrophy and glomerular fibrosis, which may result in a decreased glomerular filtration rate.

Treatment of lead poisoning involves cessation of exposure and treatment with therapeutic chelators, such as ethylenediaminetetraacetic acid (EDTA) and dimercaptosuccinic acid. These substances are capable of removing lead from soft tissue and bone by forming low-molecular-weight, high-affinity complexes that can be cleared by renal filtration. The efficacy of this therapy is monitored by determining the urinary concentration of lead.

The assessment of total body burden of lead poisoning is best evaluated by the quantitative determination of lead concentration in whole blood.[17] The use of urine is also valid but correlates closer to the level of recent exposure. Care must be taken during specimen collection to ensure that the specimen does not become contaminated from exogenous sources, and lead-free containers (tan-top K_2EDTA tube) are recommended for this purpose.

Several methods can be used to measure lead concentration. Screening methods such as erythrocyte protoporphyrin lack the sensitivity and can only detect toxic levels, usually ≥ 25 µg/dL of lead. Chromogenic reactions methods have been used but lack sufficient analytical sensitivity. Point-of-care units have employed these methodologies to screen for lead exposure in children or adults in the workplace. X-ray fluorescence is used to measure environmental levels of lead in nonbiological samples such as soil and foods. For biological specimens, graphite furnace AAS and anodic stripping voltammetry (ASV) have been used to confirm whole BLLs, but a new standard of measurement has been set with quantitative ICP-MS. AVS has a sensitivity of 3 µg/dL, while ICP-MS can detect levels as low as 0.1 µg/dL.[18]

Mercury

Mercury is a metal that exists in three forms: elemental, which is a liquid at room temperature; as inorganic salts; and as a component of organic compounds. Inhalation and accidental ingestion of inorganic and organic forms in industrial settings is the most common reason for toxic levels. Consumption of contaminated foods is the major source of exposure in the general population. Each form of mercury has different toxic characteristics. Elemental mercury (Hg^0) can be ingested without significant effects, and inhalation of elemental mercury is insignificant due to its low vapor pressure. Cationic mercury (Hg^{2+}) is moderately toxic; whereas, organic mercury, such as methylmercury (CH_3Hg^+), is extremely toxic.

Considering that the most common route of exposure to mercury is via ingestion, the primary factor determining toxicity is gastrointestinal absorbance. Elemental mercury is not absorbed well because of its viscous nature. Inorganic mercury is only partially absorbed but has significant local toxicity in the gastrointestinal tract. The portion that is absorbed distributes uniformly throughout the body. The organic forms of mercury are rapidly and efficiently absorbed by passive diffusion and partitions into hydrophobic compartments. This results in high concentrations in the brain and peripheral nerves.[19] In these lipophilic compartments, organic mercury is biotransformed to its divalent state, allowing it to bind to neuronal proteins. Elimination of systemic mercury occurs primarily via renal filtration of bound, low-molecular-weight species or the free, ionized state. Considering that most mercury is bound to protein, the elimination rate is slow, and chronic exposure, therefore, exerts a cumulative effect.

Mercury toxicity is a result of protein binding, which leads to a change of protein structure and function. The most significant result of this interaction is the inhibition of many enzymes. After ingestion of inorganic mercury, binding to intestinal proteins results in acute gastrointestinal disturbances. Ingestion of moderate amounts may result in severe bloody diarrhea because of ulceration and necrosis of the gastrointestinal tract. In severe cases, this may lead to shock and death. The absorbed portion of ingested inorganic mercury affects many organs. Clinical findings include tachycardia, tremors, thyroiditis, and most significantly, a disruption of renal function. The renal effect is associated with glomerular proteinuria and loss of tubular function. Organic mercury may also have a renal effect at high levels of exposure; however, neurological symptoms are the primary toxic effects of this hydrophobic form. Low levels of exposure cause tremors, behavioral changes, mumbling speech, and loss of balance. Higher levels of exposure may result in hyporeflexia, hypotension, bradycardia, and renal dysfunction and can lead to death. Analysis of mercury is by AAS using whole blood or an aliquot of a 24-hour urine specimen or anodal stripping voltammetry. Analysis of mercury by AAS requires special techniques as a result of the volatility of elemental mercury.

Pesticides

Pesticides are substances that have been intentionally added to the environment to kill or harm an undesirable life form. Pesticides can be classified into several categories including insecticides, herbicides, fungicides, and rodenticides. These agents are generally applied to control vector-borne disease and urban pests and to improve agricultural productivity. Pesticides can be found in occupational settings and in the home; therefore, there are frequent opportunities for exposure. Contamination of food is the major route of exposure for the general population. Inhalation, transdermal absorption, and ingestion as a result of hand-to-mouth contact are common occupational and accidental routes of exposure.

Ideally, the actions of pesticides would be target specific. Unfortunately, most are nonselective and result in toxic effects to many nontarget species, including humans. Pesticides come in many different forms with a wide range of potential toxic effects. The health effects of short-term, low-level exposure to most of these agents have yet to be well elucidated. Extended low-level exposure may result in chronic disease states. Of primary concern, though, is high-level exposure, which may result in acute disease states or death. The most common victims of acute poisoning are people who apply pesticides without taking appropriate safety precautions to avoid exposure. Ingestion by children at home is also common, and ingestion of pesticides is also a common in suicide attempts.

There is a wide variation in the chemical configuration of pesticides, ranging from simple salts of heavy metals to complex high-molecular-weight organic compounds. Insecticides are the most prevalent of pesticides. Based on chemical configuration, the organophosphates, carbamates, and halogenated hydrocarbons are the most common insecticides. Organophosphates are the most abundant and are responsible for about one-third of all pesticide poisonings. Organophosphates and carbamates function by inhibiting acetylcholinesterase, an enzyme present in both insects and mammals. In mammals, acetylcholine is a neurotransmitter found in both central and peripheral nerves. It is also responsible for the stimulation of muscle cells and several endocrine/exocrine glands. The actions of acetylcholine are terminated by the actions of membrane-bound, postsynaptic acetylcholinesterase. Inhibition of this enzyme results in the prolonged presence of acetylcholine on its receptor, which produces a wide range of systemic effects. Low levels of exposure are associated with salivation, lacrimation, and involuntary urination and defecation. Higher levels of exposure result in bradycardia, muscular twitching, cramps, apathy, slurred speech, and behavioral changes. Death due to respiratory failure may also occur.

Absorbed organophosphates bind with high affinity to several proteins, including acetylcholinesterase. Protein binding prevents the direct analysis of organophosphates; thus, exposure is evaluated indirectly by the measurement of acetylcholinesterase inhibition. Inhibition of this enzyme has been found to be a sensitive and specific indicator of organophosphate exposure, but because acetylcholinesterase is a membrane-bound enzyme, serum activity is low. Erythrocyte acetylcholinesterase is commonly analyzed for organophosphate testing instead of serum because erythrocytes have a high surface activity of acetylcholinesterase, which increases the analytical sensitivity of this test. Evaluation of erythrocytic acetylcholinesterase activity for detection of organophosphate exposure, however, is not routinely performed in many clinical laboratories because of low demand and the lack of an automated method.

An alternative test that has become more widely available is the measurement of serum pseudocholinesterase (SChE) activity. SChE is inhibited by organophosphates in a similar manner to the erythrocytic enzyme. Unlike the erythrocytic enzyme, however, changes in the serum activity of SChE lack sensitivity and specificity for organophosphate exposure. SChE is found in the liver, pancreas, brain, and serum, though its biological function is not well defined. Decreased levels of SChE can occur in acute infection, pulmonary embolism, hepatitis, and cirrhosis. There are also several variants of this enzyme that demonstrate diminished activity; therefore, decreases in SChE are not specific to organophosphate poisoning. The reference range for SChE is between 4000 and 12,000 U/L with intraindividual variation (the degree of variance within an average individual) of about 700 U/L. Symptoms associated with organophosphate toxicity occur at about a 40% reduction in activity. An individual whose normal SChE is on the high side of the reference range and who has been exposed to toxic levels of organophosphates may still have SChE activity within the reference range. Because of these factors, determination of SChE activity lacks sensitivity in the diagnosis of organophosphate poisoning and should only be used as a screening test. Immediate antidotal therapy can be initiated in cases of suspected organophosphate poisoning with a decreased activity of SChE; however, continuation of therapy and documentation of

such poisoning should be confirmed by testing of the erythrocytic enzyme. Analytical methods for the evaluation and quantification of organophosphates, including metabolites, in biological matrices include the utilization of GC or LC-MS.[20]

Toxicology of Therapeutic Drugs

In some cases, toxicity is the result of accidental or intentional overdosage of pharmaceutical drugs. All drugs are capable of toxic effects when improperly administered. This discussion focuses on the therapeutic drugs most commonly tested for in the clinical laboratory.

Salicylates

Acetylsalicylic acid, or aspirin, is a commonly used analgesic, antipyretic, and anti-inflammatory drug. It functions by decreasing thromboxane and prostaglandin formation through the inhibition of cyclooxygenase. At recommended doses, there are several noteworthy adverse effects, including interference with platelet aggregation and disruption to gastrointestinal function. There is also an epidemiologic relationship between aspirin, childhood viral infections (e.g., varicella, influenza), and the onset of Reye's syndrome.

Acute ingestion of high doses of aspirin is associated with various toxic effects through several different mechanisms.[21] Because it is an acid, excessive salicylate ingestion is associated with metabolic acidosis. Salicylate is also a direct stimulator of the respiratory center, and hyperventilation can result in respiratory alkalosis. In many instances, the net result of the combined effects is an immediate mixed acid–base disturbance. Salicylates also inhibit the Krebs cycle, resulting in excess conversion of pyruvate to lactic acid. In addition, at high levels of exposure, salicylates stimulate mobilization and use of free fatty acid, resulting in excess ketone body formation. These factors all contribute to nonrespiratory (metabolic) acidosis that may lead to death. Treatment for aspirin overdose involves neutralizing and eliminating the excess acid and maintaining the electrolyte balance.

Correlations have been established between serum concentrations of salicylates and toxic outcomes. The therapeutic range for pain relief (analgesic) is 2–10 mg/dL, while the level is 10–30 mg/dL for anti-inflammatory effects. Toxicity occurs when levels exceed 50 mg/dL. Several methods are available for the quantitative determination of salicylate in serum. GC and liquid chromatography (LC) methods provide the highest analytical sensitivity and specificity, but are not widely used due to high equipment expense and required technical skill. A commonly used method is a chromogenic assay known as the Trinder reaction; in this reaction, salicylate reacts with ferric nitrate to form a colored complex that is then measured spectrophotometrically. An enzymatic method featuring the enzyme salicylate hydroxylase is also commonly used in the clinical laboratory and has minimal interference for the quantification of salicylate levels in patient specimens.[22] The oxidation of NADH (substrate) to NAD^+ (end product) is measured by a decrease in absorbance at 340 nm. The decrease in absorbance is proportional to the concentration of salicylate in the patient sample.

Acetaminophen

Acetaminophen (Tylenol), either solely or in combination with other compounds, is a commonly used analgesic drug. In healthy subjects, therapeutic dosages have few adverse effects. Overdose of acetaminophen, however, is associated with severe hepatotoxicity (**Figure 26.4**).[23]

Absorbed acetaminophen is bound with high affinity to various proteins, resulting in a low free fraction. Thus, renal filtration of the parent drug is minimal, and most is eliminated by hepatic uptake, biotransformation, conjugation, and excretion. Acetaminophen can follow several different pathways through this process, each forming a different product. The pathway of major concern is the hepatic

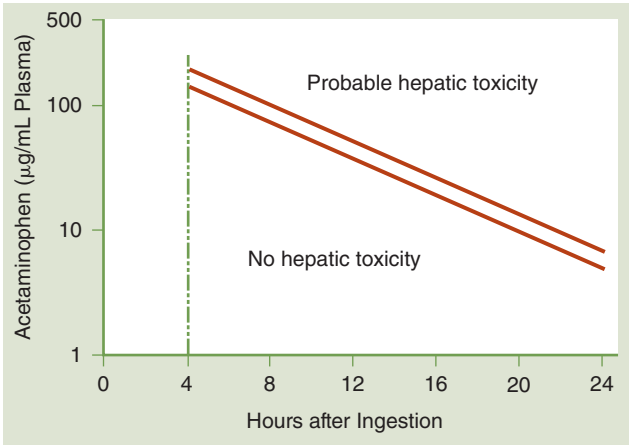


Figure 26.4 Rumack-Matthew nomogram. Prediction of acetaminophen-induced hepatic damage based on serum concentration.

Reprinted from Rumack BH, Matthew H. Acetaminophen poisoning and toxicity. *Pediatrics*. 1975;55:871, with permission.

mixed-function oxidase (MFO) system. In this system, acetaminophen is first transformed to reactive intermediates, which are then conjugated with reduced glutathione. In overdose situations, glutathione can become depleted, yet reactive intermediates continue to be produced. This results in an accumulation of reactive intermediates inside the cell. Because some intermediates are free radicals, this results in a toxic effect to the cell, leading to necrosis of the liver, the organ in which these reaction occur.

The time frame for the onset of hepatocyte damage is relatively long. In an average adult, serum indicators of hepatic damage do not become abnormal until 3 to 5 days after ingestion of a toxic dose. The initial symptoms of acetaminophen toxicity are vague and not predictive of hepatic necrosis. The serum concentration of acetaminophen that results in depletion of glutathione has been determined for an average adult. Unfortunately, acetaminophen is rapidly cleared from serum, and determinations are often made many hours after ingestion, making this information of little utility. To aid in this situation, nomograms (Figure 26.4) are available that predict hepatotoxicity based on serum concentrations of acetaminophen at a known time after ingestion. It is also worth noting that chronic, heavy consumers of ethanol metabolize acetaminophen at a faster rate than average, resulting in a more rapid formation of reactive intermediates and increased possibility of depleting glutathione. Therefore, alcoholic patients are more susceptible to acetaminophen toxicity, and using the nomogram for interpretation in these patients is inappropriate.

The reference method for the quantitation of acetaminophen in serum is high-performance liquid chromatography (HPLC). However, due to expense and expertise, this method is not widely used in clinical laboratories. Immunoassays are the most common analytical methods used for serum acetaminophen determinations, with competitive enzyme immunoassay being most frequently used. This assay is based on competition between drug in the sample and drug labeled with the enzyme for antibody binding sites. Enzyme activity decreases upon binding to the antibody, so the drug concentration in the sample can be measured in terms of enzyme activity. Active enzyme reduces NAD^+ to NADH, resulting in an absorbance change that is measured spectrophotometrically. Therapeutic levels for acetaminophen range from 10 to 30 µg/mL, and toxicity occurs at levels greater than 150 µg/mL, 4 hours after ingestion.

Toxicology of Drugs of Abuse

Assessment of substance abuse is of medical interest for many reasons. In a drug overdose, it is essential to identify the responsible agent to ensure appropriate and timely treatment. In a similar manner, identification of substance abuse in non-overdose situations provides a rationale for the treatment of addiction. For these reasons, testing for **drugs of abuse (DOA)** is commonly performed. This typically involves screening a single urine specimen for many substances by qualitative test procedures. In most instances, this procedure only detects recent drug use; therefore, with abstinence of relatively short duration, many patients engaging in substance abuse can avoid detection. In addition, a positive drug screen cannot discriminate between a single, casual use and chronic substance abuse. Identification of chronic substance abuse usually involves several positive test results in conjunction with clinical evaluation. In a similar manner, a positive drug screen does not determine the time frame or dose of the drug taken. Substance abuse or overdose can occur with prescription, over-the-counter, or illicit drugs. The focus of this discussion is on substances with addictive potential.

The use of drugs for recreational or performance enhancement purposes is relatively common, with approximately 24% of the adult population (18 years old or older) having used an illicit drug. Substance abuse testing has become commonplace in professional, industrial, and athletic settings. The potential punitive measures associated with this testing may involve or result in civil or criminal litigation. Therefore, the laboratory must ensure that all results are accurate and all procedures have been properly documented as these results may be used in court as evidence. This requires the use of analytic methods that have been validated as accurate and precise. It also requires scrupulous documentation of specimen security, especially for medicolegal specimens. Protocols and procedures must be established to prevent and detect specimen adulteration. Measurement of urinary temperature, pH, specific gravity, and creatinine is commonly performed to ensure that specimens have not been diluted or treated with substances that may interfere with testing.[24] Specimen collection should be monitored and a chain of custody established to guard against specimen exchange or tampering.

Testing for DOA can be performed by several methods. A two-tiered approach of screening and confirmation is typically utilized. Screening

procedures should be simple, rapid, inexpensive, and capable of being automated. They are often referred to as *spot tests*.[25] In general, DOA screening procedures such as lateral flow immunoassays have good analytic sensitivity with marginal specificity, meaning that a negative result can rule out an analyte with a reasonable degree of certainty.[6] These methods usually detect classes of drugs based on similarities in chemical configuration, which allows the detection of parent compounds and congeners that have similar effects. Considering that many "designer" drugs are modified forms of established DOA, these methods increase the scope of the screening process. A drawback to this type of analysis is that it may also detect chemically related substances that have no or low abuse potential; therefore, interpretation of positive test results requires integration of clinical context and further testing. Confirmation testing must use methods with high sensitivity and specificity. Many of these tests provide both quantitative as well as qualitative information. Confirmatory testing requires the use of a method different from that used in the screening procedure. GC-MS is the reference method for confirmation of most analytes.

There are several general analytic procedures commonly used for the analysis of DOA. Chromogenic reactions are occasionally used for screening procedures, but immunoassays are more widely used for screening and confirmatory testing. In general, immunoassays offer a high degree of sensitivity and are easily automated. A wide variety of chromatographic techniques are used for qualitative identification and quantitation of drugs of abuse.[26] GC and LC allow complex mixtures of drugs to be separated and quantitated, but these methods are generally labor-intensive, time-consuming, and not well-suited to STAT screening.

Trends in substance abuse vary geographically and between different socioeconomic groups. For a clinical laboratory to provide effective service requires knowledge of the substance or drug groups likely to be found within the patient population it serves. Fortunately, the process of selecting which drugs to test for has been aided by national studies that have identified the drugs of abuse most commonly seen in various populations (**Table 26.4**). This provides the basis for test selection in most situations. The following discussion focuses on select drugs with a high potential for abuse.

As with the alcohols discussed earlier, screening panels are available for drugs of abuse. These panels typically include amphetamines, barbiturates, benzodiazepines, cocaine, opioids, THC (a chemical in marijuana), and tricyclic antidepressants. These results are reported as either negative or positive; a negative result means that the drug concentration is below the cutoff level of detection, while a positive result means that the drug concentration is above the cutoff level of detection. These are presumptive positive results until they can be confirmed by a more specific method, such as mass spectrometry.

Amphetamines

Amphetamine and methamphetamine are therapeutic drugs used for narcolepsy and attention deficit disorder. These drugs are stimulants with a high abuse potential, as they produce an initial sense of increased mental and physical capacity along with a perception of well-being. These initial effects are followed

CASE STUDY 26.1, PART 4

Recall Emily. Her urine was positive for the presence of ethyl glucuronide, a metabolite of ethanol that is present in urine for up to three days after ingestion of ethanol. When discussing the laboratory results with her physician, Emily finally admitted that she has been drinking excessively after recently ending a long-term relationship. Emily then has a referral appointment with a psychiatrist, and Emily is diagnosed with major depressive disorder. The psychiatrist arranges for Emily to have biweekly meetings with a therapist, in order to treat Emily's alcohol use disorder and depression. The psychiatrist also writes a prescription for Emily to take sertraline, a selective serotonin reuptake inhibitor (SSRI) used to treat major depressive disorder.

© Amos Morgan/Thinkstock/Getty Images.

7. Emily's psychiatrist wants to monitor her compliance with taking sertraline. What laboratory test should the psychiatrist order to monitor Emily's adherence?

8. Assume that no immunoassay screening tests are available in the toxicology laboratory to monitor adherence for sertraline. What other methodologies could be used instead?

Table 26.4 Prevalence of Use of Common Drugs of Abuse

Substance	Prevalence (%)
Alcohol	55.1
Marijuana	22.1
Inhalants	0.4
Methamphetamine	0.3
Hallucinogens	1.7
LSD	0.7
MDMA	0.7
Cocaine	1.5
Crack cocaine	0.1
Heroin	0.2
Tranquilizers	1.2
Psychotherapeutics (nonmedical use)	3.7
Pain relievers	1.4
Sedatives	0.1
Stimulants	1.7

LSD, Lysergic acid diethylamide. MDMA, methylenedioxymethamphetamine.

Data from the National Survey on Drug Use and Health (NSDUH), which is sponsored by the Substance Abuse and Mental Health Services Administration (SAMHSA), an agency in the U.S. Department of Health and Human Services (HHS). The percentage values estimate the trends in prevalence of use of various drugs in individuals aged 18–25, who by survey claim to have used drugs within the past month in 2018. Adapted from the National Institutes of Health, National Institute on Drug Abuse. *Results from National Survey on Drug Use and Health: Trends in Prevalence of Various Drugs for Ages 12 or Older, Ages 12 to 17, Ages 18 to 25, and Ages 26 or Older; 2016-2018 (in percent).* Retrieved March 25, 2021, from https://www.drugabuse.gov/drug-topics/trends-statistics/national-drug-early-warning-system-ndews/national-survey-drug-use-health

by restlessness, irritability, and possibly psychosis. Abatement of these late effects is often countered with repeated use; drug tolerance and psychological dependence develop with chronic use. Overdose, although rare in experienced users, results in hypertension, cardiac arrhythmias, convulsions, and possibly death. Various compounds chemically related to amphetamines are components of over-the-counter medications, including ephedrine, pseudoephedrine, and phenylpropanolamine. These amphetamine-like compounds are common in allergy and cold medications.

Identification of amphetamine abuse involves analysis of urine for the parent drugs. Immunoassay systems are commonly used as the screening procedure. Because of variable cross-reactivity with over-the-counter medications that contain amphetamine-like compounds, a positive result by immunoassay is considered presumptive positive only. Immunoassays for amphetamines are prone to false-positives due to cross-reactivity due to these structural analogues.[26] Confirmation of immunoassay-positive tests is most commonly made with LC-MS or GC-MS.

Sedatives–Hypnotics

Many therapeutic drugs can be classified as sedatives, hypnotics, or tranquilizers, and all members of this class are CNS depressants. They have a wide range of approved therapeutic roles, but they also have abuse potential ranging from high to low. These drugs often become available for illegal use through diversion from approved sources. Barbiturates and benzodiazepines are the most common types of sedative–hypnotics abused. There are many individual drugs within the barbiturate and benzodiazepine classification. Although barbiturates have a higher abuse potential, benzodiazepines are more commonly found in abuse and overdose situations; this appears to be a result of greater availability. Overdose with sedatives–hypnotics initially presents with lethargy and slurred speech, which can rapidly progress to coma. Respiratory depression is the most serious toxic effect of most of these agents, although hypotension can occur with barbiturates as well. The toxicity of many of these agents is potentiated by ethanol use.

Immunoassay is the most common screening procedure for both barbiturates and benzodiazepines. Broad cross-reactivity within members of each group allows for the detection of many individual drugs. GC or LC methods, coupled to mass spectrometry, are used for confirmatory testing.

Barbiturates

Barbiturates represent a class of drugs that were originally introduced as sleep inducers. Secobarbital, pentobarbital, and phenobarbital are the more commonly abused barbiturates. These drugs are commonly abused as "downers" to induce sleep after an amphetamine- or cocaine-induced "high." The presence of barbiturates in a urine drug screen indicates use of one of these drugs or another barbiturate, resulting in a positive screening test result. Most of the barbiturates are fast-acting, and their presence indicates use within the past 3 days. Phenobarbital is an exception: it is a long-acting barbiturate and has a very long half-life. The presence of phenobarbital in urine indicates that the patient has used the drug sometime within the past 30 days.

Benzodiazepines

Benzodiazepines are any of a group of compounds having a common molecular structure and acting similarly as depressants of the CNS. As a class of drugs, benzodiazepines are among the drugs most commonly prescribed because of their efficacy, safety, low addiction potential, minimal side effects, and high public demand for sedative and anti-anxiety agents. Diazepam (Valium), chlordiazepoxide (Librium), lorazepam (Ativan), and alprazolam (Xanax) are the most commonly abused benzodiazepines. Benzodiazepines are metabolized rapidly, and the parent compounds are not detected in urine, only their metabolites. The presence of a long-acting benzodiazepine indicates exposure within a 5- to 20-day interval preceding specimen collection. Following a dose of diazepam, the drug and its metabolites appear in the urine within 30 minutes. Peak urine output is reached between 1 and 8 hours.

Cannabinoids: Tetrahydrocannabinol (THC)

Cannabinoids are a group of psychoactive compounds found in marijuana. Of these, THC is the most potent and abundant. Marijuana, or its processed product, hashish, can be smoked or ingested. A sense of well-being and euphoria are the subjective effect of exposure. It is also associated with an impairment of short-term memory and intellectual function. Effects of chronic use have not been well established though tolerance, and a mild dependence may develop over time. THC overdose has not been associated with any specific adverse effects. THC is a lipophilic substance that is rapidly removed from circulation by passive distribution into hydrophobic compartments, such as the brain and fat. This results in slow elimination as a result of redistribution back into circulation and subsequent hepatic metabolism.

The half-life of THC in circulation is 1 day after a single use and 3 to 5 days for chronic, heavy consumers. Hepatic metabolism of THC produces several products that are primarily eliminated in urine. The major urinary metabolite is 11-nor-tetrahydrocannabinol-9-carboxylic acid (THC-COOH). This metabolite can be detected in urine for up to 5 days after a single use or up to 4 weeks following chronic, heavy use. Immunoassay tests for THC-COOH are used to screen for marijuana consumption, and GC-MS is used for confirmation. Both methods are sensitive and specific, and because of the low limit of detection of both methods, it is possible to find THC-COOH in urine as a result of passive inhalation. Urinary concentration standards have been established to discriminate between passive and direct inhalation. Analytical methods based on LC-MS/MS with improved sensitivity for THC and THC metabolites, including glucuronidated cannabinoids, are also currently available.[27]

Cocaine

Cocaine is an effective local anesthetic with few adverse effects at therapeutic concentrations. At higher circulating concentrations, it is a potent CNS stimulator that elicits a sense of excitement and euphoria. Cocaine is an alkaloid salt that can be administered directly (e.g., by insufflation or intravenous injection) or inhaled as a vapor when smoked in the free-base form known as crack. The half-life of circulating cocaine is brief—approximately 30 minutes to 1 hour. Acute cocaine toxicity is associated with hypertension, arrhythmia, seizure, and myocardial infarction. Because of its short half-life, maintaining the subjective effects over a single extended period requires repeated dosages of increasing quantity; therefore, correlations between serum concentration and the subjective or toxic effects cannot be established. Because the rate of change is more important than the serum concentration, a primary factor that determines the toxicity of cocaine is the dose and route of administration. Intravenous administration presents with the greatest hazard, closely followed by smoking.

Cocaine's short half-life is a result of rapid hepatic hydrolysis to inactive metabolites, which is the major route of elimination for cocaine. Only a small portion of the parent drug can be found in urine after an administered dose, and the primary product of hepatic metabolism is benzoylecgonine, which is largely eliminated in urine. The half-life of benzoylecgonine is 4 to 7 hours; however, it can be detected in urine for up to 3 days after a single use. The presence of this metabolite in urine is a sensitive and specific indicator of cocaine use. In chronic heavy abusers, it can be detected in urine for up to 20 days after the last dose. The primary screening test for detection of cocaine use is measurement of benzoylecgonine in urine by immunoassay. Confirmation testing is most commonly performed using GC-MS.

Opioids

Opioids are a class of substances capable of analgesia, sedation, and anesthesia. All are derived from or chemically related to substances derived from the opium poppy. The naturally occurring substances, referred

to as *opiates*, include opium, morphine, and codeine. Heroin, hydromorphone (Dilaudid), and oxycodone (Percodan) are chemically modified forms of the naturally occurring opiates. Meperidine (Demerol), methadone (Dolophine), propoxyphene (Darvon), pentazocine (Talwin), and fentanyl (Sublimaze) are the common synthetic opioids. Opioids have a high abuse potential, and chronic use leads to tolerance with physical and psychological dependence. Acute overdose presents with respiratory acidosis due to depression of respiratory centers, myoglobinuria, and possibly an increase in serum indicators of cardiac damage (e.g., CK-MB, troponin). High-level opioid overdose may lead to death caused by cardiopulmonary failure. Treatment of overdose includes the use of the antagonist naloxone.

Laboratory testing for opioids usually involves initial screening by immunoassay. Most immunoassays are designed to detect the presence of morphine and codeine; however, cross-reactivity allows for detection of many of the opioids. GC-MS is the method of choice for confirmation testing. LC-MS methods have also become more prominent in confirmatory testing for opioids, with high levels of sensitivity and specificity.[28]

Tricyclic Antidepressants

Tricyclic antidepressants (TCAs) are used to treat depression and mood disorders. Imipramine was the first tricyclic antidepressant used for depression. It is metabolized to desipramine. The total therapeutic concentration for the parent drug and metabolite is 175–300 ng/mL. Higher levels lead to toxicity, which is characterized by ventricular tachycardia. Amitriptyline, another TCA, is metabolized to nortriptyline. The therapeutic concentration for the parent drug and metabolite is 80–200 ng/mL. Both amitriptyline and nortriptyline can cause cardiac toxicity when the concentration is above 500 ng/mL, which is characterized by ventricular tachycardia. Doxepin is another common tricyclic antidepressant.

The total therapeutic concentration for doxepin and nordoxepin is 50–150 ng/mL. Toxicity of doxepin and its metabolite nordoxepin is expressed as cardiac dysrhythmias, which occur at concentrations greater than 500 ng/mL. Although commonly prescribed to treat depression, TCAs are commonly used in suicidal attempts. They block the reabsorption of serotonin and norepinephrine, increasing the levels of these two neurotransmitters in the brain. They were named for their three-ring chemical structure, hence the name tricyclic.

Methylenedioxymethamphetamine

Methylenedioxymethamphetamine (MDMA) is an illicit amphetamine derivative commonly referred to as "ecstasy."[29] Although it was strongly associated with club culture in the 1990s, its use has continued to grow. There are as many as 200 "designer" analogues that have been developed to produce effects comparable to those of MDMA. MDMA and its analogues are primarily administered orally in tablets of 50 to 150 mg. Other, less frequent routes of administration are inhalation, injection, or smoking. MDMA has a circulating half-life of approximately 8 to 9 hours. The majority of the drug is eliminated by hepatic metabolism, although 20% is eliminated unchanged in the urine.

The onset of effect is 30 to 60 minutes, and duration is about 3.5 hours. The desired effects include hallucination, euphoria, empathic and emotional responses, and increased visual and tactile sensitivity. Adverse effects include headaches, nausea, vomiting, anxiety, agitation, impaired memory, violent behavior, tachycardia, hypertension, respiratory depression, seizures, hyperthermia, cardiac toxicity, liver toxicity, and renal failure. The presenting symptoms along with patient behavior and history must be taken into account, as routine drug screening by immunoassay of a urine specimen will usually not test positive. Further analysis and confirmation of MDMA are generally performed using GC-MS or LC-MS.

Phencyclidine

Phencyclidine (PCP), also known as angel dust, is an illicit drug with stimulant, depressant, anesthetic, and hallucinogenic properties. Adverse effects, which are commonly noted at doses that produce the desired subjective effects, include agitation, hostility, and paranoia. Overdose is generally associated with stupor and coma. PCP can be ingested or inhaled by smoking PCP-laced tobacco or marijuana. It is a lipophilic drug that rapidly distributes into fat and brain tissue. Elimination is slow as a result of redistribution into circulation and hepatic metabolism. Approximately 10% to 15% of an administered dose is eliminated and unchanged in urine, which allows for identification of the parent drug in urine. In chronic, heavy users, PCP can be detected up to 30 days after abstinence. Immunoassay is used as the screening procedure with GC-MS as the confirmatory method, and confirmatory testing with LC-MS/MS has been previously demonstrated.[30]

Anabolic Steroids

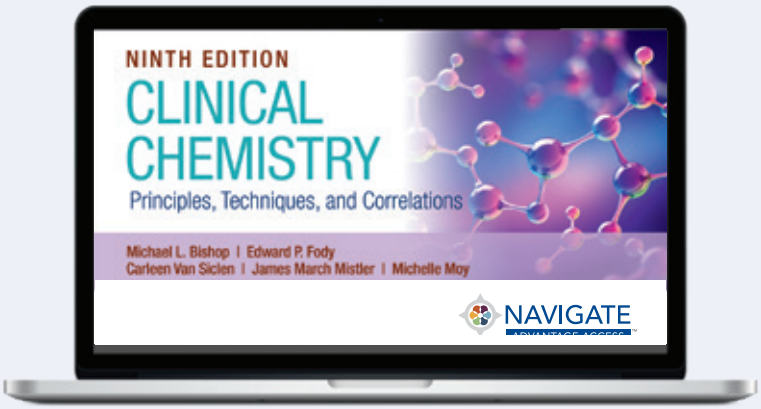
Anabolic steroids are a group of compounds that are chemically related to the male sex hormone testosterone. These artificial substances were developed in the 1930s as a therapy for male hypogonadism, although it was soon discovered that the use of these compounds in healthy subjects increases muscle mass. In many instances, this results in an improvement in athletic performance. Recent studies have reported that 6.5% of adolescent boys and 1.9% of adolescent girls reported the use of anabolic steroids without a prescription.

Most illicit steroids are obtained through the black market from underground laboratories and foreign sources. The quality and purity of these drugs are highly variable. In most instances, the acute toxic effects of these drugs are related to inconsistent formulation, which may result in high dosages and impurities. A variety of both physical and psychological effects have been associated with steroid abuse. Chronic use has been associated with toxic hepatitis as well as accelerated atherosclerosis and abnormal aggregation of platelets, both of which predispose individuals to stroke and myocardial infarction. In addition, steroid abuse causes an enlargement of the heart. In this condition, heart muscle cells develop faster than the associated vasculature, which may lead to ischemia of heart muscle cells. This predisposes the individual to cardiac arrhythmias and possible death.[31] In males, chronic steroid use is associated with testicular atrophy, sterility, and impotence. In females, steroid abuse causes the development of masculine traits, breast reduction, and sterility.

Evaluation of anabolic steroid use can be challenging. Until recently, the primary forms abused were animal-derived or synthetic forms. There are several well-established methods for the detection of the parent drug and its metabolite for the majority of these; however, the newer forms may be difficult to detect. To address this and related issues, the ratio of testosterone to epitestosterone is commonly used as a screening test; high ratios are associated with exogenous testosterone administration.[32]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Gummin D, Mowry J, Beuhler M, et al. 2019 Annual Report of the American Association of Poison Control Centers' National Poison Data System (NPDS): 37th Annual Report. *Clin Toxicol*. 2020;58(12):1360–1541.
2. Thanacoody R. Principles of assessment and diagnosis of the poisoned patient. *Medicine*. 2019;48(3):156–159.
3. Skolnik A, Monas, J. The crashing toxicology patient. *Emerg Med Clin North Am*. 2020;38(4):841–856.
4. Morrison EE, Sandilands EA. Principles of management of the poisoned patient. *Medicine*. 2019;48(3):160–164.
5. Dinis-Oliveira RJ, Nuno Vieira D, Magalhaes T. Guidelines for collection of biological samples for clinical and forensic toxicological analysis. *Forens Sci Res*. 2016;1(1):42–51.
6. DePriest AZ, Black DL, Robert TA. Immunoassay in healthcare testing applications. *J Opioid Manag*. 2015;11(1):13–25.
7. Seger C, Salzmann L. After another decade: LC-MS/MS became routine in clinical diagnostics. *Clin Biochem*. 2020;82:2–11.
8. Johnson-Davis KL, Farnsworth C, Law C, Parker R. Method validation for a multi-element panel in serum by

inductively coupled plasma mass spectrometry (ICP-MS). *Clin Biochem*. 2020;82:90–98.
9. Otvos JD, Shalaurova I, Wolak-Dinsmore J, et al. GlycA: a composite nuclear magnetic resonance biomarker of systemic inflammation. *Clin Chem*. 2015;61(5):714–723.
10. Zhong Y, Huang K, Luo Q, et al. The application of a desktop NMR spectrometer in drug analysis. *Int J Anal Chem*. 2018;2018:3104569. doi: 10.1155/2018/3104569
11. Sacks JJ, Gonzales KR, Bouchery EE, et al. 2010 national and state costs of excessive alcohol consumption. *Am J Prev Med*. 2015;49(5):e73–e79.
12. Kelly AT, Mozayani A. An overview of alcohol testing and interpretation in the 21st century. *J Pharm Pract*. 2012;25(1):30–36.
13. Rooney SL, Ehlers A, Morris C, et al. Use of a rapid ethylene glycol assay: a 4-year retrospective study at an academic medical center. *J Med Toxicol*. 2016;12:172–179.
14. Masters T, Willenbring B, Westgard B, et al. Availability of bedside and laboratory testing for carbon monoxide poisoning in the upper midwestern United States. *West J Emerg Med*. 2019;20(3):506–511.
15. Villalba V, Osborn ZT, Derickson PR, et al. Diagnostic performance of carbon monoxide testing by pulse oximetry in the emergency department. *Respir Care*. 2019;64(11): 1351–1357.
16. Hendry-Hofer T, Ng PC, Witeof AE, et al. A review on ingested cyanide: risks, clinical presentation, diagnostics, and treatment challenges. *J Med Toxicol*. 2019;15:128–133.
17. Bornhorst JA, McMillin, GA. Trace and toxic element testing in the clinical laboratory. *Lab Med*. 2006;37(11):690–695.
18. Haefliger P. *Brief Guide to Analytical Methods for Measuring Lead in Blood*. Geneva, Switzerland: World Health Organization; 2011:4–6.
19. Ye BJ, Kim BG, Jeon MJ, et al. Evaluation of mercury exposure level, clinical diagnosis and treatment for mercury intoxication. *Ann Occup Envir Med*. 2016;28(5):18.
20. Santos C, Oppolzer D, Goncalves A, et al. Determination of organophosphorous pesticides in blood using microextraction in packed sorbent and gas chromatography-tandem mass spectrometry. *J Anal Toxicol*. 2018;42:321–329.
21. Palmer BF, Clegg DJ. Salicylate toxicity. *N Engl J Med*. 2020;382(26):2544–2555.
22. Biary R, Kremer A, Goldfarb DS, Hoffman RS. Falsely elevated salicylate concentration in a patient with hypertriglyceridemia. *Toxicol Commun*. 2018;2(1):12.
23. Saccamano SJ. Acute acetaminophen toxicity in adults. *Nurs Crit Care*. 2019;14(5):10–17.
24. Kirsh KL, Christo PJ, Heit H, et al. Specimen validity testing in urine drug monitoring of medications and illicit drugs: clinical implications. *J Opioid Manag*. 2015;11(1):53–59.
25. Philp M, Fu S. A review of chemical 'spot' tests: a presumptive illicit drug identification technique. *Drug Test Anal*. 2018;10(1):95–108.
26. Moeller KE, Kissack JC, Atayee RS, Lee KC. Clinical interpretation of urine drug tests: what clinicians need to know about urine drug screens. *Mayo Clin Proc*. 2017;92(5):774–796.
27. Karschner EL, Swortwood-Gates MJ, Huestis MA. Identifying and quantifying cannabinoids in biological matrices in the medical and legal cannabis era. *Clin Chem*. 2020;66(7): 888–914.
28. Mahajan G. Role of urine drug testing in the current opioid epidemic. *Anesth Analg*. 2017;125:2094–2104.
29. Richards JR, Placone TW, Wang CG, et al. Methamphetamine, amphetamine, and MDMA use and emergency department recidivism. *J Emerg Med*. 2020;59(2):320–327.
30. Poklis JL, Guckert B, Wolf CE, Poklis A. Evaluation of a new phencyclidine enzyme immunoassay for the detection of phencyclidine in urine with confirmation by high-performance liquid chromatography-tandem mass spectrometry. *J Anal Toxicol*. 2011;35:481–486.
31. Lusetti M, Licata M, Silingardi E, et al. Pathological changes in anabolic androgenic steroid users. *J Forensic Leg Med*. 2015;33:101–104.
32. Sottas PE, Robinson N, Rabin O, Saugy M. The athlete biological passport. *Clin Chem*. 2011;57(7):969–976.

Trace Elements, Toxic Elements, and Vitamins

Michelle R. Campbell

CHAPTER OUTLINE

KEY TERMS

Atomic absorption spectroscopy
Atomic emission spectroscopy
Cofactor
Emission spectrum
Essential element
Hypervitaminosis
Hypovitaminosis
Mass spectrometry
Mass-to-charge ratio
Metalloenzyme
Metalloprotein
Nonessential elements
Parenteral nutrition
Trace element
Ultratrace element
Vitamin

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define metalloprotein, metalloenzyme, cofactor, trace element, ultratrace element, and vitamin.
- State the biological functions of trace elements, toxic elements, and vitamins.
- Explain the absorption, transport, and excretion of trace and toxic elements.
- Distinguish between essential and nonessential elements.
- Discuss the clinical significance of deficiency and toxicity of trace elements.
- Determine specimen collection and handling for trace elements and vitamins.
- List common laboratory methods for measuring trace elements and vitamins.
- Delineate the biochemical roles of vitamins.
- Correlate alterations in vitamin status with circumstances of increased metabolic requirements, age-related physiologic changes, or pathologic conditions.
- Describe drug–nutrient interactions that influence vitamin status.
- Apply knowledge of clinical chemistry to answer case study questions in the chapter.

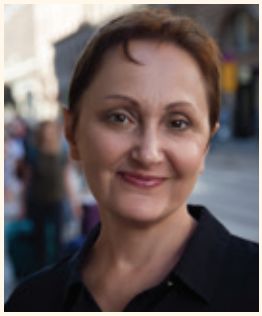
CASE STUDY 27.1, PART 1

Jason is a 27-year-old male. He has been seen regularly by his physician for skin lesions and increasing peripheral neuropathy over a period of several months. This coincides with a home project using reclaimed wood.

© Roman Samborskyi/Shutterstock.

CASE STUDY 27.2, PART 1

Nora is a 42-year-old woman with a history of diabetes. She was recently seen by her physician for weight loss, anorexia, and general fatigue. As part of the physical examination, both "bronze" skin pigmentation and an enlarged liver were noted.



© The Art of Pics/Shutterstock.

CASE STUDY 27.3, PART 1

Martin is a 42-year-old male. He presented to his physician with complaints of mild memory loss. During his visit, Martin mentioned that he exercises at least 4 days a week and maintains a strict vegetarian diet.

© AJP/Shutterstock.

Trace and Toxic Elements

Almost half of the elements listed in the periodic table have been found in the human body.[1] An element is considered an **essential element** if a deficiency impairs a biochemical or functional process and replacement of the element corrects this impairment. Decreased intake, impaired absorption, increased excretion, and genetic abnormalities are all examples of conditions that could result in deficiency of one or several **trace elements** (mg/dL concentration) or **ultratrace elements** (µg/dL concentration). The World Health Organization has established the dietary requirement for nutrients as the smallest amount of the nutrient needed to maintain optimal function and health. Any element that is not considered essential is classified as **nonessential elements**. Nonessential trace elements are of medical interest primarily because many of them are toxic.

The trace and toxic elements included in this chapter all have biochemical importance, whether minor or major. The essential trace elements are often associated with an enzyme (**metalloenzyme**) or another protein (**metalloprotein**) as a **cofactor**. Deficiencies typically impair one or more biochemical function, and excess concentrations are associated with at least some degree of toxicity.

The first section of this chapter provides an introduction to common testing methodologies, as well as an overview of the trace and toxic elements frequently encountered in the clinical laboratory, in alphabetical order (i.e., they are not listed in order of essential versus nonessential, or based on toxicity). The absorption, transport, distribution, metabolism (if relevant), and elimination will be described and related to the clinical significance of disease states or toxicity. Some reference ranges for trace and toxic elements have been included in the chapter text; however, several excellent sources for toxic element thresholds and trace element reference ranges are included in the bibliography. As always, the use of published thresholds and reference ranges as more than general guidelines must be done with caution, as variations in geographical location, testing methodologies, and population differences can often compromise their validity. The final section

provides an overview to fat-soluble and water-soluble vitamins.

Instrumentation and Methods

For many years, the most commonly used instrumentation for trace and toxic metal analysis has been the atomic absorption spectrometer, either with flame (FAAS) or flameless (i.e., graphite furnace, GFAAS) atomization. Atomic emission spectrometry is also useful for some elements, particularly if used in the form of inductively coupled plasma atomic emission spectroscopy (ICP-AES) for atomization and excitation. Inductively coupled plasma **mass spectrometry** (ICP-MS) is now more widely used because of its increased sensitivity, wide range of elements covered, and relative freedom from interferences. There is no single technique that is best for all purposes. A summary of the relative advantages and disadvantages of the main techniques is given in **Table 27.1**.

Sample Collection and Processing

Specimens for analysis of trace elements must be collected with scrupulous attention to details such as anticoagulant, collection apparatus, and specimen type (urine, whole blood, serum, plasma, hair, and nails).[2] Because of the low concentration in biologic specimens and the ubiquitous presence of trace metals in the environment, extraordinary measures are required to prevent contamination of specimens, which include using special collection tubes (royal blue stopper with or without EDTA additive). Specimen handling is also important; do not insert a pipette into the specimen; instead, transfer specimen by pouring. In addition, do not ream the serum with a wooden applicator stick. When submitting hair and nails specimens, submit in separate containers to avoid contamination.

The laboratory procedures to prevent environmental contamination are just as important. The water, reagents, pipettes, and sample cups must be carefully evaluated for use in trace and ultratrace analyses. In addition, the laboratory environment must be carefully controlled. Recommended measures include placing the trace elements laboratory in a separate room incorporating rigorous contamination control features, such as sticky mats at doors, non-shedding ceiling tiles, carefully controlled airflow to minimize particulate contamination, disposable booties worn over shoes, and particle monitoring equipment. Many useful measures are borrowed from those employed in semiconductor clean rooms.

Atomic Emission Spectroscopy

The simplified principle of the **atomic emission spectroscopy** (AES) instruments is presented in **Figure 27.1**.

The three most important components of AES are as follows:

1. A source, in which the sample is atomized at a sufficient temperature to produce an excited-state species. Those species will emit radiation upon relaxation back to the ground state.
2. A wavelength-selecting device (monochromator), for the spectral dispersion of the radiation and separation of the analytical line from other radiation.
3. A detector permitting measurement of radiation intensity.

Table 27.1 Relative Advantages and Disadvantages of Main Techniques for Elemental Analysis

	FAAS	GFAAS	ICP-AES	ICP-MS
Sensitivity	Moderate	Excellent	Moderate	Excellent
Selectivity	Excellent	Good	Poor	Good
Elemental coverage	Moderate	Good	Good	Excellent
Speed for one analyte	Fast	Slow	Fast	Fast
Multielement capabilities	No	No	Yes	Yes
Initial cost of instrument	Low	Moderate	Moderate	High
Cost of consumables	Very low	Very high	Low	Moderate
Ease of operation	Excellent	Poor	Moderate	Moderate

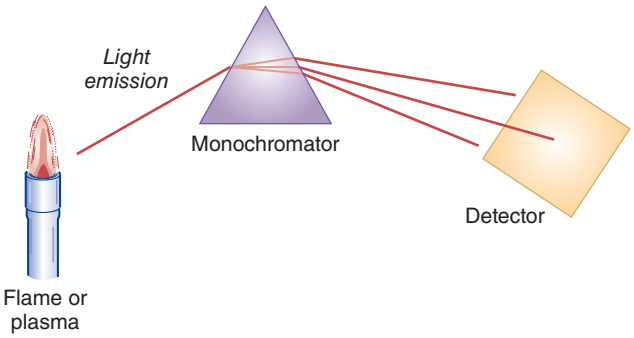


Figure 27.1 Simplified schematic of AES.

A liquid sample, containing element(s) of interest, is converted into an aerosol and delivered into the source, where it receives energy sufficient to emit radiation. The intensity of the emitted radiation is correlated to the concentration of an analyte and is the basis for quantitation. The most commonly used sources in AES are flame and inductively coupled plasma. Flames are capable of producing temperatures up to 3000 K. Typical fuel gases include hydrogen and acetylene, while oxidant gases include air or nitrous oxide. The gases are combined in a specially designed mixing chamber. A sample is also introduced into the mixing chamber using a nebulizer that converts liquid into a fine spray. The mixing chamber and burner assembly are shown in **Figure 27.2**. The same assembly can be used for atomic absorption spectroscopy (AAS) instrumentation.

In AES, both atomic and ionic excited states can be produced (depending on the element and the source), which leads to the production of complicated emission spectra. The "**emission spectrum**" of an element is composed of a series of very narrow peaks (sometimes known as "lines"), with each line at a different wavelength and matched to a specific transition. Each element has its own characteristic emission spectrum. For example, sodium can be detected by tuning the monochromator to a wavelength of 589 nm. Ideally, each emission line of a given element would be distinct from all other emission lines of other elements. However, there are many cases where emission lines from distinct elements overlap, resulting in interferences. The choice of interference-free wavelength (atomic or ionic line) may be challenging. While there are several possible wavelengths for a given element, wavelengths producing suitable analytical performance, such as limit of quantitation, freedom from interferences, and works, and robustness, are selected.

The first detectors in AES used photographic film. Contemporary AES instruments feature photomultiplier tubes or array-based detector systems.

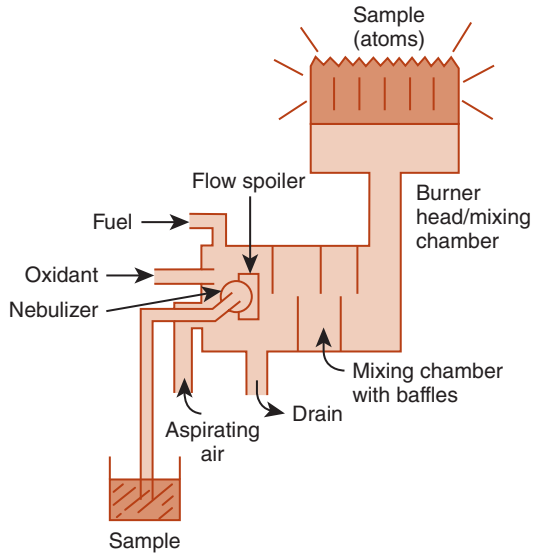


Figure 27.2 Mixing chamber burner for flame AAS.

Atomic Absorption Spectroscopy

Atomic absorption spectroscopy (AAS) is an analytical procedure for the quantitative determination of elements through the absorption of optical radiation by free atoms in the gas phase. The spectra of the atoms are line spectra that are specific for the absorbing elements.

Absorption is governed by the Beer-Lambert law:

$$A = -\log_{10}\left(\frac{I_1}{I_0}\right)\varepsilon L C_g \qquad \text{(Eq. 27.1)}$$

where A is the absorbance of the sample, I_0 is the incident light intensity, I_1 is the transmitted light intensity, ε is the molar absorptivity of the target analyte for the wavelength being used, L is the path length, and C_g is the gas-phase concentration of the target analyte. Under some simplifying assumptions, this equation takes the form:

$$A = KC \qquad \text{(Eq. 27.2)}$$

where K is a constant determined by calibration, and C is the solution phase concentration of the analyte.

The simplified principle of the AAS instruments is presented in **Figure 27.3**.

The four most important components of the AA spectrophotometer are:

1. Radiation (light) source, which emits the spectrum of the analyte element
2. Atomizer, in which the atoms of the element of interest in the sample are formed

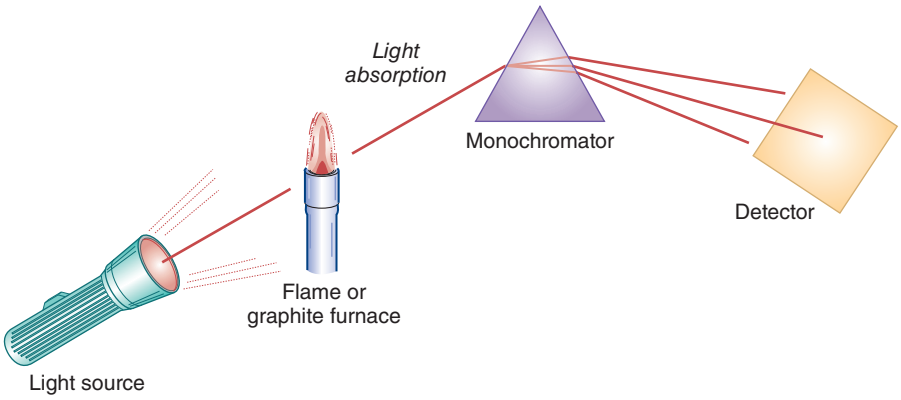


Figure 27.3 Simplified schematic of AAS instrumentation.

3. Monochromator, for the spectral dispersion of the radiation and separation of the analytical line from other radiation
4. Detector permitting measurement of radiation intensity

Typical light sources for AAS are hollow cathode lamps (HCLs) and electrodeless discharge lamps (EDLs). The HCL contains a quantity of the target element in the form of a hollow cylinder. During operation, a small quantity of the target element is vaporized, and some of the gas-phase atoms of the target element become electronically excited and emit photons with the right wavelength to be absorbed by atoms of the target element in the atomizer. While HCLs are an ideal source for determining most elements by atomic absorption, for volatile elements the use of EDLs is recommended.

The most common sources in AAS are flame (FAAS) and graphite furnace (GFAAS, also called flameless or electrothermal AAS). The mixing chamber burner, which produces laminar flames of high optical transparency, was already described in the section on AES in this chapter. Copper, iron, and zinc are often measured by FAAS.

The graphite tubes are the most commonly used atomizers in flameless AAS. Tubes are made of high-purity polycrystalline electrographite and coated with pyrolytic graphite, and they can be heated to a high temperature by an electrical current. A small aliquot (approximately 20 µL) of liquid sample is placed in the tube at the ambient temperature. The heating program (specifying the temperatures and times) is designed to first dry the sample, then pyrolyze, vaporize, and atomize the sample, followed by a cleaning step.

Selenium, cadmium, and lead are often measured by GFAAS. GFAAS allows for measurements of both liquid and solid samples. A common problem in GFAAS is that analyte volatility depends on the molecular form of the analyte and the sample matrix. To overcome this limitation, chemical modifiers (palladium nitrate, magnesium nitrate, or a mixture of both) are frequently added to samples, calibrators, and controls.

Inductively Coupled Plasma Mass Spectrometry

Inductively coupled plasma mass spectrometry (ICP-MS) is a newer and exceptionally sensitive analytical technique for elemental analysis. The term *plasma* in ICP refers to an ionized gas (typically argon), in which a certain proportion of electrons are free.

Like other mass spectrometers, the ICP-MS measures the **mass-to-charge ratio** (molecular mass divided by ionic charge [*m*/*z*] of selected analyte ions) and includes the following components: (1) an ion source, (2) a mass analyzer, and (3) an ion detector. A simplified schematic of an ICP-MS is given in **Figure 27.4**.

The argon plasma induced by commercial ICP instruments (both ICP-AES and ICP-MS) generates temperatures ranging from 6000 K to 10,000 K and serves several purposes. First, it dries and then vaporizes the droplets produced by the nebulizer. This step is followed by atomization of any molecular species. Finally, atoms are thermally ionized, at which point they are ready for introduction into the mass spectrometer.

Nearly all ICP torches consist of three concentric quartz tubes surrounded by a coil carrying radio-frequency (RF) power. The middle tube of the torch carries the argon (Ar) that forms the plasma.

Quantitative analysis for clinical samples is best performed with the use of an internal standard. All

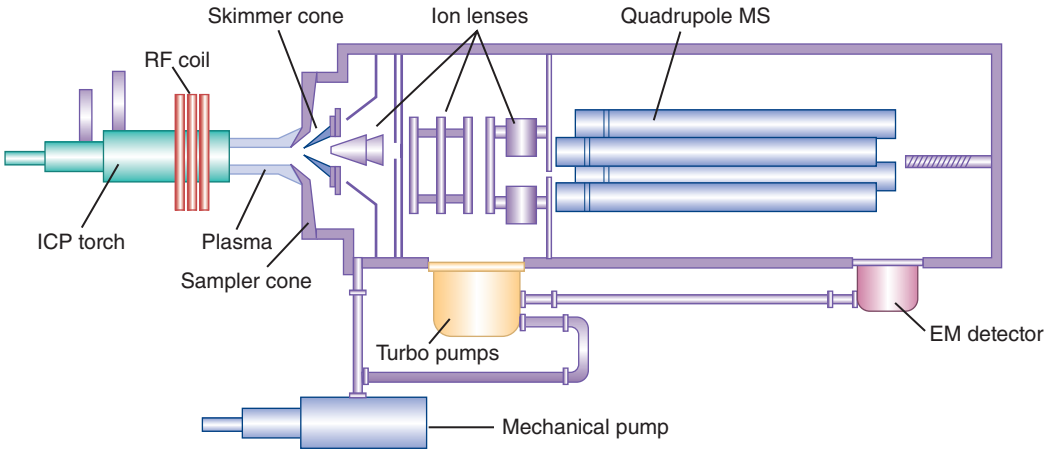


Figure 27.4 Simplified schematic of ICP-MS instrumentation.

patient samples, calibrators, and controls are diluted with an internal standard, usually a solution of an uncommon element such as yttrium (^{89}Y). Rather than using the raw signal level of the target elements as the basis for quantitation, the signal for each of the target elements is divided by the signal of the internal standard to give signal ratios (i.e., normalized intensities).

Quadrupole Mass Spectrometers. The typical mass spectrometer used for ICP-MS is a quadrupole mass spectrometer. The analyzer consists of four parallel conducting rods arranged in a square array. Applying RF and constant (direct current; DC) voltages to the rods, the instrument can be tuned so that only ions of a specific *m/z* ratio can pass through the device to reach the detector. This type of instrument tends to be relatively simple to use and maintain, but the resolution (the ability to discriminate between closely spaced *m/z* values) is limited, being able to well resolve peaks separated by one *m/z* unit but not able to resolve peaks separated by a small fraction of an *m/z* unit.

High-Resolution Mass Spectrometers. Other ICP-MS instruments incorporate high-resolution mass spectrometers. These are usually "double focusing sector field" instruments. Such instruments separate ions of different *m/z* values via deflection in a magnetic field, with ions of greater *m/z* being deflected to a lesser degree than those of lower *m/z*. The magnetic field is adjusted so as to allow only ions of a selected *m/z* to reach the detection system at any given point in time. A second device, known as an electrostatic analyzer, corrects for certain non-ideal effects, allowing the instrument to achieve high resolution. Commercially available high-resolution ICP-MS instruments are capable of a resolution of 10,000 (10% valley). This is enough to resolve, for example, $^{75}As^+$ from $^{75}ArCl^+$, both nominally 75 *m/z* units but which differ by 1×10^3 units when viewed at high resolution. However, magnetic sector instruments are not able to resolve elemental isobaric interferences such as $^{115}Sn/^{115}In$ or $^{40}Ca/^{40}Ar$, which would require resolution much higher than 10,000.

Interferences

In general, the interferences in elemental analysis are classified as spectroscopic or nonspectroscopic.

Spectroscopic. Spectral interferences generally result from a spectral overlap with the spectrum of the target analyte. For example, in AAS, certain molecular species may have broad absorption spectra that can overlap the line spectra of the elements of interest, leading to false elevations of the target element concentrations. A much less common occurrence would be for the absorption spectrum of one element to overlap with that of another.

Various strategies are used to deal with spectral interferences in AAS. A continuum source background corrector may be included in the instrument design at the cost of some instrument complication. Another alternative is Zeeman background correction, which relies on shifting the atomic spectral lines by the application of a magnetic field.

In ICP-MS, spectral interferences include polyatomic species whose *m/z* may overlap the *m/z* of the target analyte. For example, $^{56}(ArO^+)$ has the same nominal *m/z* as $^{56}Fe^+$. The argon oxide ion, which can be a significant component of plasma generated by an ICP torch, can potentially interfere with iron analysis

by ICP-MS. Another well-known polyatomic interference is argon chloride ion $^{75}(ArCl)^{+}$ on determination of $^{75}As^{+}$. An extensive table of polyatomic interferences in ICP-MS has been published.[3]

Several approaches are used to deal with polyatomic interferences in ICP-MS. One applies algebraic equations, together with relative isotopic abundance information, to mathematically correct for interferences.[4] Another approach interposes a reaction cell or collision cell between the main ion lenses and the mass analyzer.[5] A small amount of a gas such as helium or ammonia introduced into the cell removes interferences, either by chemical reaction or by an energy filtering process, using the fact that polyatomic species, with their larger collisional cross sections, lose energy faster than atomic ions. A recent tandem mass spectrometer that has been introduced to the market is similar to conventional tandem mass spectrometers, while the design includes two mass filtering quadrupoles separated by an octopole collision/reaction cell. High-resolution mass spectrometers provide a third way to remove interferences as discussed in an earlier section.

A second source of spectral interferences in ICP-MS arises from nearby elements in the periodic table. For example, tin (Sn) and cadmium (Cd) both have isotopes at 114 Da (amu), meaning they could potentially interfere with each other if the instrument is set to measure 114 *m/z*. This can usually be handled by using a different isotope for the analysis. For example, cadmium also has an isotope at 111 Da that is free from isobaric elemental interferences.

A third source of spectral interferences in ICP-MS comes from doubly charged ions. For example, $^{136}Ba^{2+}$ appears at the same *m/z* as $^{68}Zn^{+}$ (136/2 is equal to 68/1). These are relatively uncommon and can typically be avoided by choosing a different isotope for analysis or tuning the torch to reduce multiply charged ions.

Nonspectroscopic. Matrix interferences involve the bulk physical properties of the sample to be analyzed. The aqueous samples may behave differently from organic and biological specimens, depending upon the technology used and the analyte of interest. The properties of significance are viscosity, presence of easily ionized elements, and presence of carbon. Matrix matching of the calibrators, controls, and specimens helps to overcome matrix interferences. Dilution of the specimens helps to minimize matrix effect, but it is only applicable to certain analytical techniques and to the determination of analytes with higher concentrations.

Anything that could interfere with atomization of the sample could be classified as a nonspectral interference. For example, in AAS, a flame may not be hot enough for efficient atomization. Difference in sample viscosity between standards and unknown samples, resulting in differing rates of sample introduction, is another example of a nonspectral interference. In AES, anything that would prevent the efficient excitation or emission of spectral lines used for the analysis would constitute a nonspectral interference.

Elemental Speciation

The toxicity of elements may depend on their chemical forms. For example, arsenobetaine is a relatively nontoxic form of arsenic. Methylated forms of arsenic are intermediate in toxicity, and inorganic arsenic, such as As^{3+} and As^{5+}, is highly toxic. In the medical evaluation of patients, it can be important to know whether an elevated arsenic level is due to relatively innocuous forms, such as arsenobetaine, perhaps from a seafood meal ingested up to 3 days before the specimen collection, or due to dangerous forms, such as inorganic arsenic. In addition, the concentrations of methylated forms may be useful information for monitoring recovery from toxic exposure. The methodologies for elemental analysis discussed in previous sections are generally not capable of specifying the chemical form of the target elements.

So-called hyphenated techniques allow for speciation determinations. In a hyphenated analysis, the combination of two or more complementary analytical techniques is used to measure the specific form of an analyte. A classic example of this approach is liquid chromatography-ICP-MS (LC-ICP-MS). The sample is injected into a liquid chromatography system, which separates the different chemical forms of the analyte, producing a characteristic retention time. Concurrently, the eluting sample is continuously analyzed by a mass spectrometer. The retention time partially identifies the analytes, and the mass spectrometer further identifies the element. In some cases, AAS may substitute for MS in elemental speciation schemes. Methods for elemental speciation are becoming increasingly common. Despite clinical matrices being among the most difficult for speciation, several applications have been reported.

Alternative Analytical Techniques

Voltametric methods, such as anodic stripping voltammetry (ASV) and adsorptive stripping voltammetry (ADSV), can be used in determination of selected

metals and are the basis for some point-of-care devices.[6] Additional alternative techniques employed in trace element analysis include ion chromatography (IC), gas chromatography-mass spectrometry (GC-MS), and laser ablation ICP-MS (LA-ICP-MS).

Aluminum

Aluminum (Al) is a silver-white, crystalline, ductile metal. Aluminum is the most abundant metal in the earth's crust (~8%). It is always found combined with other elements such as oxygen, silicon, and fluorine. Aluminum as an elemental metal is obtained from aluminum-containing minerals.

Due to its good conductivity of heat and electricity, ease of welding, tensile strength, light weight, and corrosion resistant oxide coat, aluminum is applicable to a wide variety of industrial and household uses. Aluminum is used for beverage cans, pots and pans, airplanes, siding and roofing, and foil. Aluminum is often mixed with small amounts of other metals to form aluminum alloys, which are stronger and harder than aluminum alone.

Aluminum compounds have many different uses, for example, as alums in water treatment and alumina in abrasives and furnace linings. They are also found in consumer products such as antacids, astringents, buffered aspirin, food additives, cosmetics, and antiperspirants.

Absorption, Transport, and Excretion

The average adult in the United States ingests about 5 to 10 mg aluminum per day in food, all of which is excreted.[7] The human organism can absorb aluminum and its compounds orally, through inhalation and parenterally. There is no indication of dermal absorption. Approximately 1.5% to 2% of inhaled and 0.01% to 5% of ingested aluminum are absorbed. The absorption efficiency is dependent on chemical form, particle size (inhalation), and concurrent dietary exposure to chelators such as citric acid or lactic acid.[8,9]

After a relatively quick uptake of aluminum into the intestinal walls, its passage into the blood is much slower. In plasma, aluminum is bound to carrier proteins such as transferrin.[10] Aluminum binds to various ligands in the blood and distributes to every organ, with highest concentrations ultimately found in bone (~50% of the body burden) and lung tissues (~25% of the body burden). Aluminum levels in lungs increase with age.[8] Urine accounts for 95% of aluminum excretion, with 2% eliminated in the bile.

Health Effects and Toxicity

The mechanisms by which aluminum applies its toxicity are not well understood, though aluminum has been shown to interfere with a variety of enzymatic processes,[11] and administration of aluminum to experimental animals is known to produce encephalopathy similar to that seen in Alzheimer's disease in humans.[12] Although aluminum-containing over-the-counter oral products are considered safe in healthy individuals at recommended doses, adverse effects have been observed following long-term use in some individuals. Workers who breathe large amounts of aluminum dusts may develop lung problems, such as coughing and difficulty breathing, as well as changes that are evident in chest x-rays.

Signs and symptoms of aluminum toxicity include encephalopathy (stuttering, gait disturbance, myoclonic jerks, seizures, coma, abnormal electroencephalogram [EEG]), osteomalacia or aplastic bone disease (painful spontaneous fractures, hypercalcemia, and tumorous calcinosis), proximal myopathy, increased risk of infection, microcytic anemia, increased left ventricular mass, and decreased myocardial function.[9] Aluminum toxicity occurs in people with renal insufficiencies who are treated by dialysis with aluminum-contaminated solutions or oral agents that contain aluminum. The clinical manifestations of aluminum toxicity include anemia, bone disease, and progressive dementia with increased concentrations of aluminum in the brain. Prolonged intravenous feeding of preterm infants with solutions containing aluminum is associated with impaired neurologic development.

Laboratory Evaluation of Aluminum Status

Aluminum is primarily measured using ICP-MS and may also be measured by GFAAS.[7,13] Accurate measurements are often complicated by the increased risk for environmental contamination of specimens.[13] Rubber stoppers on standard evacuated blood collection devices are commonly made of aluminum silicate, and their use results in measurable contamination of blood samples. For this reason, special, metal-free collection tubes must be used for samples intended for aluminum measurement.[7] Urine and serum levels are useful in screening for aluminum toxicity, monitoring exposure over time in dialysis patients, monitoring metallic prosthetic implant wear, as well as monitoring chelation therapy.[13]

Reference Range

0–6 ng/mL

Arsenic

Arsenic (As) is a ubiquitous element displaying both metallic and nonmetallic properties. Its content in the earth's crust is estimated at 1.5 to 2.0 mg/kg. For most people, food is the largest source of arsenic exposure (about 25 to 50 micrograms per day [µg/d]), with lower amounts coming from drinking water and air.[14] Anthropogenic sources of arsenic (production of metals, burning of coil, fossil fuels, timber and its use in agriculture) release three times more arsenic than natural sources. A common current use of arsenic is as a wood preservative. Other current and past uses of arsenic include pesticides, pigments, poison gases, ammunition manufacturing, semiconductor processing, and medicines.[15]

Health Effects and Toxicity

The clinical signs and symptoms of arsenic exposure depend on the duration and extent of the exposure to inorganic and methylated species of arsenic, as well as the underlying clinical status of the patient. For acute arsenic exposure, the symptoms may affect gastrointestinal (nausea, emesis, abdominal pain, rice water diarrhea), bone marrow (pancytopenia, anemia, basophilic stippling), cardiovascular (electrocardiogram [ECG] changes), central nervous (encephalopathy, polyneuropathy), renal (renal insufficiency, renal failure), and hepatic (hepatitis) systems. For chronic arsenic exposure, systems affected and associated symptoms may be dermatologic (Mees' lines, hyperkeratosis, hyperpigmentation, alopecia), cardiovascular (hypertension, peripheral vascular disease), central nervous system ("socks and glove" neuropathy, tremor), hepatic (cirrhosis, hepatomegaly), and malignancies (squamous cell skin, hepatocellular, bladder, lung, renal). Chronic arsenic exposure has been shown to cause blackfoot disease, a severe form of peripheral vascular disease that leads to gangrenous changes.

The white powder of arsenic trioxide is odorless, tasteless, and one of the most common poisons in human history. Doses of 0.01 to 0.05 g produce toxic symptoms. The lethal dose is reported to be between 0.12 and 0.30 g; however, recoveries from higher doses have been reported.[15] Immediate treatment of expected exposure consists of lavage and use of activated charcoal to reduce arsenic absorption. The most effective antidotes for arsenic poisoning are the following chelating agents: dimercaprol (a.k.a., British anti-Lewisite [BAL]), penicillamine, and succimer.[16]

In the year 2000, the U.S. Food and Drug Administration (FDA) approved the use of arsenic trioxide for the treatment of acute promyelocytic leukemia,[17] which is diagnosed in approximately 1500 people in United States every year.

Absorption, Transport, and Excretion

The main routes of exposure are ingestion of arsenic-containing foods, water and beverages, or inhalation of contaminated air; however, arsenic toxicity is a complex topic. Organic forms of arsenic such as arsenocholine and arsenobetaine are commonly found in fish and seafood, are considered relatively nontoxic, and are cleared rapidly (1 to 2 days).[18] Inorganic species of arsenic are highly toxic and occur naturally in rocks, soil, and groundwater. They are also found in many synthetic products, poisons, and industrial processes. Methylated species are intermediate in toxicity and arise primarily from metabolism of inorganic species, but small amounts may arise directly from food. Organic methylated arsenic compounds such as monomethylarsonic acid (MMA) and dimethylarsinic acid (DMA) are formed by hepatic metabolism of As^{3+} and As^{5+}. The methylated inorganic forms are considered less toxic than As^{3+} and As^{5+}, however, and they are eliminated slowly (1 to 3 weeks). The Biological Exposure Index established by the American Conference of Governmental Industrial Hygienists for the sum of inorganic plus metabolites of arsenic in urine is 35 µg/L. However, clinical symptoms may not be evident at 35 µg/L.[16]

Laboratory Evaluation of Arsenic Status

Arsenic is primarily measured using ICP-MS, GFAAS, or HGAAS, which only requires the hydride generation (HG) module. In many cases, arsenic is best detected in urine due to the short half-life of arsenic in

CASE STUDY 27.1, PART 2

Remember Jason. He is working with reclaimed wood that has been treated with chromated copper arsenate (approximately one-third inorganic arsenic).

1. Describe the metabolism of inorganic arsenic after ingestion.

© Roman Samborskyi/Shutterstock.

blood. Hair and nail analysis is also used to establish a timeline of arsenic exposure, as arsenic demonstrates high affinity binding to keratin composing hair and nails.[7] When arsenic speciation is desired (typically following a high total urine arsenic result reported by ICP-MS or AAS), a separation method is employed prior to elemental analysis.

Reference Range

<13 ng/mL (blood)
<35 µg/L (urine)

Cadmium

Cadmium (Cd) is a soft, bluish-white metal, which can be easily cut with a knife. While cadmium serves no biological purpose in humans,[7] principal industrial uses of cadmium include manufacture of pigments and batteries, as well as in the metal-plating and plastics industries.[19] Workers in the automotive industry are at risk for exposure to cadmium if spraying organic-based paints without the use of proper breathing apparatuses.[7] In the United States, the burning of fossil fuels such as coal or oil and the incineration of municipal waste materials constitute the largest sources of airborne cadmium exposure, along with zinc and lead smelters in some locations. Cadmium-containing waste products and soil contamination, primarily as a result of human activity, are becoming a concern. The U.S. Environmental Protection Agency (EPA) has established loading rates of 20 kilograms per hectare (kg/ha) for high cation-exchange-capacity soils with a pH of 6.5, and a lower rate of 5 kg/ha for acid soils.[19]

Absorption, Transport, and Excretion

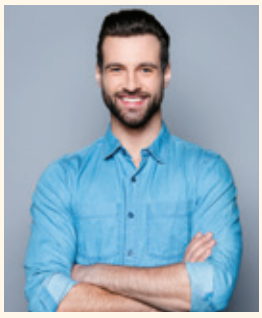
Based on renal function, the reference dose for cadmium in drinking water is 0.0005 mg per kg per day (mg/kg/d), and the dose for dietary exposure to cadmium is 0.001 mg/kg/d.[20] Absorption of cadmium is higher in females than in males due to differences in iron stores. The absorption of inhaled cadmium in air is 10% to 50%, with gastrointestinal absorption of cadmium estimated to be 5%. The absorption of cadmium in cigarette smoke is 10% to 50%, and smokers of tobacco products have about twice as much cadmium abundance in their bodies as nonsmokers. For nonsmokers, the primary exposure to cadmium is through ingested food.[21]

Health Effects and Toxicity

Cadmium has no known role in normal human physiology. Toxicity is believed to be a result of protein-Cd adducts causing denaturation of the associated proteins, resulting in a loss of function.[22] Ingestion of high amounts of cadmium may lead to a rapid onset of severe nausea, vomiting, and abdominal pain.[22] Renal dysfunction is a common presentation for chronic cadmium exposure, often resulting in slow-onset proteinuria. Acute effects of inhalation of fumes containing cadmium include respiratory distress due to chemical pneumonitis and edema and can cause death. Breathing of cadmium vapors can also result

CASE STUDY 27.1, PART 3

Remember Jason. His physician suspects arsenic toxicity from his home project and had Jason provide a urine sample for analysis. The initial total arsenic urine level was found to be 47 µg/L (reference range is <35 µg/L in urine).



© Roman Samborskyi/Shutterstock.

2. What is the likely cause of the elevated total arsenic result?
3. What reflex testing may be performed?

Results of reflexive arsenic speciation indicated:

- 13 µg/L As, Total Inorganic
- 33 µg/L As, Total Methylated
- <10 µg/L of As, Organic

4. Do these results support your prediction for the likely cause of the elevated total arsenic result? Why or why not?

Based on his clinical presentation, history of exposure to chromated copper arsenate, and laboratory results, Jason's physician diagnosed him with chronic arsenic toxicity.

5. Would blood arsenic levels have helped in the diagnosis? Why or why not?

in nasal epithelial damage and lung damage similar to emphysema.[20] Cadmium exposure can affect the liver, bone, immune, blood, and nervous systems.[21,23] Ethylenediaminetetraacetic acid (EDTA) can be used as a chelating agent in cadmium poisoning.

Laboratory Evaluation of Cadmium Status

Cadmium is usually quantified by ICP-MS. GFAAS is also used.[24] In blood, absorbed cadmium is found mostly (80%) in the red blood cells.[7] Cadmium in blood reflects the average uptake during the past few months and can be used for monitoring purposes but does not accurately reflect recent exposure. Urinary excretion is about 0.001% and 0.01% of the body burden per 24 hours. At low exposure, urine cadmium reflects the total accumulation.[24]

Reference Range

<5.0 ng/mL

Chromium

Chromium (Cr), from the Greek word *chroma* ("color"), makes rubies red and emeralds green. Chromium is the 21st most abundant element in the earth's crust and is used in the manufacturing of stainless steel. Occupational exposure to chromium occurs in wood treatment, stainless steel welding, chrome plating, the leather tanning industry, and the use of lead chromate or strontium chromate paints. Chromium exists in two main valency states: trivalent and hexavalent.

Absorption, Transport, and Excretion

Cr^{6+} is better absorbed and much more toxic than Cr^{3+}.[24] Both transferrin and albumin are involved in chromium absorption and transport.[25] Transferrin binds the newly absorbed chromium, while albumin acts as an acceptor and transporter of chromium if the transferrin sites are saturated. Other plasma proteins, including β- and γ-globulins and lipoproteins, also bind chromium.[26]

Health Effects, Deficiency, and Toxicity

Cr^{3+} is an essential dietary element, serving to heighten the action of insulin.[27] Cr^{3+} plays a role in maintaining normal metabolism of glucose, fat, and cholesterol. The estimated dietary chromium intake for adults is approximately 20 to 30 µg/d.[2] Foods such as breads, fish, fresh vegetables, meats, and even beer are common dietary sources of Cr^{3+}.[28]

Dietary chromium deficiency is relatively uncommon, and most cases occur in persons with specific clinical situations such as total **parenteral nutrition** or malnutrition. Chromium deficiency is characterized by insulin-resistant glucose intolerance, diabetes, hypercholesterolemia, and cardiovascular disease.[2,29]

Cr^{6+} compounds are powerful oxidizing agents and are more toxic systemically than Cr^{3+} compounds, given similar amounts and solubilities. At physiological pH, Cr^{6+} forms CrO_2^- and readily passes through cell membranes due to its similarity to essential phosphate and sulfate oxyanions. Intracellularly, Cr^{6+} is reduced to reactive intermediates, producing free radicals and oxidizing DNA, both potentially inducing cell death.[30] Severe dermatitis and skin ulcers can result from contact with Cr^{6+} salts. Up to 20% of chromium workers develop contact dermatitis. Allergic dermatitis with eczema has been reported in printers, cement workers, metal workers, painters, and leather tanners. Data suggest that a Cr^{3+}–protein complex is responsible for the allergic reaction.[29]

When inhaled, Cr^{6+} is a respiratory tract irritant, resulting in airway irritation and airway obstruction, and Cr^{6+} is an identified carcinogen associated with lung cancer.[2,31] The target organ of inhaled chromium is the lung; the kidneys, liver, skin, and immune system may also be affected.[2]

Low-dose, chronic chromium exposure typically results only in transient renal effects. Elevated urinary β_2-microglobulin levels (an indicator of renal tubular damage) have been found in chrome platers, and higher levels have generally been observed in younger persons exposed to higher Cr^{6+} concentrations.[29]

In May 2011, the FDA issued orders for postmarket surveillance studies to manufacturers of metal-on-metal (MoM) hip replacement systems in response to an increasing concern over failed implant devices reported in Europe.[32] In the United States, no guidelines have been established for the assessment of metal ions in asymptomatic patients due to a lack of knowledge regarding the prevalence of adverse events in the U.S. population, and no clear threshold levels have been associated with an adverse event. Although it is not an FDA requirement, measuring chromium levels is useful when monitoring metallic prosthetic implant wear.

Laboratory Evaluation of Chromium Status

Trivalent chromium (+3) is the only chromium valence state found in biological specimens. Chromium is most commonly measured by ICP-MS and

GFAAS and may be determined by neutron activation analysis (NAA). Plasma, serum, and urine do not indicate the total body status of the individual, whereas urine levels may be useful for metabolic studies and to evaluate individuals for recent occupational or environmental exposure to Cr.[2,25] In the setting of suspected failure of metal-on-metal (MoM) hip implants that use a cobalt–chrome alloy femoral head, serum is the preferred specimen type for both chromium and cobalt analysis.[33]

Reference Range

<1.0 ng/mL

Copper

Copper (Cu) is a relatively soft yet tough metal with excellent electrical and heat-conducting properties. Copper is widely distributed in nature both in its elemental form and in compounds. Copper forms alloys with zinc (brass), tin (bronze), and nickel (cupronickel, widely used in coins). Copper is an essential trace element found in four oxidation states, Cu(0), Cu(+1), Cu^{2+}, and Cu^{3+}, with Cu(+1) and Cu^{2+} present in biological systems.[2,34] Copper is an important cofactor for several metalloenzymes and is critical for the reduction of iron in heme synthesis.

Absorption, Transport, and Excretion

The copper content in a normal adult human is 50 to 120 mg. Copper is distributed through the body, with the highest concentrations found in the liver, brain, heart, and kidneys. Hepatic copper accounts for about 10% of the total copper in the body.[35] Copper is also found in the cornea, spleen, intestine, and lung. The amount of copper absorbed from the intestine is 50% to 80% of ingested copper.[29] The average daily intake in the United States is approximately 1.0 to 1.6 µg/d through dietary sources including organ meats, seafood, and grains.[36] Copper is absorbed and transported primarily in the small intestine and, to a lesser degree, the stomach.[2] Copper is transported to the liver bound to albumin, transcuprein, and low-molecular-weight components in the portal system. In the liver, copper is incorporated into ceruloplasmin for distribution throughout the body.[34] Ceruloplasmin is an α_2-globulin, and each 132,000-molecular-weight molecule contains six atoms of copper.[35]

In a normal physiological state, 98% of copper excretion is through the bile, forming feces, with copper losses in the urine and sweat constituting approximately <3% of dietary intake.[2,29]

Health Effects, Deficiency, and Toxicity

Copper is a component of several metalloenzymes, including ceruloplasmin, cytochrome C oxidase, superoxide dismutase, tyrosinase, metallothionein, dopamine hydroxylase, lysyl oxidase, clotting factor V, and an unknown enzyme that cross-links keratin in hair.

Copper deficiency is observed in premature infants, and copper absorption is impaired in severe diffuse diseases of the small bowel, lymph sarcoma, and scleroderma.[35] Copper deficiency is related to malnutrition, malabsorption, chronic diarrhea, hyperalimentation, and prolonged feeding with low-copper, total-milk diets. Signs of copper deficiency include (1) neutropenia and hypochromic anemia in the early stages, (2) osteoporosis and various bone and joint abnormalities that reflect deficient copper-dependent cross-linking of bone collagen and connective tissue, (3) decreased pigmentation of the skin and general pallor, and (4) in the later stages, possible neurologic abnormalities (hypotonia, apnea, psychomotor retardation).[29]

Subclinical copper depletion contributes to an increased risk of coronary heart disease. An extreme form of copper deficiency is seen in Menkes disease, an X-linked recessive disease of copper metabolism, with symptoms usually appearing at the age of 2 to 3 months and death typically occurring by the age of 3. This invariably fatal, progressive brain disease is characterized by kinky or steely hair (pili torti) and failure to thrive. Clinical signs include progressive mental deterioration, coarse feces, disturbance of muscle tone, seizures, and episodes of severe hypothermia.[2,37,38]

Copper toxicity has been associated with living near copper-producing facilities, using water from copper pipes, cooking with copper-lined vessels, or exposure to algaecides, herbicides, pyrotechnics, ceramic glazes, electrical wiring, or welding supplies. Because of its redox potential, copper is an irritant to epithelia and mucous membranes and can cause hepatic and renal damage with hemolysis.[39] Copper-induced emesis has a characteristic blue-green color.

Wilson's disease is an autosomal recessive disorder of copper accumulation that usually presents between the ages of 6 and 40 years. Clinical findings include neurologic disorders, liver dysfunction, and

Kayser-Fleischer rings (green-brown discoloration) in the cornea caused by copper deposition. Early diagnosis of Wilson's disease is important because complications can be effectively prevented, and in some cases, the disease can be halted with the use of zinc acetate or chelation therapy.[29] Serum ceruloplasmin levels and the direct measurement of free copper are key diagnostic steps in the diagnosis of Wilson's disease.[40]

Laboratory Evaluation of Copper Status

Copper is routinely measured by ICP-MS, although FAAS, ICP-AES, and ASV may be used. Serum copper and urine copper are used to monitor for nutritional adequacy and subacute management of copper toxicity. Direct measurement of free copper and ceruloplasmin in serum is used to screen for Wilson's disease. Common trends in laboratory testing seen in various diseases states are summarized in **Table 27.2**.

Reference Range

75–145 mcg/dL

Iron

Iron (Fe) is fourth most abundant element in the earth's crust and the most abundant transition metal. Methods of extracting iron from ore have been practiced for centuries. The physical properties of iron alloys can be varied over an enormous range by appropriate alloying and heat-treating methods, giving a range of strength, hardness, toughness, corrosion resistance (in the form of stainless steels), and magnetic properties and ability to take and hold a sharp edge.

Although highly abundant in the earth's crust, iron is classified as a trace element in the body. Iron ions can participate in redox chemistry in both the ferrous [Fe^{2+}] and ferric [Fe^{3+}] states, allowing iron to fill numerous biochemical roles as a carrier of other biochemically active substances (e.g., oxygen) and as an agent in redox and electron transfer reactions (e.g., via various cytochromes). Iron's high activity is a double-edged sword, and free iron ions in the body also participate in destructive chemistry, primarily in catalyzing the formation of toxic free radicals. As a result, very little free iron is normally found in the body.

Table 27.2 Interpretation of Copper Testing Results

	Serum Copper	Urine Copper
Nutritional deficiency	↓	↓
Acute copper toxicity	↑ or ↑↑	↑
Chronic copper toxicity	↑	↑
Wilson's disease	N or ↓	↑ or ↑↑
Menkes disease	↓	↑

N, normal; ↓, decreased; ↑, increased; ↑↑, significantly increased.

Data from https://labtestsonline.org/tests/copper and Mayo Clinic Laboratories: https://www.mayocliniclabs.com/test-catalog/Clinical and Interpretive/8612

Absorption, Transport, and Excretion

Absorption of iron from the intestine is the primary means of regulating the amount of iron within the body. Typically, only about 10% of the approximately 10 to 20 mg/d of dietary iron is absorbed.[41] To be absorbed by intestinal cells, iron must be in the ferrous oxidation state and bound to protein. Because Fe^{3+} is the predominant form of iron in foods, it must first be reduced to Fe^{2+} by agents such as vitamin C or ferric reductases present in the intestinal epithelium before it can be absorbed. In the intestinal mucosal cell, Fe^{2+} can be bound by ferritin for storage and eliminated after sloughing off or be exported to the basolateral side. From there, iron is oxidized to Fe^{3+} and bound by apotransferrin for transport throughout the body. The peptide hormone hepcidin regulates iron absorption in the upper gastrointestinal tract by modulating the export of iron from cells by ferroportin.[42] After about 120 days in circulation, red cells are degraded by the spleen, liver, and macrophages, which return iron to the circulation for reuse. Absorption and transport capacity can be increased in conditions such as iron deficiency, anemia, or hypoxia. Iron is lost primarily by desquamation of epithelia and red cell loss to urine and feces. With each menstrual cycle, women lose approximately 3 to 24 mg of iron. Approximately 1 in 10 women lose >1.4 mg/d of iron during menstruation, which may lead to iron deficiency anemia.[43]

Health Effects, Deficiency, and Toxicity

Of the 3 to 5 g of iron in the body, approximately 2 to 2.5 g of iron is in hemoglobin, mostly in RBCs and red cell precursors. A moderate amount of iron

(~130 mg) is in myoglobin, the oxygen-carrying protein of muscle. A small (~150 mg) but extremely important pool is incorporated into enzymes that require iron for full activity. These include peroxidases, catalases, and many of the Krebs cycle enzymes.[44] Iron is also stored as ferritin and hemosiderin, primarily in the bone marrow, spleen, and liver. Only 3 to 5 mg of iron is found in plasma, almost all of it associated with transferrin, albumin, and free hemoglobin.[45]

Iron deficiency affects about 15% of the worldwide population. Those with a higher-than-average risk of iron deficiency anemia include pregnant women, young children, adolescents, and women of reproductive age.[46] Increased blood loss, decreased dietary iron intake, or decreased release from ferritin may result in iron deficiency. Reduction in iron stores usually precedes both a reduction in circulating iron and anemia, as demonstrated by a decreased red blood cell count, mean corpuscular hemoglobin concentration, and microcytic RBCs.

Iron overload states are collectively referred to as hemochromatosis, whether or not tissue damage is present. Primary iron overload is most frequently associated with hereditary hemochromatosis (HH). HH is an autosomal recessive disorder leading to abnormally high iron absorption, culminating in iron overload. HH causes tissue accumulation of iron, affecting liver function, and often leads to hyperpigmentation of the skin. Some conditions associated with severe HH include diabetes mellitus, arthritis, cardiac arrhythmia or failure, cirrhosis, hypothyroidism, impotence, and liver cancer. Treatment may include therapeutic phlebotomy or administration of chelators, such as deferoxamine. Transferrin can be administered in the case of atransferrinemia.[42]

Secondary iron overload may result from excessive dietary, medicinal, or transfusional iron intake or be due to metabolic dysfunction. Hemosiderosis has been used to specifically designate a condition of iron overload as demonstrated by an increased serum iron and total iron-binding capacity (TIBC) or transferrin, in the absence of demonstrable tissue damage.

Iron may play a role as a pro-oxidant, contributing to lipid peroxidation,[47,48] atherosclerosis,[48,49] deoxyribonucleic acid (DNA) damage,[47–49] carcinogenesis,[50,51] and neurodegenerative diseases.[52,53] Fe^{3+}, released from binding proteins, can enhance production of free radicals to cause oxidative damage. In iron overloaded individuals with thalassemia who are treated with chelators to bind and mobilize iron, intake of ascorbic acid may actually promote the generation of free radicals.[54]

Laboratory Evaluation of Iron Status

Disorders of iron metabolism are evaluated primarily by packed cell volume, hemoglobin, red cell count and indices, total iron and TIBC, percent saturation, transferrin, and ferritin.[55] Iron studies with expected lab results seen in various diseases states are summarized in **Table 27.3**.

Serum Iron. Measurement of serum iron concentration refers specifically to the Fe^{3+} bound to transferrin and not to the iron circulating as free hemoglobin in serum. The specimen may be collected as serum without anticoagulant or as plasma with heparin. Oxalate, citrate, and EDTA binds Fe ions, and all are unacceptable anticoagulants. Early morning sampling is preferred because of the diurnal variation in iron concentration. Specimens with

CASE STUDY 27.2, PART 2

Recall Nora. Her physician ordered general chemistry testing to include albumin, AST, ALT, total bilirubin, and serum iron.

1. Based on Nora's clinical presentation, what possible diagnoses do you suspect?
2. What do you expect the results of this testing to be?

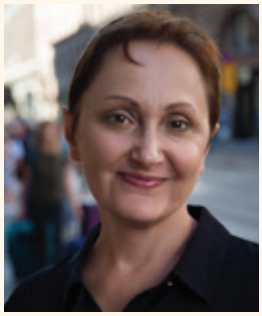


Table 27.3 Laboratory Markers of Iron Status in Several Disease States

Condition	Serum Iron	Transferrin	Ferritin	Percent Saturation	TIBC
Iron deficiency	↓	↑	↓	↓	↑
Iron overdose	↑	↓	↑	↑	↓
Hereditary Hemochromatosis	↑	Slightly ↓	↑	↑	↓
Malnutrition	↓	↓	↓	Variable	↓
Chronic infection	↓	↓	↑	↓	↓
Acute liver disease	↑	Variable	↑	↑	Variable
Chronic anemia	↓	N or ↓	N or ↑	↓	N or ↓

N, normal; ↓, decreased; ↑, increased.

Data from Iron | Lab Tests Online: Accessed at https://labtestsonline.org/tests/iron; Mayo Clinic Laboratories: Accessed at https://www.mayocliniclabs.com/test-catalog/Clinical and Interpretive/34624; and Jacobs DS, DeMott WR, Oxley DK. *Jacobs & Demott Laboratory Test Handbook*. 5th ed. Hudson, OH: Lexi-Comp.; 2001:1031

visible hemolysis should be rejected. Under acidic conditions, iron is liberated from transferrin. Ascorbate reduces the released Fe^{3+} ions to Fe^{2+} ions, which then react with ferrozine to form a colored complex. The color intensity is directly proportional to the iron concentration and can be measured photometrically.

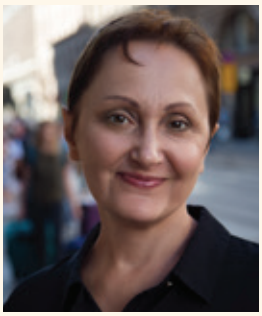
Transferrin. Transferrin is the primary plasma iron transport protein, which binds iron. Transferrin is saturated with approximately 25% to 30% iron. The additional amount of iron that can be bound is the unsaturated iron-binding capacity (UIBC). The serum transferrin method uses anti-transferrin antibodies, which react with the antigen in the sample to form an antigen/antibody complex. Following agglutination, this is measured turbidimetrically. Addition of polyethylene glycol increases sensitivity. Transferrin is increased in iron deficiency and decreased in iron overload and HH. Transferrin may also be decreased in chronic infections and malignancies.

Total Iron-Binding Capacity (TIBC). TIBC refers to the theoretical amount of iron that could be bound if transferrin and other minor iron-binding proteins present in the serum or plasma sample were saturated. The total iron-binding capacity (TIBC) can be indirectly determined using the sum of the serum iron and UIBC. TIBC is calculated (TIBC = transferrin × 1.18). Typically, only one-third of the iron-binding sites on transferrin are saturated.

CASE STUDY 27.2, PART 3

Recall Nora. Nora's initial chemistry panel showed the following results:

Patient Results	Reference Range
Albumin	3.7 g/dL (3.8–5.0)
AST	180 U/L (14–50)
ALT	200 U/L (10–60)
Total bilirubin	2.5 mg/dL (0.2–1.2)
Serum iron	180/dL (45–150)



© The Art of Pics/Shutterstock.

3. Based on these results, do you suspect iron deficiency or toxicity?

Percent Saturation. The percent saturation, also called the transferrin saturation, is the ratio of serum iron to TIBC. The normal range for this is approximately 20% to 50%, but it varies with age and sex. Percent saturation (100 × serum iron/TIBC) is usually normal or decreased in persons who are iron deficient, pregnant, or are taking oral contraceptives.

Ferritin. One molecule of ferritin is capable of binding more than 4000 atoms of iron, making ferritin the major iron storage protein for the body. The concentration of ferritin is directly proportional to the total iron stores in the body. One ferritin assay uses an immunoenzymatic sandwich assay. The patient's sample is added with antiferritin alkaline phosphatase conjugate, and particles coated with antiferritin complexes. Serum ferritin binds to the immobilized monoclonal antiferritin on the solid phase, while the antiferritin enzyme conjugate reacts with different antigenic sites on the ferritin molecules. After incubation, materials bound to the solid phase are captured, while unbound materials are washed away. A chemiluminescent substrate is added, and light generated by the reaction is measured with a luminometer. The light production is directly proportional to the concentration of ferritin in the sample. Serum ferritin is a much more sensitive and reliable test for demonstration of iron deficiency. Ferritin is decreased in iron deficiency anemia and increased in iron overload and HH. It is an acute-phase reactant and may be elevated in patients with inflammation. Besides HH, increases are seen in sideroblastic anemia and in patients receiving multiple blood transfusions.

A liver biopsy sample can be digested and analyzed for iron by AAS or ICP-MS as a follow-up to abnormal blood tests consistent with a HH diagnosis.[56] Iron quantification in liver specimen is not used for evaluation of acute iron toxicity. Hepcidin testing has not yet been shown to be clinically useful.

Reference Range

Serum iron:

- Males: 50–150 µg/dL
- Females: 35–145 µg/dL

TIBC: 250–400 µg/dL

Percent Saturation: 14%–50%

Ferritin:

- Males: 24–336 µg/L
- Females: 11–307 µg/L

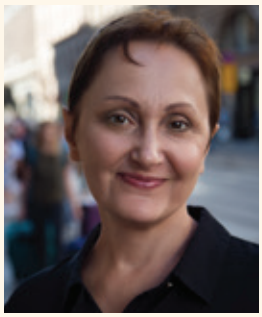
Lead

Metallic lead (Pb) is soft, bluish-white, highly malleable, and ductile. It is a poor conductor of electricity and heat and is resistant to corrosion. Lead is widely distributed in the earth's crust, and the main lead ores are galena, cerussite, and anglesite.[57] Lead is used in production of storage batteries, ammunition, solder, and foils. Tetraethyl lead was once used extensively as an additive in gasoline (petrol) for its ability to increase the fuel's octane rating and is present in many paints manufactured before 1970. The manufacture of lead-based household paints was banned in the United States in 1972, but lead is still used in paints intended for nondomestic use. Toxic concentrations of lead can be found in areas adjacent to homes painted with lead-based paints and around highways where it has accumulated from the past use of leaded gasoline. Lead plays no known role in normal human physiology.

CASE STUDY 27.2, PART 4

Recall Nora. Further testing on Nora for the elevated iron showed the following:

Patient Results	Reference Range
Serum iron	177 µg/dL (45–150)
Transferrin	195 mg/dL (200–380)
Ferritin	300 µg/L (10–250)
% Transferrin saturation	80% (20–50)



© The Art of Pics/Shutterstock.

4. Are Nora's laboratory results and symptoms typical of hemochromatosis? Could this be hereditary?

Absorption, Transport, and Excretion

Exposure to lead is primarily respiratory or gastrointestinal. Inhalation results in 30% to 40% absorption efficiency. Gut absorption depends on a variety of factors, including age and nutritional status, with enhanced gastrointestinal absorption occurring in children younger than 6 years of age. Certain substances, such as iron, calcium, magnesium, alcohol, and fat, may weaken lead absorption, while low dietary zinc, ascorbic acid, and citric acid can enhance the absorption of lead.[57] Ninety-nine percent of absorbed lead is taken up by erythrocytes, where it interferes with heme synthesis. Lead distributes to soft tissues, such as the liver, kidneys, and brain, with the skeletal lead concentrations containing greater than 90% of the body burden of lead.[58] Absorbed lead is excreted primarily in urine (76%) and feces (16%), and the remaining 8% is excreted in hair, sweat, nails, and others.[57]

Health Effects and Toxicity

The clinical presentation of lead toxicity is variable. In children, obvious symptoms are usually seen at blood levels of 60 µg/dL or higher, with 45 µg/dL the typical threshold for acute, clinical intervention with chelation therapy. IQ declines are seen in children with blood lead levels of 10 µg/dL or higher. Other central nervous system (CNS) symptoms of lead toxicity in children may include clumsiness, gait abnormalities, headache, behavioral changes, seizures, and severe cognitive and behavioral problems. Gastrointestinal symptoms include abdominal pain, constipation, and colic. Other conditions may include acute nephropathy and anemia. The U.S. Centers for Disease Control and Prevention (CDC) estimates an incidence of more than 450,000 for children in the United States with blood lead levels higher than the reference range of 3.5 µg/dL.[59] In adults, the following symptoms may be observed: peripheral neuropathies, motor weakness, chronic renal insufficiency, systolic hypertension, and anemia.

Lead exposure primarily arises in two settings: childhood exposure, usually through paint chips, and adult occupational exposure in the smelting, mining, ammunitions, soldering, plumbing, ceramic glazing, and construction industries. Other sources include lead-glazed ceramics and some traditional Chinese herbal medicines.[60] U.S. government websites contain extensive information on the health and environmental impacts of lead.[61]

Laboratory Evaluation of Lead Status

The most common specimen type is whole blood (EDTA-anticoagulated blood),[7] the result of which is commonly referred to as the blood lead level or BLL. This is preferred over plasma and serum, as circulating lead is predominantly associated with red blood cells. Elevated lead levels in capillary blood specimens should be confirmed with a venous specimen to avoid the potential contribution of external contamination.

Urine lead may be useful for detecting recent exposures to lead or to monitor chelation therapy. Other testing, such as plasma aminolevulinic acid, whole blood zinc protoporphyrin, or free erythrocyte protoporphyrins, may be useful for screening in occupational exposures. Noninvasive measurements of lead in bone may be available radiographically. Removal of further lead exposure and parental education are essential parts to management for patients with elevated blood lead levels.[39] ICP-MS is a preferred method of analysis. In addition to GFAAS and ASV[7] may be used.

Reference Range

<3.5 µg/dL

Manganese

As the twelfth most abundant element in the earth's crust, manganese (Mn) is found in over 250 minerals, of which 15 have commercial importance. Nearly all the elemental manganese is used in the production of the alloy ferromanganese widely used in steel production. Other uses of elemental manganese include a scavenger role in copper and aluminum alloys and in a production of dry cell batteries. Various manganese compounds are widely used in fertilizers, animal feeds, pharmaceutical products, dyes, paint dryers, catalysts, and wood preservatives and in production of glass and ceramics. The manganese-based compound methylcyclopentadienyl manganese tricarbonyl (MMT) is a fuel supplement used to increase the octane level of gasoline, and though banned in 1977, the ban was later lifted, and MMT is currently approved for use in the United States.[62,63]

Absorption, Transport, and Excretion

Roughly 2% to 15% of dietary manganese is absorbed in the small intestine. Dietary factors that affect manganese absorption include iron, calcium, phosphates, and fiber.[29] Manganese absorption is age-dependent,

with infants retaining higher levels of manganese than adults. Manganese is a normal component in tissue, with the highest levels found in fat and bone. Though accumulation of manganese in the healthy population has not been observed, chronic liver disease or other types of liver dysfunction can reduce manganese elimination and promote accumulation in various regions of the brain.[62] Manganese elimination occurs predominately through the bile.

Health Effects, Deficiency, and Toxicity

Manganese is biochemically essential as a constituent of metalloenzymes and as an enzyme activator. Manganese-containing enzymes include arginase, pyruvate carboxylase, and manganese superoxide dismutase in the mitochondria. Manganese-activated enzymes include hydrolases, kinases, decarboxylases, and transferases. Many of these activations are not specific to manganese, and other metal ions (magnesium, iron, or copper) can replace manganese as an activator and mask the effects of manganese deficiency.[29]

Blood clotting defects, hypocholesterolemia, dermatitis, and elevated serum calcium, phosphorus, and alkaline phosphatase activity have been observed in some subjects who underwent experimental manganese depletion.[29] Low levels of manganese have been associated with epilepsy,[9] hip abnormalities, joint disease, congenital malformation,[29] heart and bone problems, and stunted growth in children.[64] Manganese toxicity causes nausea, vomiting, headache, disorientation, memory loss, anxiety, and compulsive laughing or crying known as *locura manganica* (or "manganese madness").[65] Chronic manganese toxicity resembles Parkinson's disease with akinesia, rigidity, tremors, and masklike faces.[9]

Laboratory Evaluation of Manganese Status

Manganese is routinely measured by ICP-MS and GFAAS may also be used. Blood manganese concentrations are used as measure of exposure. Urine samples may also be measured but are prone to contamination with manganese-containing dust.[7]

Reference Range

<2.4 ng/mL

Mercury

Mercury (Hg), also called quicksilver, is a heavy, silvery metal. Along with bromine, mercury is one of only two elements that are liquid at room temperature and pressure.[7] There are three naturally occurring oxidation states of mercury: Hg^0, Hg^{1+}, and Hg^{2+}. Organic mercury refers to various forms of mercury bound to a carbon atom, with mercury usually in the 2+ oxidation state.

Mercury is released to the atmosphere as a product of the natural degassing of rock (30,000 tons per year) and through various human activities (20,000 tons per year). Mercury is used in dental amalgams, electronic switches, germicides, fungicides, and fluorescent light bulbs.[66] The use of mercury in medicine has greatly declined in all respects; however, mercury compounds are found in some over-the-counter drugs, including topical antiseptics, stimulant laxatives, diaper rash ointment, eye drops, and nasal sprays. Mercury is widely used in small concentrations in the production of eye cosmetics, especially mascara.

Absorption, Transport, and Excretion

Routes of exposure include: (1) inhalation, primarily as elemental mercury vapor but occasionally as dimethyl mercury; (2) ingestion of mercury-containing foods such as predatory fish species; (3) cutaneous absorption of methyl mercury through the skin and even through latex gloves; (4) injection of relatively inert liquid mercury and mercury-containing tattoo pigments; and (5) dental amalgams. Inhaled mercury vapor is retained in the lungs to about 80%, whereas liquid metallic mercury passes through the gastrointestinal tract largely unabsorbed.[67]

Mercury enters the food chain primarily by volcanic activity and man-made sources such as coal combustion and smelting. Most of the dietary intake comes from consumption of meat and fish products, with estimates of dietary intake varying based upon geographical location and dietary sources.[66] The kidney is the major storage organ after elemental or inorganic mercury exposure. Methyl mercury is efficiently absorbed from the gastrointestinal tract, and distribution to tissues, including the brain, appears complete in 48 hours. Movement of methyl mercury across the blood–brain barrier appears to be dependent on coupling with the amino acid cysteine.[68]

There is relatively little bioaccumulation of inorganic and elemental mercury. Half-lives vary according to the route of exposure and form of mercury, from 5 days in blood for phenylmercury to 90 days in urine for chronic exposure to inorganic mercury. Normally, the highest accumulation of mercury is in the kidney, liver, spleen, and brain. Mercury can

accumulate in the pituitary and thyroid glands, the pancreas, and the reproductive organs. The bulk of mercury accumulated in the body is eliminated in approximately 60 days; however, organic forms of mercury can accumulate in the brain and may take up to several years to be eliminated. Fecal and urinary excretions are the main elimination routes for inorganic and organic mercury. A special form of elimination is the transfer of mercury from mother to fetus through the placenta.

Health Effects and Toxicity

Mercury has no known function in normal human physiology. Toxicities have been observed following inhalation, ingestion, and dermal absorption of mercury compounds. Mercurial salts were historically used as diuretics, topical disinfectants, and laxatives before mercury toxicity was well understood. Since the 1930s, some vaccines have contained the preservative thimerosal, a mercury-containing compound metabolized into ethyl mercury. Although it was widely speculated that this mercury-based preservative could cause or trigger autism in children, scientific studies show no evidence supporting any such link. Nevertheless, thimerosal has been removed from or reduced to trace amounts in all U.S. vaccines recommended for children 6 years of age and younger, except for the inactivated influenza vaccine.

Organic mercury and elemental mercury vapor are toxic to both the central and peripheral nervous systems. Mercury attacks the CNS well before an individual shows symptoms. Elemental mercury readily vaporizes, and its inhalation can produce harmful effects on the nervous, digestive, and immune systems as well as the lungs and kidneys. The inorganic salts of mercury can affect the skin, eyes, gastrointestinal tract, and kidneys.

The toxicity of mercury is primarily through reaction with protein sulfhydryl groups, resulting in dysfunction and inactivation. Liquid elemental mercury is poorly absorbed and relatively nontoxic, while elemental mercury vapor is highly absorbed and is highly toxic. Inorganic, ionized forms of mercury are toxic. Further bioconversion to an alkyl mercury, such as methyl mercury, yields a very toxic species of mercury that is highly selective for lipid-rich mediums such as the brain.[69]

Mercury toxicity can manifest in many signs and symptoms that affect several organ systems, including headache, tremor, impaired coordination, abdominal cramps, diarrhea, dermatitis, polyneuropathy, proteinuria, and hepatic dysfunction.[70] Because many of these are relatively nonspecific signs and symptoms, laboratory testing provides a key role in assessing mercury intoxication.

Laboratory Evaluation of Mercury Status

Mercury is usually determined as total mercury levels in blood and urine without regard to chemical form. Hair analysis is also used to establish a timeline of mercury exposure, although this can be complicated by contamination concerns.[7] Analytical methods include ICP-MS and cold vapor atomic absorption spectroscopy (CV-AAS) for environmental samples.

Reference Range

$<$10 ng/mL

Molybdenum

Molybdenum (Mo) is a hard, silvery-white metal occurring naturally as molybdenite, wulfenite, and powellite. Most molybdenum is used for the production of alloys, as well as catalysts, corrosion inhibitors, flame retardants, smoke depressants, lubricants, and molybdenum blue pigments. Molybdenum is an essential trace element important in molybdenum-containing organic compounds identified in biological systems over 80 years ago.[71]

Absorption, Transport, and Excretion

Between 25% and 80% of ingested molybdenum is absorbed predominately in the stomach and small intestine,[29] with the majority of absorbed molybdenum retained in the liver, skeleton, and kidneys. In blood, molybdenum is extensively bound to α_2-macroglobulin and to red blood cell membranes.[71] Molybdenum can cross the placental barrier, and high levels of molybdenum in the diet of the mother can increase the molybdenum in the liver of the neonate.[72] Urine output of Mo correlates to dietary intake, with homeostasis being regulated by the kidneys.[2]

Health Effects, Deficiency, and Toxicity

Molybdenum is vital to human health through its inclusion in at least three enzymes: xanthine oxidase, aldehyde oxidase, and sulfite oxidase. Molybdenum is a necessary cofactor for the function of these enzymes, forming a molybdopterin complex.[2]

Dietary molybdenum deficiency is rare, with a single case reported as a result of total parenteral nutrition in a man with Crohn's disease.[71] Molybdenum cofactor deficiency is a recessively inherited error of metabolism due to a lack of functional molybdopterin. The symptoms include seizures, anterior lens dislocation, decreased brain weight, and usually death prior to 1 year of age.[9]

Molybdenum toxicity is rarely reported as there are few known cases of human exposure to excess molybdenum. High dietary and occupational exposures to molybdenum have been linked to elevated uric acid in blood and an increased incidence of gout.[29]

Molybdenum is rapidly eliminated in both urine and bile, with urine excretion predominating when intake is high.[71]

Laboratory Evaluation of Molybdenum Status

Molybdenum levels are commonly measured by ICP-MS. Both GFAAS and NAA may also be used. Levels of molybdenum in whole blood, plasma, and serum are too low to be measured to detect deficiency. Urinary output is used to monitor for changes in input.[2,71]

Reference Range

0.3–2.0 ng/mL

Selenium

Selenium (Se) is an essential metalloid with many chemical and physical properties similar to those of sulfur. Selenium functions biologically through selenoproteins.[73] Selenium is a major constituent of 40 minerals and a minor constituent of 37 others.[74] Most processed selenium is used in the electronics industry; however, other uses include nutritional supplements, pigments, pesticides, rubber production, antidandruff shampoos, and fungicides.

Absorption, Transport, and Excretion

Selenium is well absorbed from the gastrointestinal tract (~50%). Selenium exposure occurs primarily from food but can be found in drinking water, usually in the form of inorganic sodium selenate or sodium selenite. Selenium homeostasis is largely achieved by excretion via urine. Other routes or elimination include sweat and, at very high intakes, exhalation of volatile forms of selenium.[75]

Health Effects, Deficiency, and Toxicity

In the 1930s, selenium was considered a toxic element; in the 1940s, a carcinogen; in the 1950s, it was declared an essential element; and since the 1960s and especially the 1970s, it has been viewed as an anticarcinogen. Glutathione peroxidase (in the form of selenocysteine) is part of the cellular antioxidant defense system against free radicals,[29] and selenium is also involved in the metabolism of thyroid hormones (e.g., deiodinase enzymes and thioredoxin reductase).[75,76]

Selenium deficiency has been associated with cardiomyopathy, skeletal muscle weakness, and osteoarthritis. A significant negative correlation was observed between selenium intakes and the rates of leukemia and cancers of the large intestine, rectum, prostate, breast, ovary, and lungs.[74]

Keshan disease, an endemic cardiomyopathy that affects mostly children and women of childbearing age in certain areas in China, has been associated with selenium deficiency and the resultant increased virulence of the coxsackievirus.[2] Symptoms include dizziness, malaise, loss of appetite, nausea, chills, abnormal electrocardiograms, cardiogenic shock, cardiac enlargements, and congestive heart failure. Selenium supplementation has been shown to effectively control Keshan disease.[29] Kashin-Beck disease, an endemic osteoarthritis that occurs during adolescent and preadolescent years, is another disease linked to low selenium status in northern China, North Korea, and eastern Siberia.[77]

Acute oral exposure to extremely high levels of selenium may produce gastrointestinal symptoms (nausea, vomiting, and diarrhea) and cardiovascular symptoms such as tachycardia. Chronic exposure to very high levels can cause dermal effects, including diseased nails and skin and hair loss, as well as neurologic problems such as unsteady gait or paralysis.[75] In 1984, 12 cases of selenium toxicity were reported to the FDA and CDC, caused by the ingestion of selenium supplements containing levels almost 200 times higher than stated on the label. The most common symptoms reported in these cases were nausea and vomiting, nail changes, hair loss, fatigue, abdominal cramps, watery diarrhea, and garlicky breath. No abnormalities of blood chemistry were seen in 67% of the victims, and renal and liver functions were normal.[78]

The EPA has determined that one specific form of selenium, selenium sulfide, is a probable human carcinogen. Selenium sulfide is a very different chemical from the organic and inorganic selenium compounds

found in foods and in the environment.[77] In Hubei Province (China) during 1961 through 1964, almost half of the population of many villages died from chronic selenosis. The most common signs of in these cases of selenium poisoning were loss of hair and nails, skin lesions, tooth decay, and abnormalities of the nervous system.[74]

Laboratory Evaluation of Selenium Status

Selenium is most often determined by ICP-MS, GFAAS, or HGAAS. Inductively coupled plasma-triple quad mass spectrometer (ICP-QQQ-MS) are now being used for its measurement. The determination of urinary selenium is a useful measure of dietary intake. Blood selenium is the primary measure of selenium status.[2]

Reference Range

Over 18 years old: 150–241 ng/mL

Zinc

Zinc (Zn) is a bluish-white, lustrous metal that is stable in dry air and becomes covered with a white coating when exposed to moisture. Zinc is used in a production of alloys, especially brass (with copper), in galvanizing steel, in die casting, in paints, in skin lotions, as treatment for Wilson's disease, and in many over-the-counter medications. Zinc is a plentiful essential trace element in the human body. Deficiency is common throughout life, especially in individuals who do not ingest meat.

Absorption, Transport, and Excretion

The body content in a normal individual varies substantially with age and is predominantly distributed in muscle (60%) and the skeleton (30%). The remaining 10% is distributed in various other tissues with highest concentrations found in the eyes, prostate, and hair. Zinc absorption mainly occurs in the small intestine and especially in the jejunum.[79] Factors increasing zinc absorption include the presence of animal proteins[80] and amino acids in a meal,[81] intake of calcium,[82] and unsaturated fatty acids.[83] Conversely, factors decreasing zinc absorption include the intake of iron,[84,85] taking zinc on an empty stomach,[83] presence of copper at high levels,[86] and potentially age.[87] In blood, the absorbed zinc is distributed between red blood cells (80%), plasma (17%), and white blood cells (3%).[88] In normal dietary circumstances, about 90% of zinc is excreted in feces.[89]

Health Effects, Deficiency, and Toxicity

Zinc is second only to iron in abundance as an essential trace element in the body. The main biochemical role of zinc is seen in its influence on the activity of more than 300 enzymes in classes such as oxidoreductases, transferases, hydrolases, leases, isomerases, and lipases. As a result of the importance of zinc for the structure, regulation, and catalytic action of various enzymes, zinc is indirectly involved in the synthesis and metabolism of DNA and RNA, the synthesis and metabolism of proteins, the metabolism of glucose and cholesterol, membrane structure maintenance, insulin function, and growth factor effects.[79] Chronic oral zinc supplementation interferes with copper absorption and may cause copper deficiency, forming the basis for using zinc to treat Wilson's disease.

Zinc deficiency causes growth retardation and testicular atrophy, slows skeletal maturation, and reduces taste perception. Old age, pregnancy, lactation, and alcoholism are also associated with poor zinc nutrition.[29] Infants with acrodermatitis enteropathica (zinc malabsorption) first develop a characteristic facial and diaper rash. Untreated, symptoms progress and include growth retardation, diarrhea, impaired T-cell immunity, insufficient wound healing, infections, delayed testicular development in adolescence, and early death.[9] Zn deficiency in adolescents is manifested by slow growth or weight loss, altered taste, delayed puberty, dwarfism, impaired dark adaptation, alopecia, emotional instability, and tremors. In severe cases, lymphopenia and death can result from an overwhelming infection.[9]

Zinc is relatively nontoxic. Nevertheless, high doses (1 g) or repetitive doses of 100 mg/d for several months may lead to gastrointestinal tract symptoms, decrease in heme synthesis due to an induced copper deficiency, and hyperglycemia.[79] Exposure to ZnO fumes and dust may cause "zinc fume fever," with symptoms including chemically induced pneumonia, severe pulmonary inflammation, fever, hyperpnea, coughing, pains in legs and chest, and vomiting.[79]

Laboratory Evaluation of Zinc Status

Zinc is routinely measured by ICP-MS, FAAS, GFAAS, and ICP-AES.[90,91] Zinc is typically measured in serum, plasma, and urine,[2] while zinc protoporphyrin is measured using whole blood. Low urine zinc levels in the

presence of low serum zinc levels usually confirms zinc deficiency.[9] Low serum zinc in an apparently healthy (nonstressed, nonseptic) patient who has normal serum albumin levels can be used as evidence of zinc deficiency, especially if urine zinc levels are also low. Normal serum zinc cannot be interpreted as evidence of normal zinc stores. Zinc concentration in red blood cells is approximately 10 times that in serum.[29] Copper status should be monitored in patients on long-term zinc therapy.[9,87]

Reference Range

Over 11 years old: 0.66–1.10 µg/mL

Vitamins

Overview

The second section of this chapter provides an overview of vitamins that may be encountered in the clinical laboratory. Vitamins have a wide range of functions in biologic tissue, serving as **cofactors** in many enzymatic reactions, so that these enzymes have low catalytic activity in cellular reactions if vitamins are not present. These compounds and their biologically inactive precursors must be partially obtained from food sources and, in some instances, from bacterial synthesis. When cellular vitamin and activity levels from diet or intestinal absorption are inadequate, it is termed vitamin deficiency. The term **vitamin** has a historical basis in deficiency states that were relieved by specific food intake. The most notable examples are scurvy (vitamin C, sailors and lime consumption,), rickets (vitamin D in the Industrial Revolution), beriberi (thiamine in alcoholics), pellagra (niacin), night blindness (vitamin A), megaloblastic anemia (folic acid or vitamin B_{12}), spina bifida (folic acid), and pernicious anemia with neuropathy (vitamin B_{12}). Abnormal increases of metabolism requiring high supplies of one of these cofactors may be termed *vitamin insufficiency* or *vitamin dependency*, depending on the level of supply demanded for physiologic function.[92,93]

Variabilities in clinical expression of vitamin abnormalities result from differences in any of the following: specific cause, degree, and duration of vitamin inadequacy; the simultaneous presence of nutritional insufficiencies; and/or the increased metabolic demands imposed by conditions such as pregnancy, infection, and cancer. The clinical symptoms of vitamin deficiencies are usually nonspecific in early stages and in mild, chronic deficiency states. A combination of clinical findings, dietary history, nutrition-focused physical examination, and laboratory measurements is often required to diagnose a vitamin deficiency. Vitamin metabolism is complex, and vitamin supplementation of foods is common. It is not unusual to find vitamin toxicities from the inappropriate use of vitamin supplementation.

For simplicity, vitamins of diverse chemical structure are classified as either fat-soluble or water-soluble. Fat-soluble vitamins include A, D, E, and K. Those vitamins soluble in water include the B complex of vitamins—thiamine, riboflavin, niacin, vitamins B_6 and B_{12}, biotin, and folic acid—and vitamin C. Water-soluble vitamins are readily excreted in the urine and are less likely than fat-soluble vitamins to accumulate to toxic levels in the body. Vitamins, classified as fat- or water-soluble, and the symptoms usually seen in deficiency states, are shown in **Table 27.4**.[2, 94-99]

Investigating the dietary deficiency of vitamins (**hypovitaminosis**) is sustained primarily from knowledge of dietary sources and dietary practices that produce inadequate intake or absorption. In the early 1990s, the Food and Nutrition Board of the Institute of Medicine began the process of reviewing the previous nutritional recommendations known as recommended dietary allowances (RDAs) to also include the prevention of chronic diseases. In addition to the RDA, there are additional categories such as estimated average requirement, adequate intake (AI), and tolerable upper intake level (UL).[95,100,101]

Chemical determination of human vitamin states has been approached through measurement of the following:

- Active cofactors or precursors in biologic fluids or blood cells
- A biochemical function requiring the vitamin (e.g., enzymatic activity), with and without in vitro addition of the cofactor form
- Urinary metabolites of the vitamin
- Urinary excretion of vitamin or metabolites after a test load of the vitamin
- Urinary metabolites of a substance, the metabolism of which requires the vitamin after administration of a test load of the substance

Reduced serum concentrations of a vitamin do not always indicate a deficiency that interrupts cellular function. Conversely, results within the reference range do not always reflect adequate function. Interpretation of laboratory results must be done with knowledge of the biochemistry and physiology of vitamins.[102-104]

Table 27.4 Vitamins and Deficiency States

Vitamin Name	Clinical Deficiency
Fat-Soluble Vitamins	
Vitamin A	Night blindness, growth retardation, abnormal taste response, dermatitis, recurrent infections
Vitamin D	Rickets (children), osteomalacia (adult)
Vitamin E	Red cell fragility leading to hemolytic anemia (premature infants), ataxia
Vitamin K	Decreased ability to form blood clots, hemorrhage (ranging from easy bruising to massive bruising)
Water-Soluble Vitamins	
Vitamin B_1 (thiamine)	Infants: cardiac failure, cyanosis, vomiting Adults: anorexia, irritability, memory loss, peripheral neuropathy, ataxia, tachycardia, peripheral edema
Vitamin B_2 (riboflavin)	Angular stomatitis (mouth lesions), dermatitis, photophobia, neurologic changes
Vitamin B_3 (niacin/nicotinamide)	Pellagra (dermatitis, mucous membrane inflammation, weight loss, disorientation)
Vitamin B_5	Fatigue, headaches, muscle cramps, nausea, malaise
Vitamin B_6	Seizures (young), normocytic anemia, dermatitis, glossitis, depression
Vitamin B_7 (biotin)	Alopecia, dermatitis, hypotonia, seizures, ataxia, developmental delay (adolescents)
Vitamin B_9 (folic acid)	Megaloblastic anemia, neural tube defects
Vitamin B_{12}	Pernicious and megaloblastic anemia, neurologic abnormalities
Vitamin C (ascorbic acid)	Scurvy, anemia, diabetes mellitus, osteoporosis

Fat-Soluble Vitamins

Vitamin A

Vitamin A is derived directly from dietary sources, primarily as retinyl esters, or from metabolism of dietary carotenoids (provitamin A), primarily beta carotene. Major dietary sources of these compounds include animal products and pigmented fruits and vegetables (carotenoids). Vitamin A is stored in the liver and transported in the circulation complexed to retinol-binding protein (RBP) and its carrier protein transthyretin. Vitamin A and related retinoic acids are a group of compounds essential for vision, cellular differentiation, growth, reproduction, and immune system function. A clearly defined physiologic role for retinol is in vision. Retinol is oxidized in the rods of the eye to retinal, which, when complexed with opsin, forms rhodopsin, allowing dim light vision. This vitamin and vitamin D act through specific nuclear receptors in the regulation of cell proliferation. Vitamin A deficiency leads to night blindness (nyctalopia) and, when prolonged, may cause total blindness. In vitamin A deficient states, epithelial cells (cells in the outer skin layers and cells in the lining of the gastrointestinal, respiratory, and urogenital tracts) become dry and keratinized. Fruits and vegetables contain carotene, which is a precursor of retinol. Carotenes provide more than one-half of the retinol requirements in the American diet. Vitamin A deficiency is most common among children living in nonindustrialized countries and is usually a result of insufficient dietary intake. Deficiency may also occur because of chronic fat malabsorption or impaired liver function or may be associated with severe stress and protein malnutrition. Premature infants are born with lower serum retinol and RBP levels, as well as lower hepatic stores of retinol; therefore, these newborns are treated with vitamin A as a preventive measure.[104,105]

When ingested in high doses, either chronically or acutely, vitamin A causes many toxic manifestations and may ultimately lead to liver damage due to **hypervitaminosis**. High doses of vitamin A may be obtained from excessive ingestion of vitamin supplements or large amounts of liver or fish oils, which are rich in vitamin A. Carotenoids, however, are not known to be toxic because of a reduced efficiency of carotene absorption at high doses and limited conversion to vitamin A. The RDA of vitamin A is 900 µg retinol activity equivalents (RAE) per day for adult males and 700 µg RAE per day for adult females.[95] Measurement of plasma retinol is the common means of assessing vitamin A status in the clinical setting. Retinol is most commonly measured by high-performance liquid chromatography (HPLC). Toxicity is usually assessed by measuring retinyl ester levels in serum rather than retinol, which is accomplished by HPLC.[105]

Vitamin D

Vitamin D refers to a group of related metabolites used for proper skeleton formation and mineral homeostasis. Exposure of the skin to sunlight (ultraviolet light) catalyzes the formation of cholecalciferol (vitamin D_3) from 7-dehydrocholesterol. The other major form of vitamin D is ergocalciferol (vitamin D_2). Vitamin D occurs in foods as cholecalciferol or ergocalciferol. The most active metabolite of vitamin D is $1,25(OH)_2D_3$. $1,25(OH)_2D_3$ stimulates intestinal absorption of calcium and phosphate for bone growth and metabolism and, together with parathyroid hormone, stimulates bone to increase the mobilization of calcium and phosphate. $1,25(OH)_2D_3$ has an important proapoptotic effect, acting through a vitamin D hormonal system, that depends on binding of the active ligand to a vitamin D receptor. This led to important drug discovery developments in which calcium and phosphate release is minimized, and proliferative and anti-inflammatory effects of D-analogues are modulated.[106]

In northern climates, it is difficult to receive enough ultraviolet light exposure to fully meet minimum requirements. Major dietary sources of vitamin D include irradiated foods and commercially prepared milk. Small amounts are found in butter, egg yolks, liver, sardines, herring, tuna, and salmon. The RDA of vitamin D for adults is 15 to 20 µg, depending on age.[105] Absorbed in the small intestine, vitamin D requires bile salts for absorption. It is stored in the liver and adipose tissue and excreted in the bile. Severe deficiency in children causes a failure to calcify cartilage at the growth plate in metaphysical bone formation, leading to the development of rickets. In adults, the deficiency leads to under mineralization of bone matrix in remodeling, resulting in osteomalacia. Low levels of vitamin D are reported with the use of anticonvulsant drugs and in small bowel disease, chronic renal failure, hepatobiliary disease, pancreatic insufficiency, and hypoparathyroidism. Vitamin D can be toxic, especially in children. Elevated levels of vitamin D are present in hyperparathyroidism, hypophosphatemia, and during pregnancy. Excess vitamin D produces hypercalcemia and hypercalciuria, which can lead to calcium deposits in soft tissue and irreversible renal and cardiac damage.[106]

Serum 25-hydroxyvitamin D [25(OH)D] is the most commonly measured indicator of vitamin D status. Immunoassays on automated platforms are arguably the most common method of measuring 25(OH)D levels, followed by liquid chromatography coupled to tandem-mass spectrometry (LC-MS/MS).[107,108] $1,25(OH)_2D$, the active form of vitamin D, is not often measured as an indicator of vitamin D status due to its short half-life in the body (3–4 hours).[109] Serum $1,25(OH)_2D$ measurement is necessary in some circumstances including to assess vitamin D status in patients with renal disease and some forms of rickets, and for the differential diagnosis of hypercalcemia.[110] Methods or measurement also include immunoassay and LC-MS/MS.[110,111] The release from the National Standards Bureau of a 25(OH)D reference material has been predicted to improve correlation between all methods for vitamin D in the future.[112] For more information on Vitamin D, please refer to Chapter 18, *Parathyroid Function*.

Vitamin E

Vitamin E is a powerful antioxidant and the primary defense against potentially harmful oxidation that causes disease and aging, as well as protecting unsaturated lipids from peroxidation (cleavage of fatty acids at unsaturated sites by oxygen addition across the double bond and formation of free radicals). The role of vitamin E in protecting the erythrocyte membrane from oxidant stress is presently its major documented role in human physiology. It has been shown to strengthen cell membranes and augment such functions as drug metabolism, heme biosynthesis, and neuromuscular function. Tocopherols are the major form of vitamin E.[113,114] Alphatocopherol is the predominant isomer in plasma and the most potent isomer by current biologic assays. About 40% of ingested vitamin E is absorbed, affected mainly by

the amount and degree of unsaturated dietary fat, largely determining the physiologic requirement. Absorbed vitamin E is associated with circulating chylomicrons, very-low-density lipoprotein, and chylomicron remnants. Dietary sources of tocopherols include vegetable oil, fresh leafy vegetables, egg yolk, legumes, peanuts, and margarine. Diets contributing to vitamin E deficiency are those low in vegetable oils, fresh green vegetables, or unsaturated fats.

The major symptom of vitamin E deficiency is hemolytic anemia. Although the use is controversial due to the increased risk of infections, premature newborns are sometimes supplemented with vitamin E to stabilize red blood cells and prevent hemolytic anemia.[115] There is evidence for preventive roles of vitamin E in retrolental fibroplasia, intraventricular hemorrhage, and mortality of small, premature infants. Premature infants receiving vitamin E in amounts that sustain serum levels above 30 mg/L have an increased incidence of sepsis and necrotizing enterocolitis.[113]

Patients with conditions that result in fat malabsorption, especially cystic fibrosis and abetalipoproteinemia, are also susceptible to vitamin E deficiency.[116] A relationship exists between vitamin E deficiency and progressive loss of neurologic function in infants and children with chronic cholestasis.[116] Absorption of dietary vitamin E is most efficient in the small intestine, where it combines with lipoproteins and is transported through the lymphatics.[2,116] Vitamin E is stored in the liver and other tissues with high lipid content and excreted in urine and feces.[116,117] Assessment of vitamin E status is primarily indicated in premature infants, patients with fat malabsorption states, and patients with motor and sensory neuropathies.[118] Although megadoses of vitamin E do not produce toxic effects, high doses have no proven health benefit. The RDA of vitamin E is 15 mg/d for adults.[95] Tocopherols are commonly measured in the laboratory using HPLC methods.[2,113]

Vitamin K

Vitamin K (from the German word "koagulation") is the group of compounds essential for the formation of prothrombin (factor II) and at least five other coagulation proteins, including factors VII, IX, and X and proteins C and S. The quinone-containing compounds are a generic description for menadione and derivatives exhibiting this activity. Vitamin K helps convert precursor forms of these coagulation proteins to functional forms, which occurs in the liver. Dietary vitamin K is absorbed primarily in the small intestine.[2] Vitamin K is synthesized by intestinal bacteria; providing 50% of the vitamin K requirement. Major dietary sources are cabbage, cauliflower, spinach and other leafy vegetables, pork, liver, soybeans, and vegetable oils. Dietary vitamin K deficiency is considered rare in healthy children and adults.

Vitamin K deficiency may be caused by antibiotic therapy, which results from decreased synthesis of the vitamin by intestinal bacteria. When vitamin K antagonists, such as warfarin sodium (Coumadin), are used for anticoagulant therapy, anticoagulant factors II, VII, IX, and X are synthesized but nonfunctional; and may lead to a hemorrhagic episode.[119-121] Several herbal supplements (e.g., garlic, gingko, and ginseng) may enhance the effects of warfarin or interact with platelets, increasing the risk of bleeding.

Prothrombin time (velocity of clotting after addition of thromboplastin and calcium to citrated plasma) determination is an excellent index of prothrombin adequacy. Prothrombin time is prolonged in vitamin K deficiency and in liver diseases characterized by decreased synthesis of prothrombin. Vitamin K deficiency also results in prolongation of the partial thromboplastin time, but the thrombin time is within the reference range.

Toxicity from vitamin K is not commonly seen in adults. Large doses in infants may result in hyperbilirubinemia. The adult AI for vitamin K is 120 µg/d for males and 90 µg/d for females.[95] In most laboratories, vitamin K is not assayed; however, prothrombin time is used as a functional indicator of vitamin K status.[121]

Water-Soluble Vitamins

Vitamin B_1

Vitamin B_1 (thiamine) acts as a coenzyme (thiamine pyrophosphate [TPP]) in decarboxylation reactions in major carbohydrate pathways and in branched-chain amino acid metabolism. It is rapidly absorbed from food in the small intestine and excreted in the urine.[122] Free thiamine circulates in plasma, whereas TPP is stored primarily in the skeletal muscle as well as the heart, liver, kidneys, and brain.[2] The clinical condition associated with chronic thiamine deficiency is beriberi. Although usually occurring in underdeveloped countries, beriberi may be found in the United States among persons with chronic alcoholism. Decreased intake, impaired absorption, and increased requirements all appear to play a role in the development of thiamine deficiency in persons with alcoholism. The RDA of thiamine is 1.2 mg/d for adult males and 1.1 mg/d for adult females.[95] Thiamine functional

activity is best measured by erythrocyte transketolase activity, before and after the addition of TPP. Thiamine deficiency is present if the increase in activity after the addition of TPP is greater than 25%.[122]

Vitamin B_2

Vitamin B_2 (riboflavin) functions primarily as a component of two coenzymes: flavin mononucleotide and flavin adenine dinucleotide. These two coenzymes catalyze various oxidation–reduction reactions. Dietary riboflavin is absorbed in the small intestine. The body stores of a well-nourished person are adequate to prevent riboflavin deficiency for 5 months. Excess riboflavin is excreted in the urine and has no known toxicity. Foods high in riboflavin include milk, liver, eggs, meat, and leafy vegetables. Riboflavin deficiency occurs with other nutritional deficiencies, alcoholism, and chronic diarrhea and malabsorption. Certain drugs antagonize the action or metabolism of riboflavin, including phenothiazine, oral contraceptives, and tricyclic antidepressants.[123] The RDA of riboflavin is 1.3 mg/d for adult males and 1.1 mg/d for adult females.[95] Reduced glutathione reductase activity greater than 40% is an indication of deficiency.

Vitamin B_3

The requirement for vitamin B_3 (niacin) in humans is met, like many vitamins, through dietary sources including foods that contain coenzymes and by the conversion of dietary tryptophan to niacin.[2] Niacin is the generic term for both nicotinic acid and nicotinamide. Niacin functions as a component of the two coenzymes nicotinamide adenine dinucleotide (NAD^+) and nicotinamide adenine dinucleotide phosphate ($NADP^+$), which are necessary for many metabolic processes including tissue respiration, lipid metabolism, fatty acid metabolism, and glycolysis. Reduction of the coenzymes yield NADH and NADPH, which have a strong absorption at 340 nm, a feature widely used in assays of pyridine nucleotide-dependent enzymes.

Niacin is absorbed in the stomach and small intestine, and excess is excreted in the form of metabolites in the urine.[2,124] Pellagra, the clinical syndrome resulting from niacin deficiency, is associated with diarrhea, dementia, dermatitis, and death. Niacin deficiency may result from alcoholism. To decrease lipid levels, pharmacologic doses of nicotinic acid are given therapeutically. The toxicity of niacin is low. When large doses are ingested, however, as often occurs during lipid-lowering therapies, flushing of the skin and vasodilation may occur. The RDA of niacin is expressed as mg of niacin equivalents (NE) and is 16 mg NE/d for adult males and 14 mg NE/d for adult females.[95] Blood or urinary niacin levels are of value in assessing niacin nutritional status.

Vitamin B_5

Vitamin B_5 is a growth factor occurring in all types of animal and plant tissue and is also called *pantothenic acid* (from Greek for "everywhere"). Dietary sources include liver and other organ meats, milk, eggs, peanuts, legumes, mushrooms, salmon, and whole grains. Approximately 50% of pantothenate in food is available for absorption. Pantothenate is metabolically converted to 4′-phosphopantetheine, which becomes covalently bound to either serum acyl carrier protein or coenzyme A. Coenzyme A is a highly important acyl group transfer coenzyme involved in many reaction types. Deficiency in vitamin B_5 is rare and results in general symptoms such as fatigue, headaches, muscle cramps, nausea, and malaise.[98] The AI for pantothenic acid in adults is 5 mg/d.[95] Pantothenic acid can be measured in whole blood or urine and employ microbiological, RIA, and mass spectrometry-based methods.[2]

Vitamin B_6

Vitamin B_6 is ubiquitous and includes three related compounds: pyridoxine, pyridoxal and pyridoxamine. The major dietary sources of vitamin B_6 are meat, poultry, fish, potatoes, and vegetables; dairy products and grains contribute lesser amounts. Readily absorbed from the intestinal tract, vitamin B_6 is excreted in the urine in the form of metabolites.[123] Vitamin B_6 deficiency rarely occurs alone; it is more commonly seen in patients deficient in several B vitamins. Those particularly at risk for deficiency are patients with uremia, liver disease, absorption syndromes, malignancies, or chronic alcoholism. High intake of proteins increases the requirements for vitamin B_6. Deficiency is associated with hyperhomocysteinemia. Vitamin B_6 has low toxicity because of its water-soluble nature; however, extremely high doses may cause neurotoxicity. The RDA of vitamin B_6 is 1.3 to 1.7 mg/d for adult males and 1.3 to 1.5 mg/d for adult females, depending on age.[95]

Vitamin B_7

Vitamin B_7 (biotin) is a coenzyme for several enzymes that transport carboxyl units in tissue and plays an integral role in gluconeogenesis, lipogenesis, and fatty acid synthesis.[125] Dietary biotin is absorbed

in the small intestine, but it is also synthesized in the gut by bacteria. Numerous foods contain biotin, although no food is especially rich (up to 20 µg/100 g). The dietary intake of biotin, while low in the neonatal period, increases as newborns switch from colostrum to mature breast milk. Biotin deficiency can be induced by ingestion of large amounts of avidin (found in raw egg whites) that binds to biotin. Biotin deficiency has been noted in patients receiving long-term parenteral nutrition and in infants with genetic defects of carboxylase and biotinidase enzymes. Biotin deficiency often affects the skin and includes alopecia and dermatitis as well as neurological symptoms such as hypotonia, seizures, ataxia, and developmental delay in adolescents.[99] The AI for biotin is 30 µg/d.[95]

Microbiological assays to measure biotin have been performed using organisms such as *Lactobacillus plantarum*.[2] Newer methods of isotopic dilution, chemiluminescent, and photometric assays are now available but rarely used in hospital laboratories. Specimens are often sent to a reference laboratory for analysis.[126]

Vitamin B_9

Vitamin B_9 (folate) is the generic term for components nutritionally and chemically similar to folic acid. Folate functions metabolically as coenzymes involved in various one carbon transfer reactions. Folate and vitamin B_{12} are closely related metabolically. The hematologic changes that result from deficiency of either vitamin are indistinguishable. Folate from dietary sources is absorbed in the small intestine, and the excess is excreted in the urine. Large quantities of folate are also synthesized by bacteria in the colon. Structural relatives of pteroylglutamic acid (folic acid) are metabolically active compounds usually referred to as *folates*. Food folates are primarily found in green and leafy vegetables, fruits, organ meats, and yeast. Boiling food and using large quantities of water result in folate destruction. The average American diet may be inadequate in folate for adolescents and for pregnant or lactating women.[127]

The major clinical symptom of folate deficiency is megaloblastic anemia. Laboratory indices of deficiency are low serum folate, hypersegmentation of neutrophils, high urinary formiminoglutamic acid (a histidine metabolite accumulating in the absence of folate), low erythrocyte folate, macroovalocytosis, megaloblastic marrow, and anemia. Serum folate levels, although an early index of deficiency, can frequently be low despite normal tissue stores. Because most folate storage occurs after the vitamin B_{12}–dependent step, erythrocyte folate can also be reduced in deficiency of either vitamin B_{12} or folate. Despite this overlap, erythrocyte folate concentration is accepted as the best laboratory index of folate deficiency.[128] Many physicians order both serum and erythrocyte folate levels because serum levels indicate circulating folate, and erythrocyte levels better approximate stores. Urine folate levels are not typically measured to discern folate status.[2] Homocysteine elevation in serum and urine occurs in folate deficiency.[129] Total homocysteine is generally measured, which is the sum of all homocysteine species, both free and protein-bound forms.[130]

Folate requirement is increased during pregnancy and especially during lactation. The increase during lactation results, in part, from the presence of high-affinity folate binders in milk. Dietary supplementation of folate in pregnant women may reduce the incidence of fetal neural tube defects. Other instances of increased folate requirement include hemolytic anemia, iron deficiency, prematurity, and multiple myeloma. Patients receiving dialysis treatment rapidly lose folate. Clinical conditions associated with folate deficiency include megaloblastic anemia, alcoholism, malabsorption syndrome, carcinoma, liver disease, chronic hemodialysis, and hemolytic and sideroblastic anemia.[130,131] Certain anticonvulsants and other drugs that interfere with folate metabolism include sulfasalazine, isoniazid, and cycloserine. Folate deficiency of dietary origin commonly occurs in older persons. Phenytoin (Dilantin) therapy accelerates folate excretion and interferes with folate absorption and metabolism. Alcohol interferes with folate's enterohepatic circulation, and methotrexate, a chemotherapeutic agent, inhibits the enzyme dihydrofolate reductase. Low levels of serum folate can occur with use of oral contraceptives. See **Table 27.5** for a list of medications and their interactions with vitamins; note this list is not exhaustive.

There are no known cases of folate toxicity from dietary sources.[2] The RDA (in dietary folate equivalents, DFE) is 400 µg DFE/d for adult males and females.[95] In women of childbearing age, 400 µg/d of folate is recommended to prevent or reduce the incidence of neural tube defects.[132]

Folate levels may be measured in serum using a microbiologic assay with *Lactobacillus casei* or, more commonly, a competitive protein-binding assay for levels in serum and erythrocytes. When folate deficiency develops, serum levels fall first, followed by a decrease in erythrocyte folate levels and

Table 27.5 Actions of Drugs and Oral Contraceptives on Vitamins

Drug or Nutrient
Pyridoxine—antagonized by isoniazid, steroids, penicillamine
Riboflavin—antagonized by phenothiazines, some antibiotics
Folate—antagonized by phenytoin, alcohol, methotrexate, trimethoprim
Ascorbate—antagonized by (increased excretion) aspirin, barbiturates, hydantoins
Ascorbate excess—interferes with actions of aminosalicylic acid, tricyclic antidepressants, anticoagulants; may cause "rebound scurvy" on withdrawal
Oral Contraceptive Agent Cause
Increased serum vitamin A, retinol-binding protein
Decreased requirement for vitamins K and C
Decreased vitamin B_6 status indices
Decreased riboflavin use
Increased niacin pathway of tryptophan
Decreased induction of thiamin deficiency
Decreased serum folate (cycle-day dependent)
Decreased induction of cervical folate deficiency

ultimately hematologic manifestation.[130,131] Measuring both serum and erythrocyte levels is helpful because serum levels indicate circulating folate, and erythrocyte levels better approximate stores.[128]

Serum contains endogenous binding proteins that can bind folate and result in falsely low serum folate concentration measurements. Although measurement of red blood cell folate concentration has advantages over the serum assay for the diagnosis of megaloblastic anemia, analytic problems may result from the various forms of folate in erythrocytes. Folate in the serum is almost exclusively present in the monoglutamate form; however, in red blood cells, it is in the polyglutamate form and as high-molecular-weight complexes.[130,131]

Vitamin B_{12}

Vitamin B_{12} (cobalamin) refers to a large group of cobalt containing compounds. Intestinal absorption of vitamin B_{12} takes place in the small intestine and is mediated by a unique binding protein called intrinsic factor, which is secreted by the stomach. Vitamin B_{12} participates as a coenzyme in enzymatic reactions necessary for hematopoiesis and fatty acid metabolism. Excess vitamin B_{12} is excreted in the urine. Vitamin B_{12} bears a corrin ring (containing pyrroles similar to porphyrin) linked to a central cobalt atom. Different corrinoid compounds, or cobalamins, are distinguished by the substituent linked to the cobalt. The active cofactor forms of vitamin B_{12} are methylcobalamin and deoxyadenosylcobalamin. The primary dietary sources for vitamin B_{12} are from animal products (e.g., meat, eggs, and milk). Therefore, total vegetarian diets are likely to be deficient or low in vitamin B_{12}. Animals derive vitamin B_{12} from intestinal microbial synthesis. The average daily diet contains 3 to 30 µg of vitamin B_{12}, of which 1 to 5 µg is absorbed. The frequency of dietary deficiency increases with age, occurring in more than 10% to 15% of people older than 60, although the symptoms resulting from dietary deficiency are rare.

CASE STUDY 27.3, PART 2

Remember Martin. He seen by his physician who suspects a deficiency in his diet may be contributing to his neurological symptoms. His physician decides to order vitamin testing for more information.

1. What vitamin deficiency do you suspect in this patient? Why?

Most vitamin B_{12} absorption occurs through a complex with intrinsic factor, a protein secreted by gastric parietal cells. This intrinsic factor–B_{12} complex binds with specific ileal receptors. "Blocking" intrinsic factor antibodies prevents binding of vitamin B_{12} to intrinsic factor, and "binding" antibodies can combine with either free intrinsic factor or the intrinsic factor–B_{12} complex, preventing attachment of the complex to ileal receptors and intestinal uptake of the vitamin. Parietal cell antibodies have also been identified as a cause of pernicious anemia. After release from the intrinsic factor complex within the mucosal cell, vitamin B_{12} circulates in plasma bound to specific transport proteins and is deposited in liver, bone marrow, and other tissues. There is a significant enterohepatic circulation of vitamin B_{12}. Plasma contains both types of transport proteins, transcobalamins, and the three forms of vitamin B_{12} (hydroxocobalamin, methylcobalamin, and deoxyadenosylcobalamin).

The Schilling test may be used to evaluate vitamin B_{12} absorption and is often used to differentiate between pernicious anemia or malabsorption in the intestine as the cause of vitamin B_{12} deficiency.[2] In the Schilling test, the patient receives a small, oral dose of radiolabeled vitamin B_{12}. Parenteral B_{12} is given simultaneously to saturate binding sites. Serum and urine are collected at timed intervals and labeled B_{12} is measured in the specimens. Patients who cannot absorb vitamin B_{12} (for example a deficiency of intrinsic factor, as in pernicious anemia)[131] cannot absorb the labeled B_{12} and, therefore, have low levels in the blood and urine.

The term *pernicious anemia* is now most commonly applied to vitamin B_{12} deficiency resulting from lack of intrinsic factor. Antibodies to intrinsic factor and parietal cells are common in patients with pernicious anemia, their healthy relatives, and patients with other autoimmune disorders. Deficiency of B_{12} can occur in strict vegetarians because of dietary deficiency. A loss of vitamin B_{12} also occurs in individuals infected with fish tapeworm or because of malabsorption diseases, such as celiac disease. Low vitamin B_{12} levels occur with folate deficiency, and a vitamin B_{12} deficiency can be masked by large doses of folate. Toxicity of vitamin B_{12} has not been reported. The RDA of vitamin B_{12} for adults is 2.4 μg/d.[95] Methods for measurement of serum vitamin B_{12} are either microbiologic assay using *Lactobacillus leichmannii*, competitive protein binding, or immunometric.[2]

Deficiency of vitamin B_{12} causes two major disorders—megaloblastic anemia (pernicious anemia) and a neurologic disorder called combined systems disorder.[130] The neurologic manifestations are variable and may be subtle. For this reason, vitamin B_{12} deficiency should be considered a cause of any unexplained macrocytic anemia or neurologic disorder, especially in the elderly. Serum vitamin B_{12} may be used in the initial assessment.[130] Methylmalonic acid (MMA) levels may be more definitive because the lower reference limit of B_{12} is unclear. Patients with pernicious anemia usually have atrophic gastritis and have an increased incidence of gastric carcinoma.

A common method for determination of vitamin B_{12} is a competitive-binding immunoenzymatic assay. The sample is added to the reagents alkaline potassium cyanide and dithiothreitol (DTT). DTT denatures B_{12}-binding proteins and converts all forms of vitamin B_{12} to cyanocobalamin. Next, an intrinsic factor-alkaline phosphatase conjugate and a solid phase with IgG-coated monoclonal anti-intrinsic factor are added to the sample. Vitamin B_{12} in the sample binds to the intrinsic factor conjugate, preventing the conjugate from binding to the solid phase anti-intrinsic factor. After incubation, unbound materials are washed away. Then, a chemiluminescent substrate is added to the reaction well and the light generated by the chemiluminescent reaction is measured by the luminometer. The light emitted is inversely proportional to the concentration of vitamin B_{12} in the sample. This competitive-binding immunoenzymatic assay has replaced the classical radioimmunoassay (RIA) method.

Vitamin C

Perhaps the most commonly discussed vitamin, vitamin C (ascorbic acid) is a strong reducing compound that has to be acquired via dietary ingestion. Major dietary sources include fruits (especially citrus) and vegetables (e.g., tomatoes, green peppers, cabbage, leafy greens, and potatoes). Ascorbic acid is important in the formation and stabilization of collagen by hydroxylation of proline and lysine for cross-linking and conversion of tyrosine to catecholamines by dopamine β-hydrolase. It increases the absorption of certain minerals, such as iron, and is absorbed in the small intestine and distributed throughout the water-soluble compartments of the body.[133] The deficiency state, known as scurvy, is characterized by hemorrhagic disorders, including swollen, bleeding gums, impaired wound healing, and anemia.[133,134]

Although urine is the primary route of excretion, measurement of urinary ascorbate is not recommended for status assessment. Drugs known to increase urinary excretion of ascorbate include aspirin, aminopyrine, barbiturates, hydantoin, and

CASE STUDY 27.3, PART 3

Remember Martin. The result of Martin's serum vitamin B_{12} laboratory test was 190 ng/L (reference range 180–914 ng/L).

2. What are common methods for the determination of vitamin B_{12}?

Although the result falls within the reference range, Martin's physician doesn't feel confident that he can rule out a deficiency.

3. What is another (potentially more definitive) test that may be ordered?

The result of a serum methylmalonic acid (MMA) were 0.52 nmol/mL (reference range ≤ 0.40 nmol/mL). Given the moderately elevated MMA levels, Martin's physician feels confident diagnosing him with vitamin B_{12} deficiency.

4. Why are these results more definitive than the vitamin B_{12} results?

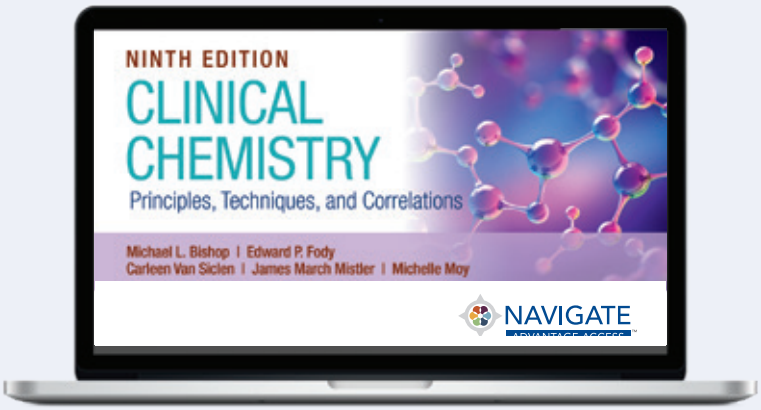
paraldehyde. Ascorbic acid requirements are increased with acute stress injury and chronic inflammatory states, but are also increased with pregnancy and oral contraceptive use. Excessive intake may interfere with vitamin B_{12} metabolism and drug actions (e.g., aminosalicylic acid, tricyclic antidepressants, and anticoagulants).[133,134] The RDA for vitamin C is 90 mg/d for adult males and 75 mg/d for adult females.[135]

Vitamin C is commonly measured in plasma and urine. Special care must be taken with plasma samples, as ascorbic acid is oxidized to dehydroascorbic acid (DHA) if not treated with a metal-chelating acid (example: EDTA) and protein-precipitating acid, and frozen at −80°C soon after collection.[2,136]

One method for analysis is the 2,4-dinitrophenylhydrazine method. In this procedure, ascorbic acid is first oxidized to dehydroascorbic acid and 2,3-diketogulonic acid with the formation of a colored product that absorbs at 520 nm. This method measures the total vitamin C content of the sample because ascorbic acid, dehydroascorbic acid, and diketogulonic acid are also measured, and it is subject to interference from amino acids and thiosulfates. HPLC has also been developed to provide increased sensitivity and specificity.[134] Leukocyte ascorbic acid levels may also be measured as an indicator of tissue stores of vitamin C, although this is not widely measured in routine clinical laboratories.[2,137]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Rodushkin I, Odman F, Branth S. Multielemental analysis of whole blood by high resolution inductively coupled plasma mass spectrometry. *Fresenius J Anal Chem.* 1999;364: 338–346.
2. Roberts NB, Taylor A, Sodi R. Vitamins and trace elements. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier, Inc. 2018: 639–718.

3. May T, Wiedmeyer R. A table of polyatomic interferences in ICP-MS. *Atom Spectrosc*. 1998:19.
4. Vaughan M, Horlick G. Correction procedures for rare earth analysis in inductively coupled plasma-mass spectrometry. *Appl Spectrosc*. 1990;44:587–593.
5. Thomas R. *Practical Guide to ICP-MS: a Tutorial for Beginners*. 2nd ed. Boca Raton, FL: CRC Press; 2008:xxv, 347.
6. Sobin C, Parisi N, Schaub T, et al. A Bland–Altman comparison of the Lead Care® system and inductively coupled plasma mass spectrometry for detecting low-level lead in child whole blood samples. *J Med Toxicol*. 2011;7: 24–32.
7. Strathmann FG, Blum LM. Toxic elements. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier, Inc.; 2018: 888–924.
8. Agency for Toxic Substances & Disease Registry. Aluminum. Accessed January 2012. http://www.atsdr.cdc.gov/toxfaqs/tf.asp?id=190&tid=34
9. Jacobs DS, DeMott WR, Oxley DK. *Jacobs & Demott Laboratory Test Handbook*. 5th ed. Hudson, OH: Lexi-Comp.; 2001:1031.
10. Schaller K, Letzel S, Angerer J. Aluminum. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:217–226.
11. Alfrey A. Aluminum. In: Mertz W, Underwood EJ, eds. *Trace Elements in Human and Animal Nutrition*. Vol. 1. 5th ed. St. Louis, MO: Academic Press; 1986:399–413.
12. Baselt R. Disposition of toxic drugs and chemicals in man. In: Baselt R, ed. *Disposition of Toxic Drugs and Chemicals in Man*. 8th ed. Seal Beach, CA: Biomedical Publications; 2004:39–41.
13. Sjögren B, Iregren A, Elinder C, et al. Aluminum. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:339–352.
14. U.S. Environmental Protection Agency. Arsenic compounds. Accessed January 2012. http://www.epa.gov/ttn/atw/hlthef/arsenic.html
15. Iffland R. Arsenic. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:237–253.
16. Fowler B, Chou C-HS, Jones R, et al. Arsenic. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:367–406.
17. Kamimura T, Miyamoto T, Harada M, et al. Advances in therapies for acute promyelocytic leukemia. *Cancer Sci*. 2011;102:1929–1937.
18. Anke M. Arsenic. In: Mertz W, Underwood EJ, eds. *Trace Elements in Human and Animal Nutrition*. Vol. 2. 5th ed. St. Louis, MO: Academic Press; 1986:347–372.
19. Pais I, Jones JB. *The Handbook of Trace Elements*. Boca Raton, FL: St. Lucie Press; 1997:xv, 223.
20. U.S. Environmental Protection Agency. Cadmium. Accessed January 2012. http:// www.epa.gov/iris/subst/0141.htm
21. Agency for Toxic Substances & Disease Registry. Cadmium. Accessed September 2021. https://www.atsdr.cdc.gov/toxguides/toxguide-5.pdf
22. Moyer T, Burritt M, Butz J. Toxic metals. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 4th ed. Philadelphia, PA: WB Saunders; 2006:1377–1378.
23. Herber RFM. Cadmium. In: Seiler HG, Sigel A, Sigel H, et al., eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:283–297.
24. Nordberg G, Nogawa K, Nordberg M, et al. Cadmium. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:445–486.
25. Herold DA, Fitzgerald RL. Chromium. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:321–332.
26. Offenbacher EG, Pi-Sunyer FX. Chromium in human nutrition. *Annu Rev Nutr*. 1988;8:543–563.
27. Kohlmeier M. Chromium. In: Kohlmeier M, ed. *Food Science and Technology, Nutrient Metabolism*. Amsterdam; Boston, MA: Academic Press. 2003:742–745.
28. Agency for Toxic Substances and Disease Registry. Case Studies in Environmental Medicine (CSEM). Chromium Toxicity. Accessed February 2021. https://www.atsdr.cdc.gov/csem/chromium/docs/chromium.pdf
29. Shenkin A, Baines M, Lyon T. Vitamins and trace elements. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 4th ed. Philadelphia, PA: WB Saunders; 2006:1075–1164.
30. Langård S, Costa M. Chromium. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:487–510.
31. Halasova E, Matakova T, Kavcova E, et al. Human lung cancer and hexavalent chromium exposure. *Neuro Endocrinol Lett*. 2009;30 Suppl 1:182–185.
32. U.S. Food and Drug Administration. Information for all health care professionals who provide treatment to patients with a metal-on-metal hip implant system. Accessed January 2012. http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/MetalonMetalHipImplants/ucm241744.htm
33. Walter LR, Marel E, Harbury R, et al. Distribution of chromium and cobalt ions in various blood fractions after resurfacing hip arthroplasty. *J Arthroplasty*. 2008;23: 814–821.
34. Ellingsen D, Horn N, Aaseth J. Copper. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:529–546.
35. Sarkar B. Copper. In: Seiler HG, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:339–347.
36. Copper. In: *DRI: Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc*. Washington, D.C.: National Academy Press; 2001:224-257.
37. Ramani PK, Parayil Sankaran B. Menkes kinky hair disease. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; October 28, 2020.
38. Kaler SG. Neurodevelopment and brain growth in classic Menkes disease is influenced by age and symptomatology at initiation of copper treatment. *J Trace Elem Med Biol*. 2014;28(4):427–430.
39. Holstege CP. Heavy metals. In: Holstege CP, Borloz MP, Benner JP, eds. *Toxicology Recall*. Baltimore, MD: Lippincott Williams & Wilkins; 2009:256–310.

40. McMillin GA, Travis JJ, Hunt JW. Direct measurement of free copper in serum or plasma ultrafiltrate. *Am J Clin Pathol.* 2009;131:160–165.
41. Iron: Fact sheet for professionals. NIH Office of Dietary Supplements. Accessed September 18, 2021. https://ods.od.nih.gov/factsheets/Iron-HealthProfessional/
42. Higgins T, Beutler E, Doumas B. Hemoglobin, iron, and bilirubin. In: Burtis CA, Ashwood ER, Bruns DE, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 4th ed. Philadelphia, PA: WB Saunders; 2006:1165–1208.
43. Ofojekwu M-JN, Nnanna OU, Okolie CE, et al. Hemoglobin and serum iron concentrations in menstruating nulliparous women in Jos, Nigeria. *Lab Med.* 2013;44(2):121–124.
44. Cappellini MD, Lo SF, Swinkels DW. Hemoglobin, iron, bilirubin. In: Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier, Inc. 2018: 719–775.
45. Rossi E. Hepcidin—the iron regulatory hormone. *Clin Biochem Rev.* 2005;26:47–49.
46. Beard JL, Dawson H, Piñero DJ. Iron metabolism: a comprehensive review. *Nutr Rev.* 1996;54:295–317.
47. Meneghini R. Iron homeostasis, oxidative stress, and DNA damage. *Free Radic Biol Med.* 1997;23:783–792.
48. Smith C, Mitchinson MJ, Aruoma OI, et al. Stimulation of lipid peroxidation and hydroxyl-radical generation by the contents of human atherosclerotic lesions. *Biochem J.* 1992;286(Pt 3):901–905.
49. Schwartz CJ, Valente AJ, Sprague EA, et al. The pathogenesis of atherosclerosis: an overview. *Clin Cardiol.* 1991;14:I1–I16.
50. Anghileri LJ. Iron, intracellular calcium ion, lipid peroxidation and carcino-genesis. *Anticancer Res.* 1995;15:1395–400.
51. Weinberg ED. Cellular iron metabolism in health and disease. *Drug Metab Rev.* 1990;22:531–579.
52. McCord JM. Iron, free radicals, and oxidative injury. *Semin Hematol.* 1998;35:5–12.
53. Smith MA, Perry G. Free radical damage, iron, and alzheimer's disease. *J Neurol Sci.* 1995;134(suppl):92–94.
54. Herbert V, Shaw S, Jayatilleke E. Vitamin c-driven free radical generation from iron. *J Nutr.* 1996;126:1213S–1220S.
55. Tofaletti JG. Trace elements. In: Bishop ML, Fody EP, Schoeff LE, eds. *Clinical Chemistry: Principles, Procedures, Correlations*. 5th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2004.
56. ARUP Laboratories. Hemochromatosis. Accessed September 18, 2021. https://arupconsult.com/content/hemochromatosis
57. Christensen J, Kristiansen J. Lead. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:425–440.
58. Skerfving S, Bergdahl I. Lead. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:599–643.
59. Centers for Disease Control and Prevention. CDC Response to Advisory Committee on Childhood Lead Poisoning Prevention Recommendations in "*Low Level Lead Exposure Harms Children: A Renewed Call of Primary Prevention*". Accessed September 18, 2021. https://www.cdc.gov/nceh/lead/acclpp/cdc_response_lead_exposure_recs.pdf
60. Tang G, Tu X, Feng P. Lead poisoning caused by traditional Chinese medicine: A case report and literature review. *Tohoku J Exp Med.* 2017;243(2):127–131.
61. Boffetta P, Nyberg F. Contribution of environmental factors to cancer risk. *Br Med Bull.* 2003;68:71–94.
62. Ŝaric′ M, Lucchini R. Manganese. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:645–674.
63. Williams M, Todd GD, Roney N, et al. Toxicological Profile for Manganese. Atlanta (GA): Agency for Toxic Substances and Disease Registry (US); September 2012.
64. Kaiser J. State court to rule on manganese fume claims. *Science (New York, NY).* 2003;300:927.
65. Donaldson J. The physiopathologic significance of manganese in brain: its relation to schizophrenia and neurodegenerative disorders. *Neurotoxicology.* 1987;8:451–462.
66. Berlin M, Zalups R, Fowler B. Mercury. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:675–729.
67. Drash G. Mercury. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:479–493.
68. Goyer R, Klaassen CD, Waalkes MP. Nervous system. In: Goyer RA, Klaassen CD,Waalkes MP, eds. *Metal Toxicology*. San Diego, CA: Academic Press; 1995:199–235.
69. Porter W, Moyer T. Clinical toxicology. In: Burtis CA, Ashwood ER, eds. *Tietz Fundamentals of Clinical Chemistry*. 4th ed. Philadelphia, PA: WB Saunders; 2006:427–456.
70. Ashwood E. *Clinical Testing*. New York, NY: ARUP Laboratories; 2004:101–102.
71. Turnlund J, Friberg L. Molybdenum. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:731–741.
72. Anke M, Giel M. Molybdenum. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:495–501.
73. Kurokawa S, Berry MJ. Selenium. Role of the essential metalloid in health. *Met Ions Life Sci.* 2013;13:499–534.
74. Magee R, James B. Selenium. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:551–562.
75. Högberg J, Alexander J. Selenium. In: Nordberg G, Fowler B, Nordberg M, et al., eds. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:783–807.
76. Gladyshev VN, Jeang KT, Stadtman TC. Selenocysteine, identified as the penultimate c-terminal residue in human t-cell thioredoxin reductase, corresponds to TGA in the human placental gene. *Proc Natl Acad Sci U S A.* 1996;93:6146–6151.
77. Agency for Toxic Substances & Disease Registry. Selenium. Accessed January 2012. http://www.atsdr.cdc.gov/substances/toxsubstance. asp?toxid=28
78. Levander O. Selenium. In: Mertz W, Underwood EJ, eds. *Trace Elements in Human and Animal Nutrition*. Vol. 1. 5th ed. St. Louis, MO: Academic Press; 1986:209–279.
79. Thunus L, Lejeune R. Zinc. In: Seiler H, Sigel A, Sigel H, eds. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:667–674.
80. Solomons NW. Biological availability of zinc in humans. *Am J Clin Nutr.* 1982;35:1048–1075.
81. Sandstrom B, Cederblad A. Zinc absorption from composite meals. II. Influence of the main protein source. *Am J Clin Nutr.* 1980;33:1778–1783.

82. Sandstrom B, Arvidsson B, Cederblad A, et al. Zinc absorption from composite meals. I. The significance of wheat extraction rate, zinc, calcium, and protein content in meals based on bread. *Am J Clin Nutr.* 1980;33:739–745.
83. Cunnane SC. Maternal essential fatty acid supplementation increases zinc absorption in neonatal rats: relevance to the defect in zinc absorption in acrodermatitis enteropathica. *Pediatr Res.* 1982;16:599–603.
84. Solomons NW. Competitive interaction of iron and zinc in the diet: consequences for human nutrition. *J Nutr.* 1986;116:927–935.
85. Sandström B, Davidsson L, Cederblad A, et al. Oral iron, dietary ligands and zinc absorption. *J Nutr.* 1985;115: 411–414.
86. Lönnderdal B. Iron-zinc-copper interactions. In: *Micronutrient Interactions Impact on Child Health and Nutrition*. Washington, DC: International Life Sciences Institute; 1996:3–10.
87. Bales CW, Steinman LC, Freeland-Graves JH, et al. The effect of age on plasma zinc uptake and taste acuity. *Am J Clin Nutr.* 1986;44:664–669.
88. Fisher GL. Function and homeostasis of copper and zinc in mammals. *Sci Total Environ.* 1975;4:373–412.
89. Hambridge K, Casey C, Krebs N. Zinc. In: Mertz W, Underwood EJ, eds. *Trace Elements in Human and Animal Nutrition*. Vol. 1. 5th ed. St. Louis, MO: Academic Press; 1986:1–137.
90. Taylor A. Measurement of zinc in clinical samples. *Ann Clin Biochem Int J Lab Med*. 1997;34(2):142–150.
91. Stevens BJ, Hare DJ, Volitakis I, Cherny RA, Roberts BR. Direct determination of zinc in plasma by graphite furnace atomic absorption spectrometry using palladium/magnesium and EDTA matrix modification with high temperature pyrolysis. *J Anal Atomic Spectrom.* 2017;32(4):843–847. doi:10.1039/c7ja00033b
92. Briggs MH, ed. *Vitamins in Human Biology and Medicine*. Boca Raton, FL: CRC Press; 1981.
93. Food and Nutrition Board. *Recommended Dietary Allowances*. 10th ed. Washington, DC: National Academy of Science; 1989.
94. Otten JJ, Hellwig JP, Meyers LD, eds. *Dietary Reference Intakes: The Essential Guide to Nutrient Requirements.* Washington, DC: Institute of Medicine, National Academies Press; 2006. Accessed August 1, 2012. http://www.nap.edu/catalog.php?record_ id=11537
95. Boosalis MG. Vitamins. In: Matarese LE, Gottschlich MM, eds. *Contemporary Nutrition Support Practice: A Clinical Guideline*. Philadelphia, PA: WB Saunders; 1998:145–162.
96. Wiley KD, Gupta M. Vitamin B1 thiamine deficiency. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; June 22, 2020.
97. Brown MJ, Ameer MA, Beier K. Vitamin B6 deficiency. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; February 11, 2021.
98. Sanvictores T, Chauhan S. Vitamin B5 (pantothenic acid). In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; September 27, 2020.
99. Saleem F, Soos MP. Biotin deficiency. In: *StatPearls*. Treasure Island (FL): StatPearls Publishing; April 20, 2020.
100. Boosalis MG. Micronutrients. In: Boosalis MG, ed. *The Science and Practice of Nutrition Support: A Case-Based Core Curriculum*. Dubuque, IA: Kendall/Hunt Publishing; 2001:85–106.
101. Otten JJ, Hellwig JP, Meyer LD, eds. *Dietary Reference Intakes: The Essential Guide to Nutrient Requirements.* Washington, DC: National Academies Press; 2006:5–19.
102. Ensminger AH, Ensminger ME, Konlande JE, et al. *Foods and Nutrition Encyclopedia*. 2nd ed. Boca Raton, FL: CRC Press; 1994:2257–2260.
103. Sauberlich HE. *Laboratory Tests for the Assessment of Nutritional Status*. 2nd ed. Boca Raton, FL: CRC Press; 1999:11–280.
104. Book LS. Fat soluble vitamins and essential fatty acids in total parenteral nutrition. In: Lebenthal E, ed. *Total Parenteral Nutrition: Indications, Utilization, Complications*. New York, NY: Raven Press; 1986:59–81.
105. Underwood BA. Methods for assessment of vitamin A status. *J Nutr.* 1990;120(Suppl 11):1459–1463.
106. Institute of Medicine, Food and Nutrition Board. *Dietary Reference Intakes for Calcium and Vitamin D.* Washington, DC: National Academies Press; 2010. Accessed February 28, 2012. http://books.nap.edu/openbook.php?record_id=13050&page=1106.
107. Dirks NF, Ackermans MT, Lips P, et al. The when, what & how of measuring vitamin D metabolism in clinical medicine. *Nutrients*. 2018;10(4):482.
108. Institute of Medicine (US) Committee to Review Dietary Reference Intakes for Vitamin D and Calcium, Ross AC, Taylor CL, Yaktine AL, Del Valle HB, eds. *Dietary Reference Intakes for Calcium and Vitamin D*. Washington, DC: National Academies Press; 2011.
109. Lukaszuk JM, Luebbers PE. 25(OH)D status: Effect of D3 supplement. *Obes Sci Pract.* 2017;3(1):99–105.
110. Test ID: DHVD 1,25-Dihydroxyvitamin D, Serum. DHVD - Overview: 1,25-Dihydroxyvitamin D, Serum. Accessed March 2021. https://www.mayocliniclabs.com/test-catalog/Overview/8822
111. Vitamin D, 1, 25-Dihydroxy: ARUP Lab Test Directory. Accessed March 2021. https://ltd.aruplab.com/Tests/Pub/0080385
112. Farrell CJL, Martin S, McWhinney B, et al. State-of-the-art vitamin D assays: a comparison of automated immunoassays with liquid chromatography-tandem mass spectrometry methods. *Clin Chem*. 2012;58:3, 531–542.
113. Bieri JG, Evarts RP, Thorp S. Factors affecting the exchange of tocopherol between red cells and plasma. *Am J Clin Nutr.* 1977;30:686.
114. Das Gupta S, Suh N. Tocopherols in cancer: an update. *Mol Nutr Food Res.* 2016;60(6):1354–1363.
115. Vitamin E supplementation for the prevention of morbidity and mortality in preterm infants. World Health Organization. Published February 11, 2019. Accessed February 2021. https://www.who.int/elena/titles/vitamine_preterm/en/
116. Institute of Medicine (US) Panel on Dietary Antioxidants and Related Compounds. *Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids.* Washington, DC: National Academies Press; 2000.
117. Trumbo P, Yates AA, Schlicker S, Poos M. Dietary reference intakes: vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc. *J Am Diet Assoc*. 2001;101(3):294–301.

118. Test ID: VITE - Clinical: Vitamin E, Serum. Accessed March 2021. https://www.mayocliniclabs.com/test-catalog/Clinical+and+Interpretive/42358
119. Suttie JW. Vitamin K. In: Diplock AD, ed. *Fat-Soluble Vitamins: Their Biochemistry and Applications*. London: Heinemann; 1985:225–311.
120. Hazell K, Baloch KH. Vitamin K deficiency in the elderly. *Gerontol Clin*. 1970;12:10–17.
121. Sitren H. Vitamin K. In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:411–430.
122. Sitren HS, Bailey LB, Cerda JJ, et al. Thiamin (vitamin B1). In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:431–449.
123. Bailey LB. Pyridoxine (vitamin B6). In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:521–539.
124. Sitren HS, Bailey LB, Cerda JJ, et al. Niacin (vitamin B3). In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:469–486.
125. Roth KS. Biotin in clinical medicine: a review. *Am J Clin Nutr*. 1981;34:1967.
126. Zempleni J, Mock DM. Advanced analysis of biotin metabolites in body fluids allows a more accurate measurement of biotin bioavailability and metabolism in humans. *J Nutr*. 1999;129:494S–497S.
127. Bailey LB. Folic acid. In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:573–590.
128. Bailey LB. Folate status assessment. *J Nutr*. 1990;120(Suppl 11):1508–1511.
129. Ueland PM, Refsum H, Stabler SP, et al. Total homocysteine in plasma or serum: methods and clinical applications. *Clin Chem*. 1993;39:1764–1779.
130. Steinkamp RC. Vitamin B12 and folic acid: clinical and pathophysiological considerations. In: Brewster MA, Naito IIK, eds. *Nutritional Elements and Clinical Biochemistry*. New York, NY: Plenum; 1980:169–240.
131. Bailey LB. Cobalamine (vitamin B12). In: Baumgartner TG, ed. *Clinical Guide to Parenteral Micronutrition*. New York, NY: Fujizawa USA, Inc.; 1991:541–572.
132. Recommendations: Women & Folic Acid. Centers for Disease Control and Prevention. Published August 13, 2019. Accessed March 2021. https://www.cdc.gov/ncbddd/folicacid/recommendations.html#:~:text=All%20women%20of%20reproductive%20age,neural%20tube%20defects%20(NTDs)
133. Englard S, Seifter S. The biochemical functions of ascorbic acid. *Annu Rev Nutr*. 1986;6:265–304.
134. Jacob RA. Assessment of human vitamin C status. *J Nutr*. 1990;120(Suppl 11):1480–1485.
135. Office of Dietary Supplements - Vitamin C. NIH Office of Dietary Supplements. Accessed February 2021. https://ods.od.nih.gov/factsheets/VitaminC-HealthProfessional/
136. Pullar JM, Bayer S, Carr AC. Appropriate handling, processing and analysis of blood samples is essential to avoid oxidation of Vitamin C to dehydroascorbic acid. *Antioxidants (Basel)*. 2018;7(2):29.
137. Mitmesser SH, Ye Q, Evans M, Combs M. Determination of plasma and leukocyte vitamin C concentrations in a randomized, double-blind, placebo-controlled trial with Ester-C(®). *Springerplus*. 2016;5(1):1161. Published 2016 Jul 25.

Bibliography

Holstege CP. Heavy metals. In: Holstege CP, Borloz MP, Benner JP, eds. *Toxicology Recall*. Baltimore, MD: Lippincott Williams & Wilkins; 2009:256–310.

Jacobs DS, DeMott WR, Oxley DK. *Jacobs & DeMott Laboratory Test Handbook*. 5th ed. Lexi-Comp's Clinical Reference Library. Hudson, OH: Lexi-Comp.; 2001:1031.

Nordberg G, Fowler B, Nordberg M, et al. *Handbook on the Toxicology of Metals*. 3rd ed. Amsterdam; Boston, MA: Academic Press; 2007:xlvii, 975.

Pais I, Jones JB. *The Handbook of Trace Elements*. Boca Raton, FL: St. Lucie Press; 1997:xv, 223.

Seiler HG, Sigel A, Sigel H. *Handbook on Metals in Clinical and Analytical Chemistry*. New York, NY: Marcel Dekker; 1994:xx, 753.

Thomas R. *Practical Guide to ICP-MS: a Tutorial for Beginners*. 2nd ed. Practical spectroscopy. Boca Raton, FL: CRC Press; 2008:xxv, 347.

CHAPTER 28

Tumor Markers

Carleen Van Siclen

CHAPTER OUTLINE

KEY TERMS

Cancer
Neoplasm
Oncofetal antigen
Oncogene
Tumor marker

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Discuss the incidence of cancer in the United States for both males and females.
- Explain the role of tumor markers in cancer management.
- Identify the characteristics or properties of an ideal tumor marker.
- State the clinical usefulness of each tumor marker in this chapter.
- Name the major tumor types and their associated markers.
- Describe the use of enzymes and hormones as tumor markers.
- Specify the definitive laboratory test for making a diagnosis of cancer.
- Compare and contrast methods of analysis and clinical application of tumor markers.
- Correlate tumor marker results with associated cancers.
- Interpret tumor marker results.

CASE STUDY 28.1, PART 1

Mark, a 33-year-old man with a history of chronic hepatitis, presents with edema, abdominal pain, and a recent 10-lb weight loss. He denies being on a diet.

© Lesia Kapinosova/Shutterstock.

CASE STUDY 28.2, PART 1

Tatenda, a 65-year-old African American man, presents to his primary care physician with the chief complaint of tarry-colored stools that started a week ago. He has had gastrointestinal discomfort and has felt fatigued during the past 2 months. Tatenda is asked to schedule a colonoscopy before leaving the office.

© SBS Studios/Shutterstock.

CASE STUDY 28.3, PART 1

Marcus, a 28-year-old man, was diagnosed with testicular cancer 10 months ago.

© StockLite/Shutterstock.

CASE STUDY 28.4, PART 1

Drew, a 55-year-old man, presented for his annual physical, where a digital rectal examination (DRE) was performed and blood was drawn for PSA. On DRE, asymmetric nodules were detected. He was scheduled for a transrectal biopsy and a follow-up appointment.

© Oleg Golovnev/Shutterstock.

Cancer remains the second leading cause of mortality in developed countries, accounting for 23% of deaths in the United States and approximately 3 million deaths/year globally. Approximately 42% of males and 38% of females will develop invasive cancer in their lifetime; males have a lifetime risk of dying from cancer of 23%, whereas females have a 19% risk.[1] These rates are higher in individuals from marginalized groups such as those from racial, ethnic, and sexual orientation communities. Cancer is a broad term used to describe more than 200 different malignancies that affect more than 50 tissue types. Despite considerable efforts to reduce the incidence of malignancies, it is estimated that there were 1.6 million new cases of cancer in the United States in 2020 (**Figure 28.1**; see additional current global cancer statistics at http://globocan.iarc.fr/factsheet.asp).

Cancer is a disease caused by the uncontrolled growth of cells that often form a new solid mass or tumor (**neoplasm**) and may spread to other areas of the body. A complex combination of inherited and acquired genetic mutations lead to tumor formation (*tumorigenesis*) and spreading (*metastasis*) (for comprehensive reviews, see references[2,3]). During tumorigenesis, mutations activate growth factors (e.g., epidermal growth factor) and **oncogenes** (e.g., *KRAS*), which are cancer-causing genes, in combination with inhibition of apoptosis, tumor suppressor, and cell cycle regulation genes (e.g., *BRCA1*, *p53*, and cyclins). As cancer progresses toward metastasis, additional genetic changes are required, such as loss of cell adhesion proteins (e.g., β-catenin and E-cadherin) and activation of angiogenesis genes (e.g., vascular endothelial growth factor) (**Figure 28.2**). An understanding of these genetic mechanisms is the basis for many current and future cancer treatments. A combination of factors determines cancer severity and is used to classify its stage. Depending on the type of cancer, these factors include tumor size, histology, regional lymph node involvement, and presence of metastasis. For most solid tumors (e.g., breast, lung, and kidney), cancer is broadly classified (using roman

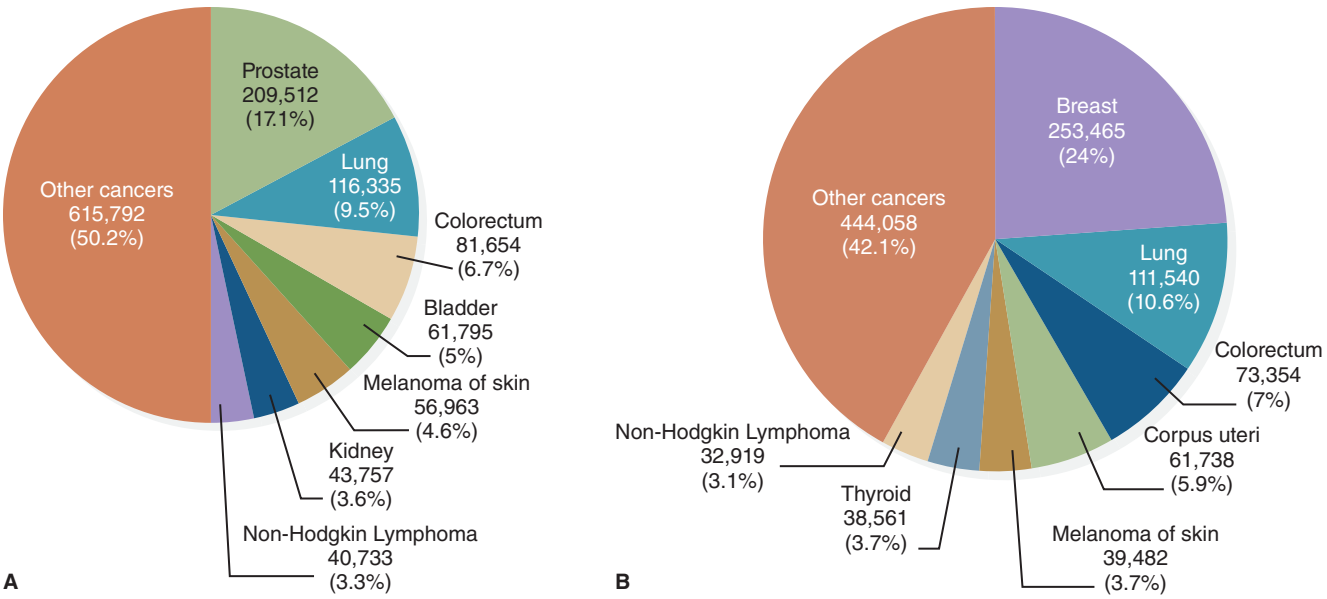


Figure 28.1 Estimated number of new cases of cancer in the United States (2020). **(A)** Males. **(B)** Females.

Reproduced from Howlader N, Noone AM, Krapcho M, et al., eds. SEER Cancer Statistics Review, 1975-2018, National Cancer Institute. Bethesda, MD, https://seer.cancer.gov/csr/1975_2018/, based on November 2020 SEER data submission, posted to the SEER web site, April 2021.

numerals I to IV) into four stages (**Figure 28.3**). Stages correlate with disease severity, where higher stages are indicative of larger tumors and/or significant metastasis (spreading) and severe systemic disease. With disease progression, both proliferation and metastasis occur at the expense of normal organ processes, which is usually the ultimate cause of cancer-associated morbidity and mortality.

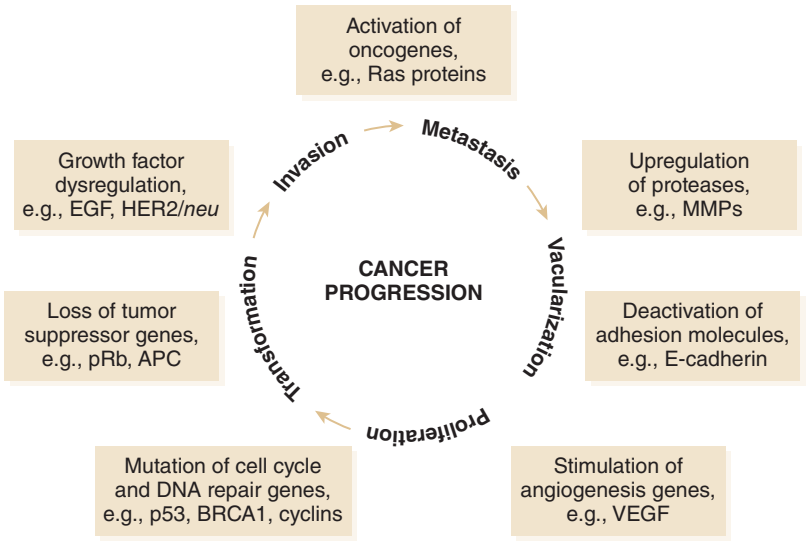


Figure 28.2 Genetic changes associated with cancer. A combination of acquired and/or hereditary defects causes tumor formation and metastasis. These processes begin with unregulated proliferation and transformation, followed by invasion and loss of cellular adhesion. A rich vascular supply of oxygen and nutrients is necessary to facilitate growth of a tumor larger than 100 to 200 µm. APC, familial adenomatous polyposis coli, mutated in colorectal cancers; *BRCA1*, breast cancer susceptibility gene; E-cadherin, adhesion molecule; EGF, epithelial growth factor; MMP, matrix metalloproteinase; *p53*, cell cycle regulator, mutated in 50% of cancers; *pRb*, retinoblastoma protein, mutated in many cancers; *Ras*, small G protein, mutated in many cancers; TIMP, tissue inhibitor of metalloproteinase; VEGF, vascular endothelial growth factor, drug target for inhibition of angiogenesis.

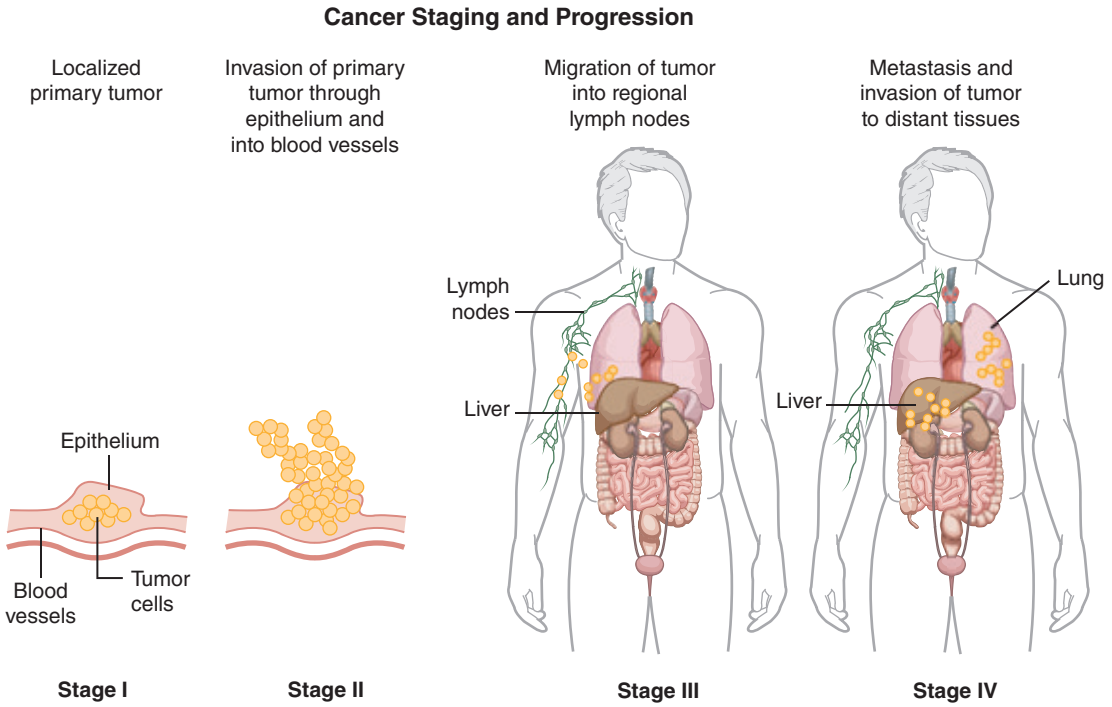


Figure 28.3 Generalized cancer staging and progression. Numerous factors are used in combination to define cancer stage; these include tumor size, extent of invasion, lymph node involvement, metastasis, and histologic assessments (basis for the TNM staging system). In this simplified diagram, stage is presented as a function of invasion and spreading regionally and to other tissues; the primary tumor is not shown. TNM, tumor, nodes, and metastases.

Types of Tumor Markers

Cancer can be detected and monitored using tumor markers. A **tumor marker** is a biomarker found in the blood or tissue and when elevated is linked to cancer. Tumor markers are produced either directly by the tumor or as an effect of the tumor on healthy tissue (*host*). Tumor markers encompass an array of diverse molecules such as oncofetal antigens, hormones, metabolites, receptors, and enzymes. An **oncofetal antigen** is a protein produced during fetal development and elevated in individuals with cancer. A variety of enzymes are elevated nonspecifically in tumors (**Table 28.1**). These elevated enzymes

Table 28.1 Enzyme Tumor Markers

Tumor Marker	Tumor Type	Method	Specimen	Clinical Utility
Prostate-specific antigen	Prostate cancer	IA	Serum	Prostate cancer screening, therapy monitoring, and recurrence
Lactate dehydrogenase	Hematologic malignancies	EA	Serum	Prognostic indicator; elevated nonspecifically in numerous cancers
Alkaline phosphatase	Metastatic carcinoma of bone, hepatocellular carcinoma, osteosarcoma, lymphoma, leukemia	EA	Serum	Determination of liver and bone involvement; nonspecific elevation in many bone-related and liver cancers
Neuron-specific enolase	Neuroendocrine tumors	IA	Serum	Prognostic indicator and monitoring disease progression for neuroendocrine tumors

EA, enzyme assay; IA, immunoassay.

Table 28.2 Serum Protein Tumor Markers

Tumor Marker	Tumor Type	Method	Specimen	Clinical Utility
Serum M-protein	Plasma cell dyscrasias	SPE/IFE	Serum	Diagnosis, therapeutic monitoring of plasma cell malignancies
Serum-free light chains	Plasma cell dyscrasias	IA	Serum	Diagnosis, therapeutic monitoring of plasma cell malignancies
β_2-Microglobulin	Hematologic malignancies	IA	Serum	Prognostic marker for lymphoproliferative disorders

IA, immunoassay; IFE, immunofixation electrophoresis; SPE, serum protein electrophoresis.

are largely a result of the high metabolic demand of these proliferative cells. Accordingly, enzyme levels tend to correlate with tumor burden, making them clinically useful for monitoring the success of therapy. Serum proteins, such as β_2-Microglobulin and immunoglobulins, are also used to monitor cancer therapy (**Table 28.2**). β_2-Microglobulin is found on the surface of all nucleated cells and can, therefore, be used as a nonspecific marker of the high cell turnover common in tumors. In hematologic malignancies such as multiple myeloma, immunoglobulins provide a relatively specific measure of plasma cell production of monoclonal proteins. In endocrine malignancies, hormones and hormone metabolites are widely used as specific markers of secreting tumors (**Table 28.3**). Hormones can be valuable in diagnosing neuroblastomas, as well as pituitary and adrenal adenomas. One of the first

Table 28.3 Endocrine Tumor Markers

Tumor Marker	Tumor Type	Method	Specimen	Clinical Utility
ACTH	Pituitary adenoma, ectopic ACTH-producing tumor	IA	Serum	Diagnosis of ectopic ACTH-producing tumor
ADH	Posterior pituitary tumors	IA	Serum	Diagnosis of SIADH
C-peptide	Insulin-secreting tumors	IA	Serum	Diagnosis of insulinoma
Calcitonin	MTC and neuroendocrine tumors	IA	Serum	Screening,[a] response to therapy, and monitoring recurrence of MTC
Chromogranin A	Pheochromocytoma, neuroblastoma, carcinoid tumors, small cell lung cancers	IA	Serum	Aid in diagnosis of carcinoid tumors, pheochromocytomas, and neuroblastomas
Cortisol	Adrenal tumors	IA	Serum or urine	Diagnosis of Cushing's syndrome, adrenal adenoma
Gastrin	Neuroendocrine tumor	IA	Serum	Zollinger-Ellison syndrome; gastrinoma
GH	Pituitary adenoma, ectopic GH-secreting tumors	IA	Serum	Diagnosis and post monitoring of acromegaly
HVA	Neuroblastoma, pheochromocytoma, paraganglioma	HPLC	24-h urine	Diagnosis of neuroblastoma[b]
5-HIAA	Carcinoid tumors	HPLC	24-h urine	Diagnosis of carcinoid tumors
Metanephrines (fractionated)[c]	Pheochromocytoma, paraganglioma, neuroblastoma	HPLC	24-h urine or plasma	Screening and diagnosis of pheochromocytoma

(continues)

Table 28.3 Endocrine Tumor Markers *(continued)*

Tumor Marker	Tumor Type	Method	Specimen	Clinical Utility
PTH	Parathyroid adenoma	IA	Serum	Diagnosis and postsurgical monitoring of parathyroid adenoma
PRL	Pituitary adenoma	IA	Serum	Diagnosis and postsurgical monitoring of prolactinoma
VMA	Pheochromocytoma, paraganglioma, neuroblastoma	HPLC	24-h urine	Diagnosis of neuroblastoma[b]

5-HIAA, hydroxyindoleacetic acid; ACTH, adrenocorticotropic hormone; ADH, antidiuretic hormone; ELISA, enzyme-linked immunosorbent assay; GH, growth hormone; HPLC, high-performance liquid chromatography; HVA, homovanillic acid; IA, immunoassay; LC-MS/MS, liquid chromatography-TANDEM mass spectrometry; MTC, medullary thyroid carcinoma; PTH, parathyroid hormone; PRL, prolactin; RIA, radioimmunoassay; SIADH, syndrome of inappropriate antidiuretic hormone secretion; VMA, vanillylmandelic acid.

[a] Screening family members for MTC.
[b] HVA and VMA are used in combination for diagnosis of neuroblastomas.
[c] Metanephrine, normetanephrine.

classes of tumor markers discovered was the oncofetal antigens. Oncofetal antigens such as carcinoembryonic antigen (CEA) and α-fetoprotein (AFP) are expressed transiently during normal development and are then turned on again in the formation of tumors (see **Table 28.4** for use of oncofetal antigens). Other tumor markers include monoclonal defined antigens identified from human tumor extracts and cell lines. These antibodies are directed toward specific carbohydrate or cancer antigens and are best used for monitoring treatment of tumors that secrete these epitopes (**Table 28.5**). Finally, receptors are used to classify tumors for therapy (**Table 28.6**). These "non-serologic" markers are outside the scope

Table 28.4 Use of Oncofetal Antigens for Testicular Cancer Classification

Pathology	Germ Cell Tumor	AFP	hCG
Nonseminomatous tumors	Yolk sac tumor	Increased	Normal
	Choriocarcinoma	Normal	Increased
	Embryonal carcinoma	Increased	Increased
	Teratoma	Normal	Normal
Seminoma		Normal	Increased

AFP, α-fetoprotein; hCG, human chorionic gonadotropin.

Table 28.5 Carbohydrate and Cancer Antigen Tumor Markers

Tumor Marker	Tumor Type	Methodology	Clinical Application
CA 19-9	Gastrointestinal cancer and adenocarcinoma	IA	Monitoring response to therapy
CA 15-3	Metastatic breast cancer	IA	Response to therapy and detecting recurrence
CA 27-29	Metastatic breast cancer	IA	Response to therapy and detecting recurrence
CA-125	Ovarian cancer	IA	Monitoring response to therapy

Table 28.6 Receptor Tumor Markers

Tumor Marker	Tumor Type	Method	Specimen	Clinical Application
Estrogen receptor	Breast cancer	IHC	Biopsy	Hormonal therapy indicator
Progesterone receptor	Breast cancer	IHC	Biopsy	Hormonal therapy indicator
Her-2/*neu*	Breast, ovarian, GI	IHC, FISH	Biopsy	Prognostic and hormonal therapy indicator
Epidermal growth factor receptor	Head, neck, ovarian, cervical cancers	IHC	Biopsy	Prognostic indicator

FISH, fluorescence in situ hybridization; IHC, immunohistochemistry.

of the chapter, but they are an important example of the diversity of tumor markers. Prototypic examples of such a marker are estrogen and progesterone receptors and growth factors (HER-2), which are used to choose between endocrine and cytotoxic therapy; endocrine therapy, such as tamoxifen, typically is more effective in patient with ER- and PR-positive patients.

Tumor markers are an invaluable set of tools that healthcare providers can use for a variety of clinical modalities. Depending on the marker and the type of malignancy, tumor markers for screening, diagnosis, prognosis, therapy monitoring, and detecting recurrence applications are in routine clinical use (**Figure 28.4**).

Applications of Tumor Marker Detection

The ideal tumor marker would be tumor specific, absent in healthy individuals, and readily detectable in body fluids. Unfortunately, all of the presently available tumor markers do not fit this ideal model. However, numerous tumor markers have been identified that have a high enough specificity and sensitivity to be used on a targeted basis for aiding diagnosis, prognosis, detection of recurrence, and/or monitoring the response to treatment (Figure 28.4). Clinically, tumor markers are used in combination with clinical signs, symptoms, and histology to facilitate clinical decision making.

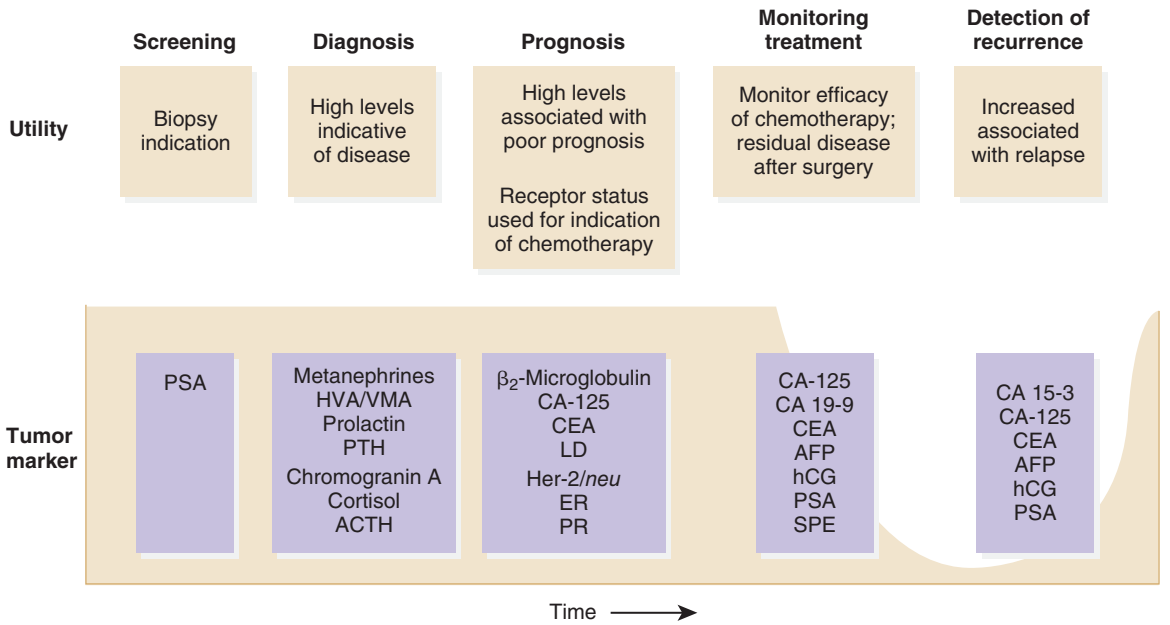


Figure 28.4 Tumor markers are used for screening, prognosis, treatment monitoring, and detecting recurrence of several types of cancer. Whereas few markers are used for screening, many are used to monitor therapy. Endocrine and hormone metabolite markers are often used to aid in diagnosis of secreting tumors. The list is not comprehensive but provides examples of the most commonly used markers. Note: PSA screening remains controversial; see text.

Screening and Risk Assessment

Currently, strategies focus on informed decision-making support to enable patients to weigh the benefits of detecting disease early against the harms of overtreatment. With that said, no tumor marker identified to date can be used to effectively screen asymptomatic populations. This is because most of the clinically used tumor markers are found in normal cells and benign conditions in addition to cancer. Screening asymptomatic populations would therefore result in detection of false positives (patients without disease with detectable tumor marker), leading to undue alarm and risk to patients (e.g., unnecessary imaging, biopsy, and surgery). Presently, only a few tumor markers are used to screen populations with high incidence (targeted screening).

Susceptibility to cancer can be determined using molecular diagnostics in patients with breast, ovarian, or colon cancer by identifying germ-line mutations in patients with a family history of these diseases. Screening for susceptibility to breast and ovarian cancers is done by identifying germline *BRCA1* and *BRCA2* mutations. Similarly, familial colon cancers can be identified by the presence of the adenomatous polyposis coli (*APC*) gene. Since greater than 99% of people with familial APC develop colon cancer by the age of 40 years, prophylactic colectomy is routinely performed on APC+ patients. While gene testing can be done from blood samples, these are not really considered circulating tumor markers and therefore not discussed further.

Prognosis

Tumor marker concentration generally increases with tumor progression, reaching its highest levels when tumors metastasize. Therefore, serum tumor marker levels at diagnosis can reflect the aggressiveness of a tumor and help predict the outcome for patients. High concentrations of a serum tumor marker at diagnosis might indicate the presence of malignancy and possible metastasis, which is associated with a poorer prognosis. In other instances, the mere presence or absence of a particular marker may be valuable. Such is the case with some of the receptors used to base chemotherapeutic treatment in breast cancer as described above, where endocrine therapy is indicated only in the presence of given markers.

Monitoring Effectiveness of Therapy and Disease Recurrence

One of the most useful and common applications of tumor markers is monitoring therapy efficacy and detecting disease recurrence. After surgical resection, radiation, or drug therapy of cancer (chemotherapy), tumor markers are routinely followed serially. In patients with elevated tumor markers at diagnosis, effective therapy results in a dramatic decrease or disappearance of the tumor marker. If the initial treatment is effective, the appearance of circulating tumor markers can then be used as a highly sensitive marker of recurrence. Many tumor markers have a lead time of several months before disease would be detected by other modalities (e.g., imaging), allowing for earlier identification and treatment in cases of relapse.

Laboratory Considerations for Tumor Marker Measurement

The unique characteristics and concentrations of tumor makers require special laboratory considerations. Two major considerations are the size and variability of the tumor marker concentration between different manufacturers due to the lack of harmonization and standardization. Lack of standardization makes comparison of serial patient results using different assays treacherous. There are multiple reasons why these assays are not comparable, including differences in antibody specificity, analyte heterogeneity, and assay design; lack of standard reference material, calibration, and kinetics; and variation in reference ranges. To accurately monitor tumor marker concentrations in a patient, it is important to use the same methodology. It is also important to monitor quality control after changing lot numbers. This includes a careful comparison of QC material and patient samples because detection of tumor markers can vary widely between reagent lot numbers. This is particularly a concern when polyclonal antibodies are used as reagents (e.g., serum free light chains).

The other main consideration for tumor marker measurement is the wide range of concentrations encountered clinically. Tumor markers often vary in concentration by orders of magnitude, making accurate measurement challenging compared with routine chemistry analytes (e.g., concentration extremes for sodium are between 120 and 160 mmol/L, whereas

human chorionic gonadotropin [hCG] may vary between 10 and 10,000,000 mIU/mL!). Handling these ranges requires careful attention to dilution protocols and the risk of antigen excess. These considerations are discussed in the context of specific methodology in the following sections.

Immunoassays

Immunoassays are commonly used method to measure tumor markers. There are many advantages to this method, such as the ability to automate testing and relative ease of use. Many tumor markers are amenable to automation and relatively rapid analysis using large immunoassay or integrated chemistry test platforms. However, there are some unique factors to be considered when using immunoassays to measure tumor markers including assay linearity, antigen excess (hook effect), and the potential for heterophile antibodies.

Linearity

The *linear range* is the span of analyte concentrations over which a linear relationship exists between the analyte and signal. Linearity is determined by analyzing (in replicates) specimens spanning the reportable range. Guidelines for this determination are outlined in the Clinical Laboratory Improvement Amendments (CLIA) guidelines for linearity.[4] Samples exceeding the linear range, which is much more likely to occur in the detection of tumor markers, need to be systemically diluted to determine values within the reportable linear range. Dilutions must be done with careful consideration of the diluent and awareness of the risk of error if using manual calculations (it is common practice to have manual calculations reviewed by another individual). Excessively high tumor marker concentrations can result in falsely low measurements, a phenomenon known as *antigen excess* or *hook effect*.

Hook Effect. When analyte concentrations exceed the analytical range excessively, there is potential for *antigen excess* or *hook effect*. When very high antigen concentrations are present, capture and/or label antibodies can be saturated, resulting in a lack of "sandwich" formation and thus in a significant decrease in signal. The name hook effect refers to the shape of the concentration–signal curve when the reagents are saturated with excess antigen (**Figure 28.5**). But the practical understanding of the hook effect is that it causes the actual tumor marker concentration to be grossly underestimated. If clinical suspicion is high for an elevated tumor marker, it can be identified by the laboratory with dilution and repeat testing. Samples displaying hook effect will yield higher (accurate) values on dilution (Figure 28.5). This phenomenon typically only affects sandwich-type immunoassays.

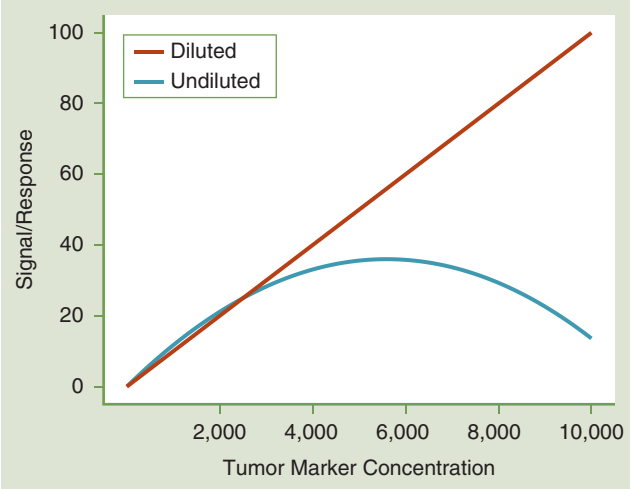


Figure 28.5 Hook effect (antigen excess) can occur with tumor markers because they may be found at very high concentrations. When reagents are depleted by excess antigen (the tumor marker), falsely low results may occur (represented by the "neat" curve). Dilution of samples can be used to detect and account for hook effect (represented by the "diluted" line).

Heterophile Antibodies. Significant interference can be seen in immunoassays if an individual has circulating antibodies against human or animal immunoglobulin reagents. A subset of heterophilic antibodies are human anti-animal antibodies and human antimouse antibodies (HAMAs). HAMAs may be encountered in patients who have been given mouse monoclonal antibodies for therapeutic reasons or who have been exposed to mice,[5] but they may be idiopathic. In patients, these antibodies can cause false-positive or, less commonly, false-negative results by cross-linking the capture/label antibody (see Chapter 4, *Analytic Techniques*). To confirm the presence of heterophilic antibodies, samples need to be diluted and the dilutions analyzed (similar to eliminating hook effect). Patient samples with heterophilic antibodies do not give linear results upon dilution. The presence of anti-animal immunoglobulins can also be detected directly with commercial reagents. Non-immune animal serum is often added

to immunoassays to minimize the effects of heterophilic antibodies, and there are commercial blocking reagents that can be used to remove HAMAs. Many monoclonal therapeutic agents are now derived to include only fragments of an antibody so that patients do not develop heterophilic antibodies to the full antibody. In the laboratory, heterophile antibodies can be detected by investigating results that are inconsistent with history and clinical scenario.

Common Analytical Concerns Applied to Tumor Marker Immunoassays. Immunoassays for tumor markers can be affected by interference from icterus, lipemia, hemolysis, and antibody cross-reactivity in the same manner as other immunoassays. As with all automated tests, the potential for carryover with high levels of tumor marker analytes can also be a concern, leading to falsely elevated levels in patients if adequate washing steps are not included between patient samples.

High-Performance Liquid Chromatography

High-performance liquid chromatography (HPLC) is commonly used to detect small molecules, such as endocrine metabolites. With respect to tumor markers, HPLC is used to detect catecholamine metabolites in plasma and urine. Generally, there is an extraction process, by which the analytes of interest are separated from either plasma or urine. Extractions are applied to a column where they are separated by their physical characteristics (charge, size, and polarity). Catecholamines and catecholamine metabolites are used to help diagnose carcinoid tumors, pheochromocytoma, and neuroblastoma. Neuroblastoma is a common malignant tumor occurring in approximately 7.5% of children under 15 years of age. Neuroblastoma is diagnosed by the detection of high levels of plasma epinephrine, norepinephrine, and dopamine (catecholamines). Pheochromocytoma, a rare tumor associated with hypertension, is diagnosed by detecting elevated plasma metanephrines (along with urine vanillylmandelic acid and free catecholamines). Carcinoid tumors are serotonin-secreting tumors that arise from the small intestine, appendix, or rectum leading to a host of symptoms (carcinoid syndrome), including pronounced flushing, bronchial constriction, cardiac valve lesions, and diarrhea. The diagnosis of carcinoid tumors involves the detection of 5-hydroxyindoleacetic acid, which is a serotonin metabolite. In all of these cases, HPLC is used to detect the hormones and metabolites secreted by these tumors for diagnosis, therapeutic monitoring, and recurrence. HPLC is not subject to hook effect, lot-to-lot antibody variation, and heterophile antibodies but is more labor intensive and requires more experience and skill than automated immunoassays.

Immunohistochemistry and Immunofluorescence

While not found in circulation, it is important for laboratorians to be familiar with solid tissue tumor markers. These are identified in tissue sections typically from a fine-needle aspirate or biopsy samples. Specific antibodies (and the proper control antibodies) are incubated with tissue sections to detect the presence (or absence) of antigens using colorimetric or fluorescent secondary antibodies. In many ways, this is similar to detection by immunoassay, but the added value is the ability to determine whether the antigen in question is in a particular cell type (such as a tumor) in the specific subcellular location. A good example of the use of a tumor marker that is detected by immunohistochemistry is the identification of estrogen and progesterone receptors in breast cancer. When breast tumors are positive for estrogen and progesterone receptors at the cell surface, they tend to respond to hormonal therapy, while tumors lacking these receptors are treated with other chemotherapeutic modalities.[6]

Enzyme Assays

The detection of elevated circulating enzymes generally cannot be used to identify a specific tumor or site of tumor. One key exception to this is the PSA, which is in fact an enzyme; PSA is a serine protease of the kallikrein family, found in both diseased and benign prostate glands (it is also found in low concentrations in amniotic fluid, breast milk, and some other cancers). Before the widespread use of immunoassays and the discovery of oncofetal antigens, enzyme detection use was widespread. When cells die (autolysis/necrosis) or undergo changes in membrane permeability, enzymes are released from intracellular pools into circulation where they are readily detected. Examples of enzymes that have been used as tumor markers include alkaline phosphatase (bone, liver, leukemia, and sarcoma), lactate dehydrogenase (liver, lymphomas, leukemia, and others), and of course PSA (prostate). Enzyme activity assays are used to quantify these enzymes. PSA is measured by immunoassay.

Tumor Marker Tests

α-Fetoprotein

AFP is an abundant serum protein normally synthesized by the fetal liver that is re-expressed in certain types of tumors. This re-expression during malignancy classifies AFP as a carcinoembryonic protein. AFP is often elevated in patients with hepatocellular carcinoma (HCC) and germ cell tumors.

Regulation and Physiology

AFP is a 70-kD glycoprotein related to albumin that normally functions as a transport protein. Like albumin, it is involved in regulating fetal oncotic pressure. During development, AFP peaks at approximately one-tenth the concentration of albumin at 30 weeks of gestation. The upper normal limit for serum AFP is approximately 15 ng/mL (reference ranges are method dependent) in healthy adults. Infants initially have high serum AFP values that decline to adult levels at an age of 7 to 10 months.[7]

Clinical Usefulness and Interpretation

AFP is used for the diagnosis, staging, prognosis, and treatment monitoring of HCC. Also known as hepatoma, HCC is a tumor that originates in the liver, often due to chronic diseases, such as hepatitis and cirrhosis. Patients with HCC frequently have elevated serum AFP. However, as with most tumor markers, AFP is not completely specific. For example, AFP can also be increased in benign conditions such as pregnancy and other nonmalignant liver disease, as well as other types of malignancies (e.g., testicular cancer—see below).

AFP has been used to detect HCC in populations with high disease prevalence. When used for screening high-risk populations, AFP has a sensitivity ranging from 40% to 65% and specificity of 80% to 95% (at cutoffs ranging from 20 to 30 ng/mL).[8] The AFP reference range is less than 10 ng/mL, and values vary by test method. Very high levels of AFP (>500 ng/mL) in high-risk individuals are considered diagnostic of HCC. Several expert groups, including the National Comprehensive Cancer Network, National Academy of Clinical Biochemistry, and the British Society of Gastroenterology, now recommend that AFP be used in conjunction with ultrasound imaging every 6 months in patients at high risk for developing HCC. This includes patients with hepatitis B virus–induced and/or hepatitis C virus–induced liver cirrhosis.[9]

High levels of AFP in HCC are associated with poor prognosis and are exemplified in individuals who do not respond to therapy or have residual disease following surgery. Correspondingly, a decrease in circulating AFP levels after treatment is associated with prolonged survival rates. It is therefore recommended that serial measurements of AFP be used to monitor treatment and post-surgery in patients with HCC.

The other major use for AFP as a tumor marker is for classification and monitoring therapy for testicular cancer. Testicular cancer includes several subtypes broadly classified into seminomatous and nonseminomatous tumors. Seminomatous tumors form directly from malignant germ cells, whereas nonseminomatous tumors differentiate into embryonal carcinoma, teratoma, choriocarcinoma, and yolk sac tumors (endodermal sinus tumor).[10] AFP is used in combination with β-human chorionic gonadotropin (β-hCG) to classify nonseminomatous tumors (Table 28.4). Serum AFP is also useful for tumor staging; AFP is increased in 10% to 20% of stage I tumors, 50% to 80% of stage II tumors, and 90% to 100% of stage III nonseminomatous testicular cancer. As with HCC, AFP can be used serially to monitor therapy efficacy and disease progression, where increases are indicative of relapse or resistance.

Methodology

AFP is measured using any of a variety of commercially available automated immunoassays. These are typically sandwich immunoassays relying on monoclonal or polyclonal antibodies directed toward different regions of AFP. Serial monitoring of AFP should be done using the same laboratory and assay method to ensure changes (or lack of change) are due to the tumor and not assay variation. As with other glycoproteins, AFP displays some heterogeneity where certain isoforms are preferentially produced by malignant cells; AFP isoforms differ in their glycosylation and sialyation. Antibodies against these isoforms produced by malignant cells may in the future be used to improve the specificity of AFP immunoassays.

Clinical Application

The primary applications of AFP as a tumor marker are for HCC and nonseminomatous testicular cancer. AFP is typically used as a marker to monitor therapy, detect residual tumor, or detect relapse; AFP is also used as part of maternal serum screening for neural tube defects and chromosomal abnormalities.

CASE STUDY 28.1, PART 2

Remember Mark. His laboratory results reveal a low platelet count, hypoalbuminemia, a prolonged prothrombin time, and activated partial thromboplastin time.

1. What tumor marker may aid in Mark's diagnosis?
2. What additional laboratory test is required to make a definitive diagnosis?
3. After Mark's surgery, explain if additional tumor marker testing will be necessary.

Cancer Antigen 125 (CA-125)

Cancer antigen 125 (CA-125) was first defined by a murine monoclonal antibody raised against a serous ovarian carcinoma cell line.[10] CA-125 may be useful for detecting ovarian tumors at an early stage and for monitoring treatments without surgical restaging.

Regulation and Physiology

CA-125 is expressed in the ovary, in other tissues of Müllerian duct origin, and in human ovarian carcinoma cells. The CA-125 gene encodes a high molecular weight (200,000 to 1,000,000 kDa) mucin protein containing a putative transmembrane region and a tyrosine phosphorylation site.[11] Although it is not usually found in serum, CA-125 may be elevated in patients with endometriosis, during the first trimester of pregnancy, or during menstruation.

Clinical Usefulness and Interpretation

CA-125 is a serologic marker of ovarian cancer. Ovarian cancer accounts for approximately 3% of the newly diagnosed malignancies in women and is among the top five causes of cancer-related death. Ovarian cancer includes a broad range of categories, including sex cord tumors, stromal tumors, germ cell tumors, and, most commonly, epithelial cell tumors. As with most other tumor markers, CA-125 should not be used to screen for ovarian cancer in asymptomatic individuals. However, CA-125 *is* elevated in a high percentage of ovarian tumors and is recommended as an annual test for women with a family or prior history of ovarian cancer. CA-125 levels also correlate with ovarian cancer stage. CA-125 is elevated in 50% of patients with stage I disease, 90% of patients with stage II, and more than 90% of patients with stage III or IV.

Methodology

CA-125 can be detected by immunoassays that use OC125 and M11 antibodies. These monoclonal antibodies recognize distinct nonoverlapping regions of the CA-125 epitope. CA-125 is available on many automated platforms. However, results from different platforms are not interchangeable due to differences between reagent detection methods.

Clinical Application

CA-125 is predominantly used to monitor therapy and to distinguish benign masses from ovarian cancer.[12] It is important in evaluating the patients' response to ovarian cancer therapy as well as predicting recurrent ovarian cancer. In postmenopausal women with a palpable abdominal mass, a high level (>95 U/mL) of CA-125 has a 90% positive predictive value for ovarian cancer. For therapy monitoring, CA-125 is useful both for predicting the success of surgery (debulking procedures) and for determining the efficacy of chemotherapy. Therefore, patients with elevated CA-125 following either treatment modality have a poor prognosis. The upper normal range for serum CA-125 is typically 35 U/mL.

Carcinoembryonic Antigen (CEA)

CEA was discovered in the 1960s and is a prototypical example of an oncofetal antigen. It is expressed during development and then re-expressed in tumors. CEA is the most widely used tumor marker for colorectal cancer and is also frequently elevated in lung, breast, and gastrointestinal tumors. CEA can be used to aid in the diagnosis, prognosis, and therapy monitoring of colorectal cancer. Although high levels of CEA (>10 ng/mL) are frequently associated with malignancy,

high levels of CEA are not specific for colorectal cancer, and therefore, CEA is not used for screening.

Regulation and Physiology

CEA is a large heterogeneous glycoprotein with a molecular weight of approximately 200 kDa. It is part of the immunoglobulin superfamily and is involved in apoptosis, immunity, and cell adhesion. Because of its role in cell adhesion, CEA has been postulated to be involved in metastasis. Akin to other serologic tumor markers, CEA may be elevated nonspecifically because of impaired clearance or through increased production. Increased CEA concentrations have been observed in heavy smokers and in some patients following radiation treatment and chemotherapy. CEA may also be elevated in patients with liver damage due to prolonged clearance. The upper normal range for serum CEA is 3 to 5 ng/mL depending on the assay.

Clinical Usefulness and Interpretation

The main clinical use of CEA is as a marker for colorectal cancer. In colon cancer, CEA is used for prognosis, in post-surgery surveillance, and to monitor response to chemotherapy. For prognosis, CEA can be used in combination with histology and the TNM (see definition box) staging system to establish the need for adjuvant therapy (addition of chemotherapy or treatment after surgery). Adjuvant therapy is indicated in patients with stage II disease (i.e., tumor has spread beyond immediate colon but not to lymph nodes) who have high levels of CEA.[13]

Methodology

CEA can be detected by immunoenzymatic sandwich assays that use two mouse monoclonal anti-CEA antibodies that react with different epitopes of CEA. CEA is available on numerous commercial automated platforms. Due to the high heterogeneity of CEA, it is essential that the same assay test method is used for serial monitoring of the patient.

Clinical Application

Before surgical resection, baseline CEA values are typically obtained to confirm successful removal of the tumor burden. After surgery and during chemotherapy, it is recommended that CEA levels be serially monitored every 2 to 3 months to detect recurrence and determine therapy efficacy; the half-life of CEA is approximately 2 to 8 days depending on the assay and the individual. If treatment is successful, CEA levels should drop into the reference ranges in 1 to 4 months. CEA is not recommended for screening asymptomatic individuals for colorectal cancer. While there are no specific guidelines recommending the use of CEA in other types of cancer, it may be of value for detecting recurrence of antigen-positive breast and gastrointestinal cancers and medullary thyroid carcinoma and to aid in the diagnosis of non–small cell lung cancer.

TNM Staging System

T—tumor size and involvement/invasion of nearby tissue; scale 0–4

N—regional lymph nodes involvement; scale 0–3

M—metastasis; extent of tumor spreading from one tissue to another; scale 0–1

Example grading of a tumor:

T1 N0 M0 = small tumor, no nodal involvement, and no metastasis

CASE STUDY 28.2, PART 2

Remember Tatenda. During Tatenda's colonoscopy, a biopsy was obtained from a mass in the sigmoid colon. The pathology report indicated adenocarcinoma.

1. Is the CEA test useful as a screening test for colon carcinoma?
2. What other conditions can result in elevated CEA levels?
3. How will CEA be used to monitor Tatenda after his colorectal surgery?

Human Chorionic Gonadotropin (hCG)

Human chorionic gonadotropin (hCG) is a dimeric hormone normally secreted by trophoblasts to promote implantation of the blastocyst and the placenta to maintain the corpus luteum through the first trimester of pregnancy. Some types of tumor invasion mimic uterine implantation, except that implantation in pregnancy is regulated and limited. hCG is elevated in trophoblastic tumors, mainly choriocarcinoma, and germ cell tumors of the ovary and testis.

Regulation and Physiology

hCG is a 45-kD glycoprotein consisting of α- and β-subunits. A unique aspect of hCG is that it is degraded into multiple fragments. In serum, this results in the presence of the intact molecule, nicked hCG, the free β-subunit (β-hCG), and a hyperglycosylated intact form. Either intact hCG or the free β-subunit may be elevated in malignancies, and most assays detect multiple fragments of hCG.

Clinical Usefulness and Interpretation

hCG has several clinical applications as a tumor marker. It is a prognostic indicator for ovarian cancer, a diagnostic marker for classification of testicular cancer, and the most useful marker for detection of gestational trophoblastic diseases (GTDs).[2] GTDs include four distinct types of tumors (hydatidiform mole, persistent/invasive gestational trophoblastic neoplasia, choriocarcinoma, and placental site trophoblastic tumors) that are classified by clinical history, ultrasound, histology, and hCG levels. hCG is invariably elevated in women with GTDs[14] and is often found at higher levels than are observed in normal pregnancy (i.e., >100,000 mIU/mL). It is particularly a helpful marker for monitoring GTD therapy, as levels of hCG correlate with tumor mass and prognosis; hCG is not actually cleared by the FDA for use as a tumor marker despite its widespread utility.

Methodology

hCG can be measured by using any of a variety of widely available automated immunoassays. Typical assays use monoclonal capture and tracer antibodies targeted toward epitopes in the β-subunit and intact hCG. Total β-hCG assays are the most useful assays because they detect both intact hormone and free β-hCG. Due to the variability in hCG assays,[15] it is imperative that patients be monitored with the same technique. It is also important for laboratories to be aware of the relative cross-reactivity of their assay with different hCG isoforms; because hCG assays are designed to detect pregnancy, they are not all equivalent for application as tumor markers.

Clinical Application

In testicular cancer, the free β-hCG subunit is elevated in 60% to 70% of patients with nonseminomas. hCG can be used in combination with AFP and biopsies to diagnose subtypes of testicular cancer (Table 28.4). Ectopic β-hCG is also occasionally elevated in ovarian cancer and some lung cancers. In practice, free β-hCG is sensitive and specific for aggressive neoplasms; the free β-hCG is not detectable in the serum of healthy subjects.

Prostate-Specific Antigen (PSA)

PSA is a 28-kD glycoprotein produced in the epithelial cells of the acini and ducts of the prostatic ducts in the prostate. It is a serine protease of the kallikrein gene family. It functionally regulates seminal fluid viscosity and is instrumental in dissolving the cervical mucus cap, allowing sperm to enter.

Regulation and Physiology

In healthy men, low circulating levels of PSA can be detected in the serum. There are two major forms of PSA that are found circulating in the blood: (1) free and (2) bound. Most of the circulating PSA is bound to α_1-antichymotrypsin or α_2-macroglobulin. Chemistry assays detect total PSA and free PSA. While the detection of total PSA has been used in screening for and monitoring of prostate cancer, there is also evidence that detecting free PSA as a fraction of total is useful.[16] Patients with malignancy have a lower percentage of free PSA. As with other tumor markers, PSA is not entirely specific. Men with benign prostatic hyperplasia (BPH) and prostatitis can also have high PSA levels. Additional markers, such as prostate cancer gene-3 (*PCA-3*), are starting to be used to address this lack of specificity. At this time, prostate cancer tumor markers remain controversial and are actively researched to improve the sensitivity and specificity.

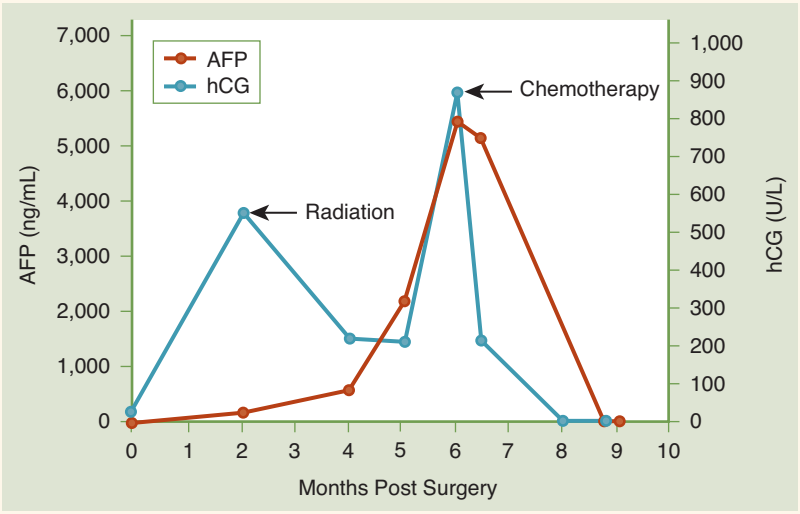
Clinical Usefulness and Interpretation

Several large clinical trials have shown that although monitoring PSA levels can reduce mortality from prostate cancer, it is only a small reduction in overall risk.[17,18] Moreover, harm from screening in the form

CASE STUDY 28.3, PART 2

Two months after Marcus's surgery, he started radiation treatments, and 6 months after surgery he began chemotherapy. Both β-hCG and AFP were monitored regularly during treatment. **See Case Study Figure A.**

© StockLite/Shutterstock.



A

1. Based on Marcus's tumor markers results, what type of cancer did he have that required surgery?
2. Explain the pattern of AFP and hCG observed in this Case Study Figure A.
3. Can a final diagnosis be made based only on the tumor marker findings? Explain.

of biopsies and treatment may outweigh the benefits. Thus, current recommendations focus on informed decision making, where the risks and benefits are outlined and individual patients decide whether or not to be screened.[19–21] Conversations about screening should begin at age 50 (American Cancer Society and American Urological Association).[22] Men at higher risk include those who have a first-degree relative with prostate cancer or who are Black/African American. Black/African American men are more likely to be diagnosed with prostate cancer and 2.5 times more likely to die from it than other men, in part due to lack of access to quality medical care. These higher-risk individuals should consider screening starting between 40 to 45 years of age. Serial monitoring is appropriate with a 2-year screening interval for men with PSA levels less than 2.0 ng/mL. Screening should always include a digital rectal examination (DRE). Screening utility decreases with age and is not appropriate for men with less than a 10-year life expectancy. In addition to the use of standard cutoff values of total PSA (<4 ng/mL is generally considered normal), other measurements of PSA have been used to test for prostate malignancy; age-adjusted cutoff values of PSA, PSA velocity (rate of rise over time), and free PSA/total PSA ratios have been used to increase the accuracy of PSA testing.[23] The reference range for PSA is age dependent. A cutoff value of 4 ng/mL is used. If the total PSA is 4.0–10 ng/mL, then a free PSA test should be performed and a free PSA/total PSA ratio calculated. A ratio ≤ 0.1 indicates 50% risk of prostate cancer.[16] It is recommended that men with a total PSA greater than 4 ng/mL and/or a clinical suspicion of cancer by DRE undergo biopsy to confirm the presence of prostate cancer.

Prostate infection, irritation, and enlargement (also known as benign prostatic hyperplasia [BPH]) can result in increased PSA levels. Moreover, recent ejaculation or DRE can also lead to increases in circulating PSA.[24] It has also been recognized that cancer can be present at all concentrations of PSA (**Table 28.7**).[25,26] Taken together, these studies urge the interpretation of PSA in the context of the clinical picture (including DRE, imaging, and family history).

Methodology

PSA is measured by immunoassay, which detects both free PSA and PSA complexed with α_1-antichymotrypsin but not α_2-macroglobulin. Most immunoassays commercially available use enzyme, fluorescence, or chemiluminescence on an automated immunoassay

CASE STUDY 28.4, PART 2

Remember Drew. At Drew's follow-up appointment, his provider shared that his PSA level was elevated (7.1 ng/mL), and the pathology report from the biopsy revealed the presence of carcinoma. A radical prostatectomy was scheduled.

1. Based on the total PSA result, should a free PSA be performed? Provide a rationale.

Table 28.7 Prevalence of Prostate Cancer at Different Prostate-Specific Antigen Concentrations

PSA Concentration (ng/mL)	Prevalence of Prostate Cancer (%)
<1	6–10
1–4	17–25
4–10	20–30
>20	>80

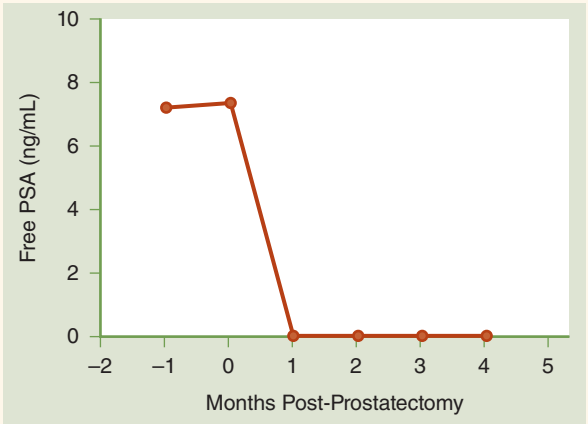
platform. Because antibodies recognizing different epitopes may recognize the multiple forms of PSA variably, there can be some discrepant PSA results between manufacturers. Known interferences that have been reported for PSA include both the hook effect[27] and HAMAs (see above).[28,29]

Clinical Application

The best clinical use and first clinical application of PSA testing were to monitor for the progression of prostate cancer after therapy. After radical prostatectomy, serum PSA should become undetectable if the cancer is localized. This use of PSA to monitor cancer

CASE STUDY 28.4, PART 3

Remember Drew. Another PSA level was drawn as Drew was being prepped for surgery (day 0). The result was reported as 7.2 ng/mL. (**See Case Study Figure A**)



A

2. Explain the PSA pattern observed in Case Study Figure A.
3. Is there any evidence of residual disease or recurrence in this patient?

progression has also been found useful after radiation or endocrine therapy.[30]

Cancer Antigen 15-3 (CA 15-3)

Regulation and Physiology

Breast cancer is the most common form of cancer in women in the United States. As a result, breast cancer research continues to try to identify more precise biomarkers for its diagnosis as well as targeted gene therapies. Mucin 1 (MUC1) is a transmembrane glycoprotein and is the most researched tumor-associated antigen.[31] It is normally expressed in the glandular or luminal epithelial cells of the mammary gland, esophagus, stomach, duodenum, pancreas, uterus, prostate, and lungs[32] and is responsible for encoding tumor-associated antigens including CA 15-3 and CA 27.29.

Clinical Usefulness and Interpretation

When used in conjunction with clinical information and ancillary diagnostic procedures, CA 15-3 is useful for managing breast cancer patients. It is important in the early detection of breast cancer recurrence in previously treated breast cancer patients. It is also useful for monitoring response to therapy in patients with metastatic breast cancer. The reference range is less than 30 U/mL. In serial testing, increasing CA 15-3 results may indicate disease recurrence requiring the need for additional ancillary testing and/or procedures. Conversely, decreasing CA 15-3 results correlate with disease regression.

Methodology

CA 15-3 can be detected by immunoenzymatic sandwich assays, which use a biotinylated monoclonal CA 15-3-specific antibody and a monoclonal CA 15-3-specific antibody. CA 15-3 in the patient's serum binds to both the biotinylated monoclonal CA 15-3-specific antibody (mouse) and the monoclonal CA 15-3-specific antibody (mouse) labeled with a ruthenium complex, forming a sandwich complex. Unbound substances are then removed using a wash step.

Clinical Application

The best clinical application of CA 15-3 serial testing is to assist in early detection of breast cancer recurrence in previously treated stage II and III breast cancer patients.

Carbohydrate Antigen 19-9 (CA 19-9)

Regulation and Physiology

The carbohydrate antigen 19-9 (CA 19-9) is a modified Lewis (Le^a) blood group antigen. Individuals who are Le (a-b-) do not express CA 19-9 because these individuals lack fucosyltransferase, an enzyme needed for CA 19-9 production.[33]

Clinical Usefulness and Interpretation

When paired with a surgical pathology biopsy of the pancreas, CA 19-9 is useful for diagnosis and monitoring of pancreatic cancer. It may be useful in differentiating patients with cholangiocarcinoma and primary sclerosing cholangitis.[34] A single specimen for the measurement of CA 19-9 has limited value, and serial monitoring should begin prior to therapy to establish a baseline value. CA 19-9 may be elevated in patients with GI malignancies such as bile duct cancer, known as cholangiocarcinoma, pancreatic cancer, and/or colon cancer. Benign conditions such as pancreatitis and cholestasis, a condition characterized by bile flow slowing and stopping, can also cause slightly increased serum CA 19-9 concentrations.

Methodology

CA 19-9 can be detected by immunoenzymatic sandwich assays involving polyclonal goat-antibiotin antibody, mouse monoclonal-biotin conjugate, and a buffered protein solution. After formation of the labeled immune complex and a separation step, which removes unbound reagents, a monoclonal-alkaline phosphatase conjugate is added. After incubation, unbound reagents are washed away. A chemiluminescent substrate is then added. Finally, the chemiluminescent reaction is measured with a luminometer. The light produced by the chemical reaction is directly proportional to the concentration of CA 19-9 in the patient's sample.

Clinical Application

CA19-9 is the best-validated biomarker for pancreatic cancer. Recent research has shown that not all Lewis antigen (-) patients with pancreatic cancer are non-secretors of CA19-9.[35]

Immunoglobulin Free Light Chains

Regulation and Physiology

Multiple myeloma is a blood cancer of plasma cells producing increased levels of immunoglobulins, which is visualized as a monoclonal gammopathy on serum protein electrophoresis. Waldenström macroglobulinemia is an example of a IgM monoclonal gammopathy. Monoclonal immunoglobulins are initially detected by serum protein electrophoresis and confirmed by immunofixation electrophoresis. However, some monoclonal light chain diseases do not produce immunoglobulin levels in high enough concentrations to be detected and quantitated by serum protein electrophoresis or immunofixation.

Clinical Usefulness and Interpretation

Serum immunoglobulin free lights chains is useful for monitoring patients with monoclonal light chain diseases without an M-spike identified on serum protein electrophoresis. There are two types of light chains measured: κ and λ. In addition to measuring each light chain, the κ and λ (K/L) light chain ratio is calculated. After diagnosis of a monoclonal light chain disease is made, the free light chain assay is used to monitor disease activity. Increases and decreases in the quantity of free light chains reflects changes in the size of the monoclonal plasma cell population.

Methodology

Serum κ or λ free light chains can be measured by automation using nephelometry. The patient sample containing the κ or λ free light chain (antigen) is added to a cuvette containing its corresponding κ or λ antibody. The beam of light passing through the cuvette is scattered as insoluble antigen-antibody immune complexes are formed. Light scatter is measured at an angle, other than 180 degrees, away from the incident light. The quantity of immune complexes formed is directly proportional to the free light chain concentration.

Clinical Application

There are three major clinical applications for the free light chain (FLC) assay in the evaluation and management of multiple myeloma and other plasma cell disorders. The serum FLC assay minimizes the need to quantitate or to perform urine protein electrophoresis. Baseline FLC measurement provides prognostic value in plasma cell dyscrasias. It also allows for monitoring of patients who had previously been deemed to have non-secretory myeloma.[36]

Human Epididymis Protein 4 (HE4)

Regulation and Physiology

The function of human epididymis protein 4 (HE4) is currently unknown. It has been shown to be overexpressed in ovarian carcinomas. In a study of 233 patients with a pelvic mass, including 67 with epithelial ovarian cancer, HE4 had a higher sensitivity for ovarian cancer detection than CA-125 at a specificity of 95%.[37] Researchers also found HE4 to be elevated in more than half of the ovarian cancer patients who did not have elevated CA-125 levels.

Clinical Usefulness and Interpretation

Human epididymis protein 4 (HE4) measurement is part of the risk of ovarian malignancy algorithm (ROMA) for women who present with a tissue mass in or near the uterus, ovaries, or fallopian tubes on imaging tests such as CT scan or MRI. The HE4 reference range is less than 140 pmol/L. Elevated levels of HE4 suggests disease progression or recurrence, while a decrease suggests disease regression or therapeutic response to the treatment plan. In pre-menopausal females, a ROMA cutoff value of 1.14 is used to differentiate low risk from high risk of finding epithelial ovarian cancer during at surgery. In postmenopausal females, a ROMA cutoff value of > 2.99 is used to differentiate low risk from high risk of finding epithelial ovarian cancer during at surgery.[38]

Methodology

HE4 can be detected by immunoenzymatic sandwich assays. Similar to other tumor marker assays, the immunoassay uses a biotinylated monoclonal HE4-specific antibody and a monoclonal HE4-specific antibody labeled with a chemiluminescent ruthenium label. The patient's specimen containing HE4 reacts with both the biotinylated monoclonal HE4-specific mouse antibody and the monoclonal HE4-specific mouse antibody labeled with ruthenium, forming an immune sandwich complex.

Clinical Application

Human epididymis protein 4 (HE4) is primarily used for ovarian cancer. HE4 offers improved specificity over CA-125, due to the presence of elevated

CA-125 found in nonmalignant conditions, such as endometriosis.[24]

Neuron-Specific Enolase (NSE)

Regulation and Physiology

Enolase is an enzyme that catalyzes the conversion of 2-phosphoglycerate to phosphoenolpyruvate. It had three isoenzymes: α, β, and γ. Neuron-specific enolase (NSE) is the predominant γ isoenzyme found in neuroendocrine tissues. Due to its presence in erythrocytes, hemolyzed samples are rejected. Measurement of NSE in serum or CSF assists in the differential diagnosis of neurodegenerative disorders. An elevated CSF concentration assists in the diagnosis of Creutzfeldt-Jakob Disease.[40]

Clinical Usefulness and Interpretation

NSE is useful as a follow-up marker in patients with NSE-secreting tumors. In conjunction with a surgical pathology biopsy specimen, NSE is an auxiliary test in the diagnosis of small cell lung carcinoma.[41] It is also an auxiliary test in the diagnosis of carcinoid tumors,[42] pancreatic islet cell tumors,[43] and neuroblastomas.[44] The NSE reference range is less than 15 ng/mL. With successful treatment, serum NSE concentrations begin to decrease within 24 hours due to its short half-life. NSE concentrations that remain approximately the same indicates unsuccessful treatment. Increasing concentrations indicate tumor metastasis or disease recurrence.

Methodology

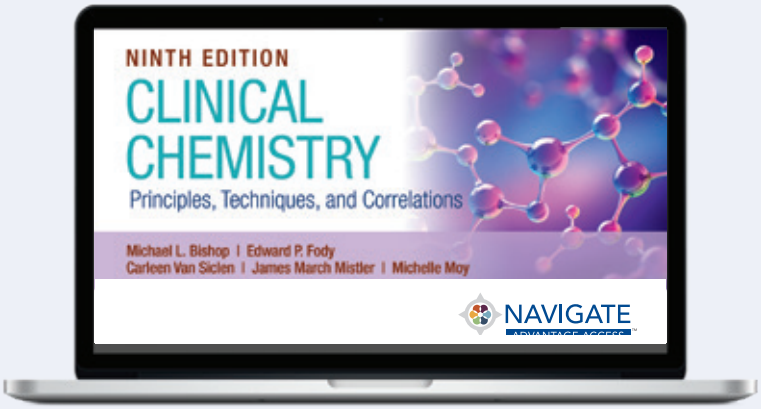
NSE may be measured by a homogeneous immunofluorescent sandwich assay. NSE is sandwiched between the two antibodies. When the immune sandwich complex is excited, the energy is emitted as fluorescence. A ratio of the fluorescent energy emitted to that emitted at the internal reference is calculated for each sample. Signal intensity is proportional to the number of fluorescent immune complexes formed, which is directly proportional to the NSE antigen concentration.

Clinical Application

NSE is elevated in patients with small cell lung carcinoma. Other neuroendocrine tumors with an elevated NSE concentration include carcinoid tumors, islet cell tumors, and neuroblastomas.[39-44]

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



Suggested Reading

Milose JC, Filson CP, Weizer AZ, et al. Role of biochemical markers in testicular cancer: diagnosis, staging, and surveillance. *Open Access J Urol*. 2011;4:1–8. Published 2011 Dec 30. doi:10.2147/OAJU.S15063

Diamandis EP, Fritsche HA, Lilja H, et al. *Tumor Markers: Physiology, Pathobiology, Technology, and Clinical Applications*. Washington, DC: AACC Press; 2002.

Wu JT. *Circulating Tumor Markers of the New Millennium: Target Therapy, Early Detection, and Prognosis*. Washington, DC: AACC Press; 2002.

References

1. Howlader N, Noone AM, Krapcho M, et al., eds. *SEER Cancer Statistics Review*. Bethesda, MD: National Cancer Institute; 1975–2017. http://seer.cancer.gov/archive csr/1975_2017/.
2. Abeloff MD, Armitage JO, Niederhuber JE, et al., eds. *Abeloff's Clinical Oncology*. 4th ed. Philadelphia, PA: Churchill Livingstone; 2008.
3. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100:57–70.
4. Clinical and Laboratory Standards Institute. *Standards and Guidelines for Clinical Laboratories*. Accessed October 20, 2008. http://www.clsi.org
5. Klee GG. Human anti-mouse antibodies. *Arch Pathol Lab Med*. 2000;124:921–923.
6. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol*. 2007;25:5287–5312.
7. Bader D, Riskin A, Vafsi O, et al. Alpha-fetoprotein in the early neonatal period—a large study and review of the literature. *Clin Chim Acta*. 2004;349:15–23.
8. Gupta S, Bent S, Kohlwes J. Test characteristics of alpha fetoprotein for detecting hepatocellular carcinoma in patients with hepatitis C. A systematic review and critical analysis. *Ann Intern Med*. 2003;139:46–50.
9. Diamandis EP, Hoffman BR, Sturgeon CM. National Academy of Clinical Biochemistry laboratory medicine practical guidelines for the use of tumor markers. *Clin Chem*. 2008;54:1935–1939.
10. Raghavan D. *Germ Cell Tumors*. London and Ontario: BC Decker; 2003.
11. Yin BW, Lloyd KO. Molecular cloning of the CA125 ovarian cancer antigen: identification as a new mucin, MUC16. *J Biol Chem*. 2001;276:27371–27375.
12. Duffy MJ, Bonfrer JM, Kulpa J, et al. CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. *Int J Gynecol Cancer*. 2005;15:679–691.
13. Duffy MJ, van Dalen A, Haglund C, et al. Clinical utility of biochemical markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines. *Eur J Cancer*. 2003;39:718–727.
14. Bower M, Rustin G. Serum tumor markers and their role in monitoring germ cell cancers of the testis. In: Vogelzang N, Scardino P, Shipley W, et al., eds. *Comprehensive Textbook of Genitourinary Oncology*. 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2000:931.
15. Cole LA, Sutton JM, Higgins TN, Cembrowski GS. Between method variation in human chorionic gonadotropin test results. *Clin Chem*. 2004;50:874–882.
16. Ito K, Yamamoto T, Ohi M, et al. Free/total PSA ratio is a powerful predictor of future prostate cancer morbidity in men with initial PSA levels of 4.1 to 10.0 ng/mL. *Urology*. 2003;61(4):760–764. doi:10.1016/s0090-4295(02)02427-5
17. Schröder FH, Hugosson J, Roobol MJ, et al. Prostate-cancer mortality at 11 years of follow-up. *N Engl J Med*. 2012;366(11): 981–990.
18. Andriole GL, Crawford ED, Grubb RL, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med*. 2009;360(13):1310–1319.
19. American Cancer Society Prostate Cancer Screening Decision Support. Accessed March 21, 2016. http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-024618.pdf.
20. American Society for Clinical Oncology Prostate Cancer Screening Decision Support. Accessed March 21, 2016. https://www.asco.org/sites/; http://www.asco.org/files/psa_pco_decision_aid:71612.pdf.
21. Centers for Disease Control Prostate Cancer Screening Decision Support. Accessed March 21, 2016. http://www.cdc.gov/cancer/prostate/basic_ info/infographic.htm.
22. Smith RA, Cokkinides V, Eyre HJ. American Cancer Society guidelines for the early detection of cancer, 2006. *CA Cancer J Clin*. 2006;56:11–25; quiz 49–50.
23. Loeb S, Catalona WJ. Prostate-specific antigen in clinical practice. *Cancer Lett*. 2007;249:30–39.
24. Tarhan F, Orcun A, Kucukercan I, et al. Effect of prostatic massage on serum complexed prostate-specific antigen levels. *Urology*. 2005;66:1234–1238.
25. Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med*. 2003;349:215–224.
26. Thompson IM, Pauler DK, Goodman PJ, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or = 4.0 ng per milliliter. *N Engl J Med*. 2004;350:2239–2246.
27. Akamatsu S, Tsukazaki H, Inoue K, et al. Advanced prostate cancer with extremely low prostate-specific antigen value at diagnosis: an example of high dose hook effect. *Int J Urol*. 2006;13:1025–1027.
28. Park S, Wians FH Jr, Cadeddu JA. Spurious prostate-specific antigen (PSA) recurrence after radical prostatectomy: interference by human antimouse heterophile antibodies. *Int J Urol*. 2007;14:251–253.
29. Descotes JL, Legeais D, Gauchez AS, et al. PSA measurement following prostatectomy: an unexpected error. *Anticancer Res*. 2007;27:1149–1150.
30. Roach M III, Hanks G, Thames H Jr, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus Conference. *Int J Radiat Oncol Biol Phys*. 2006;65:965–974.
31. Nath S, Mukherjee P. MUC1: a multifaceted oncoprotein with a key role in cancer progression. *Trends Mol Med*. 2014;20(6):332–342. doi:10.1016/j.molmed.2014.02.007
32. Jing X, Liang H, Hao C, Yang X, Cui X. Overexpression of MUC1 predicts poor prognosis in patients with breast cancer. *Oncol Rep*. 2019;41(2):801–810. doi:10.3892/or.2018.6887
33. Scara S, Bottoni P, Scatena R. CA 19-9: Biochemical and clinical aspects. *Adv Exp Med Biol*. 2015;867:247–260. doi: 10.1007/978-94-017-7215-0_15
34. Torok N, Gores GJ. Cholangiocarcinoma. *Semin Gastrointest Dis*. 2001 Apr;12(2):125–132.
35. Luo G, Fan Z, Cheng H, et al. New observations on the utility of CA19-9 as a biomarker in Lewis negative patients with pancreatic cancer. *Pancreatology*. 2018;18(8):971–976. doi:10.1016/j.pan.2018.08.003
36. Dispenzieri A, Kyle R, Merlini G, et al. International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. *Leukemia*. 2009;23(2): 215–224. doi: 10.1038/leu.2008.307
37. Ferraro S, Braga F, Lanzoni M, et al. Serum human epididymis protein 4 vs carbohydrate antigen 125 for ovarian cancer diagnosis: a systematic review. *J Clin Pathol*. 2013 Apr;66(4):273–281.

38. Karlsen MA, Sandhu N, Hogdall C, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol Oncol*. 2012;127(2):379–383.
39. Burghuber OC, Worofka B, Schernthaner G, et al. Serum neuron-specific enolase is a useful tumor marker for small cell lung cancer. *Cancer*. 1990;65:1386–1390.
40. Aksamit AJ, Preissner CM, Homburger HA: Quantitation of 14-3-3 and neuron-specific enolase proteins in CSF in Creutzfeldt-Jacob disease. *Neurology*. 2001;57:728–730.
41. Huang L, Zhou JG, Yao WX, et al. Systematic review and meta-analysis of the efficacy of serum neuron-specific enolase for early small cell lung cancer screening. *Oncotarget*. 2017;8(38):64358–64372.
42. Lamberts SW, Hofland LJ, Nobels FR: Neuroendocrine tumor markers. *Front Neuroendocrinol*. 2001;22:309–339.
43. Portela-Gomes GM, Hacker GW, Weitgasser R: Neuroendocrine cell markers for pancreatic islets and tumors. *Appl Immunohistochem Mol Morphol*. 2004;12:183–192.
44. Riley RD, Heney D, Jones DR, et al. A systematic review of molecular and biological tumor markers in neuroblastoma. *Clin Cancer Res*. 2004;10;4–12.

Point-of-Care Testing

Heather McNasby

CHAPTER OUTLINE

KEY TERMS

Connectivity
High-complexity tests
Moderate-complexity tests
Point-of-care testing (POCT)
Standardization
Waived tests

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define point-of-care testing (POCT).
- Explain the role of the laboratory in the management of a POC program.
- Explain the process of implementing a new POC test.
- State the basic principles behind common POC applications.

CASE STUDY 29.1, PART 1

Two laboratorians, Miles and Mía, are assigned the task of developing an implementation plan, training plan, and procedures for a new point-of-care (POC) test.

Assist Miles and Mía by providing a list of components required to effectively implement the new test.

Point-of-care testing (POCT) is defined as "those analytical patient-testing activities provided within the institution, but performed outside the physical facilities of the clinical laboratories."[1] POCT is also known by many names, including "near-patient testing,"[2] "extra-laboratory analyses,"[3] "ancillary testing," "bedside testing,"[4] "physician's office testing,"[5] and "alternative site testing." A convenience for the patient and physician, POCT brings the laboratory to the patient. POCT is increasingly being used not only in the emergency department, operating rooms, and intensive care units (ICUs) but also in outpatient specialty clinics, physician offices, urgent care locations, nursing homes, pediatric units, pharmacies, counseling centers, and ambulances. Advances in medical, analytical, and engineering technologies in the last two decades led to the appearance of a large number of portable measurement devices for a variety of different analytes. These include tests for electrolytes, glucose, creatinine, hemoglobin A1c (HbA1c), urinalysis, pregnancy, drugs of abuse, therapeutic drug monitoring, occult blood, blood gases, coagulation tests, enzymes, and cardiac markers. POCT is also available for infectious diseases such as HIV, gonorrhea, syphilis, streptococcus, influenza, respiratory viruses such as coronavirus, fungal infections, and tuberculosis.

POCT offers several advantages compared with central laboratory testing (**Table 29.1**). The major advantage of POCT is faster delivery of results. POCT is designed to work with the flow of patient management and care, allowing physicians to order and perform the test and immediately make a medical decision.

Smaller sample volume allows POCT use in neonatal and pediatric population and is preferred for those undersized patients requiring frequent testing. The overall cost of patient care is arguably lower due to improved patient workflows. Testing near the patient requires fewer steps than transporting an appropriate specimen to the laboratory, processing, aliquoting, testing, and communicating results back to clinical staff. POCT therefore aids in the reduction of preanalytical and postanalytical laboratory-related errors. Portability of POCT allows for increased access to testing in a wider variety of sites. Examples of these include rural areas, accident sites, areas with limited infrastructure and personnel, and locations with underserved populations. One example of the latter is the use of a portable POCT for tuberculosis in at risk populations in countries where it is endemic.[6]

Although POCT offers many advantages, there are some negative aspects of POCT (Table 29.1). On a cost-per-test basis, POCT is more expensive than central laboratory testing due to higher consumable and reagent costs associated with POCT analyzers. End users may be able to counterbalance these increased costs by gaining efficiencies in improved patient workflows and increased patient satisfaction. Additionally, quality of results is a concern because POCT is usually performed by nonlaboratory personnel who may include nurses, physicians, emergency medical technicians, paramedics, pharmacists, patients, and medical office assistants. End users with multiple responsibilities and limited laboratory training may not appreciate the value of laboratory quality control (QC) in ensuring quality test results. Compliance with QC requirements continues to challenge many POC coordinators.

Laboratory Regulations

Accreditation

In the United States, laboratory testing for patient care (with limited exceptions) requires a Clinical Laboratory Improvement Amendments (CLIA) certificate. CLIA was enacted by Congress and established quality standards for all laboratory testing, thereby ensuring the accuracy, reliability, and timeliness of patient test results regardless of where the test was performed.[7] The requirements are based on the complexity of the test and not the type of laboratory where the testing is performed (POC or central laboratory). All laboratories performing POCT must be certified under one of the five types of CLIA certificates listed in **Table 29.2**. The CLIA certificate must be appropriate for the testing that is performed in the laboratory (i.e., appropriate complexity). In addition to the federal program, state departments of public health or public organizations may apply for "deemed" status, which allows these organizations to perform laboratory accreditation and inspections.

POCT Complexity

Tests are classified based on their complexity. The Food and Drug Administration (FDA) uses several criteria to assign complexity to any test, and the three categories are **waived tests**, **moderate-complexity tests**, and **high-complexity tests**. Waived tests are a category of tests defined by CLIA, such as dipstick tests, urine pregnancy tests, and blood glucose monitoring devices, that are subject to the lowest level of regulation and are cleared by the FDA for home

Table 29.1 Potential Advantages and Disadvantages of POCT[a]

Advantages	Disadvantages
Convenience for both the attending clinician and patient	POCT is significantly more expensive than the cost of central laboratory testing due to the higher consumable and reagent costs associated with POCT analyzers
Reduced turnaround time for test results, expediting medical decision making in operating rooms, emergency departments, and intensive care units	Maintenance of quality control and quality assurance of test results are difficult. A multitude of health care providers perform POCT, and results can be reported without quality control in the absence of QC lockout in some POCT analyzers
Reduction in clinic visits, hospital admissions, and length of hospital stay due to faster turnaround times for laboratory services	Management of POCT is challenging—there are numerous operators to train, multiple sites to manage, and hundreds of POC tests to validate
Better patient management due to improved visibility and interaction with a patient	Preanalytic, analytic, and postanalytic issues are not recognized easily as tests are performed by nonlaboratory individuals
Decreased personnel associated with test requesting and reporting, especially for patients necessitating tests several times a day (e.g., glucose determination)	Interindividual variability in POCT results may be greater when compared with central laboratory testing due to different test methodologies
Fingerstick POCT is less traumatic for a patient as smaller sample volume is required	Difficulties with documentation of test results, billing, and regulatory compliance
Reduced risk of preanalytical errors including specimen collection and labeling, due to absence of transporting a specimen to a central laboratory, processing and aliquoting prior to testing. Reduced risk of sample deterioration	Proper integration of test result into the patient's electronic medical record may be more difficult if tested on a device that does not directly cross the interface to the laboratory or hospital information systems
Improved patient outcome with immediate access to laboratory result(s)	Managing reagent lot numbers, reagent and control expiration dates at multiple sites can be problematic
Improved cost of overall patient care	Central laboratory test results and POCT results are not always comparable due to differences in specimen types, test methodologies, interfering substances, etc.
Wide menu of POC analytes for specimens that do not require processing	
Availability to a wider variety of sites (e.g., rural areas, areas with limited infrastructure/personnel, and sites with underserved populations)	

[a] POCT, point-of-care testing.

uses. They employ methodologies that are so simple to perform that they render the likelihood of erroneous results negligible and pose almost no risk of harm to the patient if the test is performed incorrectly. A full list of currently waived tests can be found on the FDA website (www.fda.gov). Laboratories performing non-waived tests (moderate-complexity tests and high-complexity tests) must fulfill all the requirements for personnel qualifications, proficiency testing, inspections, maintenance, quality assurance, quality control, and patient test management.

The fourth category of POCT is *provider-performed microscopy procedures (PPMPs)*, which is a subcategory of moderate-complexity testing. These tests involve the use of a microscope, limited to bright-field or phase-contrast microscopy. Generally, the specimens are labile and cannot survive transport to a clinical laboratory. Only licensed physicians, dentists, and midlevel practitioners (i.e., nurse practitioner or physician assistant) may perform PPMP. There are usually no QC materials available for PPMP; however, the individual performing PPMP must participate in the

Table 29.2 Types of CLIA Certificates[a]

Certificate of Waiver	Issued to a laboratory to perform only waived tests
Certificate of Registration	Issued to a laboratory that enables the entity to conduct moderate- or high-complexity laboratory testing or both until the entity is determined by survey to be in compliance with CLIA regulations
Certificate for PPMPs	Issued to a laboratory in which a physician, midlevel practitioner, or dentist performs no tests other than the microscopy procedures Permits the laboratory to also perform waived tests
Certificate of Compliance	Issued to a laboratory after an inspection that finds the laboratory to be in compliance with all applicable CLIA requirements
Certificate of Accreditation	Issued to a laboratory on the basis of the laboratory's accreditation by an accreditation organization approved by the Health Care Finance Administration

[a] CLIA, Clinical Laboratory Improvement Amendments; POCT, point-of-care testing; PPMP, provider-performed microscopy procedure.

Data from Clinical Laboratory Improvement Act. How to Obtain a CLIA Certificate of Waiver, http://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/downloads//howobtaincertificateofwaiver.pdf

training and certification process, as well as participate in proficiency testing.

To determine to which category POC test or procedure belongs, refer to the FDA website for an FDA-approved test menu: www.fda.gov.

Implementation

Establishing Need

When establishing a POCT site, the approach and steps in the implementation will be different in nearly every case insofar as organizational culture and the financial situation of any given institution may vary greatly. However, some common guidelines to implementation apply to nearly every case. Support of the organization must be present as well as the willingness and readiness to evaluate the requests for POCT. Typically, a multidisciplinary committee consisting of laboratory staff, physicians, nursing representatives, and hospital/laboratory administration is formed to create a structure for receiving and establishing criteria for approving requests for POCT. Before implementing a POCT service, an interdisciplinary committee should address the questions similar to those listed in **Box 29.1**. The decision to establish a POCT program is made after justifying the clinical need, cost comparisons, and the analytic performance requirements. Other factors should also be considered, such as space, previous track record of regulatory compliance, testing personnel, laboratory personnel who will be managing the training and overseeing the POCT program, and measurable quality metrics to determine if the POCT implementation is successful.

POCT Implementation Protocol

After making the decision to implement POCT, the responsible party selects the method of choice and begins validation. Desirable characteristics of POCT analyzers include ease of use, method accuracy, comparability to the central laboratory method, portability, durability, low maintenance, simple QC, QC lockout features, operator lockout features, simple sample handling requirements, barcode patient and operator identification capabilities, and the ability to interface with a laboratory information system (LIS) and/or electronic medical record (EMR). The manufacturer should aid in providing minimum acceptability requirements, instrument manuals, package inserts for reagents and QCs, materials safety data sheet, and training materials. The manufacturer also typically assists with instrumentation setup and validation studies. Method validation should confirm the manufacturer's specifications. A *procedure/policy* should be written for each test with the assistance of the manufacturer. This should not be simply a collection of materials obtained from a manufacturer; rather, these documents should aid in the development of an easy-to-follow, simple, and concise procedure that aligns with the location implementing the test. The procedure should include information on the principle of the method; personnel qualifications; specimen type and handling, reagents, supplies, and equipment requirements; QC and calibrations; patient-testing procedure; troubleshooting guidelines; reference ranges; reporting limits; and interfering substances if applicable. All required regulatory or certifying standards should be in place prior to initiating patient testing (**Box 29.2**).

Box 29.1 Questions to Ask When Implementing POCT

- Which test is required and in which specific area?
- How is the service currently provided, and is this a means to resolve a problem?
- What is the expected annual test volume?
- What clinical question is being asked when requesting this test?
- What clinical decision is likely to be made and action to be taken upon receipt of the result?
- How will POCT increase patient satisfaction?
- Are personnel competent to perform the POCT? Are facilities available to perform the test and store equipment, reagents, and documentation?
- Will a change in practice be required?
- Can the central laboratory deliver the required service? How do POCT cost, reference ranges, accuracy, and precision compare to similar tests in a laboratory?
- Are available POCT analyzers user friendly? Can POCT analyzers interface with a laboratory information system (LIS) and a patient electronic medical record (EMR)?
- Are both internal and external QC materials available?
- What proficiency testing and maintenance are required, and can staff workload accommodate the requirements?
- Does the company providing the POCT analyzer(s) and reagents offer a reliable installation and support service?

Personnel Requirements

For a POC programs with any complexity tests, including for those POC programs with a Certificate of Waiver or certificate for PPMP, institutions are required to have a certain organizational and administrative structure where staff has established qualifications, competency, and experience. For instance, for POC programs, the director of laboratory is either a PhD scientist or a physician licensed to practice where the laboratory is located.[8] Responsibilities

Box 29.2 Point-of-Care Checklist

1. ***Quality management***
   The POC program has a written QM program as well as organizational system setting forth levels of authority, responsibility, and accountability. There is a documented system to address unusual patient results or instrument troubleshooting. There is a written procedure for each POCT.
2. ***Specimen handling***
   There is a documented procedure describing methods for required patient identification and preparation, as well as for specimen collection, accessioning, and preservation before testing. There is a procedure for entering POCT results into the permanent medical record of a patient. Reference ranges should be established and accompany posted POCT results.
3. ***Reagents***
   All reagents should be stored and labeled as recommended by the manufacturer.
4. ***Instruments and equipment***
   Equipment must be evaluated and scheduled for regular maintenance, as required by the manufacturer. Equipment maintenance documentation is required.
5. ***Personnel***
   The director of the POCT program is a physician or a doctoral scientist. The testing personnel have adequate, specific training to ensure competence and must be documented to meet compliance and/or accreditation standards.
6. ***Quality control and calibration***
   Calibrations and quality controls are performed at regular intervals (at least each day of patient testing). Acceptable limits are defined for control procedures, and there are documented corrective actions when control results exceed defined acceptability limits. Upper and lower limits of the analytical measurement range (AMR) for each analyte are defined.
7. ***Safety***
   There is a program to assure that the safety of patients and healthcare personnel is not compromised by POCT. Specific safety guidelines are provided by the accrediting body.

of a director are very broad and include policy making, ensuring compliance with regulatory standards, and administrative duties. Importantly, the director is responsible for the analytic performance of all tests. The director must make both technical and clinical decisions based on a constant monitoring of ongoing proficiency, quality control, accuracy, and precision. Such individuals should be a liaison between the clinicians, hospital administration, and laboratory personnel.

For laboratories performing moderate-complexity tests, there must also be a technical and a clinical consultant. The *technical consultant* is responsible for scientific oversight of the POCT, while the *clinical consultant* is required to provide clinical and medical advice. The director of a POCT program, the technical consultant, and the clinical consultant may be the same person.

Even though the implementation of the POCT program may be done differently depending on an institution, the entity responsible for quality management of the POCT program is always the laboratory. It is useful, therefore, for a laboratory to have a person supervising the POCT program, so a *POCT coordinator (POCC)* often fulfills this role. The POCC monitors day-to-day activities of testing personnel including POC test result monitoring. It is the responsibility of the POCC to coordinate POC patient testing and facilitate compliance with policies and procedures as well as regulatory requirements. The POCC develops a training program for testing personnel and ensures documentation of competency training. The POCC also oversees completion of proficiency testing programs and performs on-site review of all the following: patient testing, QC, and maintenance logs. The POCC reports problems and regulatory noncompliance to appropriate lab management personnel as necessary.

Even though the laboratory is responsible for the POCT policies and procedures, the clinical staff does the actual testing. Fostering a partnership between the operators and laboratory will help solve problems that may arise during the testing process. In this decentralized testing model, the specifics of how training is performed, how competency records are maintained, and how reagents are ordered may be different for each institution. The laboratory should monitor POCT areas and provide consistent feedback regarding compliance with any applicable regulations. This important communication might be accomplished with weekly rounding or monthly audits. However, before any testing is initiated, clarifying individual roles and expectations is key to the success of the program.

Quality Management

Accuracy Requirements

Understanding test accuracy and the need for result confirmation in the context of how the test result will be used clinically is important before any POC testing is initiated. Ideally, a POC test will provide equivalent results to those from the central laboratory, with a defined variance limit. If this were the case, depending on a clinical situation, physicians would have an option to choose either test and achieve the same clinical outcome. While technology in POCT instruments continues to advance allowing harmonization and **standardization** efforts, there are still accuracy and imprecision concerns present with some POCT. One example is the continued accuracy and precision problems with blood glucose meters. Authors of one study applied simulation modeling to relate performance characteristics of glucose analyzers to error rates in insulin dosage.[9] Interestingly, glucometers that met existing POCT quality specifications allowed a large fraction of administered insulin doses to differ from the intended doses. In addition, two landmark studies on tight glycemic controls gave completely different outcomes.[10,11] While one study showed that tight glycemic controls and following intensive insulin therapy reduce morbidity and mortality among critically ill patients in the surgical ICU,[10] authors of another study found that intensive glucose control increased mortality among adults in the ICU.[11] There has been a considerable debate in the laboratory medicine literature that the difference in these two studies' outcomes was due to the test method.[12–14] Due to complaints about glucose meter inaccuracy, in September 2020, the FDA issued new standards for accuracy in over-the-counter blood glucose meters.[15]

QC and Proficiency Testing

The purpose of a quality management program is to ensure quality test results. A thorough validation should verify the analytical performance and any applicable limitations with the assay. Ongoing daily QC will alert the operator to any reagent or instrument issues. Some devices have both internal and external QC, both of which have distinctly different roles. Internal QC, which is also referred to as onboard QC, internal checks, electronic QC, or intelligent QC,[16] is performed at specific time intervals or at least daily. Internal QC ensures the electronics of the device are performing as expected or, if a manual test, ensures the integrity of the specific test system. Some instruments automate QC and/or calibration,

which is helpful for regulatory compliance. This also ensures the instrument is ready to perform accurate testing at all times. Another important feature is to tie QC performance requirements to test availability. If QC has not been performed as required and is unsatisfactory, the instrument does not allow patient testing until corrective action has taken place; this is known as QC lockout.

The entire testing process should be checked periodically according to the manufacturer and regulatory requirements by running external QC. Here, control samples are introduced to the test system in the same manner as a patient sample. Another form of external QC is proficiency testing, where blind samples are sent to participating laboratories to perform testing. The testing of proficiency samples is performed by POCT personnel in the same manner as patient samples. In the United States, results from one laboratory are compared with the results from other laboratories using the same testing method (known as a peer group).

Quality management programs also need to ensure that the operators are competent to perform testing. After initial training, ongoing assessment of their performance is required. This is commonly called the recertification process. Careful control of personnel records is an important part of the quality management program and may be managed at the POCT site or directly by the laboratory. Some POCT programs incorporate and specifically train "super users" (supervisors, managers, or POCC) to track training renewal dates and to lock out operators whose training has expired.

The greatest source of error in POCT is preanalytic error. While the above features are necessary for any POCT program, it is more difficult to identify errors arising from improper specimen collection—for example, sample contamination, hemodilution, interfering substances (lipemia, icterus, and hemolysis), and inappropriate specimen type (arterial, venous, and fingerstick). For this reason, careful assessment of a POCT area's compliance in proper specimen collection and handling is needed prior to implementing testing.

Preventing post-analytic errors in result reporting can be achieved more reliably with **connectivity**. With the aid of barcodes, modern technology has improved patient identification and decreased transcription errors in many healthcare applications. However, linear barcode identification methods are not fail-safe.[17] Hospitals should work with laboratories, pharmacy, radiology, and admissions to standardize scanning and printing specifications across a system. Careful control and tight specifications of barcode scanning and printing equipment will minimize patient misidentification errors and increase patient safety. The most efficient barcoding is performed at the bedside and ensures laboratory results are posting to the correct patient.

Total-system POCT quality assurance is thus the combination of several QC mechanisms:

- A mechanism to perform a thorough systems validation
- Reliable, user-friendly POCT device
- Training of POCT operators
- Competency assessment of all individuals involved in POCT
- QC testing and monitoring
- Ensuring required maintenance
- Proficiency testing
- Connectivity and bar-coding technologies

POC Applications

POCT makes up nearly 30% of the total in vitro diagnostics market in the United States.[18] A wide variety of POCT devices can be loosely classified into a single-use system, handheld device, or benchtop device. POCT devices are designed with a consideration for the operators who may have little to no laboratory training. Handheld devices are getting smaller and more ergonomic with every generation. A variety of analytical principles used in a laboratory have also been implemented in POCT devices, including:

- Reflectance
- Electrochemistry, electrical impedance
- Light scattering/optical motion
- Immunoturbidimetry
- Lateral flow, flow-through, or solid phase immunoassays
- Spectrophotometry, multiwavelength spectrophotometry
- Fluorescence, time-resolved fluorescence
- Polymerase chain reaction (PCR)

Single-use devices are those that use a disposable test strip or cartridge for testing (**Figure 29.1**). These include the qualitative or semi-qualitative urine and whole blood chemistry tests, as well as those used for hCG, some infectious diseases (i.e., rapid HIV and strep tests), and cardiac markers. Certain single-use cartridges and strips can also be used with device readers for quantitative analysis and include tests such as glucose and other whole blood chemistry tests, tests for cardiac function, drug testing,

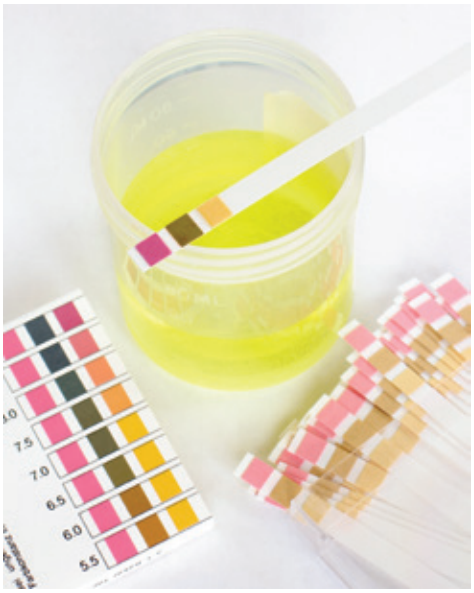


Figure 29.1 Urine sample being tested for pH using standard pH strips.

Figure 29.2 Glucose blood test analyzer.

HbA1c, as well as pH, blood gasses, and electrolytes (**Figure 29.2**). For more throughput and ease of use, handheld instruments and small benchtop devices have been manufactured for the quantitative testing of hemoglobin and bilirubin, pH and blood gases, electrolytes, cardiac markers, medications and drugs, as well as complete blood counts (**Figure 29.3**).

Contemporary POCT analyzers can deliver test results in less than a minute using a single-step protocol on a variety of unprocessed specimens such as whole blood (both capillary and arterial), cerebrospinal fluid, urine, and stool specimens. Ideally, the results of POCT should meet the analytical specifications that are "fit-for-purpose" and should be comparable to those of the central laboratory, with defined limits of variation.

Glucose testing is the highest-volume POCT. These are devices with single-use cartridges or strips that use the reflectance or electrochemistry analytical principles for glucose measurement. Hemoglobin A1c (HbA1c) POC testing is also rapidly increasing, although we have yet to see a device that measures both glucose and HbA1c. Both these tests should not be used for a tight glucose control at bedside but rather employed as monitoring or screening procedures because of the cases of hypoglycemia as well as cases of erroneous POCT measurements due to interfering substances.[19,20]

Many developed POCT areas employ tests for the measurements of urine and blood chemistry, coagulation testing, pH and blood gases, hemoglobin and bilirubin, and complete blood count. Other areas of POCT that are rapidly evolving are intraoperative immunoassay of parathyroid hormone, creatinine and cardiac markers in emergency departments, and infectious diseases.

Figure 29.3 i-STAT analyzer.

Informatics and POCT

Capturing patient ID, operator ID, reference range, and the result documentation in the permanent record can be challenging without connectivity. Due to the number of interfaces that would be required, POCT analyzers are usually connected with LIS via a docking station and/or data management (DM) system that can connect to multiple POCT devices from different manufacturers. This is done through a single POCT middleware system to the LIS/EMR for all data transmissions.

With a multitude of POCT devices, it is not practical to have a separate device–LIS interfacing system for each device. Accordingly, the Connectivity Industry Consortium (CIC), represented by more than 30 instrument manufacturers, information technology companies, and end users, has developed a set of standards to ensure POCT device bidirectionality, device connection commonality, commercial software intraoperability and security, and regulatory compliance. The developed standards govern the communication between the POCT devices and the DM as well as the actual interface between the DM and the LIS. The standards have been adopted by the Clinical and Laboratory Standards Institute.[21,22] Prior to the CIC, each vendor developed its own proprietary means of communicating data from POCT devices, with different physical connections, wiring, and even the language and communication format or protocol. Currently, the CIC standards are becoming widely accepted by vendors and by users who necessitate a more universal POCT connectivity.

According to the CLSI guidelines,[22] certain patient information should always accompany the patient test result across the POCT device, the DM or middleware, the LIS, and the EMR. These include:

- Patient, sample, operator, and device identifiers
- Date and time of both specimen collection and analysis
- Type of specimen
- Test requested and test result with appropriate units
- Error messages and action messages

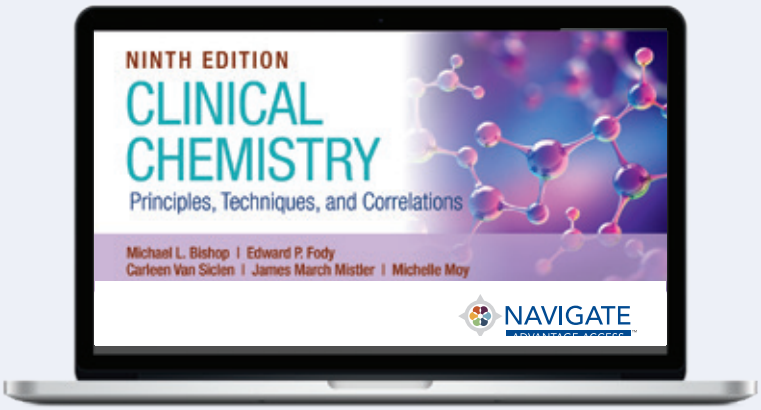
These obligatory information items might also be accompanied by additional information such as reference ranges; calibrator, reagent, or QC detail (e.g., lot number and expiration date); and specific comments and alerts.

In the course of selecting a POCT device, consideration must be given to the type of data output a POCT device provides. These include visual readings, printers, display screens, Ethernet and RS232 ports, modems, infrared beams, and radio signals. Preference is usually given to a device that is suitable for an already developed POCT connectivity system of an institution, which also makes the training of operators easier. Additionally, the connectivity communication system between a device and the DM can be unidirectional (one-way) or bidirectional (two-way). One-way connectivity allows the DM only read information from the POCT device, while the two-way connectivity also lets the DM upload data to the POCT devices. This is especially helpful in the POC system with multiple devices insofar as the DM can update information (e.g., lists of valid operators, reagent lot information, and patient information) for all POCT devices simultaneously.

Integration of POCT results with LIS and EMR allows for better management of diseases such as diabetes mellitus and hyperlipidemia, as well as in critical situations such as glycemic control in intensive care, management of heart failure, surgical case monitoring, transportation of critically ill patients to a healthcare facility, and assistance with dosing of anticoagulants. In these particular instances, the wireless connectivity of POCT devices in conjunction with docking stations is in great demand and becoming more popular.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. CAP Commission on Laboratory Accreditation, Laboratory Accreditation Program. *Point-of-Care Testing Checklist*. Northfield, IL: CAP; April 2012.
2. Crook M. *Handbook of Near-Patient Testing*. London, UK: Greenwich Medical Media Limited; 1999:1–116.
3. Price CP. Quality assurance of extra-laboratory analyses. In: Marks V, Alberti KG, eds. *Clinical Biochemistry Nearer the Patient II*. London, UK: Bailliere Tindall; 1987:166–178.
4. Oliver G. *On Bedside Testing*. London, UK: HK Lewis; 1884: 1–128.
5. Mass D. Consulting to physician office laboratories. In: Snyder JR, Wilkinson DS, eds. *Management in Laboratory Medicine*. 3rd ed. New York: Lippincott; 1998: 443–450.
6. Boehme CC, Nabeta P, Hillemann D, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med*. 2010;363:1005–1015.
7. U.S. Department of Health and Human Services. Medicare, Medicaid and CLIA Programs. Regulations implementing the Clinical Laboratory Improvement Amendments of 1988 (CLIA): final rule. *Fed Regist*. 1992;57:7002–7186.
8. CAP Commission on Laboratory Accreditation, Laboratory Accreditation Program. *Director Assessment Checklist*. Northfield, IL: CAP; August 2017.
9. Boyd JC, Bruns DE. Quality specifications for glucose meters: assessment by simulation modeling of errors in insulin dose. *Clin Chem*. 2001;47(2):209–214.
10. van den Berghe G, Wouters P, Weekers F, et al. Intensive insulin therapy in critically ill patients. *N Engl J Med*. 2001;345(19):1359–1367.
11. NICE-SUGAR Study Investigators; Finfer S, Chittock DR, et al. Intensive versus conventional glucose control in critically ill patients. *N Engl J Med*. 2009;360:1283–1297.
12. Cembrowski GS, Tran DV, Slater-MacLean L, et al. Could susceptibility to low hematocrit interference have compromised the results of the NICE-SUGAR trial? *Clin Chem*. 2010;56(7):1193–1195.
13. Mahoney JJ, Maguire P, Ellison JM, et al. Response to Cembrowski et al. regarding "Could susceptibility to low hematocrit interference have compromised the results of the NICE-SUGAR trial?" *Clin Chem*. 2010;56(10):1643; author reply 1643–1644.
14. Hoedemaekers CW, Klein Gunnewiek JM, Van der Hoeven JG. Point-of-care glucose measurement systems should be used with great caution in critically ill intensive care unit patients. *Crit Care Med*. 2010;38(1):339; author reply 339–340.
15. Food and Drug Administration. Self-monitoring blood glucose test systems for over-the-counter use: Guidance for industry and Food and Drug Administration staff. FDA-2013-D-1446. Accessed September 22, 2021. https://www.fda.gov/media/87721/download.
16. Gill JP, Shephard M. The conduct of quality control and quality assurance testing for POCT outside the laboratory. *Clin Biochem Rev*. 2010;31(3):85–88.
17. Snyder ML, Carter A, Jenkins K, et al. Patient misidentification caused by errors in standard barcode technology. *Clin Chem*. 2010;56(10):1554–1560.
18. Kalorama Information Worldwide Point of Care Diagnostics. Accessed July 30, 2012. http://www.kaloramainformation.com/.
19. Hoedemaekers CW, Klein Gunnewiek JM, Prinsen MA, et al. Accuracy of bedside glucose measurement from three glucometers in critically ill patients. *Crit Care Med*. 2008;36:3062–3066.
20. Eastham JH, Mason D, Barnes DL, et al. Prevalence of interfering substances with point-of-care glucose testing in a community hospital. *Am J Health Syst Pharm*. 2009;66: 167–170.
21. Clinical Laboratory Standards Institute. *Point-of-Care Connectivity: Approved Standard*. 2nd ed. CLSI Document POCT01-A2. Wayne, PA: CLSI; 2006.
22. Clinical Laboratory Standards Institute. *Implementation Guide of POCT01 for Healthcare Providers: Approved Guideline*. CLSI Document POCT02-A. Wayne, PA: CLSI; 2008.

CHAPTER 30

Newborn and Pediatric Clinical Chemistry

Khushbu Patel and Tracey G. Polsky

CHAPTER OUTLINE

KEY TERMS

Drug metabolism
Endocrine system
Genetic disease
Growth
Immunity
Physiologic development
Point-of-care testing (POCT)
Sexual maturation
Tandem mass spectrometry

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Define the adaptive changes that occur in the newborn.
- Describe the developmental changes that occur throughout childhood.
- Discuss the problems associated with collecting blood from small children.
- Understand the role of point-of-care testing in pediatric settings.
- Summarize the changes that occur in children with regard to electrolyte and water balance, endocrine function, liver function, and bone metabolism.
- Explain how drug treatment and pharmacokinetics differ between children and adults.
- State the procedures and test methods used to diagnose inherited metabolic diseases.
- Correlate laboratory results to pediatric disorders of the immune system.
- Interpret laboratory test results associated with genetic diseases in pediatric patients.
- Apply knowledge of clinical chemistry to answer chapter case study questions.

Developmental Changes from Neonate to Adult

Pediatric laboratory medicine provides many unique opportunities to study how the homeostatic and physiologic mechanisms that control normal human development evolve. With these opportunities to study development comes a completely new set of challenges, many based on failure of some component of the normal development process and resulting disease. With this scenario in mind, it becomes clear that the environment for the specialist pediatric laboratorian is different from the environment encountered in adult practice, in which **physiologic development**

CASE STUDY 30.1, PART 1

Meet Trevor, a 5-year-old male who is brought to his pediatrician with a 2-week history of lethargy, fatigue, irritability, and weight loss. They share that he is unusually hungry and thirsty throughout the day and recently started wetting the bed at night, which had never been an issue previously. The pediatrician performed a urinalysis and a point-of-care (POC) blood glucose.

© Ann in the uk/Shutterstock.

CASE STUDY 30.2, PART 1

Meet Tracey, a 1-month-old infant, who presented to the pediatrician with steatorrhea and persistent respiratory infections. Her older sibling also had the same clinical presentation at birth and now has a confirmed genetic condition.

© morrowlight/Shutterstock.

CASE STUDY 30.3, PART 1

Meet Jody, a female infant who was identified with a positive newborn screen within 24 hours of birth. The mother was advised to schedule an appointment with her pediatrician as soon as possible. Due to a national health emergency, the mother could not get an appointment with the pediatrician for 35 days. She continued to breast feed her baby daily.

© Karl Tapales/Moment/Getty Images.

is not a major issue. The diseases encountered in pediatric practice, therefore, differ considerably from those in adult situations. Moreover, the nature of body size and, hence, available blood volume create additional challenges for the laboratorian including minimum fill volumes, sample size, reference ranges, and critical values.

The greatest pediatric challenge relates to the birth of an infant. There is a requirement at this time for rapid adaptation from intrauterine life, in which homeostasis is maintained by maternal and placental means, to the self-maintenance needed to adapt to extrauterine life. Issues related to this adaptation are further complicated by prematurity or intrauterine growth restriction (IUGR), when many organ systems have not reached sufficient maturity to enable the newborn to adapt to the necessary changes at the time of delivery.[1]

Respiration and Circulation

At birth, the normal infant rapidly adapts by initiating active respiration. The stimuli for this process include clamping of the umbilicus, cutting off maternal delivery of oxygen, and the baby's first breath. Initiation of breathing requires the normal expression of surfactant in the lungs. Surfactant is necessary for the normal expansion and contraction of alveoli and allows gaseous exchange to take place.

Initiation of respiration and expansion of lung volume causes increased pulmonary blood flow and reduced blood pressure. This, in turn, results in closure of the ductus arteriosus and a shift in blood flow through the heart that allows newly oxygenated blood from the lungs to be directed through the left side of the heart to the body. Blood flow now goes from the right side of the heart to the lungs for oxygenation. Closure of the ductus arteriosus is essential for this process to take place.

Growth

A baby delivered at term weighs about 3.2 kg. A baby whose birth weight is below the 10th percentile for gestational age is considered small for gestational age (SGA), a condition potentially caused by IUGR. Babies are regarded as premature when born at less than 37 weeks of gestational age. In the first days of life, weight loss is a result of insensible water loss through the skin. This is generally offset by weight gain of 6 g/kg/d as feeding is initiated. An infant's body weight will double in 4 to 6 months. Premature babies tend to grow at a slower rate and often still weigh less than a term baby at the equivalent of term.

Organ Development

Most organs are not fully developed at birth. Glomerular filtration rate (GFR) of the kidney and renal tubular function mature during the first year of life, at which point laboratory markers of renal function approximate adult values. Liver function can take 2 to 3 months to fully mature. Motor function and visual acuity develop during the first year of life. This development is accompanied by changes in the electroencephalogram until the normal "adult" picture is seen. There are dramatic changes in hematopoiesis as the switch from fetal hemoglobin to adult hemoglobin takes place. This coincides with significant hyperbilirubinemia as fetal hemoglobin is broken down, coincident with immature hepatic pathways of bilirubin metabolism. Bone growth in the rapid growth phases in the first few years of life and at puberty results in cyclical changes in bone growth markers. **Sexual maturation** results in significant endocrine changes, particularly of the hypothalamic–pituitary–gonadal hormone pathway, which eventually lead to the constitutive development of adult secondary sexual characteristics and eventually to the adult.

Problems of Prematurity and Immaturity

Intrauterine development is programmed for a normal 38 to 40 weeks of gestation.[1] Many organs are not fully ready to deal with extrauterine life before this time. This organ immaturity results in many of the clinical problems that we see associated with prematurity, which include respiratory distress (lung immaturity), electrolyte and water imbalance (kidney immaturity), and excessive jaundice (liver immaturity). Infants born before their due date constitute a major burden on the laboratory. Premature infants not only have abnormal biochemical parameters that require frequent blood drawing but also have small blood volumes from which to draw on.

Pediatric Sample Collection and Handling

Phlebotomy

Blood collection from infants and young children is complicated by the patient's size and frequently by the ability of the patient to communicate with the phlebotomist. The small blood volume of small

Table 30.1 Implications of a 10-mL Blood Draw in an Infant Population

Age	Weight (kg)	Total Blood Volume (%)
26 weeks of gestation	0.9	9.0
32 weeks of gestation	1.6	5.5
34 weeks of gestation	2.1	4.0
Term	3.4	2.5
3 months	5.7	2.0
6 months	7.6	1.6
12 months	10.1	1.4
24 months	12.6	1.0

patients dictates both the number of tests that can be performed safely on the patient and the number of times that blood can be drawn safely for repeat analysis.[2] **Table 30.1** shows the percentage of total body blood volume that is drawn from an individual with a 10-mL blood draw. This volume is standard in adult laboratory medicine, but the table clearly shows that this amount of blood represents about 5% of total blood volume in a premature neonate. Clearly, frequent blood draws of this nature quickly will lead to anemia and the need for blood transfusion. **Table 30.2** shows the guidelines for blood volume collection. Infants and children have smaller veins than adults; to ensure that small veins do not collapse, narrow-gauge needles are generally used for venipuncture. Smaller needles increase the risk of hemolysis and hyperkalemia.

Frequently, good access to veins is impossible in a pediatric patient with intravenous and central lines in place. Capillary samples often are collected when suitable veins are not available. However, capillary blood obtained by skin puncture has the potential to be contaminated, at least to some extent, by interstitial fluid and tissue debris. The concentration of protein (and protein-bound constituents) is approximately three times lower in interstitial fluid than in plasma. Major differences in analyte composition between venous serum and capillary serum occur due to physiologic and collection conditions. Glucose and potassium levels are typically higher in capillary samples. The lower concentrations of protein, bilirubin, and calcium in capillary specimens likely reflect mixing (and dilution) with interstitial fluid. Sodium and chloride are also lower in capillary samples, while phosphorus and urea typically show medically insignificant variability. Capillary samples, by either heel or finger stick, should be collected by phlebotomists with pediatric expertise.

Table 30.2 Maximum Blood Draw per Kilogram of Body Weight by Patient Age

Age of Patient with Disease or Clinical Condition	Total Blood Volume (mL/kg)	Maximum draw volume/ day in mL/kg (2.5% per kg body weight)	Maximum Number of Tubes per day per kg body weight by draw volume			Maximum draw volume/ 30 days in mL/kg (5% per kg body weight)	Maximum Number of Tubes per 30 days per kg body weight by draw volume		
			2 mL draw volume	0.5 mL draw volume	0.25 mL draw volume		2 mL draw volume	0.5 mL draw volume	0.25 mL draw volume
Preterm Infant	90	2.25	1	4	9	4.5	2	9	18
Term Infant	80	2	1	4	8	4	2	8	16
1-12 months	75	1.875	1	3	7	3.75	1	7	15
1-3 years	70	1.75	1	3	7	3.5	1	7	14
Older children and teens	65	1.625	1	3	6	3.25	1	6	13

Chart compiled as an example after review from various institutions and sources. Each institution should determine appropriate guidelines for local patient population.

Data from: Howie, SRC. Blood sample volumes in child health research: review of safe limits. *Bull World Health Organ* 2011;89:46–53.

The heel should be warmed and well perfused to "arterialize" the capillaries, and the lancet puncture should be in an area of the heel away from bone. Excessive squeezing or milking of the lancet site can result in both hemolysis and factitious hyperkalemia from tissue fluid leakage.

Preanalytic Concerns

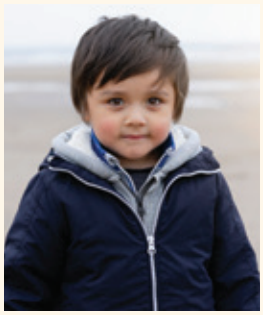
There is a growing trend toward total laboratory automation in clinical chemistry laboratories. The clear advantage with automation of sample handling is that traditional bottlenecks at sites of data entry and centrifugation are removed and turnaround times are reduced. Several issues have prevented the introduction of full-scale automation in pediatrics. A typical pediatric chemistry laboratory receives samples in tubes of many different sizes, varying from standard adult tubes to small pediatric tubes or microtainers. While several of the large chemistry analyzers can directly sample from small pediatric tubes, as of this time, no manufacturer of laboratory equipment has developed a total laboratory automation system that can handle this range of tubes. Although there have been reports in the literature from individual laboratories that have modified existing robotic instrumentation to allow complete automation of the testing process using pediatric tubes,[3,4] this is currently not routine practice.

A second important issue relates to evaporation of sample from open tubes. Most automated sample handling systems require open-topped tubes for processing. With large sample volumes, the effect of evaporation is minimal. With small volumes that have relatively large surface areas to total volume, evaporation can be significant and may affect results by as much as 10%.

CASE STUDY 30.1, PART 2

Remember Trevor. The random finger stick glucose performed at the office visit returned a result of 250 mg/dL (nonfasting).

1. Is this result significant? Why?



© Ann in the uk/Shutterstock.

Choice of Analyzer

Careful inspection and choice of analytic systems remain crucial for handling pediatric samples. Until recently, only a few analyzers performed multiple analytic procedures on small sample volumes (5 to 50 µL). Today, most analyzers can perform this function. Choice of analyzer becomes dependent on issues such as the following:

- How much dead volume is there in the system? The smaller the dead volume, the greater the number of tests that can be run.
- Does clot or bubble detection allow for salvage of sample? All analyzers can detect a clot, but not all will allow retrieval of sample.
- Is the system truly random access? This allows selectivity of menu for a given sample.

Typically, sample throughput time is less of an issue, as pediatric facilities tend to run fewer samples than busy adult services.

Point-of-Care Analysis in Pediatrics

Point-of-care testing (POCT), or near-patient testing, plays an important and expanding role in pediatric practice.[5] Testing devices that are portable and easy to use require small specimen volume, require either urine or whole blood and do not typically need sample preparation or centrifugation in order to provide rapid patient results at the bedside. To provide cost-effective and appropriate quality assurance for POCT, several factors need to be addressed.[5]

Turnaround Time

One key factor is, does the analyte really require immediate turnaround for optimal patient management? Analyzers are becoming available that measure increasing numbers of different analytes at the patient's bedside. Typically, the cost of POCT measurement is higher than the traditional laboratory measurement. The idea of instant results is so seductive that nonlaboratorian users often discount economic factors. In the author's institution, the clinical laboratory has played a leading role in determining which POCT assays will be available and clinical settings that have real value (**Table 30.3**).

Table 30.3 Important Pediatric Point-of-Care Tests and Testing Sites

Test
Blood gas
Electrolyte
Glucose
Activated clotting time
Hemoglobin
Glycated hemoglobin (hemoglobin A1c)
Pregnancy
Urinalysis
Prothrombin time
Occult blood
Rapid *Streptococcus*
Rapid testing for emerging pathogens/infectious agents (SARS-CoV-2)
Testing Sites
Coronary care unit/intensive care unit
Trauma unit/emergency department
Diabetes clinic
Primary care
Transport
Surgery
Extracorporeal membrane oxygenation sites

Evaluation of POCT Devices

A second key factor is, who chooses the POCT device? As the field of POCT expands, the number of devices on the market is also increasing. The clinical laboratory should be the setting in which POCT devices for institutional use should be first evaluated. The laboratory should make choices for instrumentation. Important features of a good POCT device should include the following:

- The ability to lock out untrained users. Only individuals accredited to use the device should be allowed access via a personal code.
- The device should not be allowed to proceed to patient sample analysis without running and validating appropriate quality assurance procedures.
- The data should be downloadable to the hospital laboratory information system (LIS) for evaluation by the hospital quality assurance officer. Downloading the data also allows for billing and data entry into patient charts—features readily lost when analyzers are used that cannot be linked to the LIS.

Device Limitations for Pediatric Use

The data generated by POCT devices have limitations. Typically, analytic performance is not as good as that with the main laboratory analyzer. POCT data are less precise and not suited to monitoring therapy in instances where small changes are important. The linear range for most POCT devices is not as broad as that of the main chemistry analyzer, and users need to be aware of these limitations. A prime example is POCT glucose analyzers, which typically have higher imprecision, sometimes coupled with a bias. Hypoglycemia is also particularly common in pediatrics and characteristically difficult to accurately quantitate using POCT devices. Low glucose levels should be confirmed using a more sensitive main laboratory analyzer. For more information, see Chapter 29, *Point-of-Care Testing*.

Regulation of Blood Gases and pH in Neonates and Infants

Primary maintenance of blood gas and pH homeostasis following birth requires that the lungs and kidneys be sufficiently mature to regulate acid and base metabolism. At about 24 weeks of gestation, the lung expresses two distinct types of cells: type 1 and type 2 pneumocytes. Type 1 pneumocytes are primarily responsible for gas exchange, while Type 2 pneumocytes are responsible for the secretion of surfactant, which contains the phospholipids lecithin and sphingomyelin. Surfactant is required for the lungs to expand and the transfer of blood gases following delivery. Oxygen crosses into the circulation, and carbon dioxide is removed and expired. Immaturity of the surfactant system as a result of prematurity or IUGR results in respiratory distress syndrome (RDS). In RDS, there is failure to excrete carbon dioxide and, as a result, carbon dioxide levels rise, causing respiratory acidosis; oxygen levels are low and result in additional oxygen requirements for the infant.

The relative amounts of lecithin and sphingomyelin are critical for normal surfactant function. Traditionally, laboratories employed thin layer chromatography to quantity components of pulmonary surfactant, including lecithin, sphingomyelin, and phosphatidylglycerol. Lamellar body count is a less labor-intensive screening test to measure surfactant production. However, these markers have fallen out of use due to poor reliability in predicting fetal and maternal outcomes. In 2019, the American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine advised against using these tests to make delivery decisions.[6] Improved care and treatment of premature infants with RDS has been one of the driving factors for this change.[7]

The trauma and relative anoxia during delivery can also induce acidosis in the newborn. This is typically a metabolic acidosis associated with increased lactic acid production. Serum bicarbonate levels are reduced in this situation compared with the respiratory acidosis in RDS. Persistent metabolic acidosis in the newborn that is difficult to correct with bicarbonate replacement is an indication for further intensive evaluation for possible inborn error of metabolism or other etiologies that require differentiation (**Table 30.4**).

Alkalosis is an unusual finding in pediatric medicine. One important cause of alkalosis is hyperammonemia, which may be secondary to a number of etiologies, including liver disease and inborn errors of metabolism (Table 30.4).

Table 30.4 Causes of Acidosis and Alkalosis in Neonates and Infants

Respiratory Acidosis	Hypoventilation/CO_2 Retention
Metabolic acidosis	Renal tubular bicarbonate wasting
	Anoxia
	Poor tissue perfusion
	Metabolic disease
Respiratory alkalosis	Hyperventilation
	High blood ammonia/metabolic disease
Metabolic alkalosis	Pyloric stenosis/loss of gastric acid
	Excessive bicarbonate administration
	Low blood potassium

Blood Gas and Acid–Base Measurement

Oxygen status can readily be measured using noninvasive transcutaneous monitoring. Good correlation has been demonstrated between the arterial pressure of oxygen and transcutaneous measurement. Transcutaneous carbon dioxide monitors are also in widespread use. The measurement of acid–base status requires blood sampling. Most blood gas/acid–base analyzers can be adapted to take the small capillary samples routinely collected in pediatric settings. It is important for the person drawing the capillary blood to do so anaerobically, which requires thorough warming of the capillary site and collection of a freely flowing blood sample from a lancet stick. The sample needs to be sealed to ensure minimal gas exchange. Analysis should be performed immediately to not compromise the sample integrity. The author's laboratory maintains a goal of a 10-minute turnaround time from the sample receipt.

Most blood gas analyzers measure pO_2, pCO_2, and pH by ion-specific electrodes and calculate bicarbonate concentration by the Henderson-Hasselbalch equation. The Henderson-Hasselbalch equation is less valid when pH is far outside the normal physiologic range (extreme acidosis or alkalosis). On these occasions, it may become important to measure the bicarbonate concentration using a direct measurement.

Many blood gas analyzers have been upgraded in recent years to measure additional analytes, including blood sodium, potassium, and chloride, using ion-specific electrodes. Several point-of-care instruments also measure lactate, urea, bilirubin, and creatinine. The major advantage of this type of analyzer in pediatrics is that whole blood can be used. The volumes are typically smaller than those required for the main chemistry analyzer, and the lack of need for centrifugation shortens the turnaround time. One disadvantage of using whole blood is that the analyst cannot detect whether a sample is hemolyzed.

Regulation of Electrolytes and Water: Renal Function

From the 35th week of gestation, the fetal kidneys develop rapidly in preparation for extrauterine life.[1] The kidneys, critical organs for the maintenance of

Table 30.5 Development of Glomerular Filtration in the Newborn

Age	Glomerular Filtration Rate (mL/min per 1.73 m² Mean)	Reference Range
1 d	24	3–38
2–8 d	38	17–60
10–22 d	50	32–68
37–95 d	58	30–86
1–2 y	115	95–135[a]

[a] Adult values.

electrolyte and water homeostasis, control the rate of salt and water loss and retention. At term, neither the glomeruli nor the renal tubules function at the normal rate. The GFR is about 25% of the rate seen in older children and does not reach full potential until age 2 years (**Table 30.5**). Tubular function also develops at a similar rate. The maximum concentrating power of the kidney is only about 78% of the adult kidney, although the tubular response to antidiuretic hormone appears to be normal. This gradual process of renal development in the newborn results in diminished filtration and impaired reabsorption of salt and water; therefore, in the newborn period and early infancy, large shifts in serum electrolyte levels can be observed. These problems are exacerbated in the preterm infant with renal function that is even less mature.

The kidneys also primarily maintain water loss and retention. However, in the newborn period, insensible water loss through the skin is also an important cause of water and electrolyte imbalance. Water loss and consequent hemoconcentration frequently result from the use of radiant heaters that are used to maintain body temperature. Increased water loss also occurs via respiration in children with RDS. Up to one-third of insensible water loss may occur through this route. The total body water content of a newborn is about 80%; 55% is intracellular fluid, and 45% is extracellular fluid. The extracellular water is 20% plasma water and 80% interstitial. During the first month of extrauterine life, the total body water content decreases to about 60%, mostly a result of loss of the interstitial component.[1]

Disorders Affecting Electrolytes and Water Balance

The causes of hypernatremia (sodium >145 mmol/L) and hyponatremia (sodium <130 mmol/L) are listed in **Table 30.6**. Both electrolyte imbalances require medical intervention due to a high risk of seizures in this population. This is a result of the shift of water out of or into brain cells, with concurrent shrinkage or expansion of these cells, respectively. Hypernatremia results from hypotonic fluid loss, and hyponatremia results from hypertonic fluid loss. Hyponatremia may also be a result of excessive body water content and needs to be distinguished from hypertonic loss. Clinical evaluation and measurement of other components, including hematocrit, serum albumin, creatinine, and blood urea nitrogen (BUN), can be used to differentiate these etiologies. These analytes will be elevated with hemoconcentration. Clinically, it is usually possible to distinguish dehydration from excessive hydration.

Treatment of electrolyte and water loss is directed at replacing the loss to regain normal physiologic levels. Care must be taken to avoid too rapid a replacement, particularly with hypertonic dehydration. If water replacement is done too quickly, a

Table 30.6 Causes of Hypernatremia and Hyponatremia

Hypernatremia
Excessive loss of water through overhead heater
Gastrointestinal fluid loss
Fluid deprivation
Renal loss of water/nephrogenic diabetes insipidus
Administration of hypertonic fluids containing sodium
Hyponatremia
Inappropriate ADH secretion due to trauma or infection
Administration of hypotonic fluids
Renal tubular acidosis
Salt-losing congenital adrenal hyperplasia (21-hydroxylase deficiency)
Cystic fibrosis
Diuretics
Renal failure

Table 30.7 Causes of Hyperkalemia and Hypokalemia

Hyperkalemia
Fluid deprivation/dehydration causing tissue leakage
Intravascular hemorrhage causing release from red cells
Trauma/tissue damage
Acute renal failure
Salt-losing adrenal hyperplasia (see *hyponatremia*)
Exchange transfusion using stored blood
Hypokalemia
Inappropriate ADH secretion
Diuretics, particularly furosemide
Alkalosis
Pyloric stenosis
Renal tubular acidosis secondary to bicarbonate loss

rapid expansion of neuronal cell volume can occur, which results in seizures.

The causes of hyperkalemia and hypokalemia are listed in **Table 30.7**. The symptoms of hyperkalemia (serum potassium >6.5 mmol/L) include muscle weakness and cardiac conduction defects that may lead to heart failure. In pediatrics, it is particularly important to recognize spurious results. For example, hyperkalemia—a result of hemolysis or hyperkalemia due to improper capillary blood collection with high potassium tissue leakage.

Because the situation regarding electrolyte and water homeostasis can change rapidly in small infants, it is important to monitor therapeutic intervention on a frequent basis.[1,2] The availability of POCT devices that use small volumes of whole blood helps with management of these imbalances, with the only caution being that it is impossible to detect hemolysis and factitious hyperkalemia on a whole blood sample.

Development of Liver Function

Physiologic Jaundice

The liver is an essential organ for many metabolic processes. Metabolic pathways and the metabolism of exogenous compounds, in particular, pharmacologic agents, proceed slower in neonates. The most striking effect of an immature liver, even in a normal-term baby, is the failure to adequately metabolize bilirubin. Bilirubin is an intermediate of the breakdown of the heme molecules, which accumulate as fetal hemoglobin is rapidly destroyed and replaced by adult hemoglobin. Normally, the liver conjugates bilirubin to glucuronic acid using the enzyme bilirubin UDP-glucuronosyltransferase. Conjugated bilirubin can be readily excreted in the bile or through the kidneys. At birth, this enzyme is too immature to complete the process, and increased levels of unconjugated bilirubin and "physiologic" jaundice result. At this time, a normal baby may have a serum bilirubin level of up to 12 mg/dL, most of which is unconjugated. This level, which would be alarming in adult practice, should fall back to baseline by about 10 days of age. Because excessive jaundice can lead to kernicterus and result in severe brain damage, the measurement of total and direct (conjugated) bilirubin has a critically important role in pediatrics. Indirect (unconjugated) bilirubin is calculated by subtracting the direct (conjugated) bilirubin from the total bilirubin. An alternative means of reducing high unconjugated circulating bilirubin levels is phototherapy with ultraviolet light, which causes bilirubin to be converted to a potentially less toxic and more readily excreted metabolite. Severe cases may require an exchange transfusion. Complete absence of the bilirubin-conjugating enzyme results in severe persistent jaundice and Crigler-Najjar disease, a rare genetic disease. It is important to differentiate Crigler-Najjar from physiologic immaturity because treatment options vary considerably.

Energy Metabolism

The liver plays an essential role in energy metabolism for the whole body (**Table 30.8**). Carbohydrates derived from the diet as disaccharides or polysaccharides form the bulk of our energy sources. They are broken down into simpler monosaccharides, which reach the liver via the portal blood system. The primary sugars in newborns and infants come from the breakdown of disaccharide lactose in milk. Lactose is broken down to glucose and galactose. When it reaches the hepatocytes, galactose is converted to glucose by a series of enzymic reactions that have unique pediatric significance. Genetic deficiency of any of the reactions results in failure to convert galactose to glucose and essentially reduce the energy content of milk by 50%. The most common cause of failure to convert galactose to glucose results

Table 30.8 Important Biochemical Pathways in the Liver

Catabolic
Transamination
Amino acid oxidation to make ketones and acetyl-CoA
Fatty acid oxidation to make ketones
Urea cycle to remove ammonia
Bilirubin metabolism (hemoglobin breakdown)
Drug and exogenous xenobiotic compounds metabolized
Anabolic
Albumin synthesis
Clotting factor synthesis
Lipoprotein synthesis, very-low-density lipoprotein
Gluconeogenesis (synthesis of glucose)
Bile acid synthesis

in galactosemia caused by a deficiency of galactose-1-phosphate uridyltransferase, a serious genetic disease of the newborn. In this disease, galactose-1-phosphate accumulates insideliver cells and causes hepatocellular damage and rapid liver failure. Other organs are also involved with this disease, including the renal tubules and the eyes. Galactose-1-phosphate accumulation causes acute renal tubular failure and tubular loss of glucose, phosphate, and amino acids. The loss of glucose incooperation with the liver damage results in severe hypoglycemia. Accumulation of galactose in the eye results in cataract formation. A simple test that directs us to the diagnosis of galactosemia is the urine-reducing substance test. This detects the presence of non-glucose-reducing sugars (galactose) in urine when a child is symptomatic. Facilities will vary in their protocols to run reducing sugar on all newborns regardless of symptoms. The clinical significance of establishing a diagnosis is clear; galactosemia is often fatal if undiscovered but completely treatable by dietary lactose restriction when diagnosis is made.

Another critical pathway of carbohydrate energy metabolism in the newborn involves the pathway of gluconeogenesis. At birth, a term baby has sufficient liver glycogen stores to provide glucose as an energy source and maintain euglycemia. If the delivery is particularly stressful, these reserves of energy may become depleted prematurely. The normal physiologic role of gluconeogenesis, which essentially converts the amino acid alanine into glucose, becomes critical in maintaining glucose homeostasis. This pathway is not always mature at birth, and suboptimal operation results in what is termed *physiologic hypoglycemia*. Newborns can survive blood glucose levels below 30 mg/dL, although adults would fall into rapid hypoglycemic coma and risk sudden death at these levels. Physiologic hypoglycemia usually corrects quickly as the enzyme systems mature or by simple intravenous glucose infusion. Persistent and severe hypoglycemia should alert the physician to a possible inborn error of metabolism, such as galactosemia, disorders of gluconeogenesis, or disorders of fatty acid oxidative metabolism.

Diabetes

Blood glucose homeostasis and hepatic metabolism of glucose are maintained by the concerted actions of several hormones.[8] Following a meal, the level of glucose in the circulation rises, which triggers increased synthesis and release of insulin by the pancreatic β cells of the islets of Langerhans. Increased levels of insulin in the circulation cause glucose to be taken up by certain cells, such as hepatocytes and muscle cells, and converted into glycogen as a future source of energy. As a result of the insulin action, blood glucose levels begin to fall to the preprandial level. Glucagon, a hormone secreted by the α cells of the islets of Langerhans, has an opposing effect to that of insulin. It is generally believed that the insulin–glucagon ratio, rather than absolute amounts of either, is the primary endocrine modulator of circulating glucose levels. Other hormones, including cortisol, epinephrine, and insulin-like growth factor (IGF), can also affect glucose levels by acting antagonistically to insulin and raising blood glucose levels. These hormones are secreted in response to stress and can affect glucose measurement when samples are collected under stressful situations.

Diabetes mellitus, a condition in which the endocrine control of glucose metabolism is abnormal, is usually related to failure of the insulin regulatory pathway. Type 1 diabetes (insulin dependent) is the most common in pediatrics. This may be caused by failure of the pancreas to secrete insulin or by the presence of circulating insulin antibodies that reduce the ability for endocrine action. This type of diabetes can be further classified as immune mediated or idiopathic. In the immune-mediated form

(the most common form), the insulin-secreting beta cells are destroyed by a T cell–mediated autoimmune response. This form of diabetes was previously termed "juvenile diabetes" because it is the predominant form in children. A patient typically presents with diabetic ketoacidosis, profound hyperglycemia, and metabolic acidosis that result from the liver increasing fatty acid metabolism and producing excess ketone bodies.

Type 2 diabetes (non–insulin dependent) is normally associated with increased resistance to normally secreted insulin in older and in obese individuals. During the past 30 years, the number of children diagnosed as being overweight has increased by >100%, and this epidemic of childhood obesity is causing children to suffer from chronic complications that were previously only seen in adults.[9] Type 2 diabetes now accounts for a considerable proportion of newly diagnosed cases of diabetes in the pediatric population.[10]

To be diagnosed with diabetes mellitus, an individual (including children) must demonstrate one of the following diagnostic criteria: fasting plasma glucose level ≥126 mg/dL; plasma glucose ≥ 200 mg/dL 2 hours after a 75-g oral glucose load; symptoms of hyperglycemia and random plasma glucose ≥200 mg/dL; or glycated hemoglobin (HbA_1c) ≥ 6.5%.[11] When type 2 diabetes is suspected, a positive result on any of the above should prompt repeated testing on a different day prior to the diagnosis of diabetes; if type 1 is suspected, blood or urine ketone testing or serum anion gap analysis should be performed to identify the onset of diabetic ketoacidosis.

Urbanization, unhealthy diets, and increasingly sedentary lifestyles have contributed to an increase in the prevalence of childhood obesity, particularly in developed countries.[12] The metabolic syndrome is a combination of medical disorders that, when occurring together, greatly increase the chance of cardiovascular disease and diabetes. Several definitions of the metabolic syndrome exist, but they generally include diabetes or impaired glucose resistance, hypertension, dyslipidemia (increased triglycerides and decreased HDL), and central obesity. The link between obesity, metabolic syndrome, and type 2 diabetes has been well characterized in adult populations but has been increasingly observed in pediatric populations. In children and adolescents, research studies have demonstrated a link between childhood metabolic syndrome and elevated cardiovascular risk in later life.[12]

It is important to recognize diabetes as a cause of hyperglycemia in children and to distinguish it from other medical causes of high blood sugar, including acute pancreatic disease or hypersecretion of counterregulatory hormones such as growth hormone (GH), cortisol, or catecholamines. Chronic hyperglycemia can be readily distinguished from acute causes by simply measuring the blood concentration of HbA_1c, a well-established marker for long-term hyperglycemia. In addition to being a diagnostic criterion for diabetes, this assay also plays a major role in monitoring treatment adherence of patients with diabetes.[11]

Nitrogen Metabolism

The liver plays a central role in nitrogen metabolism. It is involved with the metabolic interconversions of amino acids and the synthesis of nonessential amino acids. The liver synthesizes many body proteins, including most proteins found in the circulation,

CASE STUDY 30.1, PART 3

Remember Trevor. He came back to the pediatrician the next morning for a fasting glucose via venipuncture collection. The fasting plasma glucose was 201 mg/dL.

2. What is the most likely diagnosis?
3. What are the diagnostic criteria for this disease?
4. Is the POCT glucose result obtained at the pediatrician office diagnostic?

such as albumin, transferrin, and the complement clotting factors. The liver does not synthesize immunoglobulins. The liver is also responsible for complete metabolism of the breakdown products of nitrogen turnover, such as ammonia and urea through the urea cycle and creatinine and uric acid from energy stores and nucleic acids, respectively. Blood ammonia levels are higher in the newborn period than in later life, presumably due to immaturity of urea cycle enzymes and the portal circulation. A blood ammonia level of 100 µmol/L in a newborn would be regarded as less significant than the same level in a 1-year-old. Persistently elevated ammonia levels should alert the investigator to possible liver damage and secondary failure of the urea cycle. High ammonia levels suggest a possible primary defect in the urea cycle, and patients should be evaluated for such a defect.

Nitrogenous End Products as Markers of Renal Function

In contrast to the high neonatal ammonia levels, creatinine and uric acid levels are lower in newborns. Both metabolites rise eventually to normal adult ranges. Creatinine concentrations in blood increase with muscle mass and may vary with diet (i.e., meat ingestion). It is filtered at the glomerulus and not extensively reabsorbed by the renal tubules. Its measurement as a clearance ratio in blood and in a 24-hour urine sample has been used as a marker for the GFR for many years. Creatinine clearance is calculated from the creatinine concentration in the collected urine sample, the urine flow rate, and the plasma creatinine concentration. Creatine clearance calculations are not commonly performed anymore due to the difficulties in collecting a complete 24-hour urine specimen and the fact that estimated glomerular filtration rate (eGFR) using a single plasma creatinine value is more convenient and more commonly used.

While several equations exist to estimate GFR from plasma creatinine values in adults, current guidelines recommend all laboratories use the 2009 CKD-EPI creatinine equation. The equation estimates GFR from the patient's serum creatinine (mg/dL), sex, age, and race.[13]

This equation has only been validated on patients older than 18 years. In children, the most widely used and extensively validated equation is the updated or "bedside" Schwartz equation,[14] which estimates GFR from serum creatinine (mg/dL), the child's height (cm) and a constant as follows:

$$\text{eGFR (mL/min/1.73 m}^2\text{)} = (0.413 \times \text{height in cm}) / \text{serum creatinine in mg/dL} \quad \textbf{(Eq. 30.1)}$$

Use of the above equations requires the use of a creatinine assay that has been standardized and its calibration traceable to the gold standard measurement technique, isotope dilution mass spectrometry (IDMS). More information regarding the use of the eGFR can be found in Chapter 21, *Renal Function*.

Serum cystatin C, a newer and possibly more sensitive marker of GFR, is considered a potential replacement for creatinine,[15] as its concentration is independent of muscle mass and recent meat ingestion. Cystatin C (an inhibitor of cysteine protease) is a ubiquitous protein secreted by most cells in the body. Cystatin C is freely filtered at the glomerulus and reabsorbed by the tubular epithelial cells, with only a very small amount being excreted in the urine. Current guidelines recommend using cystatin C for confirmatory testing in specific circumstances where eGFR based on creatine is less accurate; however, cystatin C is used commonly for eGFR estimation in the pediatric population. Several eGFR equations that incorporate cystatin C, as well as other kidney markers, have been derived. An improved equation (Creatinine-Cystatin C-based CKID equation) for use in children with chronic kidney disease, which utilizes BUN, serum creatinine, and cystatin C, has been most recently described.[16] This equation has been validated for use in children with chronic kidney disease in a range of GFR from 15–75 mL/min per 1.73 m^2, and the applicability of this equation in children with normal stature and muscle mass remains unknown.

Liver Function Tests

As discussed, the liver is responsible for performing a large number of synthetic and catabolic processes, and normal liver function is central to maintaining body homeostasis. Several laboratory tests have emerged that are generally classified as liver function tests.

The measurements of serum albumin and total and conjugated bilirubin are true tests of liver function because they measure the synthetic and metabolic pathways for these compounds. In protein–calorie malnutrition, the reduced of availability of amino acids for synthesis of new proteins results in diminished functional synthetic rate and low levels

of newly synthesized proteins, such as albumin. Very low levels of albumin indicate a long exposure to protein restriction, and its measurement in blood is often used as a guide to nutritional status and chronic liver disease of other causes. Impaired hepatocellular function also results in reduced ability to conjugate bilirubin, with subsequent increase in the unconjugated form, which is normally barely detectable.

Other tests, such as measurement of liver enzymes, more truly reflect tests of liver cell integrity and are not strictly functional assays. Large elevations in serum aspartate transaminase (AST) and alanine aminotransferase (ALT) indicate hepatocellular damage and subsequent leakage of cellular contents into the serum, and elevated alkaline phosphatase (ALP) suggests hepatic biliary damage but gives little functional information. For more information on liver enzymes, see Chapter 19, *Liver Function*.

Calcium and Bone Metabolism in Pediatrics

Normal bone growth, which parallels body growth, requires integration of calcium, phosphate, and magnesium metabolism with endocrine regulation from vitamin D, parathyroid hormone (PTH), and calcitonin.[8] The active metabolite of vitamin D is 1,25-dihydroxy vitamin D. Hydroxylation of vitamin D from the diet takes place in the liver and in the kidneys and requires normal functioning of these organs. Absorption of vitamin D from the gastrointestinal tract, conversion to its active form in the kidney, and incorporation of calcium and phosphate into growing bone require normally active PTH. Secretion of PTH is, in turn, modulated by serum calcium and magnesium levels. Low levels of both divalent cations inhibit PTH secretion. Calcitonin has an antagonistic effect on PTH action.

Much recent attention has been given to the study of vitamin D deficiency and the role of vitamin D in maintaining optimal health. Vitamin D deficiency can result from inadequate nutritional intake of vitamin D coupled with inadequate sunlight exposure or use of sunscreens, disorders that limit vitamin D absorption, and conditions that impair the conversion of vitamin D into active metabolites including certain liver and kidney diseases. Deficiency results in impaired bone mineralization and leads to bone softening diseases including rickets in children and osteoporosis and osteomalacia in adults. According to numerous recent studies, vitamin D may be helpful in preventing other diseases, including several types of cancer, diabetes, multiple sclerosis, obesity, and hypertension.[17]

Vitamin D deficiency is common in children. The American Academy of Pediatrics recommends that all infants (beginning in the first few days of life) receive 400 IU/d of vitamin D supplementation.[17] This dose is required to prevent rickets and osteomalacia and possibly help prevent other chronic diseases. The serum concentration of 25-hydroxyvitamin D is typically used to determine vitamin D status. It reflects vitamin D produced in the skin as well as that acquired from the diet and has a fairly long circulating half-life.

Several assay manufacturers produce immunoassays that are convenient for hospital-based clinical laboratories; however, the gold standard reference method is liquid chromatography coupled with **tandem mass spectrometry**. There have been ongoing issues with variability in results of 25-hydroxyvitamin D levels measured by different methods and immunoassays produced by different manufacturers. A standard reference material is commercially available, which should allow assay manufacturers and laboratories to better standardize their assays. In addition, several new proficiency schemes have been developed to help ensure consistency between laboratories.

The rapid bone growth that occurs during infancy, and later during puberty, requires optimal coordination of mineral absorption, transport, and endocrine-controlled incorporation of the minerals into growing bone. Approximately 98% of total body calcium content is present in bone, and less than 1% is measurable in the blood. Serum calcium is present as the unbound ionized fraction (about 50% of total in blood), with the rest bound to protein (40%) or chelated to anions in the circulation, such as phosphate and citrate. Serum ionized and bound calcium levels are highly regulated and maintained within strict homeostatic limits. Abnormalities in any of the regulatory components have profound clinical effects on children.

For more information on vitamin D and PTH, see Chapter 18, *Parathyroid Function*.

Hypocalcemia and Hypercalcemia

Hypocalcemia is defined as total serum calcium below 7.0 mg/dL or ionized calcium below 3.0 mg/dL. In the newborn and particularly the immature newborn, these levels may be commonly encountered

Table 30.9 Causes of Hypocalcemia

Prematurity
Metabolic acidosis
Vitamin D deficiency
Liver disease (failure to activate vitamin D)
Renal disease (failure to activate vitamin D)
Hypoparathyroidism
Low calcium intake
High phosphorus intake
Diuretic use
Hypomagnesemia
Exchange transfusion (anticoagulants in transfused blood)

with few symptoms. However, hypocalcemia can result in irritability, twitching, and seizures. Serum calcium is usually measured in infants with seizures of unknown etiology. Prolonged hypocalcemia can result in reduced bone growth and rickets. The causes of hypocalcemia are listed in **Table 30.9**. Hypomagnesemia frequently occurs with hypocalcemia. Because low levels of serum magnesium also inhibit PTH secretion, it is important to consider the possibility of concurrent hypomagnesemia in a child with hypocalcemic seizures and to correct any abnormalities that may be identified in the magnesium status as calcium is also corrected.

Hypercalcemia is defined as total serum calcium of greater than 11.0 mg/dL. This is an unusual finding in pediatrics (**Table 30.10**) but has potentially severe clinical implications. Patients with hypercalcemia have poor muscle tone, constipation, and failure to thrive and may develop kidney stones leading to renal failure. For more information on calcium, see Chapter 11, *Electrolytes*.

Table 30.10 Causes of Hypercalcemia

Hyperparathyroidism
Acute renal failure
Excessive intake of vitamin D
Idiopathic hypercalcemia of infancy

Endocrine Function in Pediatrics

The field of endocrinology provides numerous examples of the "differences" that occur in clinical chemistry between children and adults.[2] The process of maturation from a prepubertal child into a sexually fertile adult, for example, requires a complex, endocrine-mediated, developmental process switching on during childhood. As addressed in the previous section, the bone growth that accompanies systemic **growth** also requires a complex process, which is also under endocrine control.

Hormone Secretion

The **endocrine system** relates to a group of hormones that are typically produced and secreted by one specialized cell type into the circulation, where the hormonal effect is exerted in other target cells through the binding of the hormones to specialized receptors. Some of these hormones are polypeptides; others are amino acid derivatives or steroids.

Four major endocrine systems have been described, all of which play critical roles in normal human development. These systems all involve the hypothalamus as a major higher brain control center, the pituitary gland as a major secretor of hormones, and then various end organs, which have responsive elements for the pituitary hormone and affect many metabolic and developmental functions. These end organs include the thyroid gland, adrenal cortex, liver, and gonads. Each system involves regulated secretion of a trophic hormone by the hypothalamus, which in turn is transported to and controls endocrine secretion by the pituitary and, occasionally, secondary hormonal secretion by the end organ, which then produces the appropriate cellular endocrine effect. There is feedback on the hypothalamus by the final product of the pathway (long loop feedback) and also by the endocrine product of the pituitary (short loop feedback). The feedback regulates hypothalamic control of the pathway. Clearly, there are many areas that can go wrong in each of these pathways, all of which result in disease. Certain disease conditions are uniquely pediatric, and they are discussed further.

Hypothalamic–Pituitary–Thyroid System

The hypothalamus secretes thyrotropin-releasing hormone (TRH), a 3-amino-acid peptide, into the portal blood system between the hypothalamus and anterior pituitary.[8] TRH binds to a receptor on specialized cells in the anterior pituitary, which stimulates the secretion of thyroid-stimulating hormone (TSH), a polypeptide made up of two chains (α and β). The α chain is common to human chorionic gonadotropin (hCG) and follicle stimulating hormone (FSH). Unique TSH receptors on the thyroid gland, when occupied by a TSH molecule, cause the follicular cells of the thyroid gland to synthesize and release thyroid hormones into the circulation. The synthesis of thyroid hormone involves several complex steps in which iodine is trapped within the thyroid tissue and used to iodinate tyrosine residues on a prohormone thyroglobulin, which is then cleaved to release the iodinated amino acid triiodothyronine (T_3) and tetraiodothyronine (T_4). Thyroid hormones in the circulation are greater than 99% bound to specific transporter proteins called *thyroid-binding globulins*. Free thyroid hormone, in particular free T_3, is the active form of the hormone and reacts with receptors on many peripheral tissues to cause increased metabolism and simulate normal growth and development. T_4, T_3 (long loop), and TSH (short loop) levels give feedback to the hypothalamus to regulate TRH production.

Two major areas of dysfunction in this endocrine pathway need consideration in pediatrics: primary hypothyroidism and secondary hypothyroidism. Primary hypothyroidism results from any defect that causes failure of the thyroid gland to synthesize and secrete thyroid hormone. This results in a common disease known as *congenital hypothyroidism* (CH), previously known as cretinism, which is present in 1 of 4000 births and is screened for in all newborns in the developed world. Untreated patients with this disease have severe cognitive developmental delays coupled with unusual facial appearances. Treatment by thyroid replacement therapy is usually successful when diagnosis is established. The best diagnostic test is to measure TSH levels in blood spots from newborns or in serum if CH is suspected in later childhood. In CH, TSH levels are high. In untreated newborns thyroid hormone levels—total T_4 and free T_4—are very low.

Secondary hypothyroidism is a result of the failure of the pituitary gland to secrete TSH, which results in lack of thyroid gland stimulation and subsequent low production of thyroid hormone. The differential diagnosis is established by measuring low circulating TSH levels. Because the pituitary is involved with all major endocrine systems, it is important to study the other pituitary pathways to determine if hypothyroidism is the result of an isolated TSH defect or due to panhypopituitarism involving all other pathways. Panhypopituitarism is clinically complex and may include features of hypoglycemia, salt loss, poor somatic and bone growth, failure to thrive, and, in later childhood, failure to develop secondary sexual characteristics.

Hypothalamic–Pituitary–Adrenal Cortex System

This system is essential for regulating mineral and carbohydrate metabolism.[4] The hypothalamus secretes corticotrophin-releasing hormone (CRH), a 41-amino-acid polypeptide, into the portal blood system, which binds to receptors on specialized anterior pituitary cells, resulting in the release of corticotrophin or adrenocorticotropic hormone (ACTH) into the circulation. ACTH binds to receptors on cells in the adrenal cortex, which are stimulated to secrete the steroid hormones, cortisol and aldosterone. This pathway is also stimulated by stress at the higher cerebral center. ACTH acts as a short-loop feedback control on the hypothalamus, while cortisol and aldosterone are long-loop regulators. Aldosterone functions in the kidneys to regulate salt and water homeostasis. Cortisol interacts in many peripheral tissues and has many metabolic functions involving the regulation of carbohydrate, protein, lipid, and overall energy metabolism. Cortisol also plays a poorly understood role in the provision of resistance to infection and inflammation that accounts for the therapeutic use of steroids in many clinical situations. As with all endocrine systems, diseases occur that result from hyperfunction or hypofunction of the hypothalamic–pituitary–adrenal pathway. Diseases may be primary, resulting from end organ dysfunction, or secondary, resulting from pituitary or hypothalamic disease. Pediatric diseases associated with primary disorders of the adrenal cortex are shown in **Table 30.11**. Many of the disorders listed are rare genetic diseases. However, steroid 21-hydroxylase deficiency is sufficiently common (about 1 of 5000 births) to merit whole population screening in most developed countries in the world by measurement

Table 30.11 Congenital Diseases of the Adrenal Cortex

Primary Disorder	Metabolic Profile
21-Hydroxylase[a]	↑ 17-Hydroxyprogesterone, ↓ cortisol
3β-Hydroxy dehydrogenase	↑ Dehydroepiandrosterone (DHEA)
11β-Hydroxylase	↑ 11-Deoxycortisol, ↓ cortisol
17α-Hydroxylase	↑ 17-Ketosteroids, ↓ testosterone
18-Hydroxylase	↓ Aldosterone, ↑ renin

[a] Common disorder screened for in all newborns.

of blood-spot 17-hydroxyprogesterone levels. This disorder results in failure to adequately synthesize both aldosterone and cortisol. Aldosterone deficiency results in renal salt-losing crises, and patients in the newborn period can be profoundly hyponatremic and hyperkalemic. Failure to synthesize cortisol results in stress-induced hypoglycemia. Furthermore, intermediates of steroid metabolism build up causing androgenization. Females born with this disorder frequently have ambiguous genitalia and may be first classed as males. Males may not have such pronounced abnormalities at birth but may still develop electrolyte crises.[18] Defects of the other enzymes in the steroid pathway result in a variety of clinical signs and symptoms including salt loss, androgenization, hypertension, and hypoglycemia, and diagnosis frequently requires steroid profiling.

Growth Factors

The hypothalamus secretes two regulatory hormones that affect growth. GH-releasing hormone is a 40-amino-acid polypeptide that stimulates release of GH from the anterior pituitary. GH-inhibiting factor, also known as *somatostatin*, inhibits GH secretion. Additional factors from higher cerebral centers (cerebral cortex), including catecholamines, serotonin, ghrelin, and endorphins, have a positive effect on GH secretion. Inhibition of GH secretion also occurs when infants are socially deprived. The mechanism for this reversible inhibition is not known, but neglect and potential child abuse are major differentials in infants with retarded growth.

GH is a 191-amino-acid polypeptide, with the liver as its primary site of action. GH receptors on the liver that are occupied by a GH molecule cause the liver to secrete a group of related polypeptide hormones, called *insulin-like growth factors (IGFs)*, and their binding proteins, called *IGF-binding proteins*. IGF-1 and IGF-BP3 are the most significant products of GH activity on the liver. IGF-1 and insulin have similar molecular structures. However, IGF-1 is a much more potent stimulator of linear growth and increased metabolism in infants.

This pathway probably represents the most important endocrine pathway responsible for normal childhood growth. Deficiencies of any component of the pathway are known to result in poor growth, resulting in short-statured adults.

There are difficulties in measuring GH in serum, which result from diurnal variation of secretion and stress-related effectors including catecholamines. A single, low level of GH is not sufficient to confirm GH deficiency. It is important to determine true organic deficiency, caused by hypothalamic or pituitary disease, from emotional deficiency because only organic deficiency responds to expensive GH replacement therapy, while nonorganic GH deficiency will respond to emotional lifestyle changes. Trauma to the head may also cause failure of GH secretion by the pituitary through direct anoxic damage to the very sensitive specialized GH-secreting cells. This type of growth failure will respond to GH therapy. Several stimulation tests have been devised to test the absolute capacity of the pituitary to secrete GH. These stimulation tests include direct stimulation with glucagon or inducement of hypoglycemia with insulin, arginine, or clonidine. These tests require that up to five blood samples be collected in the 2 hours poststimulation and the peak level of GH secretion determined. If this is less than 7 ng/mL, the patient has organic GH deficiency and is likely to respond to GH therapy.[19]

Downstream testing for serum IGF-1 and IGF-BP3 levels is available and shows great promise in the identification of GH deficiency because these compounds are not secreted by the liver in GH deficiency and their basal levels do not seem to have the large variation that occurs for GH. In addition, defects of both IGF and IGF-BP synthesis and secretion have been recognized as a cause of growth failure in certain infants. Individuals with these defects are unlikely to respond to GH replacement.

Endocrine Control of Sexual Maturation

The hypothalamus secretes a 10-amino-acid peptide called *gonadotropin-releasing hormone (GnRH)* into the portal blood system. GnRH bind to a specific receptor and causes the release of two larger polypeptide hormones, called *follicle-stimulating hormone (FSH)* and *luteinizing hormone (LH)*, from the anterior pituitary in both males and females.

Baseline levels of the gonadotropins FSH and LH are very low in infants and young children due to GnRH suppression. FSH and LH have different effects in males and females, both before and during puberty. In males, the target hormonal activity is directed to the testis and causes the release of androgens, primarily the steroid hormones testosterone and androstenedione that are released into the circulation. In females, the primary site of action is the ovary and results in the secretion of a different family of steroid hormones—estrogens, primarily estradiol—into the circulation. Prior to puberty, the circulating levels of androgens and estrogens are very low. The measurement of pediatric gonadotropin levels and sex steroid hormones provide excellent examples of the need for unique pediatric reference ranges for these biomarkers, as the pediatric clinical laboratory is often requested to measure these hormones when a child appears to be going into premature puberty. Testosterone is particularly difficult to measure in prepubertal children as many commercial immunoassays detect an interfering compound, which results in false elevation of the hormone level. Measurement by mass spectrometry appears to overcome these interferences.

During puberty, the GnRH suppression is removed, and there is a gradual increase in FSH and LH secretion, with concomitant increase in androgen levels in males and estrogen and progesterone levels in females. These changes result in the development of secondary sexual characteristics and onset of menarche in females. This period of childhood is also associated with a major surge in linear and bone growth until adult proportions are achieved.

Disorders of this endocrine pathway are associated with either premature or precocious puberty or delayed onset of puberty. The measurement of serum FSH, LH, testosterone, and estradiol levels are useful in evaluating disordered puberty. Often, disorders of other endocrine systems affect puberty. Congenital adrenal hyperplasia results in excess secretion of androgen-like steroids that can affect puberty. Disorders of the hypothalamus and pituitary can affect secretion of FSH and LH, which, if reduced, will result in delayed puberty.

Development of the Immune System

In pediatric healthcare facilities, most hospitalizations and admissions are related to complications arising from infectious diseases. At the same time, although the parents or caregivers may be exposed to the same infectious etiologies, they do not become so ill as to require medical attention. This is because the child does not have the same degree of **immunity** to disease at birth or during infancy.[1]

Basic Concepts of Immunity

The immune system is divided into two functional divisions: the innate immune system and the adaptive immune system. The innate immune system is the first line of defense, particularly in the newborn and infant not exposed to infection. The adaptive immune system generates a specific reaction following exposure to an infectious agent and provides greater immunity with subsequent exposure to that agent. Initially, however, the first response to exposure may be suboptimal and result in illness related to that exposure.

Components of the Immune System

Skin

The skin is normally an effective barrier to most microorganisms, although in premature babies, this barrier is less well developed and can easily become a source of infection. Many newborns with congenital infections manifest profound skin lesions. Most infectious agents enter the body by the nasopharynx, lungs, gastrointestinal tract, or genitourinary tract. Normally, various physical and biochemical defenses protect the non-orifice sites of entry. Lysozyme, an enzyme widely distributed in different secretions, for example, is capable of partially digesting a chemical bond in the membrane of many bacterial cell walls, thus reducing the infectivity. In premature babies, this pathway of defense is less effective. Surgical incisions and intravenous or central lines are also potential sites of entry for infectious agents.

Phagocytes

Phagocytes are present in many cell types. When a foreign organism such as a bacterium penetrates an epithelial surface, it encounters phagocytic cells, which are derived from bone marrow and recruited into tissue in response to the organism. These cells engulf and digest particles. Phagocytic cells include polymorphonuclear cells, which are short-lived in the circulation, and monocytes that, when exposed to a foreign particle, develop into macrophages that subsequently recognize the organism when the individual is re-exposed and mount a concerted defense, which is typically more potent than that of the first exposure.

B Cells

B cells are lymphocytes that are characterized by the presence of surface immunoglobulins. These cells can differentiate into plasma cells that are able to respond to foreign antigens in the circulation by producing neutralizing antibodies. Activation, proliferation, and differentiation of B cells are assisted by cytokine secretion from T cell lymphocytes, which do not produce antibodies. On binding an antigen, antibodies can activate a cascade involving the complement pathway, which ultimately produces lysis and cellular death of foreign organisms.

Natural Killer Cells

Natural killer (NK) cells are leukocytes capable of recognizing cell surface changes on host cells infected by viral particles. The NK cells bind to these target cells and can kill them and the virus. The NK cells respond to interferons, which are cytokine molecules produced by the host cells when infected by virus. Interferons are also part of the innate immune system capable of providing resistance to infection in host cells not virally infected.

Acute-Phase Reactants

Acute-phase reactants (APR) are defense proteins produced by the liver in response to infection, particularly bacterial infection. Certain proteins can increase in the serum by twofold to 100-fold. The most significant APR is called *C-reactive protein (CRP)* because of its ability to bind to the C protein of *pneumococci*. CRP bound to bacteria promotes the binding of complement that, in turn, aids phagocytosis. Serum CRP levels are routinely measured to determine degree of infection in pediatric patients. The required sensitivity of the CRP assay for this clinical purpose is less than that used for the high-sensitivity CRP assay used clinically as an independent risk factor for cardiac disease. In the pediatric application of this assay, rapid turnaround of results is most important. The complement system consists of at least 20 proteins, most of which are APR. They interact sequentially with each other, with antigen–antibody complexes and with cell membranes in a coordinate manner to ultimately destroy bacteria and viruses. Clinically, the complement proteins that are measured most often are C3 and C4. Low levels of either of these proteins indicate poor ability to destroy foreign particles. A more detailed description of APR can be found in Chapter 6, which discusses amino acids, proteins, and heme derivatives.

Antibody Production

Immunoglobulins (γ-globulins) are classified into five major groups, based on structure and function: IgG, IgM, IgA, IgD, and IgE. Secreted by plasma cells derived from B lymphocytes, their properties are listed in **Table 30.12**. IgG is the major immunoglobulin subclass providing antibody response in adults and representing 70% to 75% of total immunoglobulin content. IgG is further broken down into four additional subclasses: IgG 1–4. Each immunoglobulin class is built from similar structural units, based on two heavy polypeptide chains (A, G, M, D, and E) and two light chains (k and λ). The ability to recognize large numbers of foreign antigens is a result of the infinite ability of the genes for the so-called variable region of the immunoglobulin molecule to rearrange. This area recognizes foreign antigens, and a gene rearrangement with production of a unique antibody covers each new foreign antigen exposed to the body. Because of the large number of different immunoglobulin species within each subgroup, electrophoretic separation of these serum proteins on an isoelectric focusing gel during serum protein electrophoresis (SPEP) analysis generates a region that is diffuse, unlike the albumin or transferrin bands, which are distinct bands.

Neonatal and Infant Antibody Production

The human fetus synthesizes a small amount of IgM and, to a lesser degree, IgA. IgG has a lower molecular weight than IgM and readily crosses the placenta. IgG

Table 30.12 Properties of Immunoglobulin (Ig) Classes

	IgG[a]	IgA	IgM	IgD	IgE
Mass (kD)	160	160	970	184	188
Percent of total Ig	70–75	10–15	5–10	<1	Trace
Crosses placenta	Yes	No	No	No	No
In breast milk	Yes	Yes	No	Unknown	Unknown
Activates complement	Yes	Yes	Yes	No	No
In secretions	No	Yes	No	No	No
Binds to mast cells	No	No	No	No	Yes

[a] Present as four subclasses (IgG-1–IgG-4).

is also transferred from mother to baby in breast milk. Transplacental and breast milk–derived IgG offer the baby a passive immunologic protection until endogenous IgG production becomes a significant protective factor. The half-life of IgG is approximately 30 days, and, with prolonged breast-feeding, the infant can derive additional protection. The process of antibody production in infants takes several years to complete when based on total serum levels of the immunoglobulin subclasses, which take up to 4 years to be attained and provide another example of how age-related reference ranges are particularly important in pediatric laboratory medicine. Premature babies have an even greater immunoglobulin deficit than term babies because of diminished transplacental delivery of antibodies.

Immunity Disorders

Given the complexity of the immune system, there are many stages at which acquired or genetically inherited defects can result in inappropriate infectious disease in the pediatric population.[20] Transient hypogammaglobulinemia of infancy may occur because of prematurity or, in certain infants, may be a result of delayed onset of immunoglobulin production of unknown etiology. These infants eventually develop a normal immune system but will be prone to repeated bouts of severe infection until this time. At the opposite end of the spectrum, complete absence of γ-globulins occurs in males in an X-linked disorder known as *agammaglobulinemia*, or Bruton's disease. This disorder presents early in life with recurrent febrile infections. Patients do not have B cells and have low levels of all endogenous immunoglobulin subclasses. The disease process probably begins the moment that any maternally derived immunoglobulins have been lost. The most common infections are of the upper and lower respiratory tracts, causing otitis, pneumonia, sinusitis, meningitis, sepsis, and osteomyelitis. Without early γ-globulin therapy, these children die from respiratory complications. Other immune pathways are normal in these children.

Severe Combined Immune Deficiency

One of the most graphic examples of unique pediatric disease comes from infants who lack both humoral and cellular pathways for killing bacteria and viruses. These children are at risk of severe infection each time they are exposed to an infectious agent. The vivid image is of the "boy in the bubble," existing in a completely sterile environment to avoid contact with any bacteria or virus particles. Severe combined immune deficiency (SCID) may be inherited as an X-linked disorder only seen in males, or it may be autosomal recessive and females may also inherit the disease. There are several causes of SCID, including genetic diseases of purine metabolism and disorders of lymphocyte development and maturation, in which both T cells and B cells, if present, are nonfunctional. The most common purine disorder, adenosine deaminase deficiency, is responsible for 15% of SCID cases. It is diagnosed by measuring elevated levels of adenosine in body fluids. Establishing this diagnosis is important because enzyme replacement therapy using recombinant enzyme has been successfully used to treat the disorder.

Genetic Diseases

Analytic methods for the identification of **genetic disease** play an important part in the pediatric clinical chemistry laboratory.[1,21–23] The presentations of many genetic diseases are unique to the pediatric population and require physicians and clinical laboratorians who provide services to have specialized knowledge and training. Most diseases that present with clinical signs in infants and children are inherited in an autosomal recessive mode, which means that the patient has two disease-causing mutations in the gene for that disorder, one inherited maternally and one inherited paternally. Parents are almost always without any symptoms. Several examples of diseases with this inheritance pattern have already been introduced in this chapter, including galactosemia, due to inability to metabolize the milk sugar galactose, which results in severe hepatic and renal failure in the newborn, associated with congenital cataracts, and congenital adrenal hyperplasia as a result of steroid 21-hydroxylase deficiency. Treatments for these two conditions are readily available. Galactosemia responds very well to dietary restriction of milk and milk products, and hormone replacement is available for congenital adrenal hyperplasia. Certain other diseases are recessive but are inherited on the X chromosome. Typically, males inherit a mutated X chromosome from their mothers; because they do not inherit a paternal X chromosome, they show signs of disease with only one mutation. Females carrying an X-chromosome mutation may also demonstrate disease signs due to the fact that one of their X chromosomes becomes inactivated, and if the active X chromosome carries the mutation, they will have the disease or a spectrum of the disease depending upon the relative amount of X-chromosome inactivation. Dominantly inherited diseases, which can be inherited as a single mutation through either parental line, frequently tend not to present as unique childhood diseases. Examples include familial hypercholesterolemia, Huntington's disease, and factor V Leiden thrombophilia, all of which are generally regarded as diseases of the adult population, although the mutation is inherited at birth. These examples described above are diseases of DNA that replicates in the cell nucleus. Mitochondria, the organelles responsible for generating cellular energy and other important metabolic pathways, also contain DNA (mtDNA), which encodes proteins involved in energy generation. mtDNA has a high rate of spontaneous mutation and results in a number of energy-wasting diseases, many of which present with high lactate levels. Mitochondria are only inherited from the mother, so mtDNA mutations that are not spontaneous can only come from the maternal lineage.

Cystic Fibrosis

Cystic fibrosis (CF) is a common genetic disease within the Caucasian population in the United States, and pediatric clinical chemistry laboratories perform the diagnostic testing. The rate of this debilitating disease in Western populations is 1 in 2400 live births, which results from recessively inherited mutations in the cystic fibrosis transmembrane regulator (*CFTR*) gene. Patients may present in the newborn period with severe pancreatic insufficiency caused by accumulated thick mucus secretions in the pancreatic ducts, which inhibit the secretion of pancreatic digestive enzymes. These babies have steatorrhea and fail to thrive. Patients with CF do

CASE STUDY 30.2, PART 2

Remember Tracey. The laboratory determined the respiratory infections were due to a *Pseudomonas* spp. The pediatrician orders laboratory tests including a sweat chloride.

1. What condition/disease is the pediatrician trying to confirm?
2. What is the molecular mechanism of this disease?

CASE STUDY 30.3, PART 2

Remember Jody. At the next interaction with the pediatrician via a video visit, her mother reported a strange odor coming from Jody's diaper, lethargy, and not being as playful and engaged as her older brother. Blood tests ordered by the pediatrician showed elevated levels of phenylalanine.

1. What is the most likely diagnosis?
2. Why is breast feeding not advised for this baby?
3. What is the treatment for this condition?

© Karl Tapales/Moment/Getty Images.

not always develop pancreatic symptoms; however, in most patients, the thick mucus that accumulates in the lungs causes respiratory disease and makes them particularly susceptible to rare infectious diseases, such as *Pseudomonas*. Although palliative therapies, such as antibiotics and lung transplants, have improved over the past few generations because of the availability of better antibiotics, CF is still regarded as incurable.

The gold standard diagnostic test for CF has been available for many years. It involves measurement of chloride content in sweat collected after pilocarpine iontophoresis. The dynamic range for sweat chloride concentration differs from that of blood, and therefore, sweat cannot be run on a major chemistry analyzer. Analysis must be made with a more sensitive chloride meter. This type of testing is time consuming and requires specialist experience from the operator. The genetic basis for CF has been established. Although there are some mutations that are frequently encountered in the population, such as the ΔF508, about 1500 mutations are seen in CF patients. The American College of Medical Genetics and the American College of Obstetrics and Gynecology recommended heterozygote screening in selected couples.[24] This recommendation poses an analytic challenge because of the large number of mutations and is dependent upon the availability of clinically and cost-acceptable mutation-detection technology such as next-generation sequencing (see Chapter 22, *Pancreatic Function*)

Newborn Screening for Whole Populations

Certain inherited diseases are sufficiently common in the population to be considered candidates for whole population screening.[21,22] Phenylketonuria (PKU) was the first genetic metabolic disorder to be screened in every baby born in the developed world and increasingly in emerging counties. Other diseases that are readily treatable were added to the list in following years, including steroid 21-hydroxylase deficiency, sickle cell disease, and CF—and, in some states and countries, galactosemia. These genetic diseases mostly respond well to simple therapy, which is often dietary. In most U.S. states, this process takes place in the state public health laboratory. The nature of the testing procedure requires a sensitive screening test that has few false-negative results and an acceptable level of false positives. Then, a confirmatory test that is more specific is performed to rule out false-positive results.

Tandem mass spectrometry allows many different biochemical genetic diseases to be screened on a single sample at the same time, and the technology was introduced in the 1990s into many programs. This technique allows whole groups of similar compounds, in particular amino acids for the diagnosis of amino acid disorders and acylcarnitines for the diagnosis of fatty acid oxidation and some organic acid disorders, to be analyzed on small sample volumes without complex sample preparation (**Table 30.13**). The analytic time is about 2 minutes per sample, which means that it is possible to readily perform analysis for an entire state. The birth rate in the United States is approximately 4.4 million births per year, and tandem mass spectrometry has been shown to be capable of handling this workload; all states currently offer this program. This approach is increasingly being utilized for global screening of metabolic diseases.

The present status of newborn screening for metabolic diseases comprises a panel of 29 recommended conditions, referred to as the uniform panel, of which 20 can presently be diagnosed using tandem mass spectrometry. The most common of these additional

Table 30.13 Metabolic Diseases Detectable by Newborn Screening Using Tandem Mass Spectrometry[24, 25]

Amino Acids
Phenylketonuria (PKU)
Maple syrup urine disease (MSUD)
Tyrosinemia, types 1 and 2
Homocystinuria
Hypermethioninemia
Urea Cycle
Argininemia
Citrullinemia
Argininosuccinic aciduria (ASA)
Organic Acids
Propionic acidemia (PA)
Methylmalonic acidemia (MMA)
Isovaleric acidemia (IVA)
Glutaric acidemia, types 1 and 2 (GA1, GA2)
β-Ketothiolase deficiency
3-Hydroxy-3-methylglutaryl-CoA lyase (HMG-CoA lyase) deficiency
3-Methylcrotonyl-CoA carboxylase (MCC) deficiency
Malonyl-CoA decarboxylase deficiency
2-Methyl-3-hydroxybutyryl-CoA dehydrogenase*
Fatty Acids
MCAD deficiency
Short-chain acyl-CoA dehydrogenase deficiency (SCAD)* possibly a benign biochemical condition
Very-long-chain acyl-CoA dehydrogenase (VLCAD) deficiency
Carnitine palmitoyltransferase, types 1A and 2 (CPT1A, CPT2)
Carnitine-acylcarnitine translocase (CAT) deficiency
Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency (LCHAD)
Mitochondrial trifunctional protein (MTP) deficiency

* Clinical significance not established.

disorders is medium-chain acyl-CoA dehydrogenase (MCAD) deficiency, which is a defect of the pathway in which we derive energy from fatty acids during periods of fasting or high energy demand such as fever. Prior to including this defect in newborn screening programs, infants would develop fasting-induced hypoglycemia and liver failure, often leading to death or neurological damage. Subsequent to screening by measurement of blood spot octanoylcarnitine (C8) and presymptomatic diagnosis with early management, the outcomes for children with MCAD deficiency have improved greatly. In addition to the core screening diseases, a number of secondary diseases can also be identified using tandem mass spectrometry—some of which are very rare and have less data to confirm sensitivity and specificity. Many additional metabolic diseases can be diagnosed by tandem mass spectrometry; however, the current evidence to support a substantial benefit at the population level by uniform screening is less strong. Methods to diagnose additional diseases by tandem mass spectrometry are constantly being developed, and it is likely that routine population-based screening will be expanded in the future. The next group of conditions that will be included in all screening programs will be lysosomal storage diseases, which were long considered to be untreatable. However, treatment modalities are currently being investigated for many lysosomal storage diseases, and it is likely that the most success in treatment will come from early diagnosis in the newborn period. Some states in the United States have already included some lysosomal disorders in their newborn screening programs.

Diagnosis of Metabolic Disease in the Clinical Setting

At the present time, the clinical laboratory is needed to confirm the diagnosis of newborns screened positive for the disorders listed in Table 30.13[24] and also for the rest of the approximately 500 single-gene defects that result in biochemical genetic disease not presently detectable by tandem mass spectrometry. These inborn errors of metabolism can be broken down generally into two main types.

Large Molecule Diseases

Large molecule diseases have an accumulating intermediate of metabolism composed of large complex molecules; examples are listed in **Table 30.14**. Many of these diseases involve intracellular accumulation of the abnormal chemical with relatively small excretion in body fluids. In glycogen storage diseases, the glycogen accumulates in the liver and muscle but cannot be seen in blood or urine samples. The histopathologist, using microscopic examination of tissue,

Table 30.14 Examples of Large Molecule Storage Disorders

Mucopolysaccharide (MPS) or Glycosaminoglycan Storage Diseases
Hurler's disease (MPS, type I)
Hunter's disease (type II)
Morquio's disease (type IV)
Complex Lipid Storage
Gaucher's disease
Tay-Sachs disease
Niemann-Pick disease (types A, B, C)
Glycogen Storage Diseases
von Gierke (type 1)
Pompe (type 2)
McArdle (type 5)
Peptide Storage
Neuronal ceroid lipofuscinoses, types 1–10 (Batten's disease)
Posttranslational Protein Modification
Carbohydrate-deficient glycoprotein syndromes

Table 30.15 Pathways Involved with Small Molecule Metabolic Disease and Examples

Amino acids—phenylketonuria
Fatty acids—medium-chain acyl-CoA dehydrogenase deficiency
Organic acids—propionic acidemia
Urea cycle—citrullinemia
Oxidative phosphorylation—mitochondrial DNA diseases
Vitamin metabolism—pyridoxine responsive seizures
Steroid biosynthesis and breakdown—21-hydroxylase–deficient congenital adrenal hyperplasia
Cholesterol synthesis—Smith-Lemli-Opitz syndrome
Purine and pyrimidine metabolism—adenosine deaminase form of severe combined immunodeficiency
Neurotransmitter metabolism—4-hydroxybutyric aciduria
Plasmalogen synthesis—Zellweger's syndrome
Glutathione metabolism—pyroglutamic aciduria
Oxalate metabolism—hyperoxaluria types 1 and 2
Creatine metabolism—guanidinoacetic acidemia

often makes these diagnoses. A few, mostly urine, tests are available for gathering clues to large molecule diseases, including glycosaminoglycan (mucopolysaccharide [MPS]) analysis using high-voltage electrophoresis to identify unusual metabolites associated with the MPS storage diseases. These tests are relatively insensitive, and improvements made by using mass spectrometry are currently being investigated. Confirmation of the diagnosis requires measurement of deficient enzyme activity or identification of defined disease-causing mutations. Fortunately, many enzymes that result in large molecule storage diseases can be found in white blood cells, enabling confirmation to be made on a blood sample. Enzymic confirmation can be difficult to establish in small babies because large blood samples may be required for diagnostic testing.

Small Molecule Diseases

Small molecule diseases result from defects in metabolic pathways of intermediary metabolism. Usually, the abnormal compounds that are present in these diseases are low-molecular-weight compounds that are readily excreted in body fluids; the types of pathways involved are listed in **Table 30.15**. Initially, the clinical chemistry laboratory had few diagnostic tools capable of identifying the large number of metabolic intermediates that may accumulate in these diseases, and a number of simple, colorimetric urine tests, such as the dinitrophenylhydrazine (DNPH) test for ketoacids, were developed to establish diagnosis. These methods lack both sensitivity and specificity and have no role to play in the modern clinical chemistry laboratory. They have been superseded by assays with greater sensitivity and specificity, often based on separation technology and mass spectrometry.

Small molecule diseases have a variable clinical presentation and could present to almost any medical subspecialty with any organ system involved. (Table 30.15 lists certain pathways that may be involved.) Biochemical testing for these diseases is usually described in two phases. First, it is important to recognize the degree of tissue compromise at

presentation. This requires routine chemistry evaluation for blood gas status, if acidotic; anion gap measurement; liver function testing; analysis of skeletal and cardiac muscle markers, such as creatine kinase or troponin levels; lactic acid; and ammonia measurement. These analyses should be available STAT and used to monitor management once a diagnosis is established. The second phase of analysis should be to look for metabolic markers that pinpoint the site of a defect. These tests involve a form of separation technology, such as ion-exchange chromatography or ultraperformance liquid chromatography for amino acids. The preferred material for amino acid analysis is serum because the renal tubules have efficient transport systems for reabsorbing filtered amino acids. It is possible to miss an amino acid abnormality if only urine is analyzed. Urine amino acid analysis is of value if a tubular defect such as cystinuria is suspected. The most useful test for detecting abnormal metabolic intermediates is organic acid analysis. This test is performed on urine and should only be performed using gas chromatography-mass spectrometry (GC-MS). GC-MS is used to identify metabolic markers for up to 100 genetic diseases and is a form of metabolomics. As mentioned earlier, tandem mass spectrometry is another technique seeing rapid growth in the metabolic disease diagnosis field. This technique is being applied to newborn screening and is playing an increasing role in analysis of multiple different metabolites, many of which were not easily measurable prior to the introduction of this technology.

Drug Metabolism and Pharmacokinetics

There are several important differences in the way that infants and children handle pharmacologic agents compared with adults.[1] This area of pediatric laboratory medicine provides many good examples of why children should not be regarded as "small adults." It is not clinically appropriate to prorate the amount of drug prescribed to a child based on relative body weight compared with an adult dose. **Drug metabolism** depends on the following factors: absorption, circulation, volume of distribution, and metabolism and clearance. Often, the medium in which a drug is provided to a child differs from the adult version of the same drug. Syrups, for instance, provide a more rapid release of a drug and greater availability for gastrointestinal absorption than tablets, which have the drug trapped in a solid matrix that requires digestion. Children are more likely to be given medication in a palatable form, such as syrup, and to require lower doses. The pH of gastric secretions differs in infants. At birth, the gastric pH is nearly neutral, not reaching the adult level of acidity for several years. This pH difference can affect the absorption of certain drugs, including some frequently prescribed penicillins. The distribution of drugs often differs between adults and children. Lipid-soluble drugs are taken up into lipid reserves and only slowly released into the circulation. Because infants have relatively little adipose tissue, these drugs are not stored as efficiently. The overall effect is that lipid-soluble drugs reach a higher level more quickly than in individuals with sizable fat stores; however, the drug is also cleared more rapidly. It becomes appropriate for drugs to be provided in smaller, more frequent doses to optimize the effect. Hepatic metabolism of many drugs is immature in young infants. This may delay the metabolic conversion to an active drug or increase the time in which an active drug is circulating. Good hepatic function is important for clearing those drugs metabolized by the liver, and good renal function is important for clearing drugs that have water-soluble end products.

Therapeutic Drug Monitoring

The principles of therapeutic drug monitoring remain the same in adult and pediatric clinical chemistry. It is important to measure the blood levels of various drugs if that information can provide important guidance to the physician with regard to optimal dosing. This is most important if a drug has a well-defined therapeutic index. This means that the drug is known to be ineffective if the blood level is below a certain value, that there is a well-defined therapeutic range above which the drug is effective, and that there is a higher level at which the drug becomes toxic. It is important to monitor levels of drugs with these characteristics. **Table 30.16** lists drugs for which the importance of therapeutic monitoring is established.

Toxicologic Issues in Pediatric Clinical Chemistry

Issues related to the provision of a toxicologic service can be divided into two distinct groups in pediatrics. The first group involves infants and young children who unknowingly consume pharmacologic and other chemical agents. This usually involves the child finding access to medication belonging to another individual in the household and consuming the

Table 30.16 Drugs with Well-Defined Therapeutic Indices

Drug	Therapeutic Range	Toxicity
Phenytoin	10–20 mg/L	>20 does not enhance seizure control; >35 causes seizure
Phenobarbital	15–40 mg/L	>35 causes CNS depression; >65 coma
Carbamazepine	4–12 mg/L	>15 causes drowsiness
Theophylline[a]	6–13 mg/L	>20 can cause cardiac arrhythmia
Caffeine[a]	8–20 mg/L	Less toxic than theophylline
Methotrexate	<0.1 μmol/L (72 hours after dose)	High levels cause myelosuppression
Gentamicin	5–10 mg/L (peak)[b]	>10 ototoxicity, nephrotoxicity

[a] Theophylline is metabolized to caffeine in neonates but not in adults. Used to treat apnea.
[b] Peak level should be drawn 30 minutes after last dose for aminoglycoside drugs. Children are particularly prone to hearing loss at toxic levels.
Data from Wu AHB, ed. *Tietz Clinical Guide to Laboratory Tests.* 4th ed. St. Loius, MO: WB Saunders; 2006.

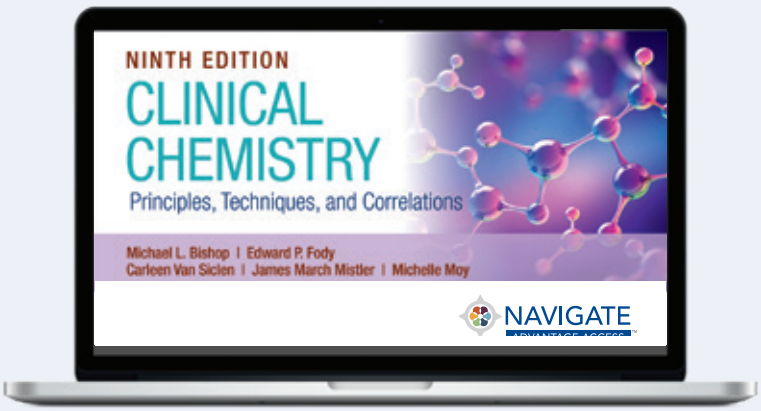
medication as if it were candy. It is relatively easy for the investigator to ascertain the nature of the medication by identifying what is available in the household. Toxicologic investigation can usually be restricted to a few specific tests.

A rare but potentially dangerous condition is that of Munchausen syndrome by proxy. In this condition, mental illness in a caregiver causes them to give unnecessary and illness-causing drugs to an otherwise healthy child. This can go unrecognized and result in multiple hospitalizations and even death of the child. Clinical suspicion of this form of child abuse should involve performing a comprehensive drug screen to identify causative agents. Because of the intermittent nature of clinical presentation of Munchausen syndrome by proxy, it can often be confused with metabolic disease. Metabolic studies, in addition to comprehensive toxicologic studies, may be necessary.

The likelihood of self-ingestion of street drugs of abuse (DOA) is present in older children; however, young children may inadvertently ingest them as well. Also, many DOA can be passed in breast milk from mother to infant. Because of this, pediatric clinical chemistry laboratories should make assays available for DOA similar to those in adult practice.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Green A, Morgan I, Gray J. *Neonatology and Laboratory Medicine*. London: ACB Venture Publications; 2003.
2. Soldin SJ, Rifai N, Hicks JMB, eds. *Biochemical Basis of Pediatric Disease*. Washington, DC: American Association for Clinical Chemistry; 1992.
3. Rautenberg MW, van Solinge WW, Heunks JJ, et al. Pediatric tube direct sampling by the Abbott Architect Integrated ci8200 Chemistry/Immunochemistry Analyzer. *Clin Chem.* 2006;52:768–770.
4. Demir AY, van Solinge WW, Kemperman H. Handling of and direct sampling from primary barcode-labeled pediatric tubes on Vitros Clinical Chemistry Analyzers integrated into an enGen Work Cell. *Clin Chem.* 2005;51:920–921.
5. Gill FN, Bennett MJ. The pediatrics unit. In: Price CP, Hicks JM, eds. *Point of Care Testing*. Washington, DC: American Association for Clinical Chemistry; 1999.
6. ACOG Committee Opinion No. 765: Avoidance of nonmedically indicated early-term deliveries and associated neonatal morbidities. *Obstet Gynecol*. 2019;133:e156–e163.
7. Yarbrough M, Stout M, Gronowski A. Pregnancy and its disorders. In Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier; 2018:1655–1696.
8. Polychronakos C, Bellin M, Wood JR, et al. Diabetes mellitus. In: Sarafoglou K, Hoffmann GF, Roth KS, eds. *Pediatric Endocrinology and Inborn Errors of Metabolism*. 2nd ed. New York, NY: McGraw-Hill Education; 2017:381–393.
9. Hannon TS, Rao G, Arslanian SA. Childhood obesity and type 2 diabetes mellitus. *Pediatrics.* 2005;116(2):299–306.
10. Tillotson CV, Bowden SA, Boktor SW. Pediatric type 2 diabetes mellitus. *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing; 2021. Accessed September 22, 2021. https://www.ncbi.nlm.nih.gov/books/NBK431046/
11. American Diabetes Association. Classification and diagnosis of diabetes: Standards of Medical Care in Diabetes—2020. *Diabetes Care*. 2020;43:S14–S31.
12. Zimmet P, Alberti K, George MM, et al.; IDF Consensus Group. The metabolic syndrome in children and adolescents—an IDF consensus report. *Pediatr Diabetes.* 2007:8:299–306.
13. National Kidney Foundation, Kidney Disease: Improving Global Outcomes (KDIGO) Work Group. KDIGO Clinical Practice Guideline for acute kidney injury. *Kidney Int.* 2012;2(Suppl 1):1–141.
14. Schwartz GJ, Work DF. Measurement and estimation of GFR in children and adolescents. *J Am Soc Nephrol.* 2009;4(11):1832–1643.
15. Newman DJ. Cystatin C. *Ann Clin Biochem.* 2002;39: 89–104.
16. Schwartz GJ, Schneider MF, Maier PS, et al. Improved equations estimating GFR in children with chronic kidney disease using an immunonephelometric determination of cystatin C. *Kidney Int*. 2012;82:445–453.
17. Wagner C, Greer FR. Prevention of rickets and vitamin D deficiency in infants, children, and adolescents. *Pediatrics.* 2008;122:1142–1152.
18. Austin A, Finkielstain GP. Adrenal disorders. In: Dietzen DJ, Bennett MJ, Wong ECC, eds. *Biochemical and Molecular Basis of Pediatric Disease*. 4th ed. Washington, DC: AACC Press; 2010:135–153.
19. Winter WE, Rosenbloom AL. Disorders of growth. In: Dietzen DJ, Bennett MJ, Wong ECC, eds. *Biochemical and Molecular Basis of Pediatric Disease*. 4th ed. Washington, DC: AACC Press; 2010:171–194.
20. Loechelt BJ. Primary immunodeficiency diseases. In: Dietzen DJ, Bennett MJ, Wong ECC, eds. *Biochemical and Molecular Basis of Pediatric Disease*. 4th ed. Washington, DC: AACC Press; 2010:409–427.
21. Valle D, Beaudet AL, Vogelstein B, et al., eds. In: *Online Metabolic and Molecular Bases of Inherited Disease (OMMBID)*. Accessed July 30, 2012. http://ommbid.com
22. Sargoglou K, Hoffman G, Roth K, eds. *Pediatric Endocrinology and Inborn Errors of Metabolism*. New York, NY: McGraw-Hill; 2009.
23. Hommes FA, ed. *Techniques in Diagnostic Human Biochemical Genetics*. New York, NY: Wiley-Liss; 1991.
24. American College of Obstetricians & Gynecologists, American College of Medical Genetics. *Preconception and Prenatal Carrier Screening for Cystic Fibrosis: Clinical and Laboratory Guidelines*. Washington, DC: American College of Obstetricians & Gynecologists; 2001.
25. Dietzen DJ, Rinaldo P, Whitley RJ, et al. National Academy of Clinical Biochemistry Laboratory Medicine Practice Guidelines: follow-up testing for metabolic diseases identified by expanded newborn screening using tandem mass spectrometry; Executive summary. *Clin Chem.* 2009;55: 1615–1626.

CHAPTER 31

Geriatric Clinical Chemistry

Laura M. Hickes and J. Marvin McBride

CHAPTER OUTLINE

KEY TERMS

Geriatrics
Gerontology
Menopause
Osteoporosis

CHAPTER OBJECTIVES

Upon completion of this chapter, the clinical laboratorian should be able to do the following:

- Evaluate the impact of SARS-CoV-2 (COVID-19) on the U.S. population in older adults.
- Define geriatrics, gerontology, menopause, and osteoporosis.
- Discuss the impact of geriatric patients on the clinical laboratory.
- Appraise the physiologic changes that occur with the aging process.
- Identify the age-related changes in clinical chemistry analytes.
- Explain the problems associated with establishing reference ranges for older adults.

- Describe the effects of medication on clinical chemistry results in older adults.
- State the effects of exercise and nutrition on chemistry results in older adults.
- Correlate age-related physiologic changes with laboratory results.
- Apply knowledge of clinical chemistry to answer case study questions throughout chapter.

The number and proportion of older individuals in the total population are increasing rapidly in the United States and other developed nations. There are a number of physiological and metabolic changes unique to aging that may affect clinical laboratory values. In addition, aging is often accompanied by multiple medical conditions requiring diverse medications, which can also affect clinical laboratory values. This chapter will summarize the issues that should be given special consideration with regard to clinical chemistry in the geriatric patient.

The Aging of America

Since 2010, there has been an almost 34% increase in the population of persons 65 and older. It is currently estimated that more than 49.2 million Americans are over the age of 65 years, which makes up 15.2% of the population. This demographic is projected to almost double to 98 million by the year 2060. No other age group has seen such a rapid increase. Within this group of seniors, the group over the age of 85 years is projected to more than double from 6.4 million in 2016 to 14.6 million by the year 2040. This reflects a 129% increase in this age group.[1] The first "baby boomers"—individuals born during the post–World War II surge from 1946 to 1964—started turning 65 in 2011. Despite this, for the first time in recent history, life expectancy has actually decreased. In 2021, the average U.S. life expectancy decreased by 1 year, such that the average female is expected to live to 80.5 years, while the average male is expected to live to 75.1 years.[2] Despite this decrease, since the start of the 21st century, the

CASE STUDY 31.1, PART 1

Carlos, a 65-year-old man in overall good health, was admitted to the hospital to have his appendix removed.

© Juanmonino/E+/Getty Images.

CASE STUDY 31.2, PART 1

Jada, a healthy 70-year-old woman, is seen by her primary care physician for an annual checkup.

© Jeremy Poland/iStock/Getty Images Plus/Getty Images.

CASE STUDY 31.3, PART 1

Gertrude, an 88-year-old female, presents at the emergency department with fever, cough, sore throat, and fatigue. Blood is draw and sent to the laboratory. Flu testing is also ordered.

© DG FotoStock/Shutterstock.

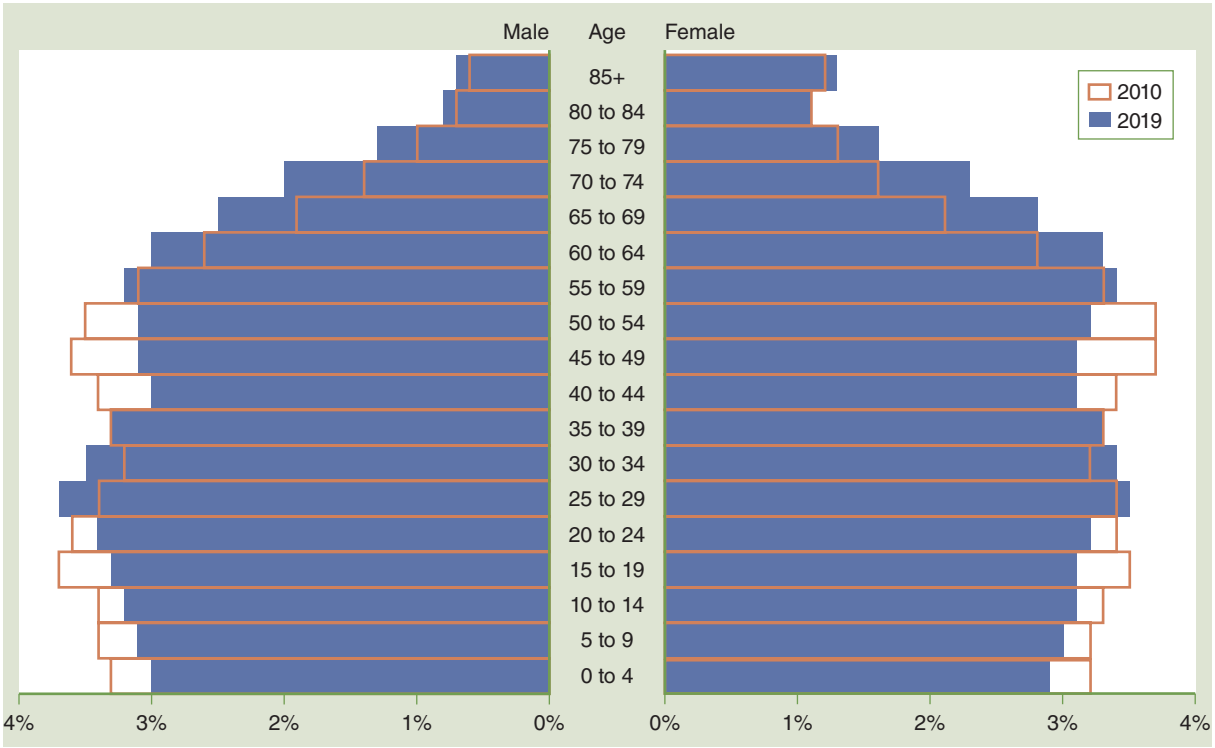


Figure 31.1 Aging of America in 2010 vs. 2019.

A Wave of Change. United States Census Bureau. June 25, 2020. https://www.census.gov/library/visualizations/2020/comm/a-wave-of-change.html. Accessed June 17, 2021.

dependency ratio, which is defined as the number of dependents age 0–14 and over the age of 65 divided by working age population, has outpaced the growth of the working-age population. Notably, there has been a growth in the 65+ population and a shrinking in the under-14 population.[3] **Figure 31.1** shows the aging of America in 2010 versus 2019. The surge in the number of Americans over the age of 65 has outpaced the growth of the working-age population in recent years.[3] These changes are largely due to the aging of the baby boomers along with immigration trends.[4] By the year 2030, all of the baby boomers will be over the age of 65. Complementary data from the 2014 Current Population Reports illustrates and characterizes the increase in the population numbers over the past 114 years.[5] Depending on immigration,[6] fertility, and longevity rates, these seniors are expected to constitute over 20% of the total U.S. population by 2020 (**Figure 31.2**).[7] Though it is unclear how COVID-19 will impact the population breakdown, it is at least a contributor to the observed decrease in life expectancy. The total population over the age of 65 is expected to reach almost 4.5% of the total population by 2050.[5] **Figure 31.3** shows a breakdown of COVID-19 deaths in the United States according to age. This shows that nearly 81% of the first ~300,000 fatalities preferentially impacted those over the age of 65.[7] While the overall mortality rate of COVID-19 has been reasonably small, especially when considering the overall case number, the fact that COVID-19 is associated with a poor outcome at a higher rate in older adults may cause projections regarding the percentage of the population that will be over the age of 65 to be less accurate. Regardless of the impact of COVID-19, with the burden of medical illness and medication use falling disproportionately on older adults, a sizable increase in the utilization of healthcare services, including clinical chemistry, is inevitable given these population trends.

Aging and Medical Science

Gerontology, the study of aging, and **geriatrics**, the subspecialty of clinical medicine that focuses on care of the aged, have much to tell us about normal aging, as well as the unique features of common illnesses in older adults. The differentiation of "normal" or "healthy" aging versus the accumulation of multiple medical problems with age is an area of active research and debate. There are several analytical measurements that change when comparing a diseased individual to a healthy one, but because a number of biochemical changes occur as a consequence

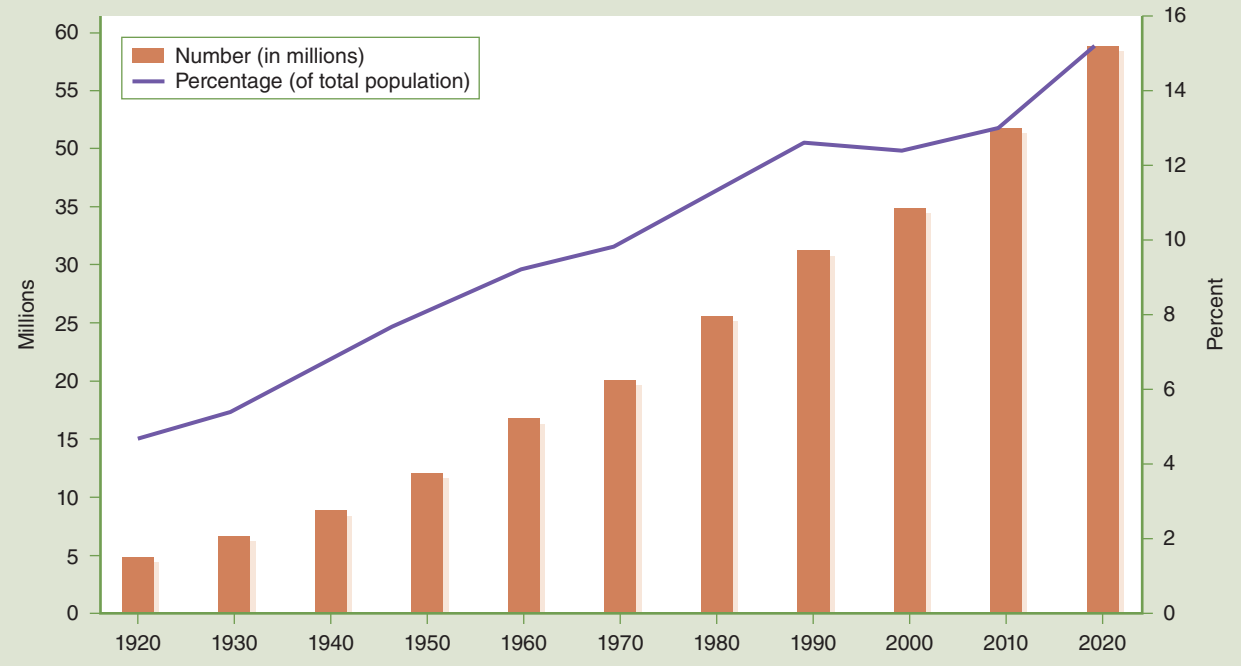


Figure 31.2 Population 65 years and older, by number and percent of total population: 1920 to 2020. This figure shows the age growth over the past 100 years in the United States. Currently, over 54 million people are over age 65. Additionally, the percentage of the population over age 65 has increased from 13% in 2010 to 15.2% in 2020.

Data from 2020 Census. United States Census Bureau. Last revised June 8, 2021. https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-main.html. Accessed June 19, 2021.

of normal aging, it can be difficult to differentiate between abnormal physiological changes and normal signs of aging in geriatric patients.

One important principle is that aging is extremely heterogeneous. The extent of differences between individuals of the same age is often quite significant. Indeed, an individual's overall health status and expected longevity are more closely related to their functional status than their chronologic age. Thus, an 86-year-old jet-setting international attorney may look very unlike a 60-year-old nursing home patient bedbound with dementia. Older adults may mistake a health symptom for a normal part of aging and may not seek medical attention as early as may be optimal,[5]

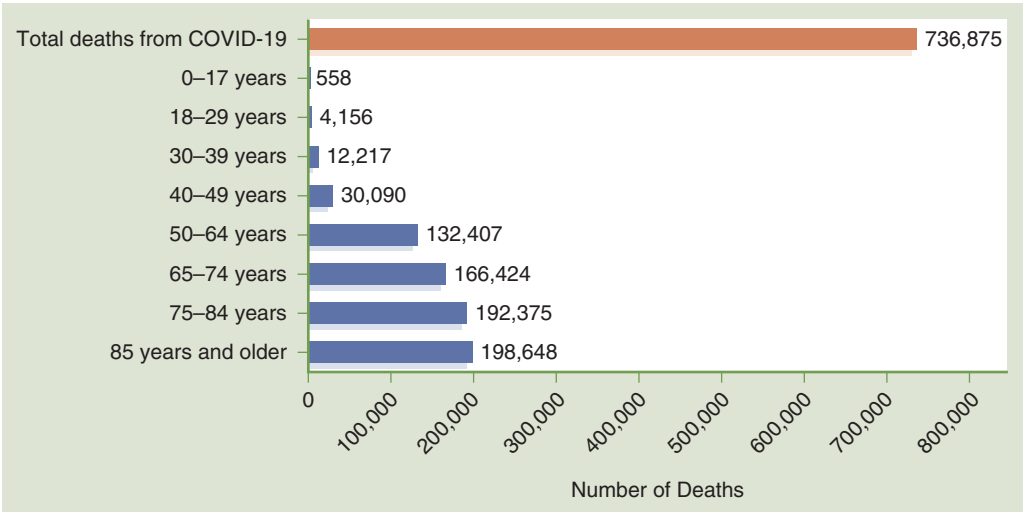


Figure 31.3 Number of SARS-CoV-2 coronavirus disease 2019 (COVID-19) deaths in the United States as of January 2, 2021, by age. While the projected percent distribution of the American population over age 65 has been predicted to increase due to the aging of the baby boomer generation, the impact of COVID-19 may alter those trends, given the prevalence of infection in the United States and the much higher mortality rate in the aged population, with 81% of all COVID-19–related deaths occurring in the age group over 65 years.[7]

Data from COVID-19 Deaths Reported in the US as of January 2, 2021, by age. Published by John Elflein, Jan 11, 2020. Source: https://www.statista.com/statistics/1191568/reported-deaths-from-covid-by-age-us/

CASE STUDY 31.1, PART 2

Remember Carlos. Prior to tomorrow morning's surgery, his laboratory results are reviewed by the anesthesiologist.

Analyte	Patient Value	Reference Range
Albumin	2.5 g/dL	3.5–5.0 g/dL
BUN	35 mg/dL	8–26 mg/dL
Creatinine	1.7 mg/dL	0.9–1.5 mg/dL
Serum osmolality	280 mOsm/kg	275–295 mOsm/kg
Sodium	140 mmol/L	135–145 mmol/L

© Juanmonino/E+/Getty Images.

1. What is Carlos's BUN/creatinine ratio?
2. What is the expected ratio, and what condition correlates with this result?
3. What are causes of abnormal plasma urea concentration in the geriatric population?
4. What is a possible explanation for Carlos's decreased albumin result?

or conversely, may unnecessarily seek medical care for normal changes of aging.

An important consideration is the concept of diminished physiological reserve. While an aged person usually has the same adaptive mechanisms to stress as a younger person, those adaptations may not be as rapid or as robust. Therefore, with diminished reserve, an aged individual may encounter adverse health effects much more quickly than a younger person. For example, an older adult undergoing diuretic therapy for hypertension would not be able to withstand a long period of vomiting and diarrhea compared with a younger person. Age-associated decrease in total body water would favor dehydration and development of hypernatremia and delirium.[8]

While age-dependent changes in concentrations of analytes are shown to be important in a number of areas of chemistry, there are also several examples of chemistry values that are unchanged as a consequence of age.[9] **Table 31.1** outlines some of the age-specific changes that occur in regard to certain chemistry analytes, and some of these changes will be specifically addressed in the coming section outlining

Table 31.1 Changes in Selected Clinical Chemistry Analytes with Age[8–28]

Increase	Decrease	Unchanged
ALP (females)	ADH	ACTH
BNP	Albumin	Calcium
Calcitonin	Aldosterone	Chloride
C-reactive protein	Vitamin B_{12}	Cortisol
EPO	LDL-C	Creatinine
Ferritin	Cholesterol	HDL-C
Fibrinogen	DHEA-S	Insulin
Folic acid	Estrogen	Sodium
FSH	Ferritin	Total T_4
Free T_4	GH	pH
GGT	IGF-1	
Homocysteine	PaO_2	
Insulin	Progesterone	
Lactate	Testosterone	
Parathyroid hormone	T_3	
Potassium	Total protein	
TBG	Uric acid	
TSH	Vitamin D	
Transferrin		
Triglycerides		

ACTH, adrenocorticotropic hormone; ADH, antidiuretic hormone; BNP, brain natriuretic peptide; DHEA-S, sulfated dehydroepiandrosterone; EPO, erythropoietin; FSH, follicle-stimulating hormone; GGT, γ-glutamyl transferase; HDL-C, high-density lipoprotein cholesterol; IGF-1, insulin-like growth factor-1; LDL-C, low-density lipoprotein cholesterol; T_3, triiodothyronine; T_4, thyroxine; TBG, thyroid binding globulin; TSH, thyroid-stimulating hormone.

physiologic changes with age. Ideally, the changes that impact laboratory results would be classified in age-specific reference ranges, but since these are not readily available, changes in analyte level over time can help to manage patient treatment.[10]

General Physiologic Changes with Aging

The aging process involves both biochemical and physiological changes and can be subdivided into various organ systems. The systems continue to function appropriately unless they are subjected to excess stress. With increasing age, an individual's ability to respond to stress decreases. This leads to an age-associated increase in the prevalence of pathological conditions. **Table 31.2** outlines some of the more common factors impacting the interpretation of clinical laboratory results.

Table 31.2 Factors Impacting Interpretation of Clinical Laboratory Results in the geriatric population[11–22]

Exercise
Duration
Type
Medications
Polypharmacy
Mobility
Immobility
Posture
Nutritional status
Personal habits
Alcohol use
Smoking
Presence of chronic or subclinical disorders
Validity of reference ranges
Specimen collection variables
Site
Trauma
Volume

Muscle

Total body muscle mass typically decreases with age, but the rate and extent of loss have a strong genetic component. This age-related decline (sarcopenia) results in a decrease in lean body mass and a decrease in total creatinine production, such that a spot serum creatinine by itself is not reliable for assessment of renal function in the aged.[23]

Bone

Total bone density and mass decrease with age in both men and women, though the changes are much more dramatic in women after **menopause**.[24] Serum calcitonin levels typically rise with age, but ionized calcium levels remain stable.[11] Parathyroid hormone (PTH) levels increase in postmenopausal women, and this increase is associated with changes in bone metabolism.[25]

Gastrointestinal System

The incidence of atrophic gastritis increases with age, with a consequent increase in vitamin B_{12} deficiency from poor absorption.[27] The incidence of achlorhydria (low or absent gastric acid production) also increases with age, which can result in decreased calcium and iron absorption, as well as an increased incidence of bacterial overgrowth in the small intestine.[26] Albumin levels may decrease, although this is more often attributed to illness, inflammation, or malnutrition than to aging per se. The incidence of malnutrition, which is associated with higher mortality rates, also increases with age.[12]

Kidney/Urinary System

The number of functional glomeruli decreases with age, resulting in a decrease in kidney size and weight. Glomerular filtration rate (GFR) declines, and renal blood flow is even more reduced, such that filtration fraction (GFR/renal plasma flow) actually increases.[23] Additionally, the kidneys concentrating ability declines. The result of all these changes is that acid/base, water, and electrolyte levels remain normal under optimal conditions, but physiologic reserve is diminished. The secretion of erythropoietin (EPO), the glycoprotein hormone that controls red blood cell production, increases with age, and the level of serum renin, an enzyme that regulates blood pressure, decreases. Renal responsiveness to atrial natriuretic peptide (ANP) decreases, and serum levels of ANP and brain natriuretic peptide (BNP) increase

CASE STUDY 31.2, PART 2

Recall Jada. Her physician asks if she had any specific concerns. She complains of muscle weakness, drowsiness, and confusion. A comprehensive metabolic panel (CMP), calcium, and a PTH are ordered and completed.

Analyte	Patient Value	Reference Range
Calcium	12.0 mg/dL	8.5–10.2 mg/dL
Parathyroid hormone	111 pg/dL	12–72 pg/mL

1. Based on the laboratory data, how is the PTH result interpreted?
2. Should any additional testing be done?

with age.[26,28] Structural, neurological, and immune changes result in an increased incidence of urinary tract infection in older adults, especially in women.[29]

Immune System

The thymus shrinks with age, causing a decrease in thymosin levels and T-cell function. B-cell function declines slowly, as well, such that cellular and humoral immune responses are less vigorous and slower in the aged than in younger individuals.[30] Autoimmune antibodies (antinuclear antibodies [ANA] and others) increase,[31] and hematopoietic stem cell numbers decrease.[31] Examples of the consequences of declining cellular immunity in older adults include reactivation of latent tuberculosis and herpes varicella zoster (HVZ) virus. HVZ is the virus responsible for causing varicella (chickenpox), often acquired in childhood (at least prior to administration of the varicella vaccine in children beginning in 1995). Reactivation of the virus, which remains dormant in the individual's nervous system after resolution of the clinical syndrome of varicella, is called herpes zoster (shingles) and presents as a painful, blistering skin rash. There are about 1 million cases of herpes zoster per year in the United States, with a lifetime risk of about 1 in 3, mostly after the age of 50 years. The incidence and morbidity of this illness led to the development of a two-dose adjuvanted recombinant HVZ vaccine in 2006 with up to 90% efficacy.[32] Recombinant zoster vaccine (RZV, Shingrix) is recommended to prevent shingles in adults 50 and older.

Endocrine System

Serum adrenocorticotropic hormone (ACTH) and cortisol levels typically do not change with age, though response to stress may be delayed. Pulsatile secretion of growth hormone (GH) typically diminishes with age, resulting in a decrease in lean body mass/fat ratio, as well as loss of overall body mass. The peak but not the basal levels of melatonin secretion decrease with age, which may contribute to sleep cycle disorders as well as diminished protection from free radicals.[9] Norepinephrine secretion generally increases, which contributes to systemic vasoconstriction and decrease in myocardial relaxation; epinephrine levels remain stable. Aldosterone levels may decline, which can contribute to orthostatic hypotension, or a drop in blood pressure upon standing.[33] Thyroid hormone levels are typically well preserved or slightly increased into very old age.[11,34]

Sex Hormones

Although **menopause** in women typically occurs prior to the earliest ages considered "geriatric," the reduction in gonadal production of estrogen and progesterone and secondary increase in hypothalamic gonadotropin-releasing hormone persist through the remainder of a woman's life.[35] Testosterone levels in men typically exhibit a gradual decrease with increasing age. The term *andropause* has been used to describe this observation in analogy to menopause, but the decline in serum levels of sex hormone is not abrupt and not as significant in degree for men as it is for women. The distinction between a normal serum testosterone level for age and hypogonadism in an aged man is a point of some contention and includes consideration of sexual function, general perception of well-being by the individual, and other clinical factors in addition to the value of the serum testosterone level.[36] Dehydroepiandrosterone (DHEA), sulfated DHEA, and pregnenolone all decrease with age.

Glucose Metabolism

Insulin secretion is unchanged, though there are age-related changes at the cellular level in insulin signaling, receptors, and glucose transporters. An individual with the genetic predisposition to type 2 diabetes mellitus is more likely to manifest clinical illness with increasing age, body mass index, and lack of exercise. The combination of an aging population, increasing obesity, and an increasing number of individuals from racial/ethnic groups at increased risk of diabetes (African American, Hispanic/Latino, Native American) has resulted in a dramatic increase in the number of older adults with diabetes.[37] Many of these individuals will also suffer some of the consequences of diabetes, including damage to the eyes, kidneys, nerves, and blood vessels.

Effects of Age on Laboratory Testing

Basic biochemical and physiological changes accompany the aging process, and these changes can impact an individual's clinical laboratory test results. It is important that laboratorians understand how these changes can impact specific tests. It has been suggested that the presence of age-specific changes may warrant separate reference ranges for older adults,[13,38] but these are not readily available. In addition, consideration of preanalytical variables must be factored into the interpretation of laboratory results. How does the aging process affect interpretations of drug levels in older adults? What are the effects of exercise and nutrition on chemistry results? **Table 31.3** describes some of the factors that laboratorians should consider when interpreting test results for older adults and defines some of the laboratory values that are impacted by these changes.

Muscle

As muscles age, they begin to decrease in number and size. Creatinine levels correlate with both muscle mass and renal function. In the geriatric population, a decrease in muscle mass coupled with the decrease in renal function may keep the creatinine level nearly the same or slightly increased, even though renal function has significantly decreased.

Bone

Osteoporosis incidence is high in the older adult population. This leads to increased risk of fracture, which can have major detrimental effects on mobility, independence, and longevity of older adults. Lack of sex hormones in a major factor contributing to osteoporosis, which is why it is so much more common in women, but it can also occur in men, especially those with hypogonadism. Adequate calcium intake and sufficient vitamin D are important in maintaining bone mass and density—many adults, including older adults, have inadequate dietary intake of calcium and insufficient vitamin D levels for multiple

Table 31.3 Effects of Age on Laboratory Testing[8–13,35–37,39–41]

	Normal Physiological Changes with Age	Laboratory Values that Correlate
Muscle	↓ Muscle mass	↓ Creatinine
Bone	↓ Mineral content of bone, ↓ cartilage	↑ PTH (females), ↓ calcitonin
GI	↓ Gastric motility, vitamin absorption, and drug absorption	↓ Vitamin B_{12}, calcium, and Fe absorption
Kidney	↓ Renal filtration	↑ Serum ANP, BNP, EPO, and creatinine, ↓ GFR and renin
Immune	↓ Hematopoietic stem cells, bone marrow activity, thymosin, and T-cell function	↑ ANAs
Endocrine	↓ Endocrine gland sensitivity to pituitary and other stimuli	↓ Aldosterone, ↑ Norepinephrine
Reproductive	↓ Sex hormones	↓ Testosterone, estrogen, progesterone, DHEA-S, and pregnenolone, ↑ GnRH

ANA, antinuclear antibody; ANP, atrial natriuretic peptide; BNP, brain natriuretic peptide; DHEA-S, sulfated dehydroepiandrosterone; EPO, erythropoietin; GFR, glomerular filtration rate; GI, gastrointestinal; GnRH, gonadotropin-releasing hormone; PTH, parathyroid hormone.

reasons, and both contribute to the risk of osteoporosis. Vitamin D helps with absorption of dietary calcium from the intestine; thus decreased vitamin D levels can decrease the amount of dietary calcium that is absorbed. Inadequate calcium absorption may eventually lead to low serum calcium levels and increased PTH levels. This increased PTH then causes increased calcium loss from the bone, which increases alkaline phosphatase levels.[41]

Gastrointestinal System

The gastrointestinal system includes the digestive tract and accessory organs including the pancreas and the liver. There are a number of age-related changes with respect to the liver analytes. C-reactive protein, an acute phase reactant, has been shown to be elevated in older adults.[42] This elevation is thought to be a nonspecific indicator of inflammation and typically is a poor prognostic indicator.[43] Levels of γ-glutamyl transferase (GGT) tend to increase with age in a gender-specific manner. Men do not show this age-related increase, but men tend to have higher levels of GGT than women. Fibrinogen, another acute phase reactant, is frequently elevated in geriatric patients, and this increase coincides with inflammatory disease, stroke, coronary dysfunction, and cancer.[41–45]

Ferritin levels can be low in the older adult population, usually due to iron deficiency anemia, similar to younger people. Transferrin levels can be reduced due to iron-deficient anemia or as a result of acute or chronic stress. Albumin levels are frequently decreased in older adults as a result of inflammation, malnutrition, and liver disease. Levels can be high in dehydration, but elevations above the upper limits of normal rarely occur in older adults due to the prevalence of other health conditions that mask alterations.[13,37] Total protein levels are also frequently decreased in older adults for the same reasons that albumin levels are low. Drugs that are highly bound to albumin and other plasma proteins, such as warfarin, digoxin, phenytoin, and valproic acid, with only the free component clinically active, may have normal total levels in serum but experience toxic clinical effects due to the increased free component.

Urinary System

Prostatic enlargement continues throughout a man's life, and this results in a gradual climb in prostate-specific antigen (PSA). Therefore, prostate-specific antigen (PSA) values are reported with the 95th percentile limits by decade of age, and reference ranges include men with benign prostatic hyperplasia. The incidence of prostate cancer also increases

CASE STUDY 31.3, PART 2

Remember Gertrude. Gertrude's laboratory results are available in the electronic health record as seen below.

© DG FotoStock/Shutterstock.

Analyte	Patient Value	Reference Range
Sodium	146 mmol/L	136–145 mmol/L
Potassium	5.1 mmol/L	3.4–5.0 mmol/L
Chloride	108 mmol/L	98–107 mmol/L
Bicarbonate	24 mmol/L	22–29 mmol/L
BUN	24 mg/dL	6–20 mg/dL
Creatinine	1.2 mg/dL	0.6–1.1 mg/dL
Calcium	8.6 mmol/L	8.6–10.0 mmol/L
Glucose	120 mg/dL	70–100 mg/dL
FLU A/B/ RSV	Negative	Negative

1. Which of the abnormal test results may be age-related?

Based on clinical presentation and laboratory results, Gertrude is diagnosed with bronchitis and discharged with a prescription for an antibiotic. She is instructed to seek medical attention if her symptoms worsen.

with age and may result in even higher PSA levels. The use of PSA testing for detection of prostate cancer in the general male population is no longer recommended, though it may be appropriate in some high-risk groups.[46] Urinary obstruction from prostatic enlargement may result in obstructive uropathy with rise in serum creatinine and other analytes. Serum creatinine levels are relatively lower in older adults due to decreased muscle mass. At the same time, a reduction in GFR can offset this decrease.

Immune System

As age increases, infection-induced morbidity and mortality rise. This is in part due to a weakened immune system. The innate immune system, commonly referred to as the "first line of defense," is the nonspecific antigen activation that confers short-term protection from a pathogen. Adaptive immune system, on the other hand, is activated by the innate system and refers to the protection from antigen that an individual develops throughout the life. Both the innate and adaptive immune systems become damaged and dysfunctional as individuals age, called "immune senescence," leading to increased risk of infection and potentially also autoimmune disease and cancer.[30,36] Gastroenteritis is more frequently observed as individuals' age, but the increased frequency is thought to be a result of a weakened immune system.[26] Pathogenic bacteria can enter and infect the digestive tract and contribute to presentation. Additionally, there is an increase in autoimmune antibodies in the older population, such that a higher titer is required to be "significant" in the diagnosis of many autoimmune disorders in the aged.[31]

Endocrine System

There are a variety of age-related changes to endocrine hormone regulation. These can be subclassified according to whether the hormone level increases or decreases.

Increased Hormone Level

Elevated levels of cortisol, though rarely present, have been reported to be associated with decreased cognitive function and memory loss.[38] Thyroid hormone levels are typically unchanged with age, but thyroid-stimulating hormone (TSH) levels may increase slightly in older adults.[34,36] True hypothyroidism also increases with age, making distinction between age-related changes in TSH, subclinical hypothyroidism, and clinically significant hypothyroidism more complex in the older adult. Follicle-stimulating hormone (FSH) levels increase with aging, but with this increase, there is a down-regulation of FSH receptors such that in menopause, there are no longer circulating cells with FSH receptors and thus no responsiveness to the high level of hormone. ANP levels increase with age, and the levels have been shown to be nearly fourfold higher in healthy elderly individuals than in younger people.[26] Anemia is common in the older adult population, though typically attributable to a variety of illnesses rather than age per se. EPO, a marker for anemia and renal insufficiency, rises in elderly patients.[36] PTH levels are slightly higher in older individuals than gender matched pairs.[36] However, the incidence of primary hyperparathyroidism also increases with age, frequently identified in older adults due to routine serum calcium testing included in chemistry panels.[47]

Decreased Hormone Level

In contrast to cortisol levels, DHEA levels have been shown to decrease by 40% to 60% with age.[11] Estrogen and progesterone levels decrease with age, and the reduction in estrogen further up-regulates FSH levels. Insulin-like growth factor (IGF-1) and GH levels decrease with age. Secretion rates and serum concentrations of aldosterone decrease with age as a consequence of decreased renin levels.[13,36] Pituitary function declines with age, and hypothalamic antidiuretic hormone (ADH) levels are increased.

Sex Hormones

Levels of expression of all sex hormones diminish with age. Testosterone levels decrease with age, as do estrogen and progesterone levels.[36]

Glucose Metabolism

Insulin sensitivity decreases with age. As a result of this decreased sensitivity, there is an increase in the prevalence of type 2 diabetes, with the incidence reaching a peak between the ages of 60 and 74.[37]

Establishing Reference Ranges for Older Adults

Most laboratory tests have gender-specific reference ranges and/or "age-specific" reference ranges. The broad categories for the age ranges are very wide, and the adult reference range includes individuals

between 18 and about 50 years. As people live longer, age-related criteria for the analysis and interpretation of test results become increasingly important. Because of the need to establish reference ranges in a healthy population, and the increased prevalence of at least one health condition in the aged, there are little data on more appropriate age-specific reference ranges for older adults. Based on this lack of data, there has been an increased interest in determining age-related reference ranges in order to more effectively identify individuals with early stage disease.[13,34,36–38] While the idea has a lot of merit, there are only a few publications establishing age-appropriate reference ranges for older adults.[11] An example of clinically well-recognized variation of laboratory values with aging that is not typically supported by specific laboratory reference ranges is TSH. The reference ranges for TSH are known to vary by age and in different trimesters of pregnancy, and clinical laboratories typically report separate reference ranges for pediatric, pregnant, and adult populations. TSH levels are known to increase slightly with advanced age, but the variation in TSH levels increases even more dramatically. When the normal range of an analyte is defined as the mean ± 2 standard deviations, the upper limit of normal for TSH in older adults may be as high as 6 or 10 μIU/mL. With an upper range of normal for adults typically published as 3.3 to 4 μIU/mL, an older adult experiencing normal aging may be diagnosed as hypothyroid and treated with thyroid replacement unnecessarily.[48,49]

Certain analyte fluctuations that are seen as individuals age are clearly the result of aging organs, but other analytes do not lend themselves to such apparent delineation. In addition, coincident medical conditions can further complicate the issue. The requirement that reference values be obtained from healthy, normal individuals unfortunately limits a large number of geriatric patients from contributing to establishing these references. Recently, a strategy for reference range data mining has been introduced to expand on age-appropriate reference ranges for all age groups. Prior to this approach, the use of exclusion criteria caused the geriatric population to be vastly underrepresented in most randomized clinical trials, which further limited the validity of age-specific reference range generation.[9] In addition, there is a wide intraindividual variability among the various analytes, which has been seen as a major obstacle for determining age-appropriate reference ranges among older adults. Instead, the current clinical approach that has been increasingly implemented is careful documentation of laboratory values and paying closer attention to changes over time instead of where the values fit in with the remainder of the population. This is not feasible among all older adults, as there is a small subset that resists treatment and thus does not see a medical professional on a regular basis.[5] A lack of baseline measurements coupled with the lack of age-specific reference ranges continues to make diagnosis challenging. Currently, most physicians who care for the geriatric patient population rely on established patient care, frequent routine examinations, and following changes in laboratory values over time as an early indicator of a problem.

Preanalytical Variables Unique to Geriatric Patients

There are a number of factors that contribute to the accuracy and validity of test results in any population, but several of these factors have a greater impact in the geriatric patient. These include sample collection, sample handing, and physiological variables. Geriatric patients can present a challenge to phlebotomists due to disease, malnutrition, or dehydration.[8] With increased age, there is a reduction in healing rate and an increased risk of acquiring infection due in part to a gradual loss in the capacity of the immune system to fight off infections.[30,31] Further, the skin and veins are less elastic and can be injured more easily during venipuncture. The decrease in muscle mass and collagen leads to a decrease in vascular stability of veins and a subsequent decrease in blood flow. Physiological changes in the patient may impact laboratory results due to unavoidable issues that arise at the time of collection. Increased hemolysis or insufficient volume can ultimately impact the validity of the result. Other factors that may influence normal laboratory values in geriatric patients include diet, medications, exercise, smoking, alcohol consumption, physical activity, and body composition. Oral biotin supplements, often taken in the hope of improving hair or nail quality, can cause artificially low TSH and thyroid hormone levels.[50] While these factors influence results independent of age, geriatric patients are more likely to have one or several of these causing variations in test results. Caution should always be exercised when reporting results in a geriatric patient, but clearly an absurd result (one not compatible with life) should be investigated to identify potential causes of the discrepancy.

Diseases Prevalent in Older Adults

There are many diseases that are especially common in older adults. **Table 31.4** outlines the prevalence of some of these conditions in Americans 65 years and older as reported in the 2017 National Council on Aging, a self-reported statistically valid sample of the U.S. population.[16] The overall prevalence of these diseases is inadequately understood due to poor documentation and the presence of multiple conditions warranting many medications. In fact, it is estimated that 80% of adults over the age of 65 have at least one of these chronic conditions, while 68% of adults over age 65 have 2 or more of these conditions. There have been attempts to classify disease prevalence in the older adult population based upon pharmacy databases.[14] Classification in this population has been difficult due to inaccuracies related to misidentification of conditions or diseases or incomplete medical records. Approaches to classify individuals at an earlier stage of disease have led clinicians and laboratorians to analyze drug combination therapies as markers of disease. A list of drugs that patients are prescribed is available in a pharmacy database, from which statistics can be generated calculating nationwide averages for disease prevalence.

The majority of the diseases noted in **Table 31.5** are degenerative conditions, leading to concerns that increased life expectancy will increase the number of years with declining health. Health is defined as the state of being free from illness or injury, and deteriorating health is the number one reason that elderly leave their homes and enter retirement facilities.[48] Unfortunately, without baseline measurement, it can be difficult to identify noteworthy changes in analyte levels.

Menopause is defined as the time when the primary function of the human ovaries is permanently terminated. Timing is variable for each female, but typically, it begins between the ages of 35 and 58 years. This wide age range is based in part on the fact that functional disorders affecting the reproductive system can speed the transition to menopause. For instance, women with cancer of the reproductive tract, polycystic ovary syndrome, or endometriosis typically transition to menopause earlier than women without these conditions.[51] The precise definition of menopause is lack of menses for greater than 1 year. This transitional phase from reproductive to nonreproductive typically lasts about a year to a year and a half, and during this period of time, there is a gradual decline in ovarian function, thus a decrease in ovarian hormone production.[52] The changes that occur are caused by reduced estrogen levels that result from loss of the granulosa and interstitial cells lining the follicles.[53] The placenta normally produces hCG during pregnancy, but in addition to placental production, hCG can also arise from the pituitary. In perimenopausal women, hCG levels are often detected in the absence of pregnancy as a result of ovarian failure.[54] Pituitary hCG is more commonly detected in women

Table 31.4 Ten Most Common Chronic Conditions for Adults Over 65 Years of Age

Chronic Disease	Estimated Prevalence
Hypertension	58%
High cholesterol	47%
Arthritis	31%
Heart disease	29%
Diabetes	27%
Chronic kidney disease	18%
Heart failure	14%
Depression	14%
Alzheimer's disease/Dementia	11%
Chronic bronchitis/Emphysema	11%

Data from National Council on Aging Annual Report. Fiscal year 2017 At https://www.ncoa.org/wp-content/uploads/FY17-NCOA-Annual-Report.pdf. Accessed February 12, 2021.

Table 31.5 Leading Causes of Death in the United States

SARS-CoV-2 (COVID-19)
Heart disease
Cancer
Alzheimer's disease
Cerebrovascular disease
Chronic lower respiratory disease
Diabete
Other respiratory diseases
Pneumonia and influenza
Renal failure

Data from Woolf SH, Chapman DA, Lee JH. COVID-19 as the Leading Cause of Death in the United States. *JAMA*. 2021;325(2):123–124. doi:10.1001/jama.2020.24865.

CASE STUDY 31.3, PART 3

Remember Gertrude. She was having severe difficulty breathing and returned to the emergency department the next evening. A chest x-ray was ordered to determine if the bronchitis developed into pneumonia. She was admitted to the ICU and STAT laboratory testing is ordered. Assist the intensive care hospitalist in interpreting her results.

© DG FotoStock/Shutterstock.

Analyte	Patient Value	Reference Range
Sodium	153 mmol/L	136–145 mmol/L
Potassium	5.1 mmol/L	3.4–5.0 mmol/L
Chloride	111 mmol/L	98–107 mmol/L
Bicarbonate	24 mmol/L	22–29 mmol/L
BUN	45 mg/dL	6–20 mg/dL
Creatinine	2.0 mg/dL	0.6–1.1 mg/dL
Calcium	8.8 mmol/L	8.6–10.0 mmol/L
Glucose	70 mg/dL	70–100 mg/dL
SARS-CoV-2	Reactive	Nonreactive
Alkaline phosphatase	95 U/L	42–98 U/L
Alanine aminotransferase	50 U/L	7–45 U/L
Aspartate aminotransferase	50 U/L	5–35 U/L
Total bilirubin	9.0 mg/dL	0.2–1.2 mg/dL
Direct bilirubin	8.9 mg/dL	0.0–0.3 mg/dL
Total Protein	6.5 g/dL	6.5–8.3 g/dL
Albumin	3.0 g/dL	3.5–5.0 g/dL
LD	225 U/L	125–220 U/L
CK	150 U/L	34–145 U/L
Procalcitonin	30 ng/mL	< 0.08 ng/mL

2. What is the most likely cause of Gertrude's symptoms?
3. How should the renal panel and liver panel be interpreted?
4. What is the clinical significance of the elevated procalcitonin result?
5. What is Gertrude's prognosis?

≥55 years, and marked elevations in this age group are most commonly associated with cancer.[36]

Osteoporosis, a bone disease that is prevalent in older adults, leads to increased risk of bone fracture. There are two types of bony tissue: cortical (compact) and trabecular (spongy). Cortical bone accounts for 80% of the total bone mass of an adult, while trabecular bone accounts for the remaining 20%. Trabecular bone is highly porous (30%–90%), while compact bone has a porosity of only 5% to 30%. The risk of fracture is highest in bones with more compact tissue, which functions in whole body support. Bone mineral density (BMD) is reduced with age, and the World Health Organization has defined osteoporosis as BMD 2.5 standard deviations below peak bone mass.[25] This decreased density and an observed reduction in a number of bone-associated proteins both contribute to the weakened condition. Classifications of osteoporosis include primary type 1, primary type 2, and secondary (also known as type 3).[41] While primary type 1 is more frequently seen in elderly females, there is a growing prevalence of osteoporosis in men. Secondary osteoporosis occurs in women or men who

are older than 70 years of age and is typically associated with decreased bone formation and decreased renal production of $1,25(OH)_2D$ (vitamin D). Vitamin D deficiency causes decreased calcium absorption, which increases the PTH levels and promotes calcium mobilization from the bones, sacrificing skeletal calcium in order to maintain normal serum calcium levels.[13] Type 3 osteoporosis is the only form of osteoporosis that affects individuals of any age. Most cases of type 3 osteoporosis are due to drug treatment or medication use, such as steroids, and type 3 is less common in older adults. Laboratory testing surrounding osteoporosis includes monitoring serum calcium, phosphate, creatinine, alkaline phosphatase, and 25-hydroxyvitamin D, and in men, testing also includes testosterone.[36,41]

Dementia is a syndrome defined as impairment of memory and at least one other cognitive domain (language, perceptual skills, attention, constructive abilities, orientation, and problem solving) that is severe enough to significantly impair day-to-day function.[53] Dementia is common in older adults—approximately 10% of individuals over 65 years suffer from the syndrome; this rises to as high as 50% of those over 85 years.[54] Mild cognitive impairment (MCI) is a term used to describe similar cognitive problems that are not severe enough to significantly impair day-to-day function, but between 6% and 25% of individuals with MCI will go on to develop dementia in 1 year.

Dementia can be described in a number of ways, including the area of the brain that is most affected (cortical vs. subcortical), whether the etiology is known or unknown (secondary vs. primary), age of onset (presenile vs. senile), or by clinical syndrome. Many types of dementia can only be definitively diagnosed by pathological examination of brain tissue, which is obviously not helpful to the individual suffering from the illness, and correlation between the clinical syndromes and pathologic findings has historically been quite poor.[55] Recent advances in imaging and neurochemistry have improved somewhat.[56] Diagnosis of dementia is primarily clinical, though hypothyroidism, hyperthyroidism, vitamin B_{12} or copper deficiency, and central nervous system infections can cause cognitive impairment. Tests for these factors are commonly obtained in the evaluation of dementia. Cerebrospinal fluid markers are increasingly being used in the diagnosis of Alzheimer's dementia, though this is still more common in a research setting than clinical practice.[57]

Alzheimer's disease (AD) is the most common form of dementia, responsible for approximately 70% of cases.[58,59] The pathologic findings of AD include amyloid plaques (made of β amyloid) and neurofibrillary tangles (made of tau protein). Individuals with AD typically have a fairly slow but steady progression of impairment.[60]

Vascular dementia is the second most common form of dementia and may be caused either by multiple clinically evident strokes (multi-infarct dementia) or by ischemic changes in the deep white matter of the brain. Individuals with vascular dementia typically have a "stuttering" or "stair-step" progression of impairment, but the speed of progression and specific manifestations may vary widely, depending on which areas of the brain are damaged.[61]

Dementia with Lewy bodies (DLB) is the third most common form of dementia and is characterized by prominent visual hallucinations, significant fluctuations in mental status, and parkinsonism (motor symptoms similar to those of Parkinson's disease). Lewy bodies are collections of α-synuclein protein inside neurons, but whether they are a cause or an effect of the illness is unclear: Lewy bodies are also seen in Parkinson's disease, as well as in some cases of Alzheimer's. DLB typically progresses more quickly than AD.[62]

There are also less common types of dementia including Pick's disease, frontotemporal dementia, Creutzfeldt-Jakob disease, normal pressure hydrocephalus, Huntington's disease, and Wernicke-Korsakoff syndrome. Typically, diagnosis of dementia is based on the patient's history and physical examination. Brain imaging and clinical laboratory testing may be performed to confirm or eliminate suspicion of specific clinical abnormalities.[63]

Age-Associated Changes in Drug Metabolism

As adults age, changes occur that impact how the body metabolizes and utilizes drug compounds. Absorption, distribution, metabolism, and excretion of the drugs are impacted by the normal aging process. The prevalence of multiple health conditions in this subset of the population has highlighted the issue of polypharmacy[16] and the importance of therapeutic drug monitoring. Such testing in the laboratory helps prevent overdose and adverse drug reactions. Calculation of the actual in situ drug level is a key factor when looking at the result and metabolic rate, and elimination of these compounds needs to be considered for optimal dosing.[62]

Absorption

The rate of drug absorption from the small intestine slows with aging, leading to a lower peak serum concentration and a decrease in the time to reach peak. Typically, the total amount of drug absorbed (bioavailability) is the same in younger and older patients.[62] An exception arises due to an age-related decrease in hepatic function. Drugs that undergo extensive first-pass metabolism (such as nitrates) tend to have greater serum concentrations or greater bioavailability than drugs involved in conjugation or acetylation reactions.

Varying disease states or other drugs administered concurrently to older adults may also cause changes in absorption. For example, calcium carbonate is frequently recommended for bone health in older adults. This calcium salt requires gastric acid for absorption and should therefore be taken with food to stimulate gastric acid secretion. Patients with achlorhydria from natural causes or from acid-suppressing medications such as proton pump inhibitors or H_2 blockers may better absorb the calcium from calcium citrate, which does not require gastric acid.[64]

Distribution

Volume of distribution (V_d) is the volume of plasma into which a fixed concentration of drug is dissolved. The larger the V_d, the greater the distribution of the compound. Older adults typically have less body water and lean body mass as a percent of total body mass—or in other words have more fat as a percent of total body mass. Drugs such as ethanol and lithium that are water soluble (*hydrophilic*) have a lower V_d in older adults. Highly water-soluble drugs tend to be more concentrated in elderly individuals due to decreased body water. Additionally, drugs such as benzodiazepines and barbiturates that are fat soluble (*lipophilic*) have higher V_d in older adults due to increased body fat and decreased lean body mass.[65] The extent to which a drug binds to plasma proteins also impacts its V_d. Many drugs bind to albumin, which is often decreased in older adults. The unbound (free) portion of the drug is the pharmacologically active portion, and many drug assays measure total (bound and unbound) drug concentrations. Elderly patients with "normal" total drug levels may have excessive levels of free drug and exhibit signs of toxicity due in part to decreased albumin levels. Digoxin and phenytoin are examples of drugs that are significantly protein bound[60]; thus geriatricians will typically aim for the low end of the therapeutic range because of this issue.[66]

Metabolism

The liver is responsible for the majority of drug metabolism, and both hepatic blood flow and hepatic mass decrease with age. Phase I pathways of drug metabolism (hydroxylation, oxidation, dealkylation, and reduction) result in metabolites that may be of lesser, equal, or greater pharmacologic effect than the parent drug. In contrast, phase II pathways (glucuronidation, conjugation, or acetylation) result in inactive metabolites. Therefore, drugs metabolized by phase II pathways are generally preferred for older adults because of decreased incidence of toxicity—for example, lorazepam is preferred over diazepam when benzodiazepine therapy is required in older adults. The elimination half-life of diazepam is dramatically lengthened by age, and as a result of this prolonged half-life, its metabolites accumulate to unexpectedly high levels in older adults with repeated dosing. Because of this, lorazepam is preferred in this patient population.[67]

Elimination

Kidney mass and blood flow decrease with age, which results in gradual decline in GFR. **Figure 31.4** graphically represents this gradual, age-related decline over time. Further, the current estimates of the prevalence of chronic kidney disease in individuals over age 60 years are nearly 40%.[68,69] This decrease in GFR impairs elimination of therapeutic drugs and increases toxicity and adverse outcomes.[70,71] Therapeutic drug monitoring is an important consideration to ensure that compounds ingested are being metabolized and excreted appropriately. The gold standard for determining GFR is a measurement of iothalamate or iohexol measurement following intake. Another means of determining GFR is the measurement of creatinine in a 24-hour urine sample, though measurement of creatinine in a random serum sample is more commonly tested due to ease of collection. A number of formulas exist to calculate GFR, and no formula is ideal for all patient populations. Three of the most common equations used to estimate renal function are the Modification of Diet in Renal Disease (MDRD) equation, the Cockcroft-Gault equation (CG), and the Chronic Kidney Disease Epidemiology Collaboration equation (CKD EPI).[72–75] All three formulas use serum creatinine levels in combination with other factors to estimate GFR. The Food and Drug Administration

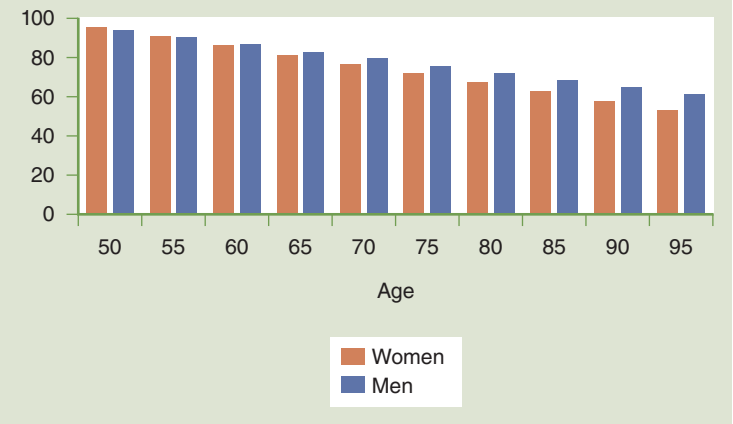


Figure 31.4 Glomerular filtration rate (GFR) in healthy aging individuals.

Data from GFR in Healthy Aging: Bjørn O. Eriksen, Runolfur Palsson, Natalie Ebert, Toralf Melsom, Markus van der Giet, Vilmundur Gudnason, Olafur S. Indridason, Lesley A. Inker, Trond G. Jenssen, Andrew S. Levey, Marit D. Solbu, Hocine Tighiouart and Elke Schaeffner. *JASN* July 2020, 31(7) 1602-1615).

traditionally uses the CG equation in its recommendations on appropriate dosing of therapies.[73] Most clinical laboratories estimate renal function using the MDRD equation because this equation requires serum creatinine, age, race, and gender but not weight. There is growing movement based upon current recommendations to use the CKD-EPI equation, which uses serum creatinine, age, race, and gender.[76] There is also interest in incorporating cystatin C as well as creatinine in an estimating equation. There have been studies to assess the best equation for different demographics and patient populations. The findings of these studies have been contradictory, and based on the mixed results, there is no consensus equation used in older adult patient population.[75] Different GFR results can change the dose of drug administered, and thus, there has been effort to find alternate methods to achieve the best estimation of GFR. There have been various methods proposed and tested, but eGFR measurement has been widely incorporated into clinical laboratory testing. Creatinine measurement alone should not be used to assess renal function. It is also important to note that current discussions of GFR calculations are leaning to remove race from the formulas due to inaccurate results in individuals who are Black/African American. For information on creatinine and GFR, see Chapter 7 *Nonprotein Nitrogen Compounds*.

Example 31.1 Diazepam

Diazepam is highly lipophilic. It is quickly and efficiently absorbed and accumulates with multiple doses. It is metabolized by phase I mechanisms to *N*-desmethyldiazepam and temazepam. Both of these metabolites are further metabolized to oxazepam—all three metabolites are pharmacologically active. Both temazepam and oxazepam are inactivated by glucuronidation (a phase II process) and eliminated via the kidney. Elimination half-life increases by approximately 1 hour for each year of age beginning with a half-life of 20 hours at 20 years of age.[65] This appears to be due to an increase in volume of distribution with age and a decrease in clearance. Consequently, older adults may have lower peak concentrations and, on multiple dosing, higher trough concentrations. It will also take longer to reach steady state, as well as for blood levels to fall after discontinuation of the drug.[64]

Example 31.2 Morphine Sulfate

Parenteral morphine has a smaller volume of distribution, higher plasma concentration, and slower clearance in elders compared with young adults. Older adults experience at least equivalent levels of pain relief with half the dose of intramuscular morphine sulfate as younger adults, and the relief lasts longer. It is not clear if this is entirely due to age-related changes in pharmacokinetics or if there are age related changes in pharmacodynamics at work here, as well. In any event, morphine should be administered at lower doses and longer intervals in older adults, at least initially, to avoid toxicity.[77]

Atypical Presentations of Common Diseases

There are a number of medical problems that can occur in any age, but these conditions may present differently in older adults. Differences in presentation or biological responsiveness to common conditions

will typically be most evident to physicians or nursing staff who interact one-on-one with patients; however, there are some unusual presentations that can be represented by the laboratory values. These are outlined below:

- Apathetic hyperthyroidism is hyperthyroidism with a paradoxical symptom of fatigue/apathy, instead of typical anxiety and hyperactivity. Thyroid disease in older adults traditionally appears in the typical manner. Apathetic hyperthyroidism, a rare form of Graves' disease, presents in an unusual way in elderly patients. Unlike younger patients with Graves', older patients generally present with low to normal TSH. Before the advent of ultrasensitive TSH testing, the moderate variation in TSH made diagnosis difficult, but today laboratories routinely use ultrasensitive TSH tests, which make diagnosis of these patients more straightforward.[78]
- Acute myocardial infarction (AMI) is the leading cause of morbidity in the United States.[79] The incidence of myocardial infarction (MI) increases with age, and manifestations of MI can be unique. While the majority of younger patients (less than age 65) present with chest pain, the most common alternate AMI symptom in an older adult patient is shortness of breath.[80] Typical age-related respiratory changes, including shallow breathing and decreased respiratory rate, as well as deconditioning, can complicate the diagnosis in some cases.

Geriatric Syndromes

There are a number of common constellations of symptoms in older adults that are considered "geriatric syndromes." This concept emphasizes that certain issues commonly observed in the older adult population typically result from multiple coexisting and interacting problems. Since clinical training outside geriatrics tends to emphasize "Occam's razor"—looking for the simplest explanation for a patient's problems—the approach to geriatric syndromes requires a change in the diagnostic logical process.[81] Multiple parallel lines of investigation, rather than a single linear approach, are often required for adequate evaluation of these syndromes. The American Geriatric Society considers the conditions in **Table 31.6** to be geriatric syndromes. The identification of a geriatric syndrome cannot be made with a specific laboratory test, but some tests can help explore the underlying cause of the symptoms that lead to patient presentation.

Table 31.6 Geriatric Syndromes

- Difficulty swallowing
- Malnutrition
- Sleep problems
- Bladder control problems
- Delirium
- Dementia
- Vision problems
- Hearing problems
- Dizziness
- Fainting
- Difficulty walking
- Falls
- Osteoporosis
- Pressure ulcers

Data from Inouye SK, Studenski S, Tinetti ME, et al. Geriatric Syndromes: Clinical, Research, and Policy Implications of a Core Geriatric Concept. *Journal of the American Geriatrics Society.* April 11, 2007. https://agsjournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1532-5415.2007.01156.x.

The Impact of Exercise and Nutrition on Chemistry Results in the Geriatric Population

Research indicates that regular exercise can increase the life span and improve the quality of life among geriatric patients.[82,83] Maintaining physical activity improves physical strength and fitness, as well as balance. Exercise helps prevent depression and manage and defend against diseases such as diabetes, heart disease, breast and colon cancer, and osteoporosis.[84,85] Physical activities also improves endurance, flexibility, sleep, mood, and self-esteem.[86] Therefore, increased physical activity may prevent a more rapid physical and mental deterioration. Fitness-related activities can reduce the risk of falling, and incorporation of weight-bearing and resistance exercise into lifestyle has been shown to be one of the best ways to improve bone mass throughout the life.[84] There have also been data linking the positive effects of exercise and long-term maintenance of cognitive function.[87] More importantly, lack of physical activity was shown to be a better predictor of all-cause mortality than being overweight or obese.[86]

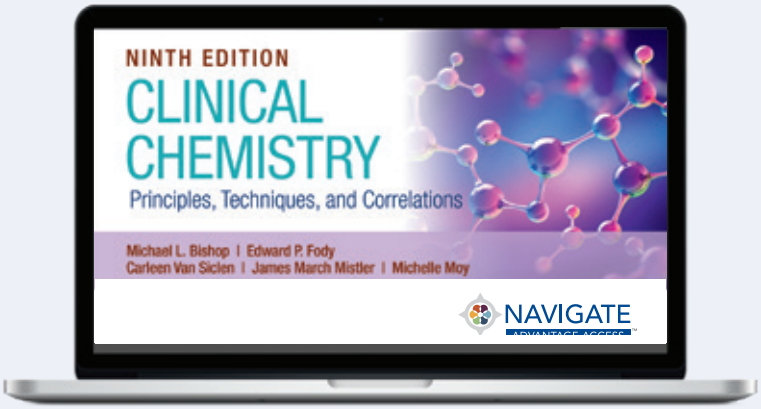
The successful management of aging requires proper nutrition in addition to regular exercise. It is reported that 10% to 40% of hospitalized older adults suffer from malnutrition, which can be defined as any disorder of nutrition status resulting from a deficiency of nutrient intake, impaired metabolism,

or overnutrition.[77] The difficulty recognizing the problem makes treatment a challenge. Some of the laboratory values that can indicate undernutrition are low albumin and low prealbumin. There are a number of issues with the identification of malnutrition in older adults. Based on this, there is increasing interest in establishing nutritional assessment tools for identification of undernourished individuals.[86]

Clearly, geriatric dietary intake is a concern in the area of undernourishment, but excessive calorie intake, resulting in obesity and increased prevalence of type 2 diabetes, is also a worry.[37] Based on increasing incidence of diabetes with aging, and the previously discussed benefits of physical activity, there is increasing encouragement for elderly people to become or continue to be active. Clinicians and laboratorians will need a better understanding of how exercise is likely to impact test results of older individuals. This will require closer attention to baseline measurements and observations of changes over time.

WRAP-UP

To support your learning, review the chapter learning objectives and complete the online activities. The Navigate 2 Advantage Access included with each new print copy of this book offers a wealth of resources. These include practical learning activities and study tools such as flashcards, math practice, an eBook with interactive questions, and more!



References

1. Administration for Community Living and the Administration on Aging. *Profile of Older Americans: 2017 Profile*. Washington, DC: U.S. Department of Health and Human Services; April 2018. Accessed September 24, 2021 at https://acl.gov/aging-and-disability-in-america/data-and-research/profile-older-americans
2. Whamsey L. American life expectancy dropped by a full year in 1st half of 2020. Accessed February 18, 2021 at https://www.npr.org/2021/02/18/968791431/american-life-expectancy-dropped-by-a-full-year-in-the-first-half-of-2020
3. Rogers L, Wilder K. Working-age population not keeping pace with growth in older Americans. United States Census Bureau. Accessed September 24, 2021 at https://www.census.gov/library/stories/2020/06/working-age-population-not-keeping-pace-with-growth-in-older-americans.html
4. Pew Research Center. Aging in the U.S. and other countries, 2010–2050. *Attitudes about Aging: A Global Perspective*. Accessed September 24, 2021 at https://www.pewresearch.org/global/2014/01/30/chapter-2-aging-in-the-u-s-and-other-countries-2010-to-2050/
5. West LA, Cole S, Goodkind, He W. *Current Population Reports: 65+ in the United States: 2010*. Washington, DC: U.S. Government Printing Office; 2014. https://www.census.gov/content/dam/Census/library/publications/2014/demo/p23-212.pdf
6. Migration Policy Institute. *Age-Sex Pyramids of U.S. Immigrant and Native-Born Populations, 1970–Present*. Washington, DC: Author. Accessed September 24, 2021 at https://www.migrationpolicy.org/programs/data-hub/charts/age-sex-pyramids-immigrant-and-native-born-population-over-time
7. COVID-19 Deaths Reported in the U.S. as of September 22, 2021, by age. Published by John Elflein, Jan 11, 2020. Source: https://www.statista.com/statistics/1191568/reported-deaths-from-covid-by-age-us/
8. Gerber A. [Navigating in evidence-poor waters — how much hydration is needed at the end of life? The recommendations of a Swiss expert group (Bigorio Group)]. *Ther Umsch*. 2012;69:91–92.
9. Faulkner WR. *Geriatric Clinical Chemistry Reference Values*. Washington, DC: AACC Press; 1997.
10. Ward KT, Reuben DB. Comprehensive geriatric assessment. *UpToDate*. Schmader KE, Givens J, eds. *UpToDate*. Waltham, MA: UpToDate Inc. https://www.uptodate.com. Last updated July 21, 2020.
11. Ryden I, Lind L, Larsson A. Reference values of thirty-one frequently used laboratory markers for 75-year-old males and females. *Ups J Med Sci*. 2012 Aug;117(3):264–272.
12. Lopez-Contreras MJ, Torralba C, Zamora S, Perez-Liamas F. Nutrition and prevalence of undernutrition assessed by different diagnostic criteria in nursing homes for elderly

people. *J Human Nutr Dietetics.* 2012;25(3):239–246. doi:10.1111/j.1365-277X.2012.01237.
13. Vásárhelyi B, Debreczeni LA. Lab test findings in the elderly. *J Int Fed Clin Chem Lab Med*. 2017;28(4):328–332.
14. National Institutes of Health National Institute on Aging. *Supporting Older Patients with Chronic Conditions*. Accessed February 24, 2021 at https://www.nia.nih.gov/health/supporting-older-patients-chronic-conditions
15. National Council on Aging. *Annual Report, Fiscal Year 2017*. Accessed February 12, 2021 at https://www.ncoa.org/wp-content/uploads/FY17-NCOA-Annual-Report.pdf.
16. Dagli RJ, Sharma A. Polypharmacy: A global risk factor for elderly people. *J Int Oral Health*. 2014;6(6):i–ii.
17. Tjia J, Field TS, Fischer SH, et al. Quality measurement of medication monitoring in the 'meaningful use' era. *Am J Manag Care*. 2011;17:633–637.
18. Garatachea N, Pareja-Galeano H, Sanchis-Gomar F, et al. Exercise attenuates the major hallmarks of aging. *Rejuv Res* 2015;18(1):59–89.
19. Sellami M, Gasmi M, Denham J, et al. Effects of acute and chronic exercise on immunological parameters in the elderly aged: can physical activity counteract the effects of aging? *Front Immunol.* 2018;9:2187. doi:10.3389/fimmu.2018.02187
20. Kraft E. Cognitive function, physical activity, and aging: possible biological links and implications for multimodal interventions. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*. 2012;19:248–263.
21. Fransen HP, May AM, Beulens JWJ, et al. Association between lifestyle factors and quality adjusted life years in the EPIC-NL cohort. *PLoS One*. 2014;9(11):e111480. doi:10.1371/journal.pone.o111480.
22. Baldwin C, Kimber KL, Gibbs M, Weekes CE. Supportive interventions for enhancing dietary intake in malnourished on nutritionally at-risk adults. *Cochrane Database Syst Rev*. 2016 Dec;2016(12):CD009840.
23. Hoenig MP, Zeidel ML. Homeostasis, the milieu interieur, and the wisdom of the nephron. *Clin J Am Soc Nephrol.* 2014;9(7);1272–1281.
24. Cosman F, de Beur SJ, LeBoff MS, et al. Clinician's guide to prevention and treatment of osteoporosis. *Osteoporos Int*. 2014;25(10):2359–2381.
25. Farr JN, Khosla S, MiyabaraY, et al. Effects of estrogen with micronized progesterone on cortical and trabecular bone mass and microstructure in recently postmenopausal women. *J Clin Endocrinol Metab.* 2013;98(2):E249–E257.
26. Song W, Wang H, Wu Q. Atrial natriuretic peptide in cardiovascular biology and disease (NPPA). *Gene* 2015;569(1):1–6.
27. Bhutto A, Morley JE. Understanding the gastrointestinal tract of the elderly to develop dietary solutions that prevent malnutrition. *Oncotarget* 2015;6(16):13858–13898.
28. Chen J, Wang Q, Wu W, Zhang SJ. Comparison of prognostic values of high sensitivity cardiac troponin T and N-terminal brain natriuretic peptide to assess mortality in elderly inpatients. *Clin Interv Aging*. 2019;14:81–90.
29. Matthews SJ, Lancaster JW. Urinary tract infections in the elderly population. *Am J Geriatr Pharmacother* 2011;9:286–309.
30. Azar A, Ballas ZK. Immune function in older adults. *UpToDate*. Waltham, MA: UpToDate Inc. https://www.uptodate.com. Last updated July 29, 2020.
31. Goronzy J, Weyand C. Immune aging and autoimmunity. *Cell Mol Life Sci*. 2012;69(10):1615–1623.
32. Lal H, Cunningham AL, Godeaux O, et al., for the ZOE-50 Study Group. Efficacy of an adjuvanted herpes zoster subunit vaccine in older adults. *N Engl J Med.* 2015;372(22):2087–2096.
33. Rockwood MRH, Howlett SE, Rockwood K. Orthostatic hypotension (OH) and mortality in relation to age, blood pressure and frailty. *Arch Gerontol Geriatr.* 2012;54(3):e255–e260. doi:10.1016/j.archger.2011.12.009
34. Bremner AP, Feddema P, Leedman PJ, et al. Age-related changes in thyroid function: a longitudinal study of a community-based cohort. *J Clin Endocrinol Metab.* 2012;97(5):1554–1562 doi:10.1210/jc.2011–3020
35. Shaw ND, Srouji SS, Histed SN, Hall JE. Differential effects of aging on estrogen negative and positive feedback. *Am J Physiol Endocrinol Metab*. 2011;301:E351–E355.
36. Van den Beld AW, Kaufman J, Killikens MC, et al. The physiology of endocrine system with ageing. *Lancet Diabetes Endocrinol.* 2018;6(8):647–658.
37. Munshi MN, Meneilly GS, et al. Diabetes in aging: pathways for developing the evidence base for clinical guidance. *Lancet Diabetes Endocrinol.* 2020;8(10):P855–867.
38. Federal Interagency Forum on Aging-Related Statistics. Older Americans 2012: Key Indicators of Well-Being. Washington, DC: U.S. Government Printing Office; 2012. https://agingstats.gov/docs/PastReports/2012/OA2012.pdf
39. Barron AM, Pike CJ. Sex hormones, aging, and Alzheimer's disease. *Front Biosci.* 2012;4:976–997.
40. Cole LA. HCG variants, the growth factors which drive human malignancies. *Am J Cancer Res.* 2012;2:22–35.
41. Lee J, Vasikaran S. Current recommendations for laboratory testing and use of bone turnover markers in management of osteoporosis. *Ann Lab Med* 2012;32:105–112.
42. McDade TW, Lindau ST, Wroblewski K. Predictors of C-reactive protein in the National Social Life, Health, and Aging Project. *J Gerontol B Psychol Sci Socl Sci.* 2011;66B(1):129–136.
43. Luan Y, Yao Y-M. The clinical significance and potential role of C-reactive protein in chronic inflammatory and neurodegenerative diseases. *Front Immunol.* 9, June 07, 2018. doi:10.3389/fimmu.2018.01302
44. Bowers LK. Failing health top reason for move to assisted living: survey. McKnight's Senior Living. October 10, 2018. Accessed February 22, 2021 at https://www.mcknightsseniorliving.com/home/news/failing-health-top-reason-for-move-to-assisted-living-survey/
45. Amin K, Cox C, Rice C, Dingel H. COVID-19 now leading cause of death in the United States. Kaiser Family Foundation. Accessed February 22, 2021 at https://www.kff.org/coronavirus-covid-19/slide/covid-19-now-leading-cause-of-death-in-the-united-states/
46. U.S. Preventive Services Task Force. Screening for Prostate Cancer: U.S. Preventive Services Task Force Recommendation Statement. *JAMA.* 2018;319(18):1901–1913. doi:10.1001/jama.2018.3710
47. Fraser WD. Hyperparathyroidism. *Lancet*. 2009;374(9684):145–158. doi:10.1016/S0140-6736(09)60507-9
48. Park SY, Kim HI, Oh HK, et al. Age- and gender-specific reference intervals of TSH and free T4 in an iodine-replete area: data from Korean National Health and

Nutrition Examination Survey IV (2013–2015). *PLoS One*. 2018;13(2):e0190738. doi:10.1371/journal.pone.0190738
49. Chen X, Zheng X, Ding Z, et al. Relationship of gender and age on thyroid hormone parameters in a large Chinese population. *Arch Endocrinol Metab*. 2020;64(1):52–58. doi:10.20945/2359-3997000000179
50. Food and Drug Administration. *UPDATE: The FDA warns that biotin may interfere with lab tests*. Updated November 5, 2019. Accessed March 20, 2021. https://www.fda.gov/medical-devices/safety-communications/update-fda-warns-biotin-may-interfere-lab-tests-fda-safety-communication
51. McConnell DS, Stanczyk FZ, Sowers MR, et al. Menopausal transition state specific changes in circulating adrenal androgens. *Menopause* 2012;19(6):658–663. doi:10.1097/gme.0b013e31823fe274.
52. Santoro N, Randolph JF Jr. Reproductive hormones and the menopause transition. *Obstet Gynecol Clin North Am*. 2011 Sep; 38(3):455–466. doi:10.1016/j.ogc.2011.05.004.
53. Hajj AE, Wardy, N. Haidar, S, et al. Menopausal symptoms, physical activity level and quality of life on women living in the Mediterranean region. *PLoS One*. 2020. doi:10.1371/journal.pone.0230515.
54. Kuhadiya, ND, Karunakaran A, Garg M, Katkar R, Mason, P. A case report of elevated hCG levels in menopause-a clinical dilemma. *Endocrinol Metab Int J*. 2017;4(2):46–48.
55. Selvackadunco S, Langford K, Shah Z, et al. Comparison of clinical and neuropathological diagnoses of neurodegenerative diseases in two centres from the Brains for Dementia Research (BDR) cohort. *J Neural Transm (Vienna)*. 2019;126(3):327–337. doi:10.1007/s00702-018-01967-w
56. Alvarez-Linera Prado J, Jimenez-Huete A. Neuroimaging in dementia. Clinical-radiological correlation. *Radiologia*. 2019 Jan–Feb;61(1):66–81.doi:10.1016/j.rx.2018.08.004
57. Ritchie C, Smailagic N, Noel-Storr AH, et al. CSF tau and the CSF tau/ABeta ratio for the diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev.* 2017;3:CD010803. doi:10.1002/14651858.CD010803.pub2
58. Gale SA, Acar D, Daffner KR, Dementia. *Am J Med*. 2018;131(10):1161–1169. doi:10.1016/j.amjmed.2018.01.022.
59. Alzheimer's Association. *Alzheimer's Disease Facts and Figures*. Accessed September 24, 2021 at https://www.alz.org/alzheimers-dementia/facts-figures
60. de la Torre JC. A turning point for Alzheimer's disease? *BioFactors*. 2012;38(2):78–83. doi:10.1002/biof.200
61. O'Brien JT, Thomas A. Vascular dementia. *Lancet*. 2015;386(10004):1698–1706. doi:10.1016/s0140-6736(15)00463-8.
62. Zupancic M, Mahajan A, Handa, K. Dementia with Lewy bodies: diagnosis and management for primary care providers. *Prim Care Companion CNS Disord* 13, 2011;.
63. Olszewska DA, Lonergan R, Fallon EM, Lynch T. Genetics of frontotemporal dementia. *Curr Neurol Neurosci Rep* 16(12):107. December 2016. doi:10.1007/s11910-016-0707-9.
64. van der Heide F. Acquired causes of Intestinal malabsorption. *Best Pract Res Clin Gastroenterol*. 2016 Apr;30(2):213–224. doi:10.1010/j.bpg.2016.03.001.
65. Ruscin JM, Linnebur SA. Pharmacokinetics in Older Adults. *Merck Manual Professional Version*. Last full revision December 2018. https://www.merckmanuals.com/professional/geriatrics/drug-therapy-in-older-adults/pharmacokinetics-in-older-adults
66. Ruiz JG, Array S, Lowenthal DT. Therapeutic drug monitoring in the elderly. *Am J Ther.* 1996;3:839–860.
67. Lavan AH, Gallagher PF, O'Mahony D. Methods to reduce prescribing errors in elderly patients with multimorbidity. *Clin Interv Aging* 2016;11:857–866. doi:10.2147/CIA.S80280.
68. Dasgupta A. Monitoring free drug concentration: Clinical usefulness and analytical challenges. In Clarke W, Dasgupta A, eds. *Clinical Challenges in Therapeutic Drug Monitoring*. Cambridge, MA: Elsevier; 2016: 71–100.
69. Almazroo OA, Miah MK, Venkataramanan R. Drug metabolism in the liver. *Clin Liver Dis*. 2017 Feb; 21(1):1–20. doi:10.1016/j.cld.2016.08.001. Epub 2016 Oct 15.
70. Wiggins JE. Aging in the glomerulus. *J Gerontol.* 2012;67(12):1358–1364. https://doi.org/10.1093/gerona/gls157.
71. Centers for Disease Control and Prevention. Chronic Kidney Disease in the United States, 2021. Chronic Kidney Disease Initiative. Accessed September 24, 2021 at https://www.cdc.gov/kidneydisease/publications-resources/ckd-national-facts.html
72. Rule AD, Glassock RJ. The aging kidney. *UpToDate*. Waltham, MA: UpToDate Inc. https://www.uptodate.com. Last updated Jun 22, 2020.
73. Portilla Franco ME, Tornero MF, Gil Gregorio P. Frailty in elderly people with chronic kidney disease. *Nefrologia* 2016;36(6):609–615. doi:10.1016/jnefro.2016.03.020.
74. Levey AS, Bosch JP, Lewis JB, et al. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of diet in renal disease study group. *Ann Intern Med.* 1999;130:61–470.
75. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976;16:31–41.
76. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. *Ann Intern Med.* 2009;150:604–612.
77. Karsch-Volk M, Schmid E, Wagenpfeil S, et al. Kidney function and clinical recommendations of drug dose adjustment in geriatric patients. *BMC Geriatr.* 2013;13;92.
78. Hebert L, Cody R, Slivka A, Sedmak D. Hypertension-induced kidney, heart, and central nervous system disease. In: Mandal AK, Jennette JC, eds. *Diagnosis and Management of Renal Disease and Hypertension*. 2nd ed. Durham NC: Carolina Medical Press; 1994.
79. Oosten AW, Abrantes, JA, Jonsson, S, et al. A prospective population pharmacokinetic study on morphine metabolism in cancer patients. *Clin Pharmacokinet.* 2017;56(7):733–746. doi:10.1007/s40262-016-0471-7
80. Boelaert K. Thyroid dysfunction in the elderly. *Nat Rev Endocrinol*. 2013 Apr;9(4):194–204. doi:10.1038/nrendo.2013.30.
81. Smolina K, Wright FL, Rayner M, Goldacre MJ. Determinants of the decline in mortality from acute myocardial infarction in England between 2002 and 2010; linked national database study. *BMJ* 2012;344:d8059.
82. Vagetti GC, Barbosa Filho VC, Moreira NB, et al. Association between physical activity and quality of life in the elderly: a systematic review, 2000–2012. *Psychiatry*. 2014 Jan-Mar; 36(1):76–88. doi:10.1590/1516-4446-2012-0895.
83. Falck RS, Davis JC, Best JR, Crockett RA, Liu-Ambrose T. Impact of exercise training on physical and cognitive function among older adults: a systematic review and meta-analysis.

Neurobiol Aging. 2019 Jul;79:119–130. doi:10.1016/j.neurobiolaging.2019.03.007. Epub 2019 Mar 26. PMID: 31051329.

84. Shah N, Wang C, Lee V, et al. Myocardial infarction in young versus older adults: an analysis of differences in proportion, risk factors, clinical demographics, angiographic findings and in-hospital outcomes. *Int J Clin Cardiol*. 2016;3:079. doi:10.23937/2378-2951/1410079
85. Drachman DA. Occam's razor, geriatric syndromes, and the dizzy patient. *Ann Intern Med*. 2000;132:403–404.
86. de Labra C, Guimaraes-Pinheiro C, Maseda A, Lorenzo T, Millán-Calenti JC. Effects of physical exercise interventions in frail older adults: a systematic review of randomized controlled trials. *BMC Geriatr*. 2015 Dec 2;15:154. doi: 10.1186/s12877-015-0155-5
87. Wiskemann J, Hummler S, Diepold C, et al. POSITIVE study: physical exercise program in non-operable lung cancer patients undergoing palliative treatment. *BMC Cancer*. 2016 Jul 19; 16:499. doi:10.1186/s12885-016-2561-1.

Appendix

Reference Ranges

Analyte	Sample	Patient	Reference Range	Chapter
17-Ketosteroids	U	Females	5–15 mg/24 h	16
	U	Males	10–20 mg/24 h	16
5-NT	S/P		3–9 U/L (37°C)	8
α-Fetoprotein	S/P		< 10 ng/mL	28
Acetaminophen	S	Toxic	> 150 µg/mL	26
	S	Therapeutic	10–30 µg/mL	26
ACP Prostatic	S/P	Males	0–3.5 ng/mL	8
ACP Total	S/P	Children	3.5–9.0 U/L (37°C)	8
	S/P	Adults	1.5–4.5 U/L (37°C)	8
ACTH	S/P	8:00 AM	7.2–63 pg/mL	16
Alanine Transaminase (ALT)	S/P	Females	7–45 U/L	8
	S/P	Males	7–55 U/L	8
Albumin	S/P	Adults	3.5–5.5 g/dL (35–55 g/L)	6
	CSF		0–35 mg/dL	6
Albumin-Globulin Ratio (A/G)	S/P	Adults	1.1–1.8	6
Alkaline Phosphatase (ALP)	S/P	Children 4–15 y	54–369 U/L	8
	S/P	Males 20–50 y	53–128 U/L	8
	S/P	Males ≥ 60 y	56–119 U/L	8
	S/P	Females 20–50 y	42–98 U/L	8
	S/P	Females ≥ 60 y	53–141 U/L	8
Aluminum	S		0–6 ng/mL	27
Amikacin	S/P		Trough < 8.0 µg/mL Peak 20.0–35.0 µg/mL	25

Analyte	Sample	Patient	Reference Range	Chapter
Ammonia	P	Adults	19–60 µg/dL (11–35 µmol/L)	7
	U, 24-h	Adults	140–1500 mg N/d (10–107 mmol N/d)	7
	S/P	Children	140–1500 mg N/d (10–107 mmol N/d)	7
Amylase	U		6.5–48.1 U/h	8
	S/P	Adults	30–220 U/L	8
	BF			
Amylase Clearance/ Creatinine Clearance Ratio	U/S		< 3.1%	22
Anion Gap		AG = Na^+ – (Cl^- + HCO_3^-)	7–16 mmol/L	11
		AG = (Na^+ + K^+) – (Cl^- – HCO_3^-)	10–20 mmol/L	11
Arsenic	WB		< 13 ng/mL	27
Aspartate Transaminase (AST)	S/P	Males	5–35 U/L	8
	S/P	Females	12–38 U/L	24
Bicarbonate	S/P	Males	22–29 mmol/L	24
	S/P	Females	22–29 mm ol/L	24
Bilirubin, Direct	S/P	Adults	0.0–0.2 mg/dL	19
Bilirubin, Indirect	S/P	Adults	0.2–0.8 mg/dL	19
Bilirubin, Total	S/P	Adults	0.2–1.0 mg/dL	19
	S/P	Premature infants at 24 h	1–6 mg/dL	19
	S/P	Premature infants at 48 h	6–8 mg/dL	19
	S/P	Premature infants at 3–5 d	10–12 mg/dL	19
	S/P	Full-term infants at 24 h	2–6 mg/dL	19
	S/P	Full-term infants at 48 h	6–7 mg/dL	19
	S/P	Full-term infants at 3–5 d	4–6 mg/dL	19
BNP	S/P	Females	≤ 150 pg/mL	20
	S/P	Males	≤ 100 pg/mL	20
Cadmium	Blood		< 5.0 ng/mL	27
Calcium, Ionized	S	Adults	4.6–5.3 mg/dL 1.15–1.33 mmol/L	11
	WB	Adults	4.6–5.1 mg/dL 1.15–1.27 mmol/L	11

Analyte	Sample	Patient	Reference Range	Chapter
Calcium, Total	S	Children	8.5–10.5 mg/dL 2.13–2.63 mmol/L	11
	S	Adults	9.0–10.1 mg/dL 2.24–2.53 mmol/L	11
Cancer Antigen 15-3	S	Females	< 30 U/mL	28
Carbamazepine, Free	S	Therapeutic	1.0–3.0 µg/mL	25
Carbamazepine, Total	S	Therapeutic	4.0–12.0 µg/mL	25
Chloride	S/P	Adults	98–107 mmol/L	11
	U (24 h)	24 h	35–285 mmol/24 h	11
Cholesterol	S/P	Adults	140–200 mg/dL (3.6–5.2 mmol/L)	10
	S, B, U	Adults	< 1.0 ng/mL	27
Clozapine	S	Therapeutic	350–420 ng/mL	25
Copper	S, U	Adults	75–145 µg/dL	27
Cortisol	S	AM	7–25 µg/dL	16
	S	PM	2–14 µg/dL	16
Cortisol, Free	U	Adults	3.5–45 µg/24 h	16
Creatine Kinase-MB	S/P	Adults	< 5 ng/mL	20
Creatine Kinase, Total (CK)	S/P	Males	46–171 U/L	8
	S/P	Females	34–145 U/L	8
Creatinine	S/P/U	Adults	0.6–1.2 mg/dL	24
CSF IgG Index	CSF/SERUM		< 0.73	23
CSF albumin/Serum albumin	CSF/SERUM		< 9	23
Cyclosporine	WB	Therapeutic	100–400 ng/mL	25
Digoxin	S	Therapeutic	0.8–2.0 ng/mL	25
Disopyramide	S	Therapeutic	3.0–7.5 µg/mL	25
Ethosuximide	S	Therapeutic	40–100 µg/mL	25
Everolimus	WB	Therapeutic	3–8 ng/mL	25
Fecal Elastase	Stool		> 200 µg/g	22
Fecal Fat	Stool		1–7 g/24 h	22
Felbamate	S	Therapeutic	25.0–60.0 µg/mL	25
Ferritin	S	Females	11–307 µg/L	27
	S	Males	24–336 µg/L	27

Analyte	Sample	Patient	Reference Range	Chapter
FSH	S	Females	1.0–10.5 mIU/mL	17
G-6-PD	WB	Adults	7.9–16.3 U/g Hgb	8
Gabapentin	S	Therapeutic	12.0–20.0 µg/mL	25
Gentamicin	S	Therapeutic	Trough < 2.0 µg/mL Peak 3.0–12.0 µg/mL	25
GGT	S/P	Males	6–55 U/L	8
	S/P	Females	5–38 U/L	8
Glucose	CSF		60%–70% that of the blood glucose	23
	S, P, WB, U	Fasting	70–100 mg/dL	9
	Synovial Fluid		< 10 mg/dL	23
HbA_1c	WB	Adults	4.0–5.7%	9
HCO_3^-	WB		22–26 mmol/L	12
HDL-C	S/P	Adults	40–75 mg/dL	
hs-CRP	S/P	Adults	≤ 0.8 mg/dL	20
Human Chorionic Gonadotropin (hCG)	S/P	Females	< 5 U/L	28
Human Epididymis Protein 4 (HE4)	S/P	Females	< 140 pmol/L	28
IgG Index	CSF		0.26–0.70	6
Iron	S	Females	35–145 µg/dL	27
	S	Males	50–150 µg/dL	27
Lactate	P, WB	Adults	0.3–2.0 mmol/L 2.7–18 mg/dL)	11
	Synovial Fluid		< 25 mg/dL	23
Lamotrigine	S	Therapeutic	2.5–15.0 µg/mL	25
LD	S/P	Adults	125–220 U/L	8
LDL-C	S/P	Adults	50–130 mg/dL (1.3–3.4 mmol/L)	10
Lead	WB		< 3.5 µg/dL	27
Levetiracetam	S	Therapeutic	12.0–35.0 µg/mL	25
Lithium	S	Therapeutic	0.5–1.2 mmol/L	25
Luteinizing Hormone (LH)	S	Females	Follicular: 1.9-14.6 IU/L Midcycle: 12.2-118.0 IU/L Luteal: 0.7-12.9 IU/L	17
Magnesium	S	Adults	(1.7–2.4 mg/dL 0.66–1.07 mmol/L	11

Analyte	Sample	Patient	Reference Range	Chapter
Manganese	S		< 2.4 ng/mL	27
Mercury	S		< 10 ng/mL	27
Methotrexate	S	Therapeutic	< 1.0 µmol/L (48 h after dose); < 0.1 µmol/L (72 h after dose)	25
Methylmalonic Acid (MMA)	S		≤ 0.40 nmol/mL	27
Microalbumin	U	Random collection	< 30 mg/g Cr	9
	U	Timed collection	< 20 µg/min	9
	U	24-h collection	< 30 mg/24 h	9
Molybdenum	U		0.3–2.0 ng/mL	27
Mycophenolic Acid	S	Therapeutic	1.0–3.5 µg/mL	25
Myoglobin	S/P	Adults	≤ 90 ng/mL	20
N-acetylprocainamide (NAPA)	S	Therapeutic	12.0–18.0 µg/mL	25
Neuron-Specific Enolase (NSE)	S, CSF		< 15 ng/mL	28
O_2Hb	WB		> 95%	12
Olanzapine	S	Therapeutic	20.0–50.0 ng/mL	25
Osmolality	S		275–295 mOsm/kg	11
	U 24 h		300–900 mOsm/kg	11
	Random U		50–1200 mOsm/kg	11
	U/S ratio		1.0–3.0	11
Osmolal Gap			5–10 mOsm/kg	11
Oxcarbazepine	S	Therapeutic	12.0–35.0 µg/mL	25
pCO_2	WB		35–45 mm Hg	12
pH	WB		7.35–7.45	12
Phenobarbital	S	Therapeutic	20.0–40.0 µg/mL	25
Phenytoin, Free	S	Therapeutic	1.0–2.0 µg/mL	25
Phenytoin, Total	S	Therapeutic	10.0–20.0 µg/mL	25
Phosphorus (Inorganic)	S/P	Neonate	4.5–9.0 mg/dL 1.45–2.91 mmol/L	11
	S/P	Children 15 y	4.0–7.0 mg/dL 1.29–2.26 mmol/L	11
	S/P	Adults	2.5–4.5 mg/dL 0.81–1.45 mmol/L	11
	U (24 h)	24 h	(0.4–1.3 g/dL 13–42 mmol/dL	11

Analyte	Sample	Patient	Reference Range	Chapter
*p*O2	WB		85–105 mmol/L	12
Potassium	S/P	Adults	3.5–5.1 mmol/L	11
	U 24 h	24 h	15–105 mmol/24 h	11
Prealbumin	S	Adults	17–34 mg/dL	24
Primidone	S	Therapeutic	5.0–12.0 µg/mL	25
Procainamide	S	Therapeutic	4.0–10.0 µg/mL	25
Prolactin	S	Females	5–25ng/mL	17
Protein, Total	S/P	Adults	6.5–8.3 g/dL (65–83 g/L)	6
Protein/Creatinine Ratio	U, 24-h	Adults	≤ 0.114 mg/mg creatinine	6
Pseudocholinesterase (SChE)	S		4,000 and 12,000 U/L	26
Quinidine	S	Therapeutic	2.0–5.0 µg/mL	25
Salicylate	S	Therapeutic	10–30 mg/dL	26
	S	Toxic	> 50 mg/dL	26
Selenium	WB	Adults	150–241 ng/mL	27
Serum Albumin Ratio	CSF		2.7–7.3	6
Serum Protein Electrophoresis Albumin	S		53–65% of total protein (3.5–5.5 g/dL)	6
Serum Protein Electrophoresis α_1-Globulins	S		2.5–5.0% of total protein (0.1–0.3 g/dL)	6
Serum Protein Electrophoresis α_2-GlobulinS	S		7.0–13.0% of total protein (0.6–1.0 g/dL)	6
Serum Protein Electrophoresis β-Globulins	S		8.0–14.0% of total protein (0.7–1.1 g/dL)	6
Serum Protein Electrophoresis γ Globulins	S		12.0–22.0% of total protein (0.8–1.6 g/dL)	6
Sirolimus	WB	Therapeutic	4–20 ng/mL	25
SO_2	WB		> 95 %	12
Sodium	S/P	Adults	135–145 mmol/L	11
	U	(24 h)	120–240 mmol/d, varies with diet	11
	CSF		136–150 mmol/L	11
T_3, Free	S/P	Adults	2.8–4.4 pg/mL	14

Analyte	Sample	Patient	Reference Range	Chapter
T_3, Total	S/P	Adults	80–200 ng/dL	14
T_4, Free	S/P	Adults	0.9–1.7 ng/dL	14
T_4, Total	S/P	Adults	4.5–11.7 µg/dL	14
Tacrolimus	WB	Therapeutic	5.0–15.0 ng/mL	25
Testosterone	S	Males	9–38 nmol/L	17
Theophylline	S	Therapeutic	10.0–20.0 µg/mL	25
Thyroglobulin Ab	S	Adults	< 4.0 IU/mL	14
Thyroperoxidase Ab	S	Adults	< 9.0 IU/mL	14
Thyrotropin Receptor Ab	S	Adults	≤ 1.75 IU/L	14
Tiagabine	S	Therapeutic	20–100 ng/mL	25
Tobramycin	S	Therapeutic	Trough < 2.0 µg/mL Peak 3.0–12.0 µg/mL	25
Topiramate	S	Therapeutic	5.0–25.0 µg/mL	25
Total CO_2 content	WB		23–27 mmol/L	12
Total Iron Binding Capacity (TIBC)	Serum	Adults	250–400 µg/dL	27
Total Protein	U, 24-h	Adults	< 150 mg/24 h	6
Total Protein, CSF	CSF	0–7 d	40–120 mg/dL	6
	CSF	8 d–1 mo	20–40 mg/dL	6
	CSF	> 1 mo	14–45 mg/dL	6
Triglycerides	S/P	Adults	60–150 mg/dL (0.7–1.7 mmol/L)	10
Troponin	S/P	Males	< 15 ng/L	20
	S/P	Females	< 10 ng/L	20
Urea Nitrogen, Blood (BUN)	S/P	Adults	6–20 mg/dL (2.1–7.1 mmol/L)	7
	U, 24-h	Adults	12–20 g/dL 0.43–0.71 mol urea/d	7
Uric Acid	Synovial Fluid		6–8 mg/dL	23
	S/P	Males	Male 3.5–7.2 mg/dL (20.21–0.43 mmol/L) Female 2.6–6.0 mg/dL (0.16–0.36 7mmol/L)	7
	S/P	Females	27.0–5.5 mg/dL (0.12–0.33 mmol/L)	7
	U	Adults	250–750 mg/dL 1.5–4.4 mmol/d	7

Analyte	Sample	Patient	Reference Range	Chapter
Urobilinogen	U		0.1–1.0 Ehrlich units every 2 h or 0.5–4.0 Ehrlich units/d (0.86–8 mmol/d);	19
	Stool		75–275 Ehrlich units/100 g of fresh feces or 75–400 Ehrlich units/24 h specimen	19
Valproic Acid, Free	S	Therapeutic	5–25 µg/mL	25
Valproic Acid, Total	S	Therapeutic	50–120 µg/mL	25
Vancomycin	S	Therapeutic	Trough 10.0–20.0 µg/mL Peak 20.0–40.0 µg/mL	25
Vitamin B_{12}	S	Adults	180–914 ng/L	27
Zinc	S/P	Age > 11 y	0.66–1.10 µg/mL	27
Zonisamide	S	Therapeutic	10–40 µg/mL	25

CSF, Cerebrospinal Fluid; P, Plasma; S, Serum; SF, Synovial Fluid; U, Urine; WB, Whole Blood

Legend

CLSI provides a guideline for defining, establishing, and verifying reference intervals in a clinical laboratory. Therefore, reference ranges vary by laboratory. The variation in reference ranges is due to variables such as age of the patient, gender of patient, assay method, reagent kit, instrument, and collection tube type. The reference ranges provided here are for general educational purposes only. Laboratory results should be interpreted based on the specific ranges and standard operating procedures of each CLIA laboratory.

Sources

1. Rifai N, Horvath AR, Wittwer CT, eds. *Tietz Textbook of Clinical Chemistry and Molecular Diagnostics*. 6th ed. St. Louis, MO: Elsevier Saunders; 2019.
2. Laposota M. *Laposata's Laboratory Medicine Diagnosis of Disease in the Clinical Laboratory*. 3rd ed. New York, NY: McGraw Hill; 2019.

Critical Values Link for Mayo Clinic Laboratories

Mayo Clinic Laboratories Critical Values/Critical Results List (mayocliniclabs.com)

Glossary

1,25-Dihydroxyvitamin D (calcitriol) Active metabolite of vitamin D; induces active absorption of calcium in the small intestine.

1_{3s} rule A data quality control rule that indicates that one data point cannot exceed three SDs. The presence of a data point beyond 3 SDs would trigger a rejection of the analytic run.

25-hydroxyvitamin D Inactive precursor of 1,25 dihydroxyvitamin D.

5-Dihydrotestosterone (DHT) An endogenous androgen sex steroid and hormone. The enzyme 5α-reductase catalyzes the formation of DHT from testosterone in certain tissues including the prostate gland, seminal vesicles, epididymides, skin, hair follicles, liver, and brain.

A

A/G Ratio The ratio of albumin present in serum in relation to the amount of globulin.

Accuracy How close the measured value is to the true value due to systematic error, which can be either constant or proportional.

Acidemia A condition in which the pH of blood is below the lower limit of the reference range (<7.35), indicating that the hydrogen-ion concentration in the blood is increased.

Activation energy The excess energy needed to form the transition state of a reaction.

Activators Inorganic cofactors, such as metal ions, needed for enzyme activity.

Active transport Use of energy to move ions or substances across cell membranes.

Acute coronary syndrome (ACS) A progression of pathologic conditions involved in ischemic heart disease, including erosion and rupture of coronary artery plaques, activation of platelets, and thrombi. This progression ranges from unstable angina to extensive tissue necrosis in acute myocardial infarction.

Acute kidney injury (AKI) A sudden, sharp decline in renal function as a result of an acute toxic or hypoxic insult to the kidneys.

Adrenocorticotropin hormone (ACTH) A peptide hormone secreted by the anterior pituitary. It stimulates the cortex of the adrenal glands to produce adrenal cortical hormones.

Affinity Attraction or force causing two substances to unite.

Airborne pathogens Any infectious agent transmissible by air, e.g., tuberculosis, virus particle, etc.

Albuminuria The presence of albumin in the urine.

Aldosterone The main mineralocorticoid steroid hormone produced by the zona glomerulosa of the adrenal cortex in the adrenal gland. This hormone controls the sodium-potassium pump, the primary mechanism for sodium reabsorption in the kidney and regulator of the blood sodium and potassium levels.

Alkalemia A condition in which blood pH is greater than the uppermost limit of the reference range (>7.45), indicating that the hydrogen-ion concentration in the blood is decreased.

Amenorrhea Temporary cessation of menstruation in a female who is past menarche but not yet in menopause.

Amines Hormones that are derived directly from amino acids.

Amino acid Simple organic compounds that serve as the building blocks of proteins; contain at least one amine functional group, one carboxyl function group, and a unique R group.

Aminoacidopathies Inborn errors of metabolism that inhibit the body's ability to metabolize specific amino acids.

Ammonia A compound consisting of nitrogen and hydrogen. Formula: NH_3 or H_3N.

Amniocentesis Puncture of the amniotic sac to obtain fluid for analysis.

Amniotic fluid (AF) A fluid in which the fetus is suspended; it provides a cushioning medium for the fetus and serves as a matrix for influx and efflux of constituents.

Amperometry The measurement of amperes. It is the unit of measure for electric current. The reduction of oxygen produces a current that is proportional to the amount of oxygen present in the sample.

Amphoteric A molecule that is both an acid and a base.

Analyte Substance of interest being measured.

Analytic Introduced during the phase of processing and assaying the specimen in the clinical laboratory.

Analytic measurement range (AMR) Also known as linear or dynamic range. Range of analyte concentrations that can be directly measured without dilution, concentration, or other pretreatment.

Analytic sensitivity Ability of a method to detect small quantities or small changes in concentration of an analyte.

Analytic specificity Ability of a method to detect only the analyte it is designed to determine, also known as cross-reactivity.

Androgen Any substance (hormone) stimulating the development of male characteristics (e.g., testosterone).

Angina pectoris Pain and a feeling of constriction around the heart caused by deficiency of oxygen to the heart muscle; may radiate down the arm and into the jaw.

Angiotensin II (AT II) The main effector molecule that causes increases in blood pressure, influences renal tubuli to retain sodium and water, and stimulates aldosterone release from adrenal gland.

Anhydrous Compound that has had water of crystallization removed.

Anion Electrolytes with negative charges.

Anion gap (AG) Difference between unmeasured anions and unmeasured cations. $AG = (Na^+ + K^+) - (Cl^- - CO_2)$.

Anosmia Inability to smell.

Anterior pituitary Also called the *adenohypophysis*; the anterior pituitary is the front-facing lobe of the pituitary, a gland located in a small cavity in the sphenoid bone of the skull called the sella turcica. The anterior pituitary produces tropic hormones that are mediated by negative feedback, which involves interaction of the effector hormones with the hypothalamus, as well as with cells of the anterior pituitary.

Antibody Glycoproteins (immunoglobulins) secreted by plasma cells, which in turn are under the control of many lymphocytes and their cytokines. Antibodies are produced in response to antigens.

Antidiuretic hormone (ADH) *See* Arginine vasopressin (AVP).

Antigen Agents that are recognized as foreign by the immune system. The immune system produces antibodies in response.

Apoenzyme The protein portion of an enzyme.

Arginine vasopressin (AVP) Also known as anti-diuretic hormone (ADH) and vasopressin; AVP is a hormone released from the anterior pituitary gland in response to water volume deficits to increase blood volume.

Arrhythmia Irregular heartbeat or action.

Arterial blood Oxygenated blood.

Arteriosclerosis A narrowing and hardening of the arteries, caused by plaque formation, leading to reduced circulation.

Ascites Excess fluid in the peritoneal cavity; the fluid is called ascitic fluid.

Atherosclerosis A disease in which there is an accumulation of lipid material in the veins and arteries, causing the arteries to narrow and harden.

Atomic absorption spectrophotometry A technique for measuring concentration by detecting the absorption of electromagnetic radiation by atoms (elements) rather than by molecules.

Atomic absorption spectroscopy An analytical procedure for the quantitative determination of elements through the absorption of optical radiation by free atoms in the gas phase.

Atomic emission spectroscopy An analytical procedure for the quantitative determination of elements through the emission of radiation by free atoms in the gas phase.

Atrial natriuretic peptide (ANP) A small peptide secreted by the heart upon atrial stretch and high systemic blood pressure. The acute effects of this potent, short-lived peptide include increased glomerular filtration and increased renal excretion of sodium and water.

Avidity Strength of bond of antigen–antibody complex; attraction.

Azotemia The presence of excess nitrogen-containing compounds in the blood.

B

β_2-microglobulin (β2-M) A polypeptide that is one of the class I major histocompatibility markers on cell surfaces.

Base excess Calculated parameter that indicates an excess of bicarbonate or relative deficit of noncarbonic acid in arterial blood gas samples.

Beer's law Mathematical expression of the relationship between analyte concentration and absorbance of light. Beer's law can be expressed as A = abc, where A is absorbance; a is the absorptivity constant for a particular compound at a given wavelength under specified conditions (such as temperature and pH); b is the length of the light path; and c is the concentration.

Bias Difference between the true value and the measured value.

Bidirectional interface The exchange of information between two computer systems—one of which can read data from the patient sample barcode and also transmit laboratory results back to the laboratory information system in electronic formats.

Bile A fluid produced by the liver and composed of bile acids or salts, bile pigments (primarily bilirubin esters), cholesterol, and other substances extracted from the blood. Total production averages about 3 L/d, although only 1 L is excreted.

Bilirubin The principal pigment in bile; derived from the breakdown of hemoglobin when aged red blood cells are phagocytized by the reticuloendothelial system, primarily in the spleen, liver, and bone marrow.

Bioaccumulation Increased body burden of a toxin or xenobiotics.

Bioavailability The fraction of the administered does of therapeutic drug that reaches its target site of action.

Biohazard Anything harmful or potentially harmful to man, other organisms, or the environment. Examples include blood or blood products and contaminated laboratory waste.

Bisphosphonates Medication that slows down or prevents bone loss.

Blastocyst An embryo in which cellular differentiation has not yet taken place.

Bloodborne pathogens Any infectious agent or pathogen transmissible by means of blood or blood products.

Blood–brain barrier A selective mechanism opposing the passage of most ions and large-molecular-weight compounds from the blood to brain tissue.

Body burden Total amount of toxin or xenobiotic present in the patient's body.

Bone turnover Regulated and coupled process of simultaneous bone formation and bone breakdown.

Buffer Weak acid or base that minimize changes in hydrogen ion concentration.

C

Calcium-sensing receptor (CSR) Respond to rising or falling calcium levels by increasing or decreasing PTH secretion, respectively.

Calibration Process that pairs an analytical signal with a concentration value of an analyte.

Cancer A disease caused by the uncontrolled growth of cells that often forms a solid mass or tumor.

Carbohydrates The major food source and energy supply for the body and are stored primarily as liver and muscle glycogen.

Carcinogens Cancer-causing agents.

Cardiac markers Diagnostic test or analytes used to assess cardiac function.

Cardiovascular disease A type of disease that affects the heart or blood vessels. The risk of certain of these diseases may be increased by smoking, high blood pressure, high cholesterol, unhealthy diet, lack of exercise, and obesity.

Carotenoids Phytochemicals, the chief precursors of vitamin A in humans found in over 600 species of plants and some fungi and algae. The six most common carotenoids in the human body account for 90% of all the carotenoids; α-carotene, β-carotene, β-cryptoxanthin, lutein, lycopene, and zeaxanthin.

Cation Electrolytes with positive charges.

Centrifugal analysis Uses the force generated by centrifugation to transfer and then contain liquids in separate cuvettes for measurement at the perimeter of a spinning rotor.

Centrifugation Process where centrifugal force separates solid matter from a liquid suspension.

Cerebrospinal fluid (CSF) Ultrafiltrate of the plasma that supplies nutrients to the nervous tissue, removes waste, and cushions the brain and spinal cord.

Chemical hygiene plan Procedures and work practices for regulating exposure of laboratory personnel to hazardous chemicals. Often referred to as exposure control plan.

Chemiluminescence The emission of light by molecules in excited states produced by chemical reactions.

Cholecystokinin (CCK) A peptide hormone of the gastrointestinal system responsible for stimulating the digestion of fat and proteins.

Cholesterol An unsaturated steroid alcohol found on the surface of lipid layers. It is an amphipathic lipid synthesized almost exclusively by animals.

Chromatography A group of techniques used to separate complex mixtures on the basis of different physical interactions between the individual compounds and the stationary phase of the system.

Chronic kidney disease (CKD) Any illness in which kidney function remains diminished for a long period of time.

Chronic lymphocytic thyroiditis Chronic autoimmune thyroiditis. Also known as Hashimoto's disease.

Chylomicron A lipoprotein that contains apo B-48 and is the largest and least dense of the lipids. Produced in the intestine from dietary lipids and accounts for the turbidity in postprandial plasma. It is mostly composed of triglycerides.

Cinacalcet Drug that binds to calcium-sensing receptor to control calcium levels in patients with PHPT.

Cirrhosis A term derived from the Greek word that means "yellow" that refers to the irreversible scarring process by which normal liver architecture is transformed into abnormal nodular architecture.

Clinically reportable range (CRR) Range of analyte that a method can quantitatively report, allowing for dilution, concentration, or other pretreatment used to extend AMR.

Closed tube sampling The ability to aspirate the patient specimen by piercing the rubber stopper of the collection tube.

Coefficient of determination (r^2 or R^2) Indicates the proportion of variation explained by one variable to predict another. Ranges from 0 to 1. Where $r^2 = 0.95$, one variable explains 95% of the variation of the other. In the case of regression, it indicates how well the line represents the data.

Coefficient of variation (CV) A method used to compare SDs with different units that reflects the SDs in percentages; CV = SD/mean × 100.

Coenzyme Organic cofactors, such as nicotinamide adenine dinucleotide (NADH), needed for enzyme activity.

Cofactor A nonprotein molecule that may be necessary for enzyme activity.

Colligative property The behavior of solute in a solution.

Compensation The body's attempt to return the pH toward normal whenever an imbalance occurs.

Competitive immunoassay An analytical technique in which the labeled Ag (Ag*) in the reagent competes with Ag in the patient sample for a limited number of antibody binding sites.

Conception The fusion of gametes that give rise to another, separate individual (offspring).

Conductivity Measure of the ability of electricity to pass through a solution.

Congestive heart failure Results from an inability of the heart to pump blood effectively. Also called congestive heart disease.

Conjugated bilirubin Bilirubin diglucuronide; it is water soluble and is secreted from the hepatic cell into the bile canaliculi and then passes along with the rest of the bile into larger bile ducts and eventually into the intestines.

Conjugated or complex protein A protein that consists of amino acids and a nonprotein prosthetic group.

Connectivity Capacity for electronic communication between devices, platforms, and systems. Can be unidirectional or bidirectional.

Constant error Type of systematic error where magnitude is constant and not dependent on the amount of analyte.

Consumables Components of a test that are "consumed" or used during analysis and then discarded such as disposable cuvettes, pipette tips, and reagents.

Continuous flow Refers to liquids (reagents, diluents, and samples) that are pumped through a system of continuous tubing.

Corpus luteum Small body that develops within a ruptured ovarian follicle; secretes progesterone.

Correlation coefficient (*r*) Defines the strength of relationship between two variables. Ranges from −1 (perfect inverse correlation) to +1 (perfect positive correlation). A value of 0 indicates no correlation or a random relationship between variables.

Corrosive chemicals Chemicals injurious to the skin or eyes by direct contact or to the tissues of the respiratory and gastrointestinal tracts if inhaled or ingested. Examples include acids (acetic, sulfuric, nitric, and hydrochloric) and bases (ammonium hydroxide, potassium hydroxide, and sodium hydroxide).

Corticotropin-releasing hormone (CRH) A peptide hormone involved in the stress response. Its main function is the stimulation of the pituitary synthesis of ACTH, as part of the HPA axis.

Cortisol A glucocorticoid synthesized by the adrenal cortex that regulates glucose metabolism.

Countercurrent multiplier system An active process occurring in the loops of Henle in the kidney, which is responsible for the production of concentrated urine in the collecting ducts of the nephrons.

Coupled enzymatic method The determination of a substance or enzyme wherein the product of the first reaction serves as the substrate of the ensuing and often more easily measurable reaction.

C-reactive protein (CRP) A marker of inflammation that was originally used as a marker of infection but currently is also used as a prognostic marker of atherosclerotic processes.

Creatine Primarily located in skeletal muscle; facilitates energy storage and serves as a high-energy source, providing a phosphate group for the regeneration of adenosine triphosphate (ATP) from ADP.

Creatine kinase (CK) An enzyme catalyzing the reversible transfer of phosphate from phosphocreatine to ADP, forming creatine and ATP; of importance in muscle contraction. Certain isozymes are elevated in plasma following myocardial infarctions.

Creatinine A naturally occurring metabolic waste product generated during creatine metabolism.

Creatinine clearance A measurement of creatinine elimination from the blood by the kidney.

Crigler-Najjar syndrome A liver disease characterized by jaundice that result in an elevations in unconjugated bilirubin.

Cryogenic material The materials (gases) used to make things cold.

Cystatin C A cysteine protease inhibitor found in the bloodstream in elevated concentrations in patients with impaired kidney function.

D

D-dimer A covalently cross-linked degradation product released from the cross-linked fibrin polymer during plasmin-mediated fibrinolysis. Laboratory measurements of this product are made using latex bead assay, or enzyme-linked immunosorbent assay can be used to identify the presence of fibrinolysis; helpful in the diagnosis of deep vein thrombosis.

Dehydroepiandrosterone (DHEA) An endogenous steroid hormone precursor; one of the most abundant circulating steroids in humans. It is produced in the adrenal glands, the gonads, and the brain. Also known as androstenolone.

Dehydroepiandrosterone sulfate (DHEA-S) An endogenous androstane steroid that is produced by the adrenal cortex; also known as androstenolone sulfate.

Deionized water Water that has had some or all ions removed.

Delta absorbance The rate of change in absorbance over time.

Delta bilirubin Conjugated bilirubin bound to albumin.

Density Mass per unit volume of a substance.

Descriptive statistics Numerical values that summarize a given data set such as mean, median, mode, and SD.

Desiccant Hygroscopic material that can remove water from the atmosphere and other materials.

Diabetes insipidus A disease of water and salt imbalance.

Diabetes mellitus Chronic metabolic syndrome of impaired carbohydrate, fat, and protein metabolism secondary to insufficiency of insulin secretion or to the inhibition of the activity of insulin.

Diagnostic sensitivity Ability of a test to detect a given disease or condition.

Diagnostic specificity Ability of a test to correctly identify the absence of a given disease or condition.

Diffusion Passive movement of ions or substances without energy use across membranes.

Dilution The amount of concentrated or stock material to the total final volume of a solution.

Disaccharides Two monosaccharide units joined by a glycosidic linkage.

Discrete analysis Refers to analyzers have the capability of running multiple tests from one sample at a time or multiple samples one test at a time.

Dispersion Spread of the data.

Distilled water Water that has been purified to remove most organic material using distillation.

Distribution The movement of a drug between blood circulation and tissues/organs and the relative proportion of the drug in the tissues.

Diurnal variation A fluctuation in the daily blood or urine concentration with peaks and troughs; circadian rhythms. An analogy would be the ocean with highest water levels seen at high tide and the lowest water levels seen at low tide. For example, cortisol is assessed as a first morning specimen at 8:00 AM. Early morning is when the hormone spikes, and the concentration tapers off throughout the day.

Dopamine (DA) A neurotransmitter, also known as prolactin-inhibiting factor, that functions primarily as a local paracrine messenger supporting several important roles in the brain and body. It is an organic chemical of the catecholamine and phenethylamine families. It controls the release of prolactin.

Dose–response relationship Comparison of the dose of a substance (i.e., drug or chemical) with its potential pathologic effects. Dose–response relationship implies that there will be an increase in toxic response with an increased dose.

Drug half-life ($T_{1/2}$) The amount of time (T) required for the blood concentration of a drug to decrease by one-half ($\frac{1}{2}$).

Drug metabolism The physiologic process by which the body breaks down and converts medication into an active chemical substance.

Drugs of abuse Drugs used illegally or inappropriately; many drugs have the potential for abuse.

Dry chemistry slide A methodology that uses dry reagents mounted on a plastic support.

D-xylose An exogenously administered simple (pentose) sugar or monosaccharide. It is not ordinarily present in the blood in significant measurable quantities because it does not require pancreatic lytic enzymes for absorption and can therefore be used to differentiate malabsorption from an intestinal etiology or exocrine pancreatic insufficiency.

Dyslipidemias Disease states associated with abnormal serum lipids usually caused by malfunctions in synthesis, transport, or catabolism of lipoproteins. These diseases may be caused by either elevated or decreased lipoprotein levels.

E

ED_{50} The dose that would be predicted to be effective or have a therapeutic benefit in 50% of the population.

Effector/trophic A hormone that causes a direct cellular response or change in the target organ, like cellular growth. These hormones elicit direct effects on the target cells.

Effusion Abnormal accumulations of pleural or pericardial fluid.

Electrochemistry The study of the relationship of electrical potential and chemical changes.

Electrolytes Ions with an electrical charge.

Electrophoresis The migration of charged solutes or particles in an electrical field.

Embden-Meyerhof pathway Pathway for anaerobic glycolysis.

Embryo An unborn offspring in the process of development from the fourth to the eighth week of gestation.

Emission spectrum Observed in atomic emission spectroscopy, a series of very narrow peaks (also known as "lines") with each line at a different wavelength and matched to a specific transition.

Endocrine A gland that secretes its products called hormones directly into the bloodstream.

Endocrine system A system of ductless glands that produces secretions (hormones) into the blood or lymph to be carried by the circulation to other parts of the body.

Endogenous pathway The lipolytic process for metabolizing VLDL that was produced by the liver. VLDL is converted to LDL which is the main lipoprotein to deliver cholesterol to peripheral cells.

Endpoint An assay in which a measurement is taken at the beginning of the reaction (the background signal) and another measurement is taken at a set time later in the reaction (plateau or endpoint signal).

Enzyme–substrate (ES) complex The first stage of an enzymatic reaction that occurs when the enzyme binds the substrate.

Epinephrine (EPI) A type of catecholamine made in the adrenal glands released in response to physical or emotional stress. Among the many functions it serves in the body are sending nerve impulses in the brain, narrowing blood vessels, and raising the heart rate. Previously called adrenaline.

Epitope Component of an antigen that functions as an antigenic determinant, allowing the attachment of certain antibodies.

Equivalent weight The gram molecular weight of a substance divided by its valence.

Erlenmeyer flasks Wide-bottom flask that tapers to smaller, short neck.

Erythropoietin A cytokine made by the kidneys that stimulates proliferation of red blood cells.

Essential element An element that, if absent or deficient, impairs a biochemical or functional process and replacement of the element corrects this impairment.

Estimated glomerular filtration rate (eGFR) Equation used to predict GFR based on serum creatinine, age, body size, and gender, without the need of a urine creatinine. The Cockcroft-Gault formula is one of the first formulas used to estimate this. This formula predicts creatinine clearance, and the results are not corrected for body surface area. This equation assumes that women will have a 15% lower creatinine clearance than men at the same level of serum creatinine.

$$\text{GFR (mL/min)} = \frac{(140 - \text{Age}) \times \text{Weight (kg)}}{72 \times S_{Cr}\,(\text{mg/dL})} \times (0.85 \text{ if female})$$

Exocrine A gland that secretes its products (enzymes) into another organ usually by way of a duct.

Exogenous pathway The lipolytic process for metabolizing dietary lipids. It begins with chylomicrons synthesized in the intestine and then transported to the liver. The chylomicrons are broken down in the liver and released as VLDL, which is then converted to LDL and transported to peripheral cells.

Exposure control plan Procedures and work practices for regulating exposure of laboratory personnel to hazardous chemicals. Often referred to as chemical hygiene plan.

Extracellular fluid (ECF) Fluid in the body that is outside of the cells; made up of intravascular, interstitial, and extravascular fluid compartments.

Exudate Accumulation of fluid in a cavity; also, the production of pus or serum. In comparison with a transudate, an exudate contains more cells and protein. Exudates demand immediate attention.

F

False negative Patients with a condition who are incorrectly identified by a test as not having the condition.

False positive Patients without a condition who are incorrectly identified by a test as having the condition.

Fatty acids Linear chains of carbon–hydrogen bonds that terminate with a carboxyl group. Most fatty acids in plasma are a constituent of triglycerides or phospholipids.

Fecal elastase An enzyme produced by the exocrine pancreas that can be measured in the feces.

Fetus An unborn offspring in the process of development from the eighth week of gestation until delivery.

Fibroblast growth factor 23 (FGF23) Powerful phosphaturic hormone secreted by osteocytes.

Filtration Separation of solid material from fluid using various mediums.

Fire tetrahedron A three-dimensional pyramid representing the element of fire; previously the fire triangle.

First-order kinetics The rate of an enzymatic reaction that occurs in which the rate is dependent on the concentration of one substrate.

Fisher projection A projection of a carbohydrate with the aldehyde or ketone at the top of the drawing.

Fluorescence Emission of light by a substance after absorption of electromagnetic radiation of a shorter wavelength.

Fluorometry A technique that uses fluorescence to determine concentration of an analyte.

Follicle stimulating hormone (FSH) A peptide hormone secreted by the anterior pituitary, which is responsible for ovarian recruitment and early folliculogenesis in women, and spermatogenesis in men.

Follicular cells One of two types of cells in thyroid, they produce and secrete thyroxine (T_4) and triiodothyronine (T_3).

Follicular phase The first half of the menstrual cycle, when estrogen effect is unopposed by progesterone; also called the proliferative phase.

Free T_3 Unbound triiodothyronine.

Free T_4 Unbound thyroxine.

Friedewald equation An equation for calculating LDL cholesterol. In this equation, estimated VLDL and HDL are subtracted from total cholesterol, which yields an estimated LDL cholesterol. This should not be used when triglycerides are over 400 mg/dL. The Friedwald equation is LDL-C = Total Cholesterol − HDL − (Trig/5).

G

Gas chromatography A chromatographic technique used to separate mixtures of compounds that are volatile or can be made volatile.

Gastrin A peptide hormone that enhances gastric growth, gastric motility, and secretion of hydrochloric acid.

Genetic disease A disorder caused by an abnormality in an individual's DNA.

Geriatrics Branch of general medicine dealing with remedial and preventable clinical problems in older adults, as well as the social consequences of such illnesses.

Gerontology Study of the aging process in the human body.

Gestation Also known as pregnancy; the process of carrying offspring in the uterus until delivery.

Glomerular filtration rate (GFR) A measurement of renal function involving flow rate and filtration capacity over time; the rate of urine formation as plasma passes through the glomeruli of the kidneys.

Glomerulonephritis A form of nephritis in which the lesions involve primarily the glomeruli. It may be acute, sub-acute, or chronic.

Glomerulus One of the capillary networks that are part of the renal corpuscles in the nephrons of the kidney.

Glucagon The primary hormone responsible for increasing glucose levels.

Glucocorticoids Hormones produced in the adrenal cortex that raise blood glucose levels. Glucocorticoids are also known as stress hormones. An example is cortisol, which ensures that the body has enough glucose available for dealing with impending stressors.

Gluconeogenesis Formation of glucose-6-phosphate from noncarbohydrate sources.

Glucose A six-carbon molecule that is the primary energy source for humans.

Glycogen Storage form of glucose.

Glycogenolysis Breakdown of glycogen to glucose for use as energy.

Glycolysis Metabolism of a glucose molecule to pyruvate or lactate for production of energy.

Glycosylated hemoglobin The formation of a hemoglobin compound produced when glucose (a reducing sugar) reacts with the amino group of hemoglobin (a protein).

Gonadotropin-releasing hormone (GnRH) A peptide hormone from the hypothalamus responsible for controlling the release of LH and FSH from the anterior pituitary.

Gout Arthritic condition characterized by the precipitation/crystallization of uric acid and subsequent deposit in joints and tissues.

Graafian follicle A follicle in which the oocyte attains its full size and is surrounded by an extracellular glycoprotein layer (zona pellucida) that separates it from a peripheral layer of follicular cells permeated by one or more fluid-filled antra; the theca of the follicle develops into internal and external layers.

Graduated cylinder Long, cylindrical tube with calibration marks along length.

Graves' disease Hyperthyroidism caused by TSH receptor autoantibody causing unregulated stimulation of thyroid.

Griffin beaker Flat-bottom beaker with straight sides, a wide opening, and a small spout.

Growth The process of increasing in physical size, weight, and development.

Growth hormone (GH) A peptide hormone released from the anterior pituitary gland that controls growth and metabolism; a direct effector hormone released in pulsatile secretions.

Growth hormone-releasing hormone (GHRH) A peptide hormone produced in the hypothalamus.

Gynecomastia Abnormally large mammary gland development in the male.

H

Half-life ($T_{1/2}$) The time needed for the blood concentration of any analyte to decrease by one-half.

Hapten (Hp) A substance that can bind with an antibody but cannot initiate an immune response unless bound to a carrier.

Hashimoto's thyroiditis Chronic autoimmune thyroiditis causing gradual loss of normal thyroid tissue.

Haworth projection A projection that represents the compound in the cyclic form.

Hazard Communication Standard Based on the fact that all employees must be informed of any health risks involving the use of chemicals; from the Hazardous Communication Standard of 1987 (Right to Know Law).

Hazardous materials A material that may potentially cause personal injury or damage if handled.

Heme Flat cyclic tetrapyrrole with a central ferrous iron atom. It is a prosthetic group found in hemoglobin, myoglobin, chlorophyll, cytochromes, and several enzymes.

Hemodialysis The use of an artificial kidney to remove urea, metabolic waste products, toxins, and excess fluid from the blood.

Hemofiltration An ultrafiltration technique to remove excess metabolic products from the blood.

Hemoglobin A_1c (HbA_1c) A glucose molecule attached to one or both N-terminal valines of the β-polypeptide chains of normal adult hemoglobin.

Hemolysis Rupture or destruction of red blood cells.

Henderson-Hasselbalch equation Mathematical description of the dissociation characteristics of weak acids.

Hepatitis Inflammation of the liver caused by a virus, bacteria, parasites, radiation, drugs, chemicals, or toxins. Among the viruses causing hepatitis, the most common are hepatitis types, A, B, C, D and E, cytomegalovirus, and Epstein-Barr, among others.

Hepatocellular carcinoma The most common form of liver cancer in adults and the fourth leading cause of cancer-related deaths worldwide.

Hepatoma Primary malignant tumors of the liver; also known as hepatocellular carcinoma or hepatocarcinoma.

Heterogeneous A type of assay that uses a physical separation step to separate free and bound labeled analyte. The labeled, unbound analyte is separated or washed away, and the remaining labeled, bound analyte is measured.

High-complexity tests Defined by the Clinical Laboratory Improvement Amendments of 1988 (CLIA) as a test requiring expertise beyond normal clinical automation.

The test may require reagent preparation, limited preevaluation, and interpretation of results.

High-density lipoprotein (HDL) The smallest and most dense lipoprotein; it is synthesized by the liver and intestine and removes excess cholesterol from peripheral cells for metabolism by the liver.

High-efficiency particulate air (HEPA) filters A filter that can remove particles as small as 0.3 microns.

High-performance liquid chromatography A chromatographic technique that uses high pressure and controlled temperature for fast separations of molecules in a mixture.

Hirsutism Excessive and abnormal hair growth.

Histograms Graphical representation of data using bars of different heights.

Holoenzyme The complete, active enzyme complex, including the apoenzyme with the coenzyme.

Homocysteine A sulfur-containing amino acid formed in plasma from the metabolic demethylation of methionine, which is derived from dietary protein.

Homogeneous A type of assay that does not require a physical separation of the free and bound labeled analyte.

Homovanillic acid A major catecholamine metabolite of dopamine that is produced by a consecutive action of monoamine oxidase (MAO).

Hook effect A phenomenon in which the measurement of immune complexes is impaired by excess concentrations of the analyte resulting in falsely low or negative results. Also called the prozone effect.

Hormone A chemical substance that sends a message to another cell in the body. This message can either regulate or control the activity of other bodily functions. These cellular messages are sent to other cells either via the bloodstream (endocrine), gastrointestinal tract (exocrine), nervous system (neurocrine), or in interstitial fluid (paracrine).

Hydrate A compound and its associated water.

Hydrolase An enzyme that catalyzes the hydrolysis of various bonds.

Hygroscopic Substance that takes up water when exposed to the atmosphere.

Hyperammonemia Elevated ammonia concentration in the blood.

Hypercalcemia Increased plasma calcium concentration.

Hypercapnia Excessive amounts of carbon dioxide in the blood.

Hypercarbia Abnormally increased arterial carbon dioxide tension.

Hyperchloremia Increased chloride levels.

Hyperglycemia An increase in plasma glucose levels.

Hyperkalemia Increased potassium levels.

Hypermagnesemia Increased magnesium levels.

Hypernatremia Increased sodium levels.

Hyperphosphatemia Increased phosphate levels.

Hyperproteinemia An increase in total plasma proteins.

Hypertension A condition in which the force of the blood against the artery walls is higher than normal; high blood pressure.

Hyperthyroidism Excess thyroid hormone in the circulation creating a clinical state.

Hyperuricemia Elevated levels of uric acid in the blood above the reference range.

Hypervitaminosis A condition that results from excessive intake of a vitamin(s).

Hypocalcemia Decreased plasma calcium concentration.

Hypoglycemia A decrease in plasma glucose levels.

Hypoglycorrhachia Decreased CSF glucose levels.

Hypogonadism Aberrant internal secretion of the gonads.

Hypokalemia Decreased potassium levels.

Hypomagnesemia Decreased magnesium levels.

Hyponatremia Decreased sodium levels.

Hypophosphatemia Decreased phosphate levels.

Hypoproteinemia A decrease in total plasma proteins.

Hypothalamic-hypophyseal portal system A system of blood vessels in that connects the hypothalamus and the pituitary gland.

Hypothalamic–pituitary–thyroid (HPT) axis The neuroendocrine system that regulates the production and secretion of hormones secreted from thyroid follicular cells.

Hypothalamus A region in the brain that controls part of the nervous system and the pituitary gland with both hormones and neurons.

Hypothyroidism Deficiency of thyroid hormone in the circulation, creating a clinical state.

Hypouricemia Lowered or deficient levels of uric acid in the blood.

Hypovitaminosis A condition that results from a lack or deficiency of vitamins in the diet.

Hypovolemia Decreased water (blood) volume.

Hypoxemia Insufficient or decreased oxygenation of the blood.

Hypoxia A deficiency of oxygen reaching the tissues of the body.

I

Icterus Yellow pigmentation in the skin or sclera (including tissues, membranes, and secretions); also known as jaundice. A result of excess bilirubin concentration in the blood.

Immunity Resistance to infection by a pathologic agent.

Inhibin A glycoprotein hormone that inhibits FSH. It is produced in the Sertoli cells in males, and in granulosa cells in females.

Insulin The primary hormone responsible for the entry of glucose into the cell.

Insulin-like growth factor (IGF) Proteins controlled by the release of GH responsible for growth and cellular development.

International unit (IU) The amount of enzyme that will catalyze the reaction of 1 µmol of substrate per minute while also including descriptions of the specified conditions of temperature, pH, substrates, and activators used in the reaction.

Interstitial fluid Fluid that surrounds cells and tissues.

Intracellular fluid (ICF) Fluid (water) inside cells.

Intravascular fluid (IVF) Fluid that makes up plasma.

Ionic strength Concentration or activity of ions in a solution or buffer.

Ion-selective electrodes Electrodes that measure electrical potential produced by the activity of free ions. Each type is designed to be sensitive toward one type of ion.

Ischemia Inadequate blood supply to an organ or tissue.

Islets of Langerhans Groups of cells from the endocrine pancreas that secrete insulin and glucagon.

Isoelectric point (pI) pH at which a substance has a net neutral charge.

Isoenzyme The forms of a particular enzyme that differ genetically.

Isoform The forms of a particular enzyme following post-translational modification.

Isomerases Enzymes that catalyze the interconversion of geometric, optical, or positional isomers.

J

Jaundice Describes the yellow discoloration of the skin, eyes, and mucous membranes most often resulting from the retention of bilirubin.

K

Ketone A functional group with the structure $R_2C{=}O$.

Kinetic A type of assay that measures the rate of complex formation, which is continuously monitored.

Kinetic assays Enzymatic reactions monitored over time, either at specified times or continuously monitored.

Kupffer cells Phagocytic macrophages capable of ingesting bacteria or other foreign material from the blood that flows through the through the sinusoids of the liver.

L

L/S ratio See *Lecithins/sphingomyelins ratio*.

Laboratory standard A rule or criterion related to the laboratory.

Lactose A sugar present in mammalian milk. It is a disaccharide composed of glucose and galactose.

LD_{50} Dose of a drug in which 50% of the individuals will result in morbidity.

Lecithins/sphingomyelins ratio (L/S ratio) A classic test that assesses the ratio of lecithins to sphingomyelins to determine fetal lung maturity.

Levey-Jennings control chart Graphical representation of observed values of a control material over time in the context of the upper and lower control limits.

Leydig cells Cells of the testicles that produce testosterone.

Ligases Enzymes that catalyze the joining of two substrate molecules, coupled with breaking of the pyrophosphate bond in adenosine triphosphate (ATP) or a similar compound.

Limit of detection (LOD) Lowest amount of analyte that can be reliably detected by a method.

Limit of quantitation (LOQ) Lowest amount of analyte that can be reported while achieving a precision target (e.g., lowest concentration at which a CV of 10% may be achieved).

Linear regression A statistical technique that provides an objective measure of the line of best fit for the data using slope, y-intercept, and r^2.

Linearity The relationship between the standard's concentration and the instrument result.

Lipemia Increased lipids.

Lipoprotein Substances composed of lipids and proteins in varying concentrations. The major lipoproteins include chylomicrons, VLDL, LDL, and HDL.

Lipoprotein a (Lp[a]) A low-density lipoprotein (LDL) particle with an added apolipoprotein(a) (apo[a]) attached to the apoplipoprotein(b) (apo[b]) component of the LDL particle via a disulfide bridge.

Loop of Henle The U-shaped portion of a renal tubule lying between the proximal and distal convoluted portions.

Low-density lipoprotein (LDL) It is formed from the lipolysis of VLDL. It is rich in cholesterol and delivers fuel to the peripheral cells. LDL = CHOL − (HDL + TRIG/5).

Luteal phase Phase in menstrual cycle; progesterone is synthesized by the corpus luteum during this phase. Also called *secretory phase*.

Luteinizing hormone (LH) A glycoprotein hormone secreted by the anterior pituitary that stimulates the final ripening of an ovarian follicle, its secretion of progesterone, its rupture to release the egg, and the conversion of the ruptured follicle into the corpus luteum.

Lyases Enzymes that catalyze the removal of groups from substrates without hydrolysis; the product contains double bonds.

M

Mass spectrometry Analytical technique based on measuring the mass-to-charge ratio of the molecular components comprising a sample.

Mass-to-charge ratio Molecular mass divided by ionic charge of selected analyte ions (m/z).

Mean Mathematical average of a data set.

Mechanical hazards Any potential danger from equipment, such as centrifuges, autoclaves, and homogenizers.

Median The middle of a data set after the data have been rank ordered. When there are an even number of values, it is calculated by adding the two middle values and dividing by 2.

Medical decision level Value for an analyte that represents the boundary between different therapeutic approaches.

Medical waste Material that is infectious or physically dangerous; includes discarded blood, tissues, fluids, body parts, sharps, etc.

Menopause The permanent cessation of menstruation.

Metalloenzyme Trace element associated with an enzyme as an essential component or cofactor.

Metalloprotein Trace element associated with a protein as an essential component or cofactor.

Michaelis-Menten constant A constant for a specific enzyme that describes the relationship between the velocity of an enzymatic reaction and substrate concentration.

Middleware Software that connects the laboratory information system and the manufacturer's hardware.

Mineralocorticoids Hormones produced in the adrenal cortex that reabsorb sodium in the distal convoluted tubule of the nephron in order to assist with the regulation of blood pressure. An example is aldosterone.

Mode The most frequently occurring value in a data set.

Moderate-complexity tests Defined by the Clinical Laboratory Improvement Amendments of 1988 (CLIA) as a test requiring basic lab knowledge and training for personnel performing the test. May require reagent preparation, limited pre-treatment of specimens, quality control, calibration, proficiency testing, some skill in troubleshooting and maintaining equipment, and some skill or judgment in interpretation of results.

Modular analyzer An instrument that combines chemistry and immunoassay capabilities in a single platform capable of meeting the needs of mid- and high-volume laboratories.

Molality (m) Amount of solute per 1 kg of solvent.

Molarity (M) Number of moles per 1 L of solution.

Monoamine oxidase (MAO) An enzyme that is involved in removing the neurotransmitters norepinephrine, serotonin, and dopamine from the brain.

Monoclonal Arising from one line of cells.

Monosaccharides Simple sugars that cannot be hydrolyzed to a simpler form.

Myeloperoxidase (MPO) An enzyme stored in azurophilic granules of polymorphonuclear neutrophils and macrophages and released into extracellular fluid in the setting of inflammatory process.

Myocardial infarction An event, commonly called a *heart attack*, that occurs when blood flow to an area of the cardiac muscle is suddenly blocked, leading to ischemia and death of myocardial tissue.

Myoglobin A heme protein found only in skeletal and cardiac muscle in humans. It can reversibly bind oxygen in a manner similar to the hemoglobin molecule, but it is unable to release oxygen except under very low oxygen tension.

N

National Fire Protection Association (NFPA) An organization whose label displays warnings for firefighters of the location of hazardous materials in the event of a fire. It includes a diamond-shaped symbol with four quadrants that indicate the relative danger level in four different areas: health (blue), fire (red), chemical stability (yellow), and other specific hazard types.

Negative acute-phase reactant A substance that decreases in the presence of an acute disease state.

Negative predictive value Chance an individual does not have a given disease or condition if the test is within the reference range.

Neoplasm A new solid mass or tumor produced by uncontrolled cell growth. They may be either benign or malignant.

Nephelometry An analytical technique that measures scattered light by placing a detector at a 90-degree angle from the incident light.

Nephrons Microscopic functional units of the kidney. Each kidney contains approximately 1 million nephrons.

Nephrotic syndrome A condition marked by increased glomerular permeability to proteins, resulting in massive loss of proteins in the urine, edema, hypoalbuminemia, hyperlipidemia, and hypercoagulability.

Neurohypophysis The posterior, or further back portion, of the pituitary gland that releases AVP and oxytocin. Also called *posterior pituitary*.

Noncompetitive immunoassay A type of assay that uses a labeled reagent Ab to detect the Ag. Also known as sandwich assay.

Nonessential element An element that, if absent or deficient, does not impair a biochemical or functional process.

Norepinephrine (NE) A catecholamine precursor to epinephrine that functions as a neurotransmitter. Previously called *noradrenaline*.

Normality Number of gram equivalent weights per 1 L of solution.

O

Occupational Safety and Health Act (OSHA) Occupational Safety and Health Act; enacted by Congress in 1970. The goal of this federal regulation was to provide all employees (clinical laboratory personnel included) with a safe work environment.

Oligoclonal bands CSF protein electrophoresis demonstrating multiple banding of the γ-globulin region representing similar, yet pathology-related, IgG bands. This observation involves a small number of clones of IgG from the same cell type with nearly identical electrophoretic properties. This occurrence is usually associated with inflammatory diseases and MS or SSPE. These types of disorders would stimulate the immunocompetent cells that produce similar bands across the γ-globulin region.

Oligosaccharides The chaining of 2 to 10 sugar units.

Oncofetal antigen A protein produced during fetal development and elevated in individuals with cancer.

Oncogene A gene that causes cancer.

One-point calibration Comparison of a known standard or calibrator concentration and its corresponding absorbance to the absorbance of an unknown value.

Osmolal gap Difference between the measured osmolality and the calculated osmolality (Measured Osmolality - Calculated Osmolality).

Osmolality Colligative property of a solution; how much of a solute is dissolved in a solvent (mmol/kg).

Osmolarity How much of a solute is dissolved in a solvent (mmol/L).

Osmometer Instrument that measures the osmolality of a substance.

Osmometry Technique used to measure the solute concentration of a solution using one of the four colligative properties, which change in proportion to osmotic pressure.

Osmotic pressure The pressure that opposes osmosis as a solvent flows through a semipermeable membrane.

Osteoblast Bone cell that forms new bone.

Osteoclast Bone cell that breaks down bone (bone resorption) to release calcium and phosphate into circulation.

Osteomalacia Abnormal mineralization of bone in adults or after completion of skeletal maturation.

Osteoporosis "Porous bone;" a bone disease that is caused by loss of bone density, usually from either bone malformation or excessive bone resorption.

Otorrhea Discharge from the ear; also, leakage of CSF from the ear.

Ovulation The periodic discharge of an ovum from the ovary.

Oxidized Substance that lost an electron.

Oxidizing agent Substance that accepts an electron in a redox reaction.

Oxidoreductase An enzyme that catalyzes oxidation-reduction reactions between two substrates.

Oxygen saturation The measure of how much hemoglobin is bound to oxygen molecules.

Oxyhemoglobin Oxygen loosely bound to hemoglobin.

Oxytocin A peptide hormone released from the posterior pituitary; responsible for lactation, labor, and childbirth.

P

Pancreatitis Inflammation of the pancreas.

Parathyroid glands Four glands adjacent to the thyroid gland. Two of these glands are found in the upper portion, and two are found near the lower portion of the thyroid gland. These four glands produce parathyroid hormone, which controls calcium and phosphate metabolism.

Parathyroid hormone (PTH) Hormone secreted by the parathyroid glands; plays a role in calcium metabolism.

Parathyroid hormone-related protein (PTHrP) A substance that has structural similarities to the N-terminal portion of human PTH molecule; secreted by cancers.

Parenteral nutrition Intense nutritional support for patients who are malnourished, or in danger of becoming malnourished, because they are unable to consume required nutrients or to take nutrients entirely. This therapy involves administering appropriate amounts of carbohydrate, amino acid, and lipid solutions as well as electrolytes, vitamins, minerals, and trace elements to meet the caloric, protein, and nutrient requirements while maintaining water and electrolyte balance.

Peak drug concentration The maximum level of a blood drug concentration.

Peptide An unbroken chain of amino acids of 50 or less.

Peptide bond A covalent bond formed between the amino group of one amino acid and the carboxyl group of another amino acid.

Percent solution The amount of solute per 100 total units of solution.

Pericardial fluid Fluid surrounding and protecting the heart. The frequency of pericardial sampling and laboratory analysis is rare.

Peritoneal fluid A clear to straw-colored fluid secreted by the cells of the peritoneum (abdominal cavity). It serves to moisten the surfaces of the viscera.

pH Negative or inverse log of hydrogen-ion concentration.

Pharmacodynamics The study of the biochemical and physiological effects of drugs and their mechanisms of action.

Pharmacogenomics The science of studying the biochemical and physiological variations and developing drug therapies to compensate for the genetic differences impacting therapy regimens.

Pharmacokinetics A study of drug movements in the body. It provides a time course of drug concentrations in the body as a function of absorption, distribution, metabolism, and excretion.

Phenylethanolamine *N*-methyltransferase An enzyme found primarily in the adrenal medulla that converts norepinephrine to epinephrine.

Pheochromocytoma Tumors of the adrenal medulla or sympathetic ganglia that produce and release large quantities of catecholamines.

Pheochromocytoma–Paraganglioma (PPGL) Highly vascular neuroendocrine tumors that arise from chromaffin cells of the adrenal medulla or neural crest progenitors located outside of the adrenal gland.

Phospholipids Lipids that contain two fatty acids and a phospholipid on a glycerol backbone. These are amphipathic lipid molecules that are found on the surface of lipoprotein molecules.

Physiologic development Growth.

Pipette Glass or plastic equipment used to transfer liquids.

Placenta A structure that resides in the uterus between the mother and fetus. Its function is to deliver nutrition and oxygen to the fetus and support the pregnancy with hormones.

Pleural fluid Essentially interstitial fluid of the systemic circulation; it is contained in a membrane that surrounds the lungs.

Point-of-care testing (POCT) Analytical testing of patient specimens performed outside the physical laboratory and at the site of patient care.

Poison Any substance that causes a harmful effect upon exposure.

Polyclonal Arising from different cell lines.

Polydipsia Increased thirst.

Polypeptide A chain of amino acids linked together by peptide bonds.

Polysaccharides Ten or more monosaccharides that are linked together.

Positive acute-phase reactant A substance that increases markedly in the presence of an acute onset of a disease state.

Positive predictive value Chance of an individual having a given disease or condition if the test is abnormal.

Postanalytic Introduced during the phase after the analysis (reporting, etc).

Posterior pituitary The posterior, or further back portion, of the pituitary gland that releases AVP and oxytocin. Also called *neurohypophysis*.

Posthepatic jaundice Extrahepatic disturbance in the excretion of bilirubin. This type of jaundice results from the impaired excretion of bilirubin caused by mechanical obstruction of the flow of bile into the intestines. This may be due to gallstones or a tumor.

Postrenal Conditions obstructing urine flow resulting in elevated urea concentrations.

Potentiometry Measures the electric potential between two electrodes, in which a change in voltage indicates the concentration of each analyte.

Preanalytic Introduced during the collection and transport of samples prior to analysis.

Precision Dispersion of repeated measurements about the mean due to analytic error.

Prehepatic jaundice Before the liver. This type of jaundice results when an excessive amount of bilirubin is presented to the liver for metabolism, such as in hemolytic anemia. This type of jaundice is characterized by unconjugated hyperbilirubinemia.

Prerenal Reduced renal blood flow resulting in azotemia.

Primary aldosteronism (PA) A type of aldosterone excess (hyperaldosteronism) resulting in secondary hypertension.

Primary standard A highly purified chemical that has an exact known concentration and purity.

Primary structure In the context of proteins, the sequence of amino acids linked by peptide bonds.

Probe A device used to aspirate either patient sample or reagent.

Proficiency test Method used to validate a particular measurement process. The results are compared with other external laboratories to give an objective indication of test accuracy.

Prolactin A peptide direct effector hormone responsible for lactation and released from the anterior pituitary.

Proportional error Type of systematic error where magnitude is dependent on analyte concentration.

Prostaglandin Any of a large group of biologically active, carbon-20, unsaturated fatty acids that are produced by the arachidonic acid through the cyclo-oxygenase pathway.

Protein Macromolecules composed of one or more polypeptides; required to maintain the structure, function, and regulation of cells, tissues, and processes throughout the body.

Protein-free filtrate A blood sample that has undergone a process of protein removal via chemical precipitation or other separation methods.

Proteinuria An increased amount of proteins in the urine.

Pulsatile secretion Hormone frequency of secretion regulated by neural modulation and released in regular patterns.

Pyrrole A heterocyclic aromatic organic compound, a five-membered ring with the formula C_4H_4NH.

Q

Quality assurance The process monitoring and assessing that the laboratory results are as accurate as possible and includes all three phases of testing; preanalytic, analytic, and postanalytic.

Quality control The process of monitoring and assessing the analytical phase of testing using materials to assess all testing parameters (i.e., reagents, calibrations, temperature, instruments, etc).

Quantal dose–response relationship Describes the change in health effects of a defined population based on changes in the exposure to the xenobiotic.

Quaternary structure Molecular association of more than one peptide/protein to form a dimer, trimer, tetramer, etc., by noncovalent bonds.

R

Radioactive materials Any material capable of emitting radiant energy (rays or particles).

Random access Enables the operator to load a STAT out of sequence so that it will be analyzed before other previously loaded specimens on an as-needed basis.

Random error Error varies from sample to sample. Causes include instrument instability, temperature variations, reagent variations, handling techniques, and operator variables.

Reabsorption Renal retention of water and solutes for return to circulation, prior to urine formation.

Reactive chemicals Substances that, under certain conditions, can spontaneously explode or ignite or that evolve heat and/or flammable or explosive gases.

Reagent-grade water Purified form of water suitable for making reagents in a clinical laboratory.

Redox potential The measure of the ability of a solution to accept or donate electrons.

Reduced Substance that gained an electron.

Reducing agent Substance that donates an electron in a redox reaction.

Reference range A pair of medical decision points that span the limits of results expected for a defined healthy population.

Renin An enzyme produced by the kidney that splits the angiotensinogen to form angiotensin I, which is then transformed to angiotensin II, which stimulates vasoconstriction and secretion of aldosterone.

Renin–angiotensin–aldosterone system (RAAS) Bodily system that regulates osmolality and water volume.

Respiratory distress syndrome (RDS) A condition that may occur upon the changeover to air as an oxygen source at birth if the proper quantity and type of phospholipid (surfactant) is not present. Also referred to as *hyaline membrane disease* because of the hyaline membrane found in affected lungs.

Resting membrane potential (RMP) Ground state cellular response.

Reverse osmosis (RO) Process that uses pressure to force water through a semipermeable membrane producing filtered water.

Rhabdomyolysis Acute destruction of muscle fibers.

Rhinorrhea Discharge from the nose; also, leakage of CSF into the nose.

Rickets A vitamin D deficiency affecting growing bones in children leading to permanent skeletal deformity in bones.

Robotics Characterized by front-end automation to "handle" specimen processing and loading prior to analysis.

Rotor A round device on some automated analyzers that holds sample cups and is capable of spinning. Also, may refer to the device for spinning samples in a centrifuge.

S

Safety Data Sheets (SDSs) A major source of safety information for employees who may use hazardous materials in their occupations. Formerly known as material safety data sheets (MSDSs).

Secondary structure In the context of proteins, amino acids near to each other interact through hydrogen bonds. This type of structure of alpha helix and beta pleated sheet is a common example.

Secretin A hormone released into the bloodstream by the duodenum to stimulate secretion by the liver and pancreas.

Secretion Renal excretion of waste products filtered by the kidney.

Serial dilution Multiple progressive dilutions ranging from more concentrated solutions to less concentrated solutions.

Serous fluid Liquid of the body similar to blood serum; in part secreted by serous membranes.

Sertoli cells Cells of the seminiferous tubules that nourish spermatids.

Serum Liquid portion of whole blood containing no cells or fibrinogen.

Sexual maturation Development of sexual processes and organs.

Shift An abrupt change in the analytic process shown by distinct changes in QC values.

Significant figures Minimum number of digits needed to express a particular value in scientific notation without loss of accuracy.

Simple protein Proteins composed only of amino acids.

Slope The ratio of vertical change over horizontal change of a line; rise over run.

Solid phase In immunoassays, solid particles, beads, trays, or tubes onto which antibody or antigen is adsorbed.

Solute Substance dissolved in a liquid.

Solution Combination of solute and solvent.

Solvent Liquid in which solute is dissolved.

Somatostatin (SS) Hormone that inhibits GH, also known as growth hormone–inhibiting hormone.

Specific gravity The ratio of the density of a material when compared with the density of pure water at a given temperature.

Spectrophotometry The measurement of light intensity at varying wavelengths.

Standard deviation The "average" distance from the center of the data (the mean) and every value in the data set.

Standard deviation index (SDI) Refers to the difference between the measured value and the mean expressed as a number of SDs. An SDI = 0 indicates the value is accurate or in 100% agreement; an SDI = 3 is 3 SDs away from the target (mean) and indicates inaccuracy. SDI may be positive or negative.

Standard precautions Guidelines that consider blood and other body fluids from all patients as infective; include hand washing, gloves, eye protection, etc. Formerly referred to as universal precautions.

Standard reference materials (SRMs) Reference material with NIST-certified chemical composition.

Standardization A set of techniques used to minimize the effects of differences in age or other confounding variables when comparing two or more populations.

Steatorrhea Increased fat in stool due to malabsorption by the intestine.

Steroids Hormones derived from a lipid, typically cholesterol. They are hydrophobic and must be transported in the blood bound to carrier proteins.

Subacute thyroiditis A condition in which there are three phases: a *thyrotoxic phase* in which thyroid hormone is leaking into the circulation, a *hypothyroid phase* when the thyroid gland is repairing itself, and a *euthyroid phase* once the gland is repaired. These phases can last weeks to months.

Subarachnoid hemorrhage Bleeding in the space that surrounds the brain (between the brain and the membrane-subarachnoid space).

Subclinical hyperthyroidism A condition in which thyroid hormone level is still within reference range while TSH is below its reference range because the pituitary gland recognizes that the current hormone level is adequate or excessive.

Subclinical hypothyroidism A condition in which thyroid hormone level is still within reference range while TSH is above its reference range because the pituitary gland recognizes that the current hormone level is insufficient.

Surfactant Surface-active agents forming a monomolecular layer over pulmonary alveolar surfaces; lipoproteins that include lecithins and sphingomyelins that stabilize alveolar volume by reducing surface tension and altering the relationship between surface tension and surface area.

Syndrome of inappropriate ADH (SIADH) Syndrome in which there is an increased AVP release and subsequent increase of fluid retention.

Synovial fluid Fluid formed by ultrafiltration of plasma across the synovial membrane of a joint. The membrane also secretes into the dialysate, a mucoprotein rich in hyaluronic acid, which causes the synovial fluid to be viscous.

Systematic error Predictable error resulting from inaccuracy in a method; results in constant or proportional bias.

Système International d'Unités (SI) International system of units based on the metric system.

T

Tandem mass spectrometry Analytical technique that allows whole groups of similar compounds to be analyzed on very small sample volumes without complex sample preparation.

TD_{50} Dose of a drug that would be predicted to produce a toxic response in 50% of the population.

Teratogens Anything that may cause abnormal development of an embryo.

Tertiary structure In the context of proteins, a three-dimensional structure due to hydrophobic effect, ionic attraction, hydrogen bonds, and disulfide bonds.

Tetany Irregular muscle spasms.

Therapeutic drug monitoring The measurement of drugs and/or their metabolites in body fluids, usually in blood, to maintain therapeutic benefits.

Therapeutic index Quantitative measurement of the relative safety of a drug.

Therapeutic range The dose (or concentration) range of a drug within which the drug produces the desired therapeutic effect.

Thermistor Electronic thermometer.

Thoracentesis Removal of fluid from the pleural space by needle and syringe after visualization by radiology.

Thyroglobulin An iodine-containing protein synthesized by the thyroid gland.

Thyroid peroxidase (TPO) An enzyme that catalyzes the process of thyroid hormone synthesis.

Thyroid-releasing hormone (TRH) Peptide hormone released from the hypothalamus that directs the anterior pituitary to release TSH.

Thyroid-stimulating hormone (TSH) A glycoprotein and peptide hormone consisting of two subunits, α and β, linked noncovalently. It is released by the anterior pituitary and controls the release of thyroxine from the thyroid gland. Also called thyrotropin.

Thyrotoxicosis A group of syndromes caused by high levels of free thyroid hormones in the circulation. Thyrotoxicosis means only that the patient is suffering the metabolic consequences of excessive quantities of thyroid hormones.

Thyrotropin See thyroid stimulating hormone (TSH).

Thyrotropin-releasing hormone (TRH) A tripeptide released by the hypothalamus. It travels along the hypothalamic stalk to the β-cells of the anterior pituitary, where it stimulates the synthesis and release of thyrotropin or thyroid-stimulating hormone (TSH).

Thyroxine (T_4) A hormone produced by the thyroid gland composed of 2 DIT residues. (contains four atoms of iodine). Quantitatively, there is more of this hormone circulating than T_3 levels in the blood, but this hormone is the least biologically active thyroid hormone. It is less potent than T_3 but also has a longer half-life than T_3.

Thyroxine-binding globulin (TBG) A protein that binds thyroid hormones in circulation.

Thyroxine-binding prealbumin (TBPA) Transport protein of thyroxine. Also known as transthyretin (TTR).

Total allowable error Error budget for a measurement or an assay to be clinically useful. It is a combination of imprecision and bias.

Total laboratory automation A comprehensive, automated system encompassing the preanalytic, analytic, and postanalytic phases of testing.

Toxicant Substances that are not produced within a living cell or microorganism and are more commonly used to describe environmental chemicals.

Toxicokinetics The properties that toxic agents exhibit related to unique absorption, distribution, metabolism, and elimination rate of change.

Toxicology The study of poisons, their actions, their detection, and the treatment of the conditions produced by them.

Toxin A substance that is a poison.

Trace element An element that occurs in biological systems at concentrations of milligram-per-kilogram amounts or less (parts per million). Typically, the daily requirement of such an element is a few milligrams per day.

Transcellular fluid Fluids, such as cerebral spinal fluid, GI, and ocular fluids, that are made by special cells.

Transferase An enzyme that transfers a group other than hydrogen from one substrate to another.

Transudate Fluid that passes through a membrane; in comparison to an exudate, it has fewer cells and is of lower specific gravity. Transudates are secondary to remote (nonpleural) pathology and indicate that treatment should begin elsewhere.

Traumatic tap The artifactual presence of blood or derivatives due to the damage of blood vessels during the lumbar puncture.

Trend A gradual change in the analytic process shown by small changes in QC values.

Triglycerides Lipids composed of fatty acid molecules attached to a glycerol backbone. They are very hydrophobic.

Triiodothyronine (T_3) A hormone produced by the thyroid gland composed of 1 MIT and 1 DIT residue (contains three atoms of iodine). Quantitatively, there is more circulating T_4 than this hormone in the blood, but this hormone is the most biologically active thyroid hormone. It is more potent than T_4 but also has a shorter half-life than T_4.

Trioses Carbohydrates that contain three carbons.

Trophic hormone A hormone that stimulates the target tissue to grow or increase in size or number.

Tropic hormone Hormones with actions that are specific for another endocrine gland.

Troponin C (TnC) A globular protein that binds to calcium to reverse the inhibitory activity of troponin I (TnI).

Troponin I (c TnI) Globular protein; specific marker for cardiac disease.

Troponin T (c TnT) Asymmetrical globular protein; cardiac marker that allows for both early and late diagnoses of AMI.

Trough drug concentration The minimum level of a blood drug concentration.

True negative Patients without a condition who are correctly identified by a test as not having the condition.

True positive Patients with a condition who are correctly identified by a test as having the condition.

TSH receptor antibodies Autoantibodies that target the TSH receptor of follicular cells. These antibodies may be present in individuals who have autoimmune thyroid diseases such as Graves' or Hashimoto's thyroiditis.

Tubular reabsorption Process by which renal tubule substances filtered by the glomerulus are reabsorbed as filtrate passes through tubules.

Tubular secretion Passage of substances from peritubular capillaries into tubular filtrate; serves two functions: (1) eliminating waste products not filtered by the glomerulus, and (2) regulating acid–base balance in the body through secretion of 90% of hydrogen ions excreted by kidneys.

Tubule A small tube or canal.

Tumor marker A biomarker found in the blood or tissue that, when elevated, is linked to cancer.

Turbidimetry Measurement of the amount of light that can pass through a sample at a 180° angle from the incident light.

U

Ultratrace element An element present in biological systems at concentrations of milligram-per-kilogram amounts or less (parts per billion) and has extremely low daily requirements (often less than 1 milligram per day).

Unconjugated bilirubin Bilirubin transported by albumin in the bloodstream to the liver.

Urea Nitrogenous compound and product of protein metabolism enzymatically formed during the urea cycle.

Urea nitrogen/creatinine ratio A calculation comparing blood urea to creatinine levels, allowing for differentiation of abnormal urea values.

Uremia or uremic syndrome A clinical condition characterized by elevated blood urea nitrogen (BUN) levels.

Uric acid A naturally occurring nitrogenous compound produced during purine nucleotide metabolism.

Uridine diphosphate glucuronosyltransferase One of the enzymes important for bilirubin metabolism.

Urobilinogen A colorless product or derivative of bilirubin formed by the action of bacteria.

V

Valence The number of units that can combine with or replace 1 mole of hydrogen ions (acids) or hydroxyl ions (bases) and the number of electrons exchanged in oxidation-reduction reactions.

Vasoactive inhibitory peptide (VIP) Also known as *vasoactive intestinal polypeptide*, this is a neuropeptide that functions as a neuromodulator and neurotransmitter. It is a potent vasodilator that regulates smooth muscle activity, epithelial cell secretion, and blood flow in the gastrointestinal tract.

Vasopressin *See* Arginine vasopressin (AVP).

Vesicular monoamine transporters (VMATs) A membrane-embedded protein that transports monoamine neurotransmitter molecules into intraneuronal storage vesicles to allow subsequent release into the synapse.

Virilization Development of masculine sex characteristics in the female.

Vitamin Organic molecules required by the body in amounts ranging from micrograms to milligrams per day for maintenance of structural integrity and normal metabolism. They perform a variety of functions in the body.

Vitamin D A steroid hormone synthesized in the skin following exposure to UVB rays from the sun that is important for calcium absorption. See also 1,25 Dihydroxyvitamin D (calcitriol).

VLDL Very-low-density-lipoprotein is a lipoprotein produced mainly by the liver. It is rich in triglycerides.

Volumetric Container that holds a precise volume of liquid.

W

Waived tests The simplest complexity listing in CLIA. It involves primarily test systems approved by the Food and Drug Administration for home use. The requirements are that there is no reasonable risk of harm to the patient if the test is performed incorrectly. The likelihood of erroneous results is negligible; the test method is simple and uncomplicated; and it is available for home use.

Whole blood The liquid portion of the blood and its cellular components.

X

Xanthomas Nodules in the skin formed as a result of increased lipids.

Xenobiotic Chemicals not normally found or produced by the body.

Y

Y-intercept Position at which a line intersects the *y*-axis in a graph.

Z

Zero-order kinetics The rate of an enzymatic reaction that occurs in which the rate is dependent only the concentration of the enzyme because the substrates are in excess.

Zollinger-Ellison syndrome A disease in which gastrinomas cause the stomach to produce too much acid, resulting in peptic ulcers.

Zona fasciculata (F-zone) The middle and also the widest zone of the adrenal cortex, sitting directly beneath the zona glomerulosa. Constituent cells are organized into bundles or "fascicles."

Zona glomerulosa (G-zone) The most superficial layer of the adrenal cortex, lying directly beneath the renal capsule. Its cells are ovoid and arranged in clusters or arches (*glomus* is Latin for "ball").

Zona reticularis (R-zone) The innermost layer of the adrenal cortex, lying deep to the zona fasciculata and superficial to the adrenal medulla. The cells are arranged as [or "in"]cords that project in different directions giving a net-like appearance.

Zwitterion A molecule that contains both positively and negatively charged functional groups, yielding a net neutral charge.

Zygote A developing cell formed when fertilized by two gametes.

Zymogen An inactive, secreted form of the enzyme.

Index

Note: Locators followed by '*f*', '*t*' and '*b*' refer to figures, tables and boxes respectively.

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Y

Z