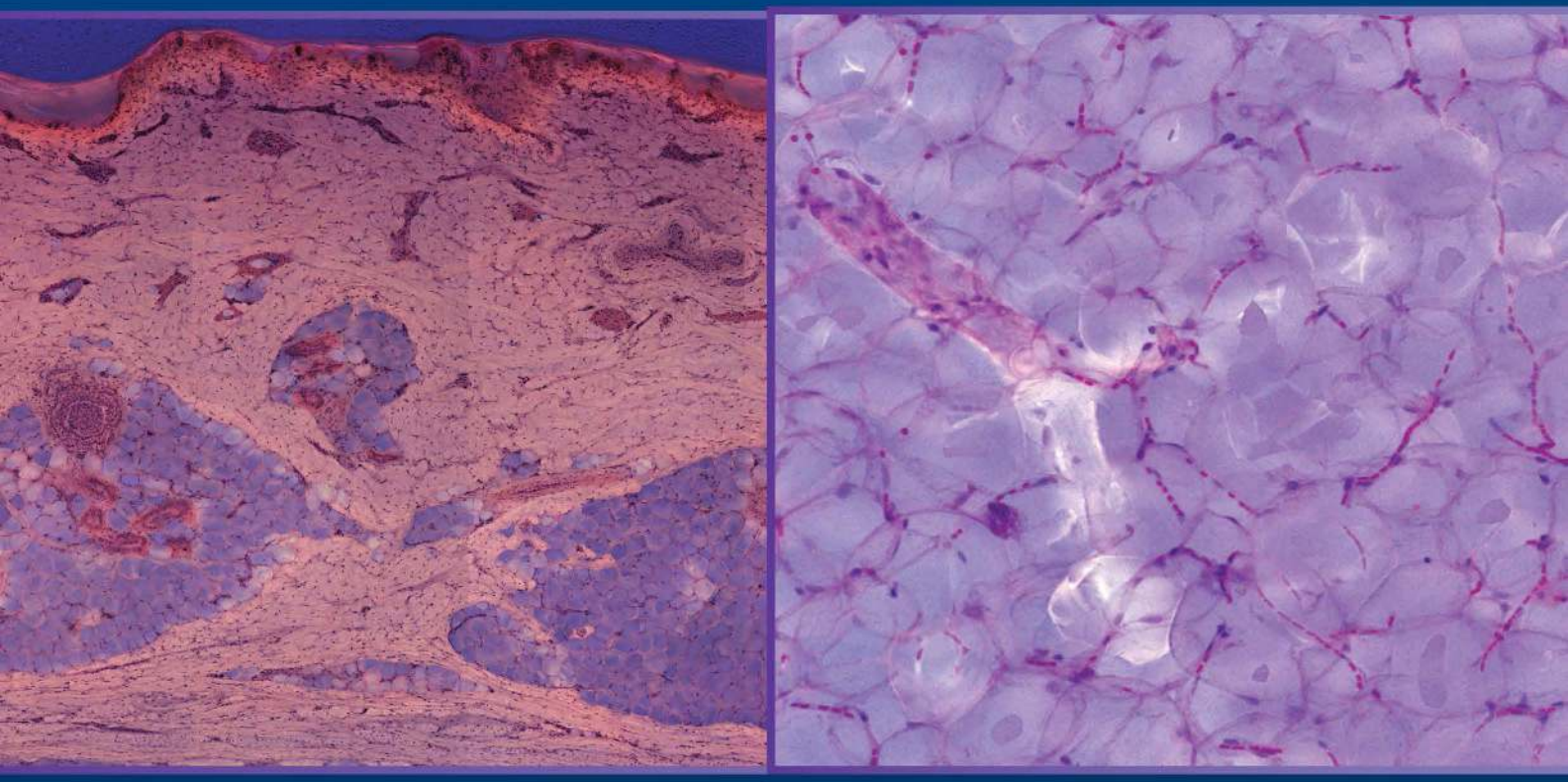
CANCER DETECTION AND DIAGNOSIS

A HANDBOOK OF EMERGING TECHNOLOGIES



Edited by **Miguel Ossandon, Ben Prickril,** and **Avraham Rasooly**

CRC Press
Taylor & Francis Group

Cancer Detection and Diagnosis

Emerging technologies for cancer detection and diagnosis are providing more and more advance warning of pathologies of clinical significance. Research devoted to cancers are revealing new ways of finding and treating these complex diseases. This volume reviews a broad array of new technologies for cancer detection and diagnosis. While there are several clinical books describing cancer diagnosis, and general molecular analytical technologies, these books are not focused on cancer detection and diagnosis. The aim of this book is to describe emerging cancer detection and diagnosis technologies.

Key Features

- Presents myriad new experimental cancer detection technologies
- Describes technology so the reader may conduct similar analyses
- Outlines clinical applications of technology for specific cancer and summarizes results
- Discusses pitfalls and limitations, future trends and potential technological developments

Miguel Ossandon has a dual background in clinical laboratory and computer science. He started working in cancer research at the Lombardi Cancer Center at Georgetown University where he also began his undergraduate training in computer science. He has been working for the National Cancer Institute since 2007. Miguel received his master's degree at the George Washington University and PhD in computer science at the University of Maryland Baltimore County. As a program director in the Diagnostic Biomarkers and Technology Branch, he manages a grant portfolio related to computational modeling and machine learning approaches for cancer diagnosis, digital image processing/analysis, and microfluidic and circulating tumor cell technology.

Ben Prickril is a former US government official working at the US National Institutes of Health, National Cancer Institute. He has a background in medically related chemistry, microbiology, immunology, patenting and intellectual property, and global health. International research development includes experience in France, Turkey, Ukraine, Czech Republic and Burkina Faso. He received his PhD from the University of Georgia.

Avraham Rasooly is in the Division of Cancer Treatment and Diagnosis, National Cancer Institute. He has been responsible for developing research programs on new approaches for cancer therapy, including technologies for microbial-based cancer therapy. He received his PhD from Michigan State University.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Cancer Detection and Diagnosis
A Handbook of Emerging Technologies

Edited by
Miguel Ossandon, Ben Prickril, and Avraham Rasooly

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Designed cover image: Miguel Ossandon, Ben Prickril, and Avraham Rasooly

First edition published 2026
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

© 2026 selection and editorial matter, Miguel Ossandon, Ben Prickril, and Avraham Rasooly; individual chapters, the contributors

Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, access www.copyright.com or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. For works that are not available on CCC please contact mpkbookspermissions@tandf.co.uk

Trademark notice: Product or corporate names may be trademarks or registered trademarks and are used only for identification and explanation without intent to infringe.

ISBN: 978-1-032-57701-2 (hbk)
ISBN: 978-1-032-58409-6 (pbk)
ISBN: 978-1-003-44994-2 (ebk)

DOI: 10.1201/9781003449942

Typeset in Times LT Std
by Apex CoVantage, LLC

Miguel R. Ossandon dedicates this book to the loving memory of his mother, Rosa Eliana Carrizo; his father, Andres Ossandon; and his brother, Ruben Alfredo Ossandon Carrizo.

Avraham Rasooly dedicates this book to the loving memory of his parents, Malah and Jury Rasooly, and his son Ilan Rasooly.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contents

Preface xiii
List of Contributors xv

Overview of Cancer Features and Impact 1
Miguel Ossandon, Ben Prickril, and Avraham Rasooly

Section 1 Optical Technologies

1 Near-Infrared Macroscopic and Mesoscopic Fluorescence Lifetime FRET Imaging to Measure Intra-Tumor Heterogeneity of Antibody-Target Engagement 9
Amit Verma, Saif Ragab, Shan Gao, Catherine Sherry, Nanxue Yuan, Vikas Pandey, Xavier Intes, and Margarida Barroso

2 Combined Reflectance Confocal Microscopy–Optical Coherence Tomography for Skin Cancer Detection and Therapy Guidance 20
Nicusor Iftimia, John Grimble, Gopi Maguluri, Aditi Sahu, Miguel Cordova, and Milind Rajadhyaksha

3 *In Vivo* Confocal Laser Endomicroscopy: An Imaging Biomarker for Risk Stratification of Precancerous Pancreatic Cystic Lesions 27
Da Yeon Ryoo, Bryn Koehler, Vanisha Patel, and Somashekar G. Krishna

4 Portable Confocal Microscopy for Aiding Diagnosis and Treatment of Skin Cancer 36
Momoka Sugimura, Jingwei Zhao, Kenneth Marcelino, Yongjun Kim, Rafael Romero, and Dongkyun Kang

5 Quantitative Fluorescence Imaging, Light-Triggering, and Monitoring of Chemodrug Release from Liposomes *In Vivo by a* Mesoscopic-Scale Theranostic Endoscope 43
Ulas Sunar, Rasel Ahmmed, Eli Kluiszo, and Semra Aygun-Sunar

6 Fluorescence Imitating Brightfield Imaging (FIBI): A Simple Slide-Free Microscopy Approach 58
Nathan Anderson, Farzad Fereidouni, and Richard Levenson

7 Fluorescently Labeled Tyrosine Kinase Inhibitors for Intracellular Protein Target Imaging 67
Allison Solanki, Jesse Korber, Summer L. Gibbs, and Lei G. Wang

8 Real-Time Cancer Metabolism Detection Using a Nanocoil Integrated Hyperpolarized Micromagnetic Resonance Spectrometer 76
Thasin Peyear, Roberta Pigliapochi, and Kayvan R. Keshari

9 Detection of TMEM Doorways and Their Activity Required for Metastasis 82
George S. Karagiannis, Robert J. Eddy, Joseph A. Sparano, Jesus Anampa Mesias, Thomas E. Rohan, Craig Branch, John S. Condeelis, David Entenberg, and Maja H. Oktay

10 Reading the Reader: Utilizing Eye Movements and Machine Learning to Enhance Accuracy during Diagnostic Visual Search 92
Gregory J. DiGirolamo, Federico Sorcini, Leo Joskowicz, Yakir Winograd, Assaf Hoog, and Max P. Rosen

11 Optical Imaging Technology for *In Vivo* Tumor Detection 103
Radka Bokorova, Duy Binh Tran, and Tohru Yamada

12 Adapting Image Foundational Model to Identify Tumor Budding from H&E Images in Colorectal Cancer Diagnostics 110
Ziyu Su, Usama Sajjad, Wendy L. Frankel, Metin N. Gurcan, Wei Chen, and Muhammad Khalid Khan Niazi

13 Quantitative Phase Imaging for Assessing Tumor Cell Adaptability 118
Rebecca G. Zitnay, Shukran Alizada, Anastasia Prokofyeva, Thomas A. Zangle, and Robert L. Judson-Torres

14 An Exo-PROS Biosensor Simultaneously Detects Tumor-Derived Exosomal Protein-MicroRNA Pairs for Lung and Breast Cancer Diagnosis 125
Chang-Chieh Hsu and Yun Wu

Section 2 Imaging Technologies

15 Epigenetic Tools for Guiding Low Dose Computerized Tomography (LDCT) Screening Decision-Making 137
Winter Philibert and Robert Philibert

16 Multi-View Models for Colorectal Polyp Detection in CT Colonography 147
Mohamed Yousuf, Samir Harb, Asem Ali, Salwa Elshazly, and Aly Farag

17 Hybrid Multi-Dimensional MRI of Prostate Cancer 155
Aritrick Chatterjee and Aytekin Oto

18 Detection and Characterization of Brain Metastases Using Quantitative Chemical Exchange Saturation Transfer MR Fingerprinting (CEST-MRF) 163
Ouri Cohen

19 Narrow-Beam CT: A Solution to the Limitations in Breast Cancer Screening 171
Peymon Ghazi

20 Broadband Coherent Anti-Stokes Raman Scattering Microscopy for Metabolic and Phenotypic Imaging in Cancer 181
Jessica Z. Dixon, Wei-Wen Chen, Rajas Poorna, Xavier Audier, and Marcus T. Ciceronex

21 Spatial Multiplexed Immunofluorescence Imaging Exploring The Tumor Immune Microenvironment of Melanoma 190
Jun Jiang, Shankar Suman, Caitlin Ward, Raymond M. Moore, Zachary C. Fogarty, Alexey Leontovich, Wendy K. Nevala, Chathu L. Atherton, Jill M. Schimke, and Svetomir N. Markovic

22 Spectroscopic Optical Coherence Tomography for *In Situ* Analysis of Colonic Epithelium 198
Jatin Roper, Wesley Kendall, and Adam Wax

23 Analysis of Circulating Tumor DNA (ctDNA) by Digital Droplet PCR (ddPCR) 206
Bryant L. McCue and George Miles

24 Noninvasively Monitoring the Delivery of DNA Alkylating Agents and Cellular Responses to These Agents Using Chemical Exchange Saturation Transfer Magnetic Resonance Imaging 213
Ethel J. Ngen

Section 3 Nucleic Acid Technologies

25 Clonal Analysis of Cancer by Mitochondrial DNA Barcoding 223
Jonathan D. Good, Ksenia R. Safina, and Peter van Galen

26 Implementing Clinical Risk Assessment for Genetic Susceptibility to Cancer: Challenges and Successes in a Statewide Initiative 231
Sarah Austin, Liliana Arida-Moody, Erika Hanson, Shayna Weiner, Matthew Demerath, Emerson Delacroix, Erika Koeppe, Ken Resnicow, Jennifer J. Griggs, and Elena M. Stoffel

27 Assessing Tumor Tissue for the Alternative Lengthening of Telomeres (ALT) Phenotype 239
C. Patrick Reynolds, Trevor A. Burrow, and Balakrishna Koneru

28 Click Chemistry-Mediated Enrichment of Tumor-Derived Extracellular Vesicles for RNA-Based Digital Scoring 247
Yazhen Zhu, Jacqueline Ziqian Yang, Anna Yaxuan Wei, Vivian Xufei Zuo, Judy Law, Sungyong You, Ju Dong Yang, Vatche G. Agopian, Edwin Posadas, Junseok Lee, and Hsian-Rong Tseng

29 Identifying Methods to Deliver Mutant p53-Based Diagnostics and Therapeutics to Target the Previously Undruggable 261
Gu Xiao, Outi Keinänen, Florencia Madorsky Rowdo, Samantha Delaney, Katherine W. Harmon, Brian M. Zeglis, and Jill Bargonetti

30 DNA Methylation Biomarkers for Preoperative Diagnostic of Thyroid Nodules 270
Byung-Wook Kim, Yuman Fong, John H. Yim, and Maria A. Hahn

31 Utilizing RADAR for the Integration of Genomic Data and Immunotherapy Targets in Precision Oncology for Multiple Myeloma 280
Alessandro Laganà, Sanjeev Sariya, David T. Melnekoff, Meghana Ram, and Samir Parekh

32 Nanopore Whole Transcriptome Sequencing Offers the Potential for Accessible Classification of Pediatric Cancers 295
Thomas B. Alexander and Jeremy R. Wang

33 Quantitative Detection of Cancer Nucleic Acid Biomarkers with a Selective Solid-State Nanopore Assay 303
Adam R. Hall, Sara Abu Jalboush, Ian D. Wadsworth, Mohamed Amin Elaguech, and Komal Sethi

34 Isolation and Characterization of Cell-Free RNA from Liquid Biopsy Taken from Cancer Patients 310
Real Sumayya Abdul Sattar, Luisa Quezada Ziesse, and Subhajyoti De

Section 4 Biomarker–Molecular Technologies

35 The Use of Dielectrophoresis to Recover Cancer-Derived Nanoparticles Straight from Undiluted Human Plasma for Cancer Detection Applications 321
Delaney Shea, Jason Ware, Michael I. Henderson, Ella Stimson, Mehrzad Sasanpour, Sarah Mitchell, Christian Ross, Shelby Nicholas, and Stuart D. Ibsen

36 Conditional Reprogramming: A Living Biomarker and Phenotypic Screening Drug Platform for Urological Cancer 331
Danyal Daneshdoust, Abdul Qawee Rani, Peng Wang, Shang-jui Wang, Akshay Sood, Debasish Sundi, Qi-en Wang, Zaibo Li, Anil V. Parwani, Nagireddy Putluri, Cheryl Lee, Jenny Li, and Xuefeng Liu

37 Proteolytic Activity Signatures as Candidate Biomarkers for Thyroid Cancer 344
Danielle E. Dettling and Giselle M. Knudsen

38 Carbohydrate Microarrays Identify the Stage-Specific Embryonic Antigen (SSEA)-0 as a Novel Oncofetal Cancer Marker 351
Denong Wang

39 Collective Attributes of Extracellular Vesicles as Biomarkers for Cancer Detection 357
Alfred Akinlalu, Tommy Gao, Sky Gao, and Dali Sun

40 Using a Gamified Mobile App Approach to Train Tobacco Control Program Implementers in Schools in India 364
Eve M. Nagler, Chuck Sigmund, Smita P. Warke, Priyanka Ghosh, Paromita Mehta, Leah C. Jones, and Mangesh Pednekar

41 Parallel Reaction Monitoring (PRM) Quantitative Analysis of Glycopeptide Biomarkers from Patient Serum 372
Komal Abhange, Yu Lin, Jie Zhang, Jianhui Zhu, Amit G. Singal, Neehar D. Parikh, and David M. Lubman

42 Point-of-Care Technologies for Molecular Subtyping of Breast Cancer in Low- and Middle-Income Countries 387
Wenting Gao, Iftak Hussain, Sally Adebamowo, David Erickson, and Clement Adebamowo

43 Development of LLM for Prostate Cancer—The Need for Domain-Tailored Training 397
Amara Tariq, Man Luo, Aisha Urooj, Avisha Das, Jiwoong Jeong, Shubham Trivedi, Haidar Abdul-Muhsin, Umer Ghaffar, Nathan Yu, Bhavik Patel, and Imon Banerjee

44 Biosample and Method Selection for Marker Studies 407
Angela J. Yoon

45 Subharmonic-Aided Pressure Estimation for Evaluating Cancer 413
Hailee Mayer, Kibo Nam, Jason Shames, and Flemming Forsberg

46 Development and Implementation of a Method for Registering Prostate-Specific Antigen (PSA) Biosensor-Based Assay Results in a Personalized QR Code 424
Maria Daoutakou1 and Spyridon Kintzios

Section 5 Organel and Cellular Technologies

47 Deep Learning-Based Colon Segmentation for Accurate Colorectal Polyps Detection 433
Samir Harb, Mohamed Yousuf, Ahmed Elsayed, Asem Ali, Salwa Elshazly, and Aly Farag

48 Emerging Blood-Based Tests for Colorectal Cancer Screening: Advances, Challenges, and Future Directions 442
Gloria D. Coronado, Jennifer L. Schneider, Cheryl A. Johnson, and Leslie Bienen

49 Microfluidic Isolation and Molecular Characterization of Circulating Tumor Cells in Prostate Cancer 449
Rea Pittie, Victor Putaturo, Shyamala Maheswaran, Daniel A. Haber, Mehmet Toner, Avanish Mishra, and David T. Miyamoto

50 Chemical Imaging for Next-Generation Histopathology 459
Kianoush Falahkheirkhah, Kevin Yeh, and Rohit Bhargava

51 GALAD Score and Extracellular Vesicle-Based Assays for Early Detection of Hepatocellular Carcinoma 469
Yi-Te Lee, Yazhen Zhu, Sungyong You, Vatche G. Agopian, Hsian-Rong Tseng, and Ju Dong Yang

52 Using Epidemiologic Research Methods to Adapt a Human Papillomavirus Assay for Cervical Cancer Screening in Low- and Middle-Income Countries 480
Louise Kuhn, Rakiya Saidu, and Lynette Denny

53 Organoid Models for Early Detection and Diagnosis of Colorectal Cancer 492
Jinghui Liang, Janvhi Suresh Machhar, and Chengyu Liang

Section 6 Computaional – Digital Technologies

54 Experimental Methods for Establishing and Maintaining Canine Bladder Cancer Organoids 505
Mohamed Elbadawy, Christopher Zdyrski, Hannah Wickham, Corey F. Saba, Savantha Thenuwara, Alison P. Mosichuk, Dipak Kumar Sahoo, Jonathan P. Mochel, and Karin Allenspach

55 Top-Down Proteomics of Cancer Cells by Advanced Capillary Electrophoresis–Mass Spectrometry 514
Tian Xu and Liangliang Sun

56 X-ray Diffraction Imaging for Cancer Detection and Diagnosis 524
Joel A. Greenberg and Anuj J. Kapadia

57 Restriction Spectrum Imaging (RSI) Models for the Detection of Female Cancers 532
Aida Haddad, Summer J. Batasin, Sabrina Kohanzad, Emily Barnard, Ana E. Rodríguez-Soto, Stéphane Loubrie, and Rebecca Rakow-Penner

Index 541

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Preface

This is the second edition of *Biosensors and Molecular Technologies for Cancer Diagnosis*, initially published in 2012 by CRC Press. Since the first edition, many new cancer technologies have been developed, which have been included to broaden the book's scope beyond biosensors. This second edition is entitled *Cancer Detection and Diagnosis: A Handbook of Emerging Technologies*.

Cancer is a major global health concern. In 2022, there were approximately 20 million new cancer cases and 9.7 million cancer-related deaths worldwide. Early cancer detection and accurate diagnoses are essential for effective treatment, and addressing cancer challenges requires improving cancer detection and diagnosis, a need that is a well-recognized imperative.

Research on new cancer technologies has expanded in recent years, and many new technologies have been published in various publications, describing novel biomedical and analytical applications for cancer detection and diagnosis. However, despite their potential, many of these technologies are absent today in clinical cancer settings.

To support the development of cancer technologies, this book describes novel experimental approaches and their potential applications to cancer detection, diagnosis, and treatment, which may help to facilitate their utilization and integration into future medical practice.

Six bioengineering analytical approaches are described in this book: (i) Optical, (ii) imaging, (iii) nucleic acid, (iv) biomarker and molecular, (v) organelle and cellular, and (vi) computational/digital technologies. Some potential readers of this technology-oriented book may not be very familiar with cancer challenges and complexities, or the terminology used in the book; for these readers we included a section on the basic elements of cancer.

The book is designed to make these technologies more accessible and understandable to molecular biologists, oncologists, pathologists, engineers, and students. The development of many of the technologies described in this book was supported by the US National Cancer Institute (NCI).

We hope the book will help bridge the gap between emerging cancer technologies and basic and clinical cancer research by providing concise and detailed technical information on these technologies, which may help to bring new biomedical technologies to cancer patients.

Miguel Ossandon, PhD

Ben Prickril, PhD

Avraham Rasooly, PhD

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Contributors

Haidar Abdul-Muhsin
Mayo Clinic
Phoenix, Arizona, USA

Komal Abhange
Department of Surgery
University of Michigan Medical Center
Ann Arbor, Michigan, USA

Clement Adebamowo
Greenebaum Comprehensive Cancer Center and the Department of Epidemiology and Public Health
University of Maryland School of Medicine
Baltimore, Maryland, USA

Sally Adebamowo
Greenebaum Comprehensive Cancer Center and the Department of Epidemiology and Public Health
University of Maryland School of Medicine
Baltimore, Maryland, USA

Vatche G. Agopian
Department of Surgery
David Geffen School of Medicine at UCLA
Los Angeles, California, USA

Rasel Ahmmed
Department of Biomedical Engineering
Stony Brook University
Stony Brook, New York, USA

Alfred Akinlalu
Biomedical Focus, Department of Electrical and Computer Engineering
Knoebel Institute for Healthy Aging
Ritchie School of Engineering and Computer Science
The University of Denver
Denver, Colorado, USA

Thomas B. Alexander
Division of Hematology and Oncology
Department of Pediatrics
Lineberger Comprehensive Cancer Center
University of North Carolina School of Medicine
Chapel Hill, North Carolina, USA

Asem Ali
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky

Shukran Alizada
Department of Chemical Engineering
University of Utah
Salt Lake City, USA

Karin Allenspach
Precision One Health Initiative
Department of Pathology
University of Georgia College of Veterinary Medicine
Athens, Georgia

Nathan Anderson
UC Davis Health
Sacramento, California, US

Liliana Arida-Moody
Department of Internal Medicine
University of Michigan School of Medicine
Ann Arbor, USA

Chathu Atherton
Department of Oncology
Mayo Clinic
Rochester, Minnesota, USA

Xavier Audier
Department of Chemistry and Biochemistry
Georgia Institute of Technology
Atlanta, USA

Sarah Austin
Rogel Cancer Center
University of Michigan
Ann Arbor, USA

Semra Aygun-Sunar
Department of Biomedical Engineering
Stony Brook University
Stony Brook, New York, USA

Imon Banerjee
Mayo Clinic
Phoenix, Arizona, USA

Jill Bargonetti
Department of Biological Sciences
Hunter College, City University of New York
New York, USA
and
The Graduate Center Biochemistry PhD Program
City University of New York
New York, USA
and
Department of Cell and Developmental Biology
Weill Cornell Medical College
New York, USA

Emily Barnard
Department of Radiology
University of California San Diego
La Jolla, USA

Margarida Barroso
Department of Cellular and Molecular Physiology
Albany Medical College
Albany, New York, USA

Summer J. Batasin
Department of Radiology
University of California San Diego
La Jolla, USA

Leslie Bienen
Independent Researcher
C3 Science
Portland, Oregon, USA

Radka Bokorova
Department of Surgery
Division of Surgical Oncology
University of Illinois College of Medicine
Chicago, USA

Craig Branch
Department of Radiology
Gruss Magnetic Resonance Research Center
Albert Einstein College of Medicine
Bronx, New York, USA

Trevor Burrow
Department of Pediatrics and Cancer Center
School of Medicine
Texas Tech University Health Sciences Center
Lubbock, USA

Aritrick Chatterjee
Department of Radiology
Sanford J. Grossman Center of Excellence in Prostate Imaging and Image Guided Therapy
University of Chicago
Chicago, USA

Wei Chen
Department of Pathology
The Ohio State University
Columbus, USA

Wei-Wen Chen
Department of Chemistry and Biochemistry
Georgia Institute of Technology
Atlanta, USA

Marcus T. Ciceronex
Department of Chemistry and Biochemistry
Georgia Institute of Technology
Atlanta, USA

Ouri Cohen
Department of Medical Physics
Memorial Sloan Kettering Cancer Center
New York, USA

John S. Condeelis
Gruss-Lipper Biophotonics Center
Integrated Imaging Program for Cancer Research
Department of Surgery; Department of Cell Biology
Albert Einstein College of Medicine
Bronx, USA
and
Tumor Microenvironment and Metastasis Program
Montefiore-Einstein Comprehensive Cancer Center
Bronx, New York, USA

Miguel Cordova
Memorial Sloan Kettering Cancer Center
New York, USA

Gloria D. Coronado
Population Sciences, Cancer Center
Epidemiology and Biostatistics
College of Public Health
University of Arizona
Tucson, USA

Danyal Daneshdoust
Comprehensive Cancer Center
Ohio State University
Columbus, USA

Maria Daoutakou
Laboratory of Cell Technology
Department of Biotechnology
Agricultural University of Athens
Athens, Greece

Avisha Das
Mayo Clinic
Phoenix, Arizona, USA

Subhajyoti De
Rutgers Cancer Institute of New Jersey
Rutgers University
New Brunswick, USA

Emerson Delacroix
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
and
Department of Internal Medicine
University of Michigan School of Medicine
Ann Arbor, USA
and
University of Michigan School of Public Health
Ann Arbor, USA

Samantha Delaney
Department of Chemistry
Hunter College, City University of New York
New York, USA

Matthew Demerath
Rogel Cancer Center
University of Michigan
Ann Arbor, USA

Lynette Denny
Department of Obstetrics and Gynaecology
Faculty of Health Sciences
University of Cape Town
Cape Town, South Africa

Danielle E. Dettling
Alaunus Biosciences, Inc.
Novato, California, USA

Gregory J. DiGirolamo
Department of Psychology
College of the Holy Cross
Worcester, USA
and
Department of Radiology
UMass Chan Medical School
Worcester, Massachusetts

Jessica Z. Dixon
Department of Chemistry and Biochemistry
Georgia Institute of Technology
Atlanta, USA

Robert J. Eddy
Integrated Imaging Program for Cancer Research
Department of Pathology
Albert Einstein College of Medicine
Bronx, New York, US

Mohamed Amin Elaguech
Virginia Tech-Wake Forest University School of Biomedical Engineering and Sciences
Wake Forest School of Medicine
Winston-Salem, North Carolina, USA

Mohamed Elbadawy
Precision One Health Initiative, Department of Pathology
University of Georgia College of Veterinary Medicine
Athens, USA
and
Department of Pharmacology
Faculty of Veterinary Medicine
Benha University
Moshtohor, Egypt

Ahmed Elsayed
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky, USA

Salwa Elshazly
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky, USA

David Entenberg
Gruss-Lipper Biophotonics Center
Integrated Imaging Program for Cancer Research
Department of Pathology
Albert Einstein College of Medicine
Bronx, USA
and
Tumor Microenvironment and Metastasis Program
Cancer Dormancy Institute
Montefiore-Einstein Comprehensive Cancer Center
Bronx, New York, USA

David Erickson
Sibley School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York, USA

Kianoush Falahkheirkhah
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
Urbana, USA

Aly Farag
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky, USA

Farzad Fereidouni
Emory University
Atlanta, Georgia, USA

Zachary Fogarty
Department of Quantitative Health Sciences
Mayo Clinic
Rochester, Minnesota, USA

Yuman Fong
Department of Surgery
City of Hope Comprehensive Cancer Center
Duarte, California, USA

Flemming Forsberg
Department of Radiology
Thomas Jefferson University
Philadelphia, Pennsylvania, USA

Wendy L. Frankel
Department of Pathology
The Ohio State University
Columbus, USA

Shan Gao
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York, USA

Sky Gao
Biomedical Focus, Department of Electrical and Computer Engineering
Knoebel Institute for Healthy Aging
Ritchie School of Engineering and Computer Science
The University of Denver
Denver, Colorado, USA

Tommy Gao
Biomedical Focus, Department of Electrical and Computer Engineering
Knoebel Institute for Healthy Aging
Ritchie School of Engineering and Computer Science
The University of Denver
Denver, Colorado, USA

Wenting Gao
Sibley School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York, USA

Umer Ghaffar
Mayo Clinic
Phoenix, Arizona, USA

Peymon Ghazi
Malcova Inc.
Newark, California, USA

Priyanka Ghosh
Healis Sekhsaria Institute for Public Health
Navi Mumbai, India

Summer L. Gibbs
Biomedical Engineering Department
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Jonathan Good
Division of Hematology
Brigham and Women's Hospital
Boston, USA
and
Broad Institute of MIT and Harvard
Cambridge, USA
and
Department of Medicine; Ludwig Center at Harvard
Harvard Medical School
Boston, Massachusetts, USA

Joel A. Greenberg
Quadridox, Inc.
Hillsborough, North Carolina, USA

Jennifer J. Griggs
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
and
Department of Internal Medicine
University of Michigan School of Medicine
Ann Arbor, USA
and
Michigan Oncology Quality Consortium (MOQC)
Ann Arbor, USA

John Grimble
Physical Sciences, Inc.
Andover, Massachusetts, USA

Metin N. Gurcan
Center for Artificial Intelligence Research
Wake Forest University School of Medicine
Winston-Salem, North Carolina, USA

Daniel A. Haber
Krantz Family Center for Cancer Research
Massachusetts General Hospital Cancer Center and Harvard Medical School
Charlestown, USA
and
Howard Hughes Medical Institute
Bethesda, Maryland, USA

Aida Haddad
Department of Radiology
University of California San Diego
La Jolla, USA

Maria A. Hahn
Department of Surgery
City of Hope Comprehensive Cancer Center
Duarte, California, USA

Adam R. Hall
Virginia Tech-Wake Forest University School of Biomedical Engineering and Sciences
Comprehensive Cancer Center
Wake Forest School of Medicine
Winston-Salem, North Carolina, USA

Erika Hanson
Michigan Oncology Quality Consortium (MOQC)
Ann Arbor, USA

Samir Harb
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky, USA

Katherine W. Harmon
Department of Biological Sciences
Hunter College
City University of New York
New York, USA

Michael I. Henderson
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering
School of Medicine
Oregon Health and Science University
Portland, USA

Assaf Hoog
School of Computer Science
Ariel University
Jerusalem, Israel

Chang-Chieh Hsu
Department of Biomedical Engineering
University at Buffalo
The State University of New York
Buffalo, USA

Iftak Hussain
Sibley School of Mechanical and Aerospace Engineering
Cornell University
Ithaca, New York, USA

Stuart D. Ibsen
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering, School of Medicine
Oregon Health and Science University
Portland, USA

Nicusor Iftimia
Physical Sciences, Inc.
Andover, Massachusetts, USA

Xavier Intes
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York, USA

Sara Abu Jalboush
Department of Cancer Biology
Wake Forest School of Medicine
Winston-Salem, North Carolina, USA

Jiwoong Jeong
Mayo Clinic
Phoenix, Arizona, USA

Jun Jiang
Department of Quantitative Health Sciences
Mayo Clinic
Rochester, Minnesota, USA

Cheryl A. Johnson
Science Program Department
Kaiser Permanente Center for Health Research
Portland, Oregon, USA

Leah C. Jones
Center for Community-Based Research
Dana-Farber Cancer Institute
Boston, Massachusetts, USA

Leo Joskowicz
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Jerusalem, Israel

Robert L. Judson-Torres
Huntsman Cancer Institute
Department of Oncological Sciences
Department of Dermatology
University of Utah
Salt Lake City, USA

Dongkyun Kang
James C. Wyant College of Optical Sciences
Department of Biomedical Engineering
University of Arizona
Tucson, USA

Anuj J. Kapadia
Quadridox, Inc.
Hillsborough, North Carolina, USA

George S. Karagiannis
Department of Microbiology and Immunology
Gruss-Lipper Biophotonics Center
Integrated Imaging Program for Cancer Research
Albert Einstein College of Medicine
Bronx, USA
and
Tumor Microenvironment and Metastasis Program
Cancer Dormancy Institute
Montefiore-Einstein Comprehensive Cancer Center
Bronx, New York, USA

Outi Keinänen
Department of Chemistry
Hunter College, City University of New York
New York, USA

Wesley Kendal
Department of Biomedical Engineering
Duke University
Durham, North Carolina, USA

Kayvan R. Keshari
Department of Radiology
Molecular Pharmacology Program
Memorial Sloan Kettering Cancer Center
New York, USA
and
Physiology, Biophysics and Systems Biology Program
Weill Cornell Medical College
New York, USA

Muhammed Khalid Khan Niazi
Center for Artificial Intelligence Research
Wake Forest University School of Medicine
Winston-Salem, North Carolina, USA

Byung-Wook Kim
Department of Surgery
City of Hope Comprehensive Cancer Center
Duarte, California, USA

Yongjun Kim
Department of Biomedical Engineering
University of Arizona
Tucson, USA

Spyridon Kintzios
Laboratory of Cell Technology
Department of Biotechnology
Agricultural University of Athens
Athens, Greece
and
Ce.B.Tec Diagnostics I.K.E.
Athens, Greece

Eli Kluiszo
Department of Biomedical Engineering
Stony Brook University
Stony Brook, New York, USA

Giselle M. Knudsen
Alaunus Biosciences, Inc.
Novato, California, USA

Bryn Koehler
Department of Internal Medicine
Ohio State University Wexner Medical Center
Columbus, USA

Erika Koeppe
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
and
Department of Internal Medicine
University of Michigan School of Medicine
Ann Arbor, USA

Sabrina Kohanzad
Department of Radiology
University of California San Diego
La Jolla, USA

Balakrishna Koneru
Department of Pediatrics and Cancer Center
School of Medicine
Texas Tech University Health Sciences Center
Lubbock, USA

Jesse Korber
Biomedical Engineering Department
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Somashekar G. Krishna
Division of Gastroenterology
Department of Internal Medicine
Ohio State University Wexner Medical Center
Columbus, USA
and
National Institutes of Health
The National Cancer Institute (NIH/NCI)
USA

Louise Kuhn
Gertrude H. Sergievsky Center
Vagelos College of Physicians and Surgeons
Department of Epidemiology
Mailman School of Public Health
Columbia University Irving Medical Center
New York, USA

Alessandro Laganà
Department of Oncological Sciences
Department of Genetics and Genomic Sciences
Tisch Cancer Institute
Icahn School of Medicine at Mount Sinai
New York, USA

Judy Law
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California
Los Angeles, USA

Cheryl Lee
Comprehensive Cancer Center
Department of Urology, Wexner Medical Center
Ohio State University
Columbus, USA

Lee, Junseok
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California
Los Angeles, USA

Yi-Te Lee
Karsh Division of Gastroenterology and Hepatology
Cedars-Sinai Medical Center
Los Angeles, California, USA

Alexey Leontovich
Department of Quantitative Health Sciences
Mayo Clinic
Rochester, Minnesota, USA

Richard Levenson
UC Davis Health
Sacramento, California, USA

Jenny Li
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Ohio State University
Columbus, USA

Zaibo Li
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Ohio State University
Columbus, USA

Chengyu Liang
Program in Molecular and Cellular Oncogenesis
The Wistar Institute
Philadelphia, Pennsylvania, USA

Jinghui Liang
Program in Molecular and Cellular Oncogenesis
The Wistar Institute
Philadelphia, Pennsylvania, USA

Yu Lin
Department of Surgery
University of Michigan Medical Center
Ann Arbor, USA

Xuefeng Liu
Comprehensive Cancer Center
Departments of Pathology, Urology and Radiation Oncology
Wexner Medical Center
Ohio State University
Columbus, USA

Stéphane Loubrie
Department of Radiology
University of California San Diego
La Jolla, USA

David M. Lubman
Department of Surgery
University of Michigan Medical Center
Ann Arbor, USA

Man Luo
Mayo Clinic
Phoenix, Arizona, USA

Janvhi Suresh Machhar
Program in Molecular and Cellular Oncogenesis
The Wistar Institute
Philadelphia, Pennsylvania, USA

Florencia Madorsky Rowdo
Englander Institute for Precision Medicine
Weill Cornell Medical College
New York, USA

Gopi Maguluri
Physical Sciences, Inc.
Andover, Massachusetts, USA

Shyamala Maheswaran
Krantz Family Center for Cancer Research
Massachusetts General Hospital Cancer Center and Harvard Medical School
Charlestown, USA

Kenneth Marcelino
James C. Wyant College of Optical Sciences
University of Arizona
Tucson, USA

Hailee Mayer
Department of Radiology
Thomas Jefferson University
Philadelphia, Pennsylvania

Bryant L. McCue
Lester and Sue Smith Breast Center
Baylor College of Medicine
Houston, Texas, USA

Paromita Mehta
ProMobile BI
San Diego, California, USA

David T. Melnekoff
Division of Hematology and Medical Oncology
Icahn School of Medicine at Mount Sinai
New York, USA

Jesus Anampa Mesias
Clinical Therapeutics Program
Montefiore-Einstein Comprehensive Cancer Center
Bronx, USA
and
Department of Oncology
Albert Einstein College of Medicine
Bronx, New York, USA

George Miles
Lester and Sue Smith Breast Center
Department of Molecular and Human Genetics
Baylor College of Medicine
Houston, Texas, USA

Avanish Mishra
Center for Engineering in Medicine and Surgery
Massachusetts General Hospital and Harvard Medical School
Boston, USA

Sarah Mitchell
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering, School of Medicine
Oregon Health and Science University
Portland, USA

David T. Miyamoto
Krantz Family Center for Cancer Research
Massachusetts General Hospital Cancer Center and Harvard Medical School
Charlestown, USA
and
Massachusetts General Hospital and Harvard Medical School
Boston, USA
and
Broad Institute of MIT and Harvard
Cambridge, USA

Jonathan P. Mochel
Precision One Health Initiative
Department of Pathology
University of Georgia College of Veterinary Medicine
Athens, USA

Raymond Moore
Department of Quantitative Health Sciences
Mayo Clinic
Rochester, Minnesota, USA

Alison Mosichuk
Department of Biomedical Sciences
College of Veterinary Medicine
Iowa State University
Ames, USA

Eve M. Nagler
Center for Community-Based Research
Dana-Farber Cancer Institute
Boston, USA
and
Department of Social and Behavioral Sciences
Harvard T.H. Chan School of Public Health
Boston, Massachusetts, USA

Kibo Nam
Department of Radiology
Thomas Jefferson University
Philadelphia, Pennsylvania, USA

Wendy Nevala
Department of Oncology
Mayo Clinic
Rochester, Minnesota, USA

Ethel J. Ngen
Russell H. Morgan Department of Radiology and Radiological Science
Johns Hopkins University School of Medicine
Baltimore, Maryland, USA

Shelby Nicholas
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Maja H. Oktay
Gruss-Lipper Biophotonics Center
Integrated Imaging Program for Cancer Research
Department of Pathology; Department of Surgery
Albert Einstein College of Medicine
Bronx, USA
and
Tumor Microenvironment and Metastasis Program
Cancer Dormancy Institute
Montefiore-Einstein Comprehensive Cancer Center
Bronx, New York, USA

Aytekin Oto
Department of Radiology
Sanford J. Grossman Center of Excellence in Prostate Imaging and Image Guided Therapy
University of Chicago
Chicago, Illinois, USA

Vikas Pandey
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York, USA

Samir Parekh
Tisch Cancer Institute
Division of Hematology and Medical Oncology
Icahn School of Medicine at Mount Sinai
New York, USA

Neehar D. Parikh
Department of Internal Medicine
University of Michigan
Ann Arbor, USA

Anil V. Parwani
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Ohio State University
Columbus, USA

Bhavik Patel
Mayo Clinic
Phoenix, Arizona, USA

Vanisha Patel
Division of Gastroenterology
Department of Internal Medicine, Wexner Medical Center
Ohio State University
Columbus, USA

Mangesh Pednekar
Healis Sekhsaria Institute for Public Health
Navi Mumbai, India

Thasin Peyear
Department of Radiology; Molecular Pharmacology Program
Memorial Sloan Kettering Cancer Center
New York, USA

Robert Philibert
Department of Biomedical Engineering
Department of Psychiatry
University of Iowa
Iowa City, USA
and
Behavioral Diagnostics LLC
Coralville, USA

Winter Philibert
Department of Biomedical Engineering
University of Iowa
Iowa City, USA

Roberta Pigliapochi
Department of Radiology
Molecular Pharmacology Program
Memorial Sloan Kettering Cancer Center
New York, USA

Rea Pittie
Krantz Family Center for Cancer Research
Massachusetts General Hospital Cancer Center and Harvard Medical School
Charlestown, USA
and
Department of Radiation Oncology
Massachusetts General Hospital and Harvard Medical School
Boston, USA

Rajas Poorna
Department of Chemical and Biomolecular Engineering
Georgia Institute of Technology
Atlanta, USA

Edwin Posadas
Division of Medical Oncology, Department of Medicine
Cedars-Sinai Medical Center
Los Angeles, California, USA

Anastasia Prokofyeva
Department of Oncological Sciences
University of Utah
Salt Lake City, USA

Victor Putaturo
Center for Engineering in Medicine and Surgery
Massachusetts General Hospital and Harvard Medical School
Boston, USA

Nagireddy Putluri
Department of Molecular and Cellular Biology
Baylor College of Medicine
Houston, Texas, USA

Saif Ragab
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York, USA

Milind Rajadhyaksha
Memorial Sloan Kettering Cancer Center
New York, USA

Rebecca Rakow-Penner
Department of Radiology
Department of Bioengineering
University of California San Diego
La Jolla, USA

Meghana Ram
Department of Oncological Sciences
Icahn School of Medicine at Mount Sinai
New York, USA

Abdul Qawee Rani
Comprehensive Cancer Center
Ohio State University
Columbus, USA

Ken Resnicow
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
and
University of Michigan School of Public Health
Ann Arbor, USA

C. Patrick Reynolds
Department of Pediatrics and Cancer Center
School of Medicine
Texas Tech University Health Sciences Center
Lubbock, USA

Bhargava, Rohit
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign Urbana, IL, USA.
and
Departments of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign Urbana, IL, USA.
and
Chemical and Biomolecular Engineering, and Chemistry
University of Illinois at Urbana-Champaign Urbana, IL, USA.
and
Bioengineering, University of Illinois at Urbana-Champaign
Urbana, IL, USA.
and
Mechanical Science and Engineering University of Illinois at Urbana-Champaign Urbana, IL, USA.
and
Cancer Center at Illinois
University of Illinois at Urbana-Champaign
Urbana, IL, USA.

Ana E. Rodríguez-Soto
Department of Radiology
University of California San Diego
La Jolla, USA

Thomas E. Rohan
Cancer Epidemiology, Prevention and Control Program
Montefiore-Einstein Comprehensive Cancer Center
Bronx, USA
and
Department of Epidemiology and Population Health
Albert Einstein College of Medicine
Bronx, New York

Rafael Romero
Department of Biomedical Engineering
University of Arizona
Tucson, USA

Jatin Roper
Division of Gastroenterology, Department of Medicine
Duke University School of Medicine
Durham, North Carolina, USA

Max P. Rosen
Department of Radiology
UMass Chan Medical School
Worcester, USA

Christian Ross
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Da Yeon Ryoo
Department of Internal Medicine
Wexner Medical Center
Ohio State University
Columbus, USA

Corey F. Saba
Department of Small Animal Medicine and Surgery
College of Veterinary Medicine
University of Georgia
Athens, Georgia, USA

Ksenia Safina
Division of Hematology
Brigham and Women's Hospital
Boston, USA
and
Broad Institute of MIT and Harvard
Cambridge, USA
and
Department of Medicine
Ludwig Center at Harvard
Harvard Medical School
Boston, Massachusetts, USA

Dipak Kumar Sahoo
Department of Biomedical Sciences
College of Veterinary Medicine
Iowa State University
Ames, USA

Aditi Sahu
Memorial Sloan Kettering Cancer Center
New York, USA

Rakiya Saidu
Department of Obstetrics and Gynaecology
Faculty of Health Sciences
University of Cape Town
Cape Town, South Africa

Usama Sajjad
Center for Artificial Intelligence Research
Wake Forest University School of Medicine
Winston-Salem, North Carolina, USA

Sanjeev Sariya
Department of Oncological Sciences
Icahn School of Medicine at Mount Sinai
New York, USA

Mehrzad Sasanpour
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Real Sumayya Abdul Sattar
Rutgers Cancer Institute of New Jersey
Rutgers University
New Brunswick, USA

Jill Schimke
Department of Oncology
Mayo Clinic
Rochester, Minnesota, USA

Jennifer L. Schneider
Science Program Department
Kaiser Permanente Center for Health Research
Portland, Oregon, USA

Komal Sethi
Virginia Tech-Wake Forest University School of Biomedical Engineering and Sciences
Wake Forest School of Medicine
Winston-Salem, North Carolina, USA

Jason Shames
Department of Radiology
Thomas Jefferson University
Philadelphia, Pennsylvania, USA

Delaney Shea
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering, School of Medicine
Oregon Health and Science University
Portland, USA

Catherine Sherry
Department of Cellular and Molecular Physiology
Albany Medical College
Albany, New York, USA

Chuck Sigmund
ProMobile BI
San Diego, California, USA

Amit G. Singal
Department of Internal Medicine
University of Texas Southwestern Medical Center
Dallas, USA

Allison Solanki
Biomedical Engineering Department
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Akshay Sood
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Department of Urology
Wexner Medical Center
Ohio State University
Columbus, USA

Federico Sorcini
Department of Radiology
UMass Chan Medical School
Worcester, USA

Joseph A. Sparano
Tisch Cancer Institute
Icahn School of Medicine at Mount Sinai
New York, USA

Ella Stimson
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering, School of Medicine
Oregon Health and Science University
Portland, USA

Elena M. Stoffel
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
and
Department of Internal Medicine
University of Michigan School of Medicine
Ann Arbor, USA

Liangliang Sun
Department of Chemistry
Michigan State University (MSU)
East Lansing, Michigan, USA

Ziyu Su
Center for Artificial Intelligence Research
Wake Forest University School of Medicine
Winston-Salem, North Carolina, USA

Momoka Sugimura
James C. Wyant College of Optical Sciences
University of Arizona
Tucson, USA

Shankar Suman
Department of Oncology
Mayo Clinic
Rochester, Minnesota, USA

Dali Sun
Biomedical Focus, Department of Electrical and Computer Engineering
Knoebel Institute for Healthy Aging
Ritchie School of Engineering and Computer Science
The University of Denver
Denver, Colorado, USA

Ulas Sunar
Department of Biomedical Engineering
Stony Brook University
Stony Brook, New York, USA

Debasish Sundi
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Department of Urology, Wexner Medical Center
Ohio State University
Columbus, USA

Markovic Svetomir
Department of Oncology
Mayo Clinic
Rochester, Minnesota, USA

Amara Tariq
Mayo Clinic
Phoenix, Arizona, USA

Savantha Thenuwara
Department of Biomedical Sciences, College of Veterinary Medicine
Iowa State University
Ames, USA

Mehmet Toner
Center for Engineering in Medicine and Surgery
Massachusetts General Hospital and Harvard Medical School
Boston, USA

Duy Binh Tran
Department of Surgery, Division of Surgical Oncology
University of Illinois College of Medicine
Chicago, USA

Shubham Trivedi
Mayo Clinic
Phoenix, Arizona, USA

Hsian-Rong Tseng
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California, Los Angeles
Los Angeles, USA

Aisha Urooj
Mayo Clinic
Phoenix, Arizona, USA

Peter Van Galen
Division of Hematology
Brigham and Women's Hospital
Boston, USA
and
Broad Institute of MIT and Harvard
Cambridge, USA
and
Department of Medicine; Ludwig Center at Harvard
Harvard Medical School
Boston, Massachusetts, USA

Amit Verma
Department of Cellular and Molecular Physiology
Albany Medical College
Albany, New York, USA

Ian D. Wadsworth
Virginia Tech-Wake Forest University School of Biomedical Engineering and Sciences
Wake Forest School of Medicine
Winston-Salem, North Carolina, USA

Denong Wang
SRI International Biosciences Division
Menlo Park, California, USA

Jeremy R. Wang
Department of Pathology and Laboratory Medicine, Department of Genetics
University of North Carolina School of Medicine
Chapel Hill, USA

Lei G. Wang
Biomedical Engineering Department
Knight Cancer Institute
Oregon Health and Science University
Portland, USA

Peng Wang
Comprehensive Cancer Center
Departments of Pathology, College of Medicine
Department of Medicine, Wexner Medical Center
Ohio State University
Columbus, USA

Qi-en Wang
Comprehensive Cancer Center
Departments of Pathology and Radiation Oncology
College of Medicine
Ohio State University
Columbus, USA

Shang-jui Wang
Comprehensive Cancer Center
Departments of Pathology and Radiation Oncology
College of Medicine
Ohio State University
Columbus, USA

Caitlin Ward
Division of Biostatistics and Health Data Science
University of Minnesota
Minneapolis, USA

Jason Ware
Cancer Early Detection Advanced Research Center
Knight Cancer Institute
Department of Biomedical Engineering, School of Medicine
Oregon Health and Science University
Portland, USA

Smita P. Warke
Center for Community-Based Research
Dana-Farber Cancer Institute
Boston, Massachusetts, USA

Adam Wax
Department of Biomedical Engineering
Duke University
Durham, North Carolina, USA

Shayna Weiner
Rogel Cancer Center
University of Michigan
Ann Arbor, USA
Michigan Oncology Quality Consortium (MOQC)
Ann Arbor, USA

Hannah Wickham
Department of Biomedical Sciences, College of Veterinary Medicine
Iowa State University
Ames, USA

Yakir Winograd
School of Computer Science and Engineering
The Hebrew University of Jerusalem
Jerusalem, Israel

Yun Wu
Department of Biomedical Engineering
University at Buffalo, The State University of New York
Buffalo, USA

Gu Xiao
Department of Biological Sciences
Hunter College, City University of New York
New York, USA

Tian Xu
Dr. Neil L. Kelleher's Research Group Northwestern University
Evanston, Illinois, USA

Tohru Yamada
Department of Surgery, Division of Surgical Oncology
University of Illinois College of Medicine
Chicago, USA
and
Richard and Loan Hill Department of Biomedical Engineering
University of Illinois College of Engineering
Chicago, USA

Ju Dong Yang
Karsh Division of Gastroenterology and Hepatology
Samuel Oschin Comprehensive Cancer Institute
Comprehensive Transplant Center
Cedars-Sinai Medical Center
Los Angeles, USA

Anna Yaxuan Wei
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California, Los Angeles
Los Angeles, USA

Kevin Yeh
Beckman Institute for Advanced Science and Technology
University of Illinois at Urbana-Champaign
Urbana, USA

John H. Yim
Department of Surgery
City of Hope Comprehensive Cancer Center
Duarte, California, USA

Angela J. Yoon
Medical University of South Carolina
Charleston, USA

Sungyong You
Departments of Urology and Computational Biomedicine
Cedars-Sinai Medical Center
Los Angeles, USA

Mohamed Yousuf
Computer Vision and Image Processing Laboratory (CVIP)
University of Louisville
Louisville, Kentucky, USA

Nathan Yu
Mayo Clinic
Phoenix, Arizona, USA

Nanxue Yuan
Department of Biomedical Engineering
Rensselaer Polytechnic Institute
Troy, New York, USA

Thomas A. Zangle
Huntsman Cancer Institute
Department of Chemical Engineering
University of Utah
Salt Lake City, USA

Christopher Zdyrski
Precision One Health Initiative
Department of Pathology
University of Georgia College of Veterinary Medicine
Athens, USA

Brian M. Zeglis
Department of Chemistry
Hunter College, City University of New York
New York, USA
and
Program in Chemistry
The Graduate Center of the City University of New York
New York, USA
and
Department of Radiology
Memorial Sloan Kettering Cancer Center
New York, USA

Jie Zhang
Department of Surgery
University of Michigan Medical Center
Ann Arbor, USA

Jingwei Zhao
James C. Wyant College of Optical Sciences
University of Arizona
Tucson, USA

Jianhui Zhu
Department of Surgery
University of Michigan Medical Center
Ann Arbor, USA

Yazhen Zhu
Department of Pathology and Laboratory Medicine
David Geffen School of Medicine at UCLA
Los Angeles, USA
and
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California, Los Angeles
Los Angeles, USA

Luisa Quezada Ziesse
Rutgers Cancer Institute of New Jersey
Rutgers University
New Brunswick, New Jersey, USA

Jacqueline Ziqian Yang
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California, Los Angeles
Los Angeles, USA

Rebecca G. Zitnay
Huntsman Cancer Institute
Department of Chemical Engineering
University of Utah
Salt Lake City, USA

Vivian Xufei Zuo
California NanoSystems Institute
Crump Institute for Molecular Imaging
Department of Molecular and Medical Pharmacology
University of California, Los Angeles
Los Angeles, USA

Overview of Cancer Features and Impact

Miguel Ossandon, Ben Prickril, and Avraham Rasooly

1. **Cancer Health Impact and Epidemiology**: Cancer is a global epidemic; worldwide in 2022, there were 20 million new cancer cases and 9.7 million cancer-related deaths, accounting for approximately 17% of annual deaths worldwide [1]. According to the US Centers for Disease Control and Prevention (CDC), cancer is the second leading cause of death in the US. In 2022, there were 1,777,566 new cancer cases and 608,366 deaths [2]. The current global number of new cases of cancer is 20 million, but it is predicted to reach 35 million by 2050 [3], resulting in 18.5 million cancer deaths projected by 2050, an 89.7% increase from the 2022 estimate of 9.7 million.

   Cancer rates are higher in developed countries: In 2022, the age-standardized rate for all cancers in areas with very high human development was 285.7 per 100,000, compared to 110.6 per 100,000 in areas with low human development.

   However, the burden of disease is shifting to less affluent countries. In Africa in the 1950s, there were a very small number of cases [4]. Although the small number may reflect a lack of monitoring, in 2020, the number of new cancer cases rose to 1.1 million, with 711,429 cancer deaths. By 2040, the number of cancer cases and deaths is expected to double [5].

   Cancer cases are expected to increase in low Human Development Index (HDI) countries: Between 2020 and 2040, cancer cases are expected to increase by 96% in low HDI countries, compared to 32.2% in very high HDI countries [6]. Approximately 47% of cancer cases each year now occur in the developing world. Cancer, a disease once associated with affluence, now places its heaviest burden on poor and disadvantaged populations [3].

   By 2050, 35.3 million cancer cases worldwide are expected, a 76.6% increase from the 2022 estimate of 20 million. Similarly, 18.5 million cancer deaths are projected by 2050, an 89.7% increase from the 9.7 million estimated in 2022.

   Cancer cases and deaths numbers are projected to nearly triple in low-HDI countries by 2050 [6], compared to a moderate increase in very high–HDI countries, 142.1% vs 41.7% for cancer cases, and 146.1% vs 56.8% for cancer deaths. Males had a higher incidence and greater number of deaths in 2022 than females, with this disparity projected to widen by up to 16.0% by 2050.

   In 2022, the mortality-to-incidence ratio (MIR) for all cancers was 46.6%, with higher MIRs observed for pancreatic cancer (89.4%), among males (51.7%), among those aged 75 years or older (64.3%), in low-HDI countries (69.9%), and in the African region (67.2%) [6].

   Cancer early detection and diagnosis is a key for cancer survival. To address global heath cancer challenges in low-HDI countries, new low-cost cancer detection and diagnosis technologies are needed; this book describes such technologies.

2. **Cancer basic features**: The US National Cancer Institute (NCI) defines cancer as a group of more than 100 diseases in which abnormal cells divide without control. Cancer cells can invade nearby tissues and spread through the bloodstream and lymphatic system to distant body parts [7]. Cancers have high levels of complexity, and each cancer has its own genetic, epigenetic, molecular, and physiological characteristics. Cancers are known to affect more than 79 human organs or tissue types [8] and represent the second leading cause of death in developed countries after cardiovascular diseases [4].

   Somatic mutations, theory of aging: A leading hypothesis about aging is that mutations accumulate in the genetic material of somatic cells over time, resulting in a decline in cellular function [9]. Mutation accumulation doesn't appear to accelerate at an older age, but it is affected by environmental factors. Cancer arises when genetic changes disrupt the normal regulatory mechanisms of cell division, leading to uncontrolled growth and the formation of tumors. Cancers are caused by mutations that may be classified as follows [10]:

   - Random mutations that arise during DNA replication (66% of cancers)
   - DNA alteration from exposure to environmental factors (29%)
   - Inherited mutations (5%)

   Once cancers arise, uncontrolled growth and cell division arise; cancer cells can then invade surrounding tissue and metastasize, spreading from the initial

DOI: 10.1201/9781003449942-1

site to remote parts of the body via the bloodstream or lymphatic system. Cancer cells often exhibit characteristics that allow them to overcome systems designed to protect the body. For example, cancers can evade the immune system, thereby avoiding elimination by B-cells, T-cells, and other immune-system components. Moreover, cancer cells have an altered profile of the secreted substances that normally regulate the position of cells in their proper locations within tissues or organs, so the cancer cells can travel to distant regions of the body. Finally, many cancer cells have modified glucose metabolism, which allows them to sustain high rates of proliferation.

Factors contributing to cancer development: An estimated 50% of all premature deaths are caused by or linked to lifestyle [11, 12]. Current estimates are that 30%–50% of cancers are preventable by reducing environmental and lifestyle risk factors [13]. Indeed, in 2019 an estimated 40% of all incident cancers in adults aged 30 years and older in the United States were attributable to lifestyle-modifiable risk factors [14], such as smoking, unhealthy diet [15], long sun exposure, and environmental carcinogens [16]. Environmental exposures caused cancer cells to become progressively dysregulated, eventually leading to the uncontrolled cell proliferation characteristic of cancer [17]. A healthy diet comprising mostly plants, limited red meat, and low alcohol consumption can significantly decrease the risk of various types of cancer including those of breast, colon, lung, kidney, and liver. This is a result of the decrease of negative epigenetic modifications such as DNA methylation by "turning off" genes associated with cancer promotion [18]. Early detection and diagnosis before the cancer metastasizes greatly increase the chances of successful treatment and survival [19], extending the life span and life quality while reducing cancer's social and economic burden. Thus, there is a need to develop technologies to aid in the detection and diagnosis of cancers.

3. **Genome-wide perspective of cancer**: Sequencing and bioinformatics analysis provide a genome-wide perspective of cancer at single-nucleotide resolution. These techniques reveal that tumor cells typically harbor an average of 3,000 more mutations compared to normal cells in the same individual, equivalent to about one mutation per million nucleotides. Of these mutations, ~300 are included in coding sequences (exons), approximately 10% of which (30–60) are capable of altering protein products [20]. On average, cancer genomes contain four to five driver mutations within coding and non-coding genomic elements. Interestingly, in ~5% of cases no driver mutations have been identified, suggesting that the cancer driver may not yet have been discovered [21]. Different cancers are associated with different levels of mutation. For example, the median number of non-synonymous mutations varies from several mutations in acute lymphoblastic leukemia to hundreds in melanoma and lung cancer, and these are correlated with mutagen exposure such as UV exposure or smoking [20].
4. **Cancer epigenetic changes**: In addition to genetic alterations, epigenetic changes are present in all human cancers [22]. Epigenetic modifications include methylation of cytosine bases in DNA, post-translational modifications of histone proteins, and the positioning of nucleosomes along the DNA.

   While cancers may involve from 50 to several hundred genes with genetic mutation (e.g., "hard drives"), epigenetics is the "software" [23] that causes these genes to express in a way that can lead to cancer development. Epigenetic changes play a significant role in cancer development by altering gene expression without changing the DNA sequence. Similar to genetic modification, environmental factors including diet and lifestyle and pharmacological factors such as the use of sodium valproate can induce epigenetic changes [23]. However, there is no precise estimate of their contribution to cancer.
5. **Factors that affect the risk of developing cancer**:

   While cancer risk factors are mainly age dependent, they vary according to cancer type. According to data from the NCI Surveillance, Epidemiology, and End Results (SEER) Program, the median age of a general cancer diagnosis is 66 years. The median age at diagnosis is 62 years for breast cancer, 67 years for colorectal cancer, 71 years for lung cancer, and 66 years for prostate cancer. In 2019, an estimated 40.0% of all incident cancers in adults aged 30 years and older in the United States were attributable to lifestyle-modifiable risk factors, including smoking, unhealthy diet, and sun exposure [14]. These factors progressively dysregulate the genome, eventually leading to the uncontrolled proliferation of cancer [17]. However, a healthy lifestyle including a balanced diet, regular exercise, and so on can significantly decrease negative epigenetic modifications such as DNA methylation by "turning off" genes associated with cancer promotion [18].
6. **Cancer-detection and diagnosis monitoring technologies**:

   New technologies have the potential to address cancer challenges, including fostering improvements in early screening and detection technologies capable of reducing or mitigating the impact of many cancers.

   The ultimate effectiveness of these technologies is bolstered by an understanding of basic cancer biology, tumor features, and potential detection targets. This book is designed to make these technologies more accessible and understandable to molecular biologists, oncologists, pathologists, engineers, and others playing major roles in the molecular diagnosis

of cancer. For readers who may be less familiar with cancer, we include in this chapter some basic information and terminology related to cancer.

The development of cancer detection and diagnostic technologies relies on:

a. Cancer's unique features, including cancer molecular markers (e.g., specific DNA mutations, protein modifications, and altered gene expression).

b. Technologies to interrogate specifically those markers. Technologies such as molecular imaging, genomics, proteomics, liquid biopsy, and artificial intelligence enable detailed cancer analysis and identification at earlier stages based on cancer-specific features. These new technologies may allow for more accurate detection and improved personalized treatment strategies.

Cancer detection, diagnosis, and monitoring techniques exploit specific tumor features such as specific biomarkers and genomic features to detect, identify, and target tumors.

7. **Cancer biomarkers and cancer detection and diagnosis**: The NCI defines a biomarker as "A biological molecule found in blood, other body fluids, or tissues that are signs of a normal or abnormal process, or a condition or disease." A biomarker may be used for cancer detection, diagnosis, prognosis, monitoring, or evaluation of treatment effectiveness as well as the body's response to treatment and disease. Samples for biomarker analysis may include blood, saliva, stool, and urine. Biomarkers can be DNA, RNA, proteins, or metabolites.
   - Examples of gene fusion cancer biomarkers include FGFR2, NTRK, ALK, RET, ROS1, and NRG.
   - Examples of protein cancer biomarkers include PSA, CEA, CA19-9, CA125, and LDH.
8. **Genomic-based cancer diagnosis**: Sequencing data can be used in cancer diagnosis. For example, Oncotype DX [24, 25] is a genomic test used for early-stage ER-positive and HER2-negative breast cancer. Oncotype DX analyzes the activity level of 21 genes in a tumor sample to determine the likelihood of recurrence within 10 years and evaluate whether chemotherapy could be beneficial to some early-stage breast cancer patients.

   Circulating tumor DNA (ctDNA) based cancer diagnosis: CtDNA comprises tumor cell DNA fragments circulating in the blood that can be utilized as a "liquid biopsy." CtDNA can be used for cancer detection, diagnosis, and prognosis. A major advantage of ctDNA over existing screening strategies is that the tests are typically minimally invasive and have the potential to detect multiple cancer types in a single test [26]. CtDNA is used in several cancer diagnostic tests including the following:
   - **Multi-cancer early detection (MCED) Galleri**: This test detects cancer-specific DNA methylation patterns covering more than 100,000 DNA regions from circulating cell-free DNA (cfDNA), which potentially regulates tumorigenesis [27]. **CancerSEEK** [28] is a test for detection of alterations in 16 genes associated with cancer, as well as the levels of nine proteins that are overproduced by some cancers, enabling high specificity (> 99.0%), although sensitivity varies depending on the stage and type of cancer. The CancerSEEK test was accurate for 626 of 1005 participants with cancer, with an overall sensitivity of 62.3%, and it correctly predicted the site of origin in 62.9% of true positives [27]. **FoundationOne®LiquidCDx** [29] is an FDA-approved companion diagnostic blood-based test that analyzes over 300 genes including blood tumor mutational burden (bTMB), microsatellite instability-high (MSI-H), and tumor fraction values (the amount of ctDNA in a liquid biopsy sample).
   - **CosmoDA** [27] is a differential abundance (DA) test that can reduce the number of false positives that are caused by correlated features. It uses penalized generalized score matching to fit models in a parsimonious way [30]. CosmoDA can accurately estimate feature interactions in populations that are heterogeneous, and it can significantly reduce the false discovery rate [31].
9. **Tumor features that can be explored for cancer detection and diagnosis**:

   Tumor cell interaction: In addition to generic features, tumor types, and associated molecules, cancer targets may include tumor interactions and tumor-associated cellular features. Tumor cells interact with themselves and with non-cancerous host cells in the tumor microenvironment (TME).
   - **Tumor microenvironment (TME)**: A non-cellular network of proteins and other molecules. The TME includes factors promoting inflammation and it is considered "the incubator of the tumor" [32]. The TME also provides nutrients for the tumor and provides space for the cancer to grow [33], making it a potential target for cancer therapeutics and diagnosis [34–37].
   - **Tumor immune macroenvironment (TIMaE)**: Distant systemic immune compartments such as the bone marrow, spleen, and circulating immune cells that interact with the tumor [38].
   - **Extracellular matrix (ECM) interactions** [39]: Complex and dynamic interactions impact many aspects of cancer biology and regulate cancer's cellular processes. The ECM interacts with adhesion receptors such as integrins to regulate many cell functions, including survival, proliferation, migration, and differentiation [40]. Some therapies also eliminate tumors; for example, tumor cells often express checkpoint molecules (e.g., PD-L1, CTLA-4) that inhibit T-cell activation and prevent immune destruction.

- **Macrophages and dendritic cells (DCs)**: These immune cells play a role in cancer targeting and can be used in the stromal immune system. For example, these include macrophages, dendritic cells, and natural killer cells [41], targeting specific markers on stromal cells such as cancer-associated fibroblasts. Current research focuses on developing drugs that specifically bind to those markers to disrupt their function or induce their destruction utilizing antibodies, small molecule inhibitors, or other targets.
- **Interactions with the immune system**: Tumor cells interact with the immune system in ways that can be both pro- and anti-tumorigenic [42]. For example, the immune system can promote cancer cell growth and metastasis by creating inflammation in the TME. However, immune cells like dendritic cells, natural killer cells, and macrophages also have anti-tumorigenic functions.
- **Non-malignant cells interacting with tumor cells**: Tumor cells interact with non-malignant cells in the TME; these non-malignant cells can be genomically stable but can also change based on interactions with tumor cells [43]. For example, fibroblasts can be rewired into cancer-associated fibroblasts (CAFs).
- **Tumor signaling**: During the interactions with TIMaE, cancer evolves and adapts to the body's environment by using complex signals to interact with other cancer cells, body tissues, and the immune system [44], leading to dysfunction and system-level failure. Such complexity determines the response to different types of anticancer treatments and drugs, which require an array of technologies for cancer detection, diagnosis, and treatment.
- **Circulating tumor cells (CTCs)**: CTCs are cancer cells that break away from a primary tumor and enter the bloodstream [45]. CTCs can travel through the blood and form new tumors in other parts of the body. CTCs can also interact with immune factors, promoting cancer growth and metastasis by creating inflammation [46]. CTC interactions with the ECM can promote tumor development but can also activate anti-tumor immune responses. Stem cell-like and epithelial-mesenchymal transition-positive (EMT-positive) cells display higher invasion and migratory potentials and display resistance to anticancer therapies. Moreover, following targeted therapy of colorectal cancer, CTCs with self-renewal ability act as metastasis-initiating cells (MICs), causing tumor relapse and loss of tissue homeostasis [47]. Tumor cells also interact with stromal cells to create an environment that supports tumor progression. CTCs may have potential as targets for cancer detection and diagnosis as a liquid biopsy.

10. **Cancer economic impact:** The overall cost of cancer diagnosis and treatment in the US in 2015 was $183 billion and is projected to increase 34% to $246 billion by 2030 [48]. More recent data suggest that the United States spends approximately $200 billion per year on cancer care, and patients spend $20 billion yearly in out-of-pocket costs [27]. In 2020, the estimated national expenditures for cancer care in the US were $208.9 billion [49]. Globally, the cost of cancer from 2020 to 2050 is estimated to be $25 trillion, which is equivalent to approximately 0.55% of the annual global GDP [50]. This data demonstrates the great negative impact of cancer on public health and health economics. New and effective early cancer detection and diagnosis technologies may reduce this impact.
11. **Cancer basic terminology**

    Cancer: A disease that causes cells to divide uncontrollably, leading to the formation of cancerous tumors. Cancer cells can spread to distant organs through the bloodstream or lymph system (metastasis).

    Tumor: An abnormal growth of tissue caused by uncontrolled cell growth. Tumors can be benign (noncancerous) or malignant (cancerous). Benign tumors often grow slowly and don't spread, while malignant tumors divide rapidly and invade other organs.

    Solid tumors: A discrete mass of cancer cells in an organ system.

    Liquid tumors: Cancers of bodily fluids. Unlike solid tumors, which form masses in organs or tissues, liquid tumors are cancerous cells that float freely in the blood, bone marrow, or lymphatic system. Blood cancers include leukemia, lymphoma, and myeloma.

    Circulating tumor cells: CTCs are cancer cells that break away from a primary tumor and enter the bloodstream.

    Metastasis: Spread of cancer cells from the original tumor to other organs or parts of the body. Cancer metastasizes through the bloodstream or lymph system forming metastatic tumors that can disrupt organ function. Metastasis is the leading cause of cancer deaths.

    Carcinomas: Tumors originating in tissues that line or cover internal organs, such as the kidney, liver, and thyroid.

    Fibro-adenoma: A common example of a mixed tumor in the breast, where both components of normal breast tissue are increased.

    Neuroblastomas: Tumors of the nerve cells that can develop in many parts of the body.

    Retinoblastomas: Childhood tumors originating in the retina at the back of the eye.

    Sarcomas: Cancerous tumors that affect bone and soft tissue.

Mixed tumor: Tumors that contain multiple types of neoplastic or abnormal cells. Mixed tumors can be benign or malignant.

Mixed benign tumor: (pleomorphic adenoma) is made of epithelial and myoepithelial cells and cells that resemble cartilage cells. It's commonly found in the salivary glands.

Mixed cancers (mixed malignant tumors): Cancers involving two classifications or subtypes, such as carcinosarcoma and adenosquamous carcinoma. For example, carcinosarcoma combines adenocarcinoma and various sarcoma cells. It's most often found in the uterus and is also known as a malignant mixed Müllerian tumor.

Mixed germ cell tumor: A rare cancer that contains at least two different types of germ cell tumors, originating in cells that produce sperm or egg cells.

REFERENCES

1. Bray, F., et al., Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2024. 74(3): p. 229–263.
2. *Cancer Data and Statistics*. 2025; Available from: https://www.cdc.gov/cancer/data/index.html.
3. WHO. *Cancer in Developing Countries: Facing the Challenge*. 2010; Available from: https://www.who.int/director-general/speeches/detail/cancer-in-developing-countries-facing-the-challenge#:~:text=According%20to%20the%20latest%20WHO%20statistics%2C%20cancer,heaviest%20burden%20on%20poor%20and%20disadvantaged%20populations.
4. Davies, J.N. and B.A. Wilson, Cancer in Kampala, 1952–1953. *East Afr Med J*, 1954. 31(9): p. 395–416.
5. https://www.wcrf.org/preventing-cancer/cancer-statistics/cancer-rates-by-human-development-index/.2025;Available from: https://www.wcrf.org/preventing-cancer/cancer-statistics/cancer-rates-by-human-development-index/.
6. Zhou, J., et al., Global burden of lung cancer in 2022 and projections to 2050: Incidence and mortality estimates from GLOBOCAN. *Cancer Epidemiol*, 2024. 93: p. 102693.
7. NIH. *Curriculum Study. NIH Curriculum Supplement Series [Internet]*. 2007; Available from: https://www.ncbi.nlm.nih.gov/books/NBK20362/.
8. Wikipedia. *List of Organs of the Human Body.* Available from: https://en.wikipedia.org/wiki/List_of_organs_of_the_human_body#:~:text=It%20is%20widely%20believed%20that,status%20as%20one%20is%20debated.
9. Kennedy, S.R., L.A. Loeb, and A.J. Herr, Somatic mutations in aging, cancer and neurodegeneration. *Mech Ageing Dev*, 2012. 133(4): p. 118–126.
10. Tomasetti, C., L. Li, and B. Vogelstein, Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*, 2017. 355(6331): p. 1330–1334.
11. Nahon, P., et al., Genetic risk markers for hepatocellular carcinoma in patients with alcoholic liver disease. *Hepat Oncol*, 2015. 2(1): p. 63–78.
12. Mensah, E.O., E.K. Danyo, and R.V. Asase, Exploring the effect of different diet types on ageing and age-related diseases. *Nutrition*, 2025. 129: p. 112596.
13. Xu, J., R. Wang, and X. Guan, Enhancing insights into global cancer statistics 2022: Implications for cancer control. *Sci China Life Sci*, 2024: p. 294–296.
14. Islami, F., et al., Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. *CA Cancer J Clin*, 2018. 68(1): p. 31–54.
15. El-Sherif, A., S. El-Sherif, A.H. Taylor, and T. Ayakannu, Ovarian cancer: Lifestyle, diet and nutrition. *Nutr Cancer*, 2021. 73(7): p. 1092–1107.
16. Yabroff, K.R., et al., Annual report to the nation on the status of cancer, part 2: Patient economic Burden associated with cancer care. *JNCI: Journal of the National Cancer Institute*, 2021. 113(12): p. 1670–1682.
17. Quan, S. and H. Huang, Epigenetic contribution to cancer. *Int Rev Cell Mol Biol*, 2024. 387: p. 1–25.
18. Alegría-Torres, J.A., A. Baccarelli, and V. Bollati, Epigenetics and lifestyle. *Epigenomics*, 2011. 3(3): p. 267–77.
19. Chen, X., et al., Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. *Nat Commun*, 2020. 11(1): p. 3475.
20. Kim, S., New and emerging factors in tumorigenesis: An overview. *Cancer Manag Res*, 2015. 7: p. 225–239.
21. Campbell, P.J., et al., Pan-cancer analysis of whole genomes. *Nature*, 2020. 578(7793): p. 82–93.
22. Baylin, S.B. and P.A. Jones, Epigenetic determinants of cancer. *Cold Spring Harb Perspect Biol*, 2016. 8(9).
23. Kanherkar, R.R., N. Bhatia-Dey, and A.B. Csoka, Epigenetics across the human lifespan. *Front Cell Dev Biol*, 2014. 2: p. 49.
24. Ameri, M.A., et al., Oncotype DX in breast cancer management: Insights and outcomes from the United Arab Emirates. *Cureus*, 2024. 16(3): p. e56535.
25. Loggie, J., et al., Is oncotype DX testing informative for breast cancers with low ER expression? A retrospective review from a biomarker testing referral center. *Breast*, 2024. 75: p. 103715.
26. Duffy, M.J. and J. Crown, Circulating tumor DNA (ctDNA): Can it be used as a pan-cancer early detection test? *Crit Rev Clin Lab Sci*, 2024. 61(4): p. 241–253.
27. Cotner, C.E. and E. O'Donnell, Understanding the landscape of multi-cancer detection tests: The current data and clinical considerations. *Life (Basel)*, 2024. 14(7).
28. McCartney, M. and D. Cohen, Galleri promises to detect multiple cancers – but new evidence casts doubt on this much hyped blood test. *BMJ*, 2024. 386: p. q1706.
29. Guo, M.Z., et al., Targeted treatment of non-small cell lung cancer: Focus on capmatinib with companion diagnostics. *Onco Targets Ther*, 2021. 14: p. 5321–5331.
30. Ostner, J., H. Li, and C.L. Muller, Score matching for differential abundance testing of compositional high-throughput sequencing data. *bioRxiv*, 2024: p. 1–46.
31. Available from: https://doi.org/10.1101/2024.12.05.627006.
32. Denk, D. and F.R. Greten, Inflammation: The incubator of the tumor microenvironment. *Trends Cancer*, 2022. 8(11): p. 901–914.

33. Jin, M.Z. and W.L. Jin, The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct Target Ther*, 2020. 5(1): p. 166.
34. Yuan, Z., et al., Extracellular matrix remodeling in tumor progression and immune escape: From mechanisms to treatments. *Mol Cancer*, 2023. 22(1): p. 48.
35. Jiang, Y., et al., Targeting extracellular matrix stiffness and mechanotransducers to improve cancer therapy. *J Hematol Oncol*, 2022. 15(1): p. 34.
36. Valkenburg, K.C., A.E. de Groot, and K.J. Pienta, Targeting the tumour stroma to improve cancer therapy. *Nat Rev Clin Oncol*, 2018. 15(6): p. 366–381.
37. Xu, Z., et al., TME-targeting theranostic agent uses NIR tracking for tumor diagnosis and surgical resection and acts as chemotherapeutic showing enhanced efficiency and minimal toxicity. *Theranostics*, 2022. 12(6): p. 2535–2548.
38. Mottini, C., et al., The cross-talk between the macro and micro-environment in precursor lesions of pancreatic cancer leads to new and promising circulating biomarkers. *J Exp Clin Cancer Res*, 2024. 43(1): p. 198.
39. He, X., B. Lee, and Y. Jiang, Cell–ECM interactions in tumor invasion. *Adv Exp Med Biol*, 2016. 936: p. 73–91.
40. Kanchanawong, P. and D.A. Calderwood, Organization, dynamics and mechanoregulation of integrin-mediated cell-ECM adhesions. *Nat Rev Mol Cell Biol*, 2023. 24(2): p. 142–161.
41. Mun, J.Y., S.H. Leem, J.H. Lee, and H.S. Kim, Dual relationship between stromal cells and immune cells in the tumor microenvironment. *Front Immunol*, 2022. 13: p. 864739.
42. Abbott, M. and Y. Ustoyev, Cancer and the immune system: The history and background of immunotherapy. *Semin Oncol Nurs*, 2019. 35(5): p. 150923.
43. Lyssiotis, C.A. and A.C. Kimmelman, Metabolic interactions in the tumor microenvironment. *Trends Cell Biol*, 2017. 27(11): p. 863–875.
44. Cao, L.L. and J.C. Kagan, Targeting innate immune pathways for cancer immunotherapy. *Immunity*, 2023. 56(10): p. 2206–2217.
45. Fabisiewicz, A. and E. Grzybowska, CTC clusters in cancer progression and metastasis. *Med Oncol*, 2017. 34(1): p. 12.
46. Miklikova, S., et al., Inflammation-based scores increase the prognostic value of circulating tumor cells in primary breast cancer. *Cancers (Basel)*, 2020. 12(5).
47. Dianat-Moghadam, H., et al., The role of circulating tumor cells in the metastatic cascade: Biology, technical challenges, and clinical relevance. *Cancers (Basel)*, 2020. 12(4).
48. Mariotto, A.B., et al., Medical care costs associated with cancer survivorship in the United States. *Cancer Epidemiol Biomarkers Prev*, 2020. 29(7): p. 1304–1312.
49. NCI. *Molecular Testing of Gastrointestinal Tumours – Current Status and Future Prospects*. Available from: https://www.cancer.gov/about-cancer/understanding/statistics.
50. Chen, S., et al., Estimates and projections of the global economic cost of 29 cancers in 204 countries and territories from 2020 to 2050. *JAMA Oncol*, 2023. 9(4): p. 465–472.

Section 1

Optical Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

1

Near-Infrared Macroscopic and Mesoscopic Fluorescence Lifetime FRET Imaging to Measure Intra-Tumor Heterogeneity of Antibody-Target Engagement

Amit Verma, Saif Ragab, Shan Gao, Catherine Sherry, Nanxue Yuan, Vikas Pandey, Xavier Intes, and Margarida Barroso

1.1 Introduction

Numerous anti-cancer targeted therapies that act on receptors overexpressed at the cell surface of cancer cells have shown great promise in preclinical and clinical settings (1). However, emergent tumor resistance and microenvironment heterogeneity can decrease drug delivery, resulting in drug failure, followed by tumor relapse and worsening patient outcomes (2). Therefore, there is a critical need to develop tumor imaging methodologies at multiscale that can visualize tumor size, morphology, and location, as well as the intra-tumor heterogeneity of targeted therapy for the optimization of drug delivery, penetration, and binding. Currently, several radiologic and nuclear medicine-based imaging modalities like X-ray computed tomography (CT), ultrasonography, magnetic resonance imaging (MRI), positron emission tomography (PET) followed by minimally invasive biopsy (needle aspirations), and invasive (surgical) biopsy coupled with histopathological examination have been utilized to establish molecular and morphological characteristics of tumors in both clinical and preclinical research. However, none of these imaging modalities highlight the location where drug binding, interaction, and target-delivery occurs across the whole tumor. Recently, immuno-PET agents like anti-^{89}Zr-labeled human epidermal growth factor receptor (HER-3) monoclonal antibody GSK 2849330 and ^{89}Zr-lumretuzumab have been shown to provide a measure of drug tumor localization and dose selection in cancer patients (3–6). However, immuno-PET involves the use of harmful radioactivity and is limited to the usage of a single radiotracer. Multiplexed immuno-PET requires complex and sophisticated approaches, restricting the ability to visualize concomitantly various metabolic, molecular, and functional features of tumors. Moreover, immuno-PET is an intensity-based imaging methodology, which requires pharmacokinetic modeling to indirectly evaluate target engagement and uptake kinetics of radiolabeled probes in tumors. For preclinical research, optical imaging offers a new paradigm for the characterization of various types of tumors by allowing multiplexing to determine the drug efficacy and drug response in a longitudinal manner (7, 8). Measurement of fluorescence intensity-based optical imaging delivers the spatiotemporal distribution of the labeled antibody-drug conjugate but is limited to revealing the local accumulation of the labeled antibody-drug conjugate. Recently, development of fluorescence lifetime-based imaging (FLI) has allowed an in-depth functional and molecular analysis of the physiological environment of proteins, cells, and tissues (9–11). Importantly, FLI as an optical imaging-based technique can generate 3D spatially resolved and longitudinal measurements ranging from microscopic to macroscopic scale. Herein, we describe the use of FLI for the direct measurement of drug-target engagement and its biological implication in basic and preclinical research using small animal models of oncological disease (7, 11–14).

1.1.1 Near-Infrared (NIR) Fluorescence Lifetime Imaging (FLI)

Fluorescence lifetime is the time that a fluorophore takes to exponentially decay from excited to ground state, through the loss of energy via fluorescence and other non-radiative processes (10). Fluorescence lifetime is an intrinsic property of a fluorophore and does not depend on factors such as the method of measurement, intensity, illumination power, and photobleaching. However, micro-environmental factors such as pH, temperature, ion-concentration, polarity, and presence of quenchers may greatly affect the fluorescence lifetime (10). FLI measures the spatial distribution of the exponential temporal fluorescence decay of fluorophores at a resolution ranging from the picosecond to nanosecond time scale using fast electronics coupled with efficient photon detectors via time domain or frequency domain (15). In the time domain FLI, a short laser pulse excites the fluorophore, and time-resolved detectors bin the photons based on their arrival time, whereas in the frequency domain FLI technique there is a continuous excitation with amplitude modulation resulting in phase and amplitude shifts that are plotted over a range of modulation frequencies. Time-correlated single photon counting (TCSPC) is one of the most common methods of measuring lifetimes due to its high accuracy and signal-to-noise ratio compared to the frequency domain. The fluorescence lifetime (τ) for a mono-exponential decay can be calculated using two equal time gates at Δt time separation to get the rapid lifetime determination (RLD), which has similar accuracy of traditional nonlinear least square fitting (16).

$$\tau = \Delta t / \ln(I1/I2)$$

DOI: 10.1201/9781003449942-3

where I1 and I2 are the intensities measured at the two gates, respectively. The benefit of the two-gate technique is fast acquisition speed, and the intensity contrast generated through radiometric measurement provides better tissue characterization (17). For mono-exponential decay, fitting/phasor analysis can be used for retrieving the fluorescence lifetime. When using several fluorophores, multiple gates and exponential fitting/phasor approaches can be used for multi-exponential lifetime calculation (18, 19).

Fluorescence lifetime microscopy (FLIM) has been used for a variety of methodological applications such as contrast gain and multiplexing. Moreover, FLIM has been used to quantify protein-protein interactions, visualize the morphology of intracellular structures, estimate absolute membrane potential, and measure pH and other cellular processes (20, 21). Most importantly, FLI can be applied to a wide range of imaging technologies across scales. FLI has been combined with optical imaging to implement *in vivo* whole-body macroscopic FLI (MFLI), which has been recently established to provide molecular information related to the morphology, pH, enzymatic activity, drug-target interaction, and vascularity of live, intact tumor xenografts (9, 13, 22, 23). In contrast to FLIM, MFLI involves the acquisition and visualization of FLI data across an entire large area (~mm^2–cm^2) with main applications in whole-body small animal preclinical imaging. Most of the FLIM use fluorophores that are emitting in the visible wavelength range (400-700nm) due to their higher brightness and longer lifetimes. However, to implement whole-body MFLI, there is a need to use fluorophores emitting in the near-infrared range (NIR; > 700 nm), to allow for higher signal-to-noise ratio at increased penetration depth. The NIR window ranging from 700 to 1000 nm facilitates the acquisition of specific signals from deep-seated tumors, due to reduced attenuation of light propagation, tissue scattering, and auto fluorescence, while still enabling multiplexing imaging (24–26). NIR MFLI has been implemented within two major approaches, fluorescence reflectance imaging and fluorescence molecular tomography (FMT). Fluorescence reflectance imaging captures 2D images directly from a time-resolved camera system or raster scanning system. FMT requires the acquisition of multiple views of the sample to solve an inverse problem that aims at retrieving the 3D biodistribution of the fluorophores. FMT can be either performed in reflectance, typically for shallow tissues (a few millimeters deep, also referred to mesoscopic regime), or in transmission geometry for deep-seated tumors (a few centimeters) (27–30).

1.1.2 NIR FLI Förster Resonance Energy Transfer (FRET) Imaging

Förster resonance energy transfer (FRET) occurs when a donor fluorophore-labeled protein is in proximity (range 2–10 nm) to the acceptor-labeled protein and there is a non-radiative transfer of energy operating through dipole-dipole coupling (31). FRET occurrence reflects the proximity between donor and acceptor proteins, indicating changes in cellular/molecular processes such as protein-protein interaction, receptor dimerization, protease-induced protein cleavage, protein conformation—pH, ion, and kinase biosensors, and others (31). FRET measurements provide a more precise indication of spatial localization of protein-protein interactions, such as receptor dimerization and/or receptor-probe binding, than do fluorescence microscopy colocalization measurements using Pearson correlation coefficient, which is based on a pixel-by-pixel intensity correlation. Nevertheless, both colocalization and FRET are proximity assays and do not necessarily indicate a physical interaction between the two target proteins (32). Although multiple techniques have been shown to measure FRET events, different techniques can influence the accuracy of the FRET measurements (9).

FLI is widely regarded as the most robust means to quantify FRET events *in vitro* as well as *in vivo*, by measuring the reduction of the fluorescence lifetime of the donor fluorophore when near the acceptor fluorophore. FLI can be used to measure FRET, independently of the methodologies, scale, or wavelength range used to acquire and analyze fluorescence lifetime. FRET events result in the quenching of donor fluorescence leading to a reduction in the quantum yield and lifetime of the donor fluorophore (7, 15). The reduction in fluorescence lifetime of the donor fluorophore is measured by calculating the energy transfer efficiency (E) as follows:

$$E = 1 - \tau_{DA}/\tau_{D}$$

Where τ_{DA} is the donor lifetime in the presence of the acceptor and τ_{D} is the donor lifetime in the absence of the acceptor.

NIR MFLI has been used to measure drug-target engagement via FRET measurement of the binding of NIR-labeled probes to their respective targets *in vivo* (Figure 1.1A). Key to performing NIR MFLI FRET is the selection of two fluorophores with spectral overlap in the NIR window. Historically, indocyanine green (ICG) has been the fluorophore of choice for NIR fluorescence imaging. Indeed, ICG is an FDA-approved NIR dye that has been used, for example, for the cardiac/hepatic function test, chorioretinal fluorescence angiography, tumor/metastasis delineation, and lymphatic architecture and function (24, 33). Ogawa and colleagues demonstrated *in vivo* labeling of tumors using anti-HER2 monoclonal antibody trastuzumab (TZM) conjugated ICG. However, this approach shows limited utility, particularly in quantifying the binding of TZM to HER2, that is, drug-target engagement (Figure 1.1A), due to the poor imaging resolution and high nonspecific fluorescence signal, which may occur since ICG fluorescence is activated after dissociation from TZM upon cell binding and uptake, leading to possible extracellular release of ICG, which would contribute to the increased fluorescence intensity (34). A variety of other NIR probes have been developed and have been commercially available over the two last decades. These include cyanine dyes and derivatives, Alexa Fluor (AF) dyes, BODIPY analogs, benzo[c]heterocycles, xanthenes, IRDye 800CW, and more. (35, 36). We have compared a variety of NIR FRET pairs and determined that currently AF700 and AF750 is the most adequate NIR FRET pair due to its strong R_0 value (78.1 Å), similar to that of visible FRET pairs (AF 488-AF 555, R_0 = 70 Å), and to the long lifetime and high quantum yield of donor Alexa Fluor 700 (8, 37, 38).

1.1.3 NIR FLI FRET Imaging to Quantify Antibody-Target Engagement in Preclinical Tumor Models

Several membrane receptors that are overexpressed in a wide variety of cancer cells have been shown to regulate key cellular processes, such as proliferation, differentiation, cell cycle arrest, and senescence. Antibodies against the extracellular domain of these oncogenic receptors have gained increasing importance in the clinic for cancer treatment due to their specificity and fewer side effects; currently, these antibody drugs are part of primary care in the prostate, breast, head, and neck cancers, (39–41). Although antibody-based targeted therapy has shown tremendous promise in the clinic, emergent resistance to antibody drugs can lead to therapy failure, contributing to cancer mortality (42). Therefore, quantitative imaging methods that can discern effective drug-target engagement as well as provide insights into the microenvironmental factors affecting delivery of therapeutic antibodies are urgently needed. However, such antibody-receptor binding could only be assessed via invasive biochemical and immunohistochemical (IHC) approaches. We have leveraged FLI's unique ability to monitor protein-protein interactions via FRET to measure the binding of the NIR-labeled antibody conjugate to its target receptor in basic and preclinical research settings (9, 31). We have characterized the NIR dye pair AF 700 (donor) and AF750 (acceptor) specifically for FLIM and MFLI FRET to quantify drug-target engagement in human breast cancer cell lines and tumor xenograft models, respectively (12, 18, 26, 43). Changes in the fluorescence lifetime of antibody-AF700 (donor) conjugate in the presence of antibody-AF750 (acceptor) indicate the binding of antibody conjugates to their respective target receptor upon its dimerization/oligomerization and/or clustering at plasma membrane and subsequent internalization and delivery to intracellular membrane trafficking pathways (8, 22, 43, 44). Recently, we have shown that IRDye QC-1 (QC-1) is an effective NIR dark acceptor for donor AF 700 when performing *in vivo* MFLI FRET imaging of the transferrin-transferrin receptor system in tumor xenografts (8). Usage of AF 700 and QC-1 donor-acceptor NIR FRET pair allows for multiplexing with the metabolic probe IRDye 800CW 2-deoxyglucose (2-DG) for tumor drug response studies (8). *In vivo* MFLI FRET measurements have been successfully validated via *ex vivo* histochemistry, establishing that *in vivo* FRET data directly report on cellular target engagement (44). Importantly, *in vivo* FRET values have been shown to correlate strongly with transferrin receptor-transferrin binding in tumor xenografts, but strikingly, not when using *ex vivo* assessment of transferrin receptor expression by IHC (44).

Human epidermal growth factor receptor 2 (HER2) receptor dimerization is necessary for eliciting the signaling pathways involved in HER2 positive cancer progression. Therapeutic agents, such as clinically relevant monoclonal antibody TZM, that bind to the extracellular domain of HER2 and inhibit the HER2 receptor dimerization have been approved by FDA for the treatment of HER2-positive breast tumors. TZM in combination with chemotherapy has been shown to improve the median progression-free survival to 7.4 months from 4.6 months and the median overall survival to 25.1 from 20.3 months (45). Despite clinical success, a significant fraction of patients shows innate or acquired resistance to TZM treatment. Several mechanisms have been proposed for TZM resistance including vascular mimicry, presence of HER2 variants (HER-2Δ16 and p95HER-2), heterogeneous HER2

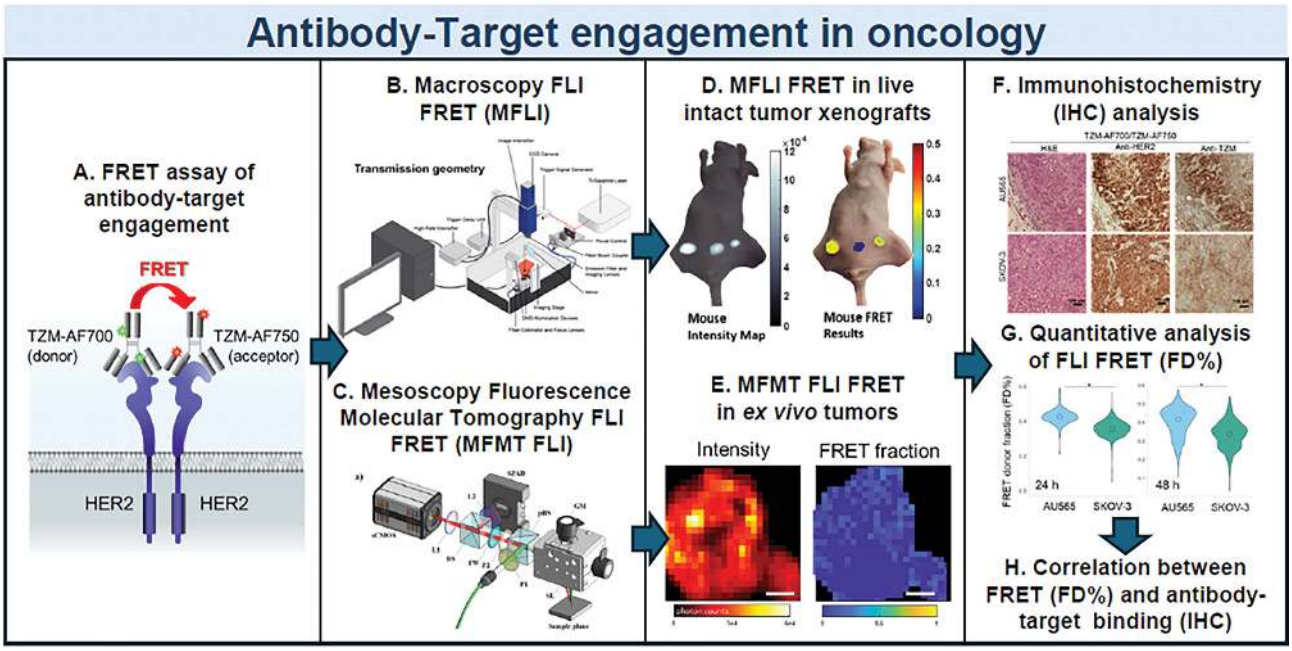


FIGURE 1.1 Visual summary of quantitative assessment of HER2 targeted drug delivery in breast cancer tumor models using MFLI and MFMT imaging. Scale bar, (E) 1 mm; (F) 100 μm.

expression and dimerization, suppression of anti-tumor antibody dependent cell cytotoxicity (ADCC), altered glucose/fatty acid metabolism, and disrupted signaling pathways (45). All these TZM resistance mechanisms are elicited due either to decreased TZM binding to HER2 or increased TZM degradation linked to endocytic internalization, and endo-lysosomal trafficking. Moreover, low tumor penetration and the immunosuppressive microenvironment can also negatively affect TZM efficacy. Recently, we demonstrated the TZM-HER2 drug-target engagement via FLI's capability to measure FRET in *in vivo* HER2 tumor xenografts using MFLI and immunohistochemistry validation (8, 20, 22, 46, 47) (Figure 1.1A–B and D–G). Moreover, we have expanded MFLI FRET imaging to assess intracellular delivery efficacy of the anti-epidermal growth factor receptor (EGFR) cetuximab (CTM) (22). NIR-labeled CTM was used to measure antibody-receptor engagement using MFLI FRET imaging of EGFR-positive tumors in live mice.

1.1.4 Macroscopic Fluorescence Lifetime (MFLI) FRET Imaging

We have used whole-body MFLI FRET for the preclinical characterization and optimization of antibody-based targeted therapy using a wide-field macroscopic fluorescence imaging approach augmented by deep learning algorithms for the analysis of complex MFLI FRET images (29, 48). NIR FLI FRET has been used to detect several receptor-probe complexes, that is, transferrin-transferrin receptor (8), CTM-EGFR (22) and TZM-HER2 receptor (46, 49) using MFLI imager with a NIR tunable pulse laser (8, 44, 49) (Figure 1.1B). The MFLI approach provides a measure of time-resolved fluorescence signals that can be further used for the estimation of the fraction of the probe undergoing FRET, for example, FRET donor fraction (FD%; A1, amplitude of short lifetime component) thereby quantitating the amount of drug-target engagement. The MFLI FRET measurements in whole living, intact animal experiments have further been validated by immuno-histological analysis, demonstrating the ability of MFLI FRET to measure drug-target engagement and intracellular drug delivery in a quantitative and longitudinal manner (8, 22) (Figure 1.1D–G). Recently, we have shown that although both intensity- and lifetime-based NIR intermolecular FRET *in vivo* imaging assays can be used to monitor receptor-ligand interactions in live intact mice, *in vivo* NIR FLI analysis is a more accurate and reliable approach for whole-body quantitative FRET imaging (43). However, such whole body macroscopic implementations, in both 2D (37) and 3D (29, 50), can only report drug-receptor binding at the organ level (43).

1.1.5 Mesoscopic Fluorescence Molecular Tomography (MFMT)

Fluorescence molecular tomography is widely used in preclinical tumor imaging of drug delivery, cell trafficking, and radiation therapy (51, 52). FMT uses a sensitive source-detector pair to perform a point-by-point scan of focal light sources over a region of interest and provides accurate visualization and quantification of distribution of targeted probes in a tissue in 3D. In FMT, a mathematical inverse algorithm is used to determine the 3D distribution of the targeted probes undergoing fluorescence emission. The major limitation of FMT is the low spatial resolution (mm range), which does not provide the probe distribution across the depth of the tissue and therefore fails to accurately detect intra-tumor heterogeneity. Recently, mesoscopic FMT (MFMT) imaging has been developed to study various biomedical applications at mm-range tissue depth and high resolution (100 μM) (27). MFMT imaging can reveal the depth-dependent morphological and vascular intra-tumor heterogeneity in breast tumor xenografts overexpressing Zs-Green and labeled with AF 594-Lectin GSII respectively (7). Recent advances in the reconstruction technique to solve the poorly defined MFMT inverse problem have produced high-resolution images (53). Interestingly, intravital mesoscopic fluorescence molecular tomography (IMFMT) has been demonstrated to noninvasively allow *in vivo* monitoring and quantification of breast cancer growth dynamics during treatment over a period of several months (54). Besides having wider applicability, MFMT imaging with fluorescence lifetime measurements capabilities is still lacking. Recently, we reported the development and characterization of time domain MFMT (TD-MFMT) for the measurement of FLI contrast (55) (Figure 1.1C). TD-MFMT system's capability to determine the FLI contrast was evaluated in *in vitro* systems (well plates and phantoms) and reconstruction of the targeted probe in *ex vivo* HER2+ ovarian tumors (55). The TD-MFMT imaging system is equipped with a pulsed laser (Mai Tai HP, Spectra Physics) as the illumination source and two detectors for (1) time-domain (TD) acquisition [16-channel SPAD array coupled with a TCSPC (Politecnico di Milano)] and (2) continuous-wave (CW) acquisition with high spatial density (Zyla 4.2 Andor Technology) (Figure 1.1C). In this chapter, we compare MFLI with TD-MFMT, which is capable of imaging tumor tissue *ex vivo* with high resolution (100μM range) at a millimeter depth (mesoscopic regime) to quantify drug-target engagement and assess intra-tumor heterogeneity in HER+ breast tumor xenograft mice model.

1.2 Materials

1.2.1 Chemical and Biological Reagents, Animals, and Instruments

AU565 cells were purchased from ATCC (CRL-2351) and grown according to the ATCC's suggested cell culture protocol and media. We used breast cancer cells passaged for less than eight to ten times to avoid cell alterations due to continued cell culture passage. PCR testing was used to check for the presence of mycoplasma and other infectious agents that may lead to significant alterations of cell physiology.

TZM was purchased from MedChemExpress (NJ, #HY-P9907). Donor TZM was labeled with NIR AF700 (Thermo Fisher Scientific, A 20010) and acceptor TZM with AF750 (Thermo Fisher Scientific, A 20011). Filter supplies include Amicon Ultra-4 centrifugal filter units (Sigma Z 648035, MWCO 30 kDa). Spectrophotometer DU 640 (Beckman Coulter, Fullerton, CA, USA) and NanoDrop Lite

spectrophotometer (Thermo Fisher Scientific, USA) were used to characterize the NIR-labeled TZM.

For animal tumor xenograft experiments, we used Cultrex BME (R&D Systems Inc., Minneapolis, MN, USA), phosphate buffered saline (Thermo Fisher Scientific, USA), and athymic nude mice (CrTac: NCr-Foxn1nu, Taconic Biosciences, Rensselaer, NY, USA). Immunohistochemistry was performed using anti-HER2 (Cell Signaling, # 2165) and rabbit monoclonal anti-TZM (R&D Systems, #MAB 95471-100) to determine HER2 and TZM tumor distributions and methyl green for cell contrast (Sigma Aldrich, #M 8884).

1.2.2 MFLI Imager

The MFLI imager setup uses an intensified charged coupled device (ICCD; (Picostar HR, LaVision GmbH) equipped with high-rate imager (HRI; Kentech Instruments, UK) for time-gated MFLI experiments (Figure 1.1B). The time-resolved data was collected by the ICCD camera as 12-bit images at 1392 X 1040 pixel resolution at each time gate with a gate width of 300 ps. The light source is a Ti Sapphire Mai Tai HP (Spectra-Physics, CA) with the following characteristics: 690–1040 nm emission, 100 fs width, 80 MHz repetition rate, power range 1 W to 3 W. Multimode optical fiber (QP 200-2-VIS-NIR, Ocean Optics, FL, USA) was used for light transfer from laser to a digital micromirror device (DMD, D2D module, Texas Instruments Inc.). For MFLI FRET studies, all the images were captured in reflectance geometry (2D imaging). The wide field illumination pattern at the excitation wavelength (700 nm) was achieved using the DMD over an 8 cm × 6 cm area with 256 grayscale levels encoding at each pixel. The DMD was controlled by Microsoft PowerPoint or LabView to ensure optimal signal detection over the whole body of mice (56–58). The instrument response function (IRF) was captured through a neutral density (ND) filter using plain white paper. The IRF and fluorescence signals were collected with a gate delay of 40 ps over an 8.0 ns time window in reflectance setup and at the same height. The total number of gates acquired was 201, and the maximum range detection was 4096 photons per pixel per gate. The micro-channel plate (MCP) employed for signal amplification was set to 560 V for fluorescence lifetime imaging. The laser excitation for AF 700 was set at 700 nm and emission filters were 720 ± 6.5 nm (FF01-720/13–25, Semrock, IL, Rochester, NY, USA) and 715 nm long pass filter (Semrock, FF01-715/LP-25). The emission signals from the sample were collected using a specific bandpass filter through a photographic lens (AF Nikkor 50mm f/1.8D, Nikon, Tokyo, Japan).

1.2.3 TD-MFMT Imager

The optical components for MFMT consist of two cameras: an sCMOS camera (Zyla 4.2, Andor Technology, UK) which is used to measure fluorescence intensities and a customized SPAD detector (Politecnico di Milano, Italy) that is coupled with a TCSPC module, which is used for time resolved measurements (Figure 1.1C). A tunable Ti:Sapphire pulse laser (MaiTai, Spectral Physics) delivers a light beam through a polarizer that is then reflected by a polarizing beamsplitter (pBS) (48–54548–545, Edmund Optics). A galvanomirror (GM) (GVSM002/M, Thorlabs) raster-scans the laser beam over the sample. The emitted backscattered and reflected light is collected by the scanning lens (SL) (LSM03-VIS, Thorlabs) and then returns through the same path as the illumination. A second polarizer and a filter wheel (CFW6, Thorlabs) are used to reduce specular reflection and for spectral separation (Figure 1.2). Finally, the emission signal passes through a beam splitter (BSW10R, Thorlabs) that splits the signal into

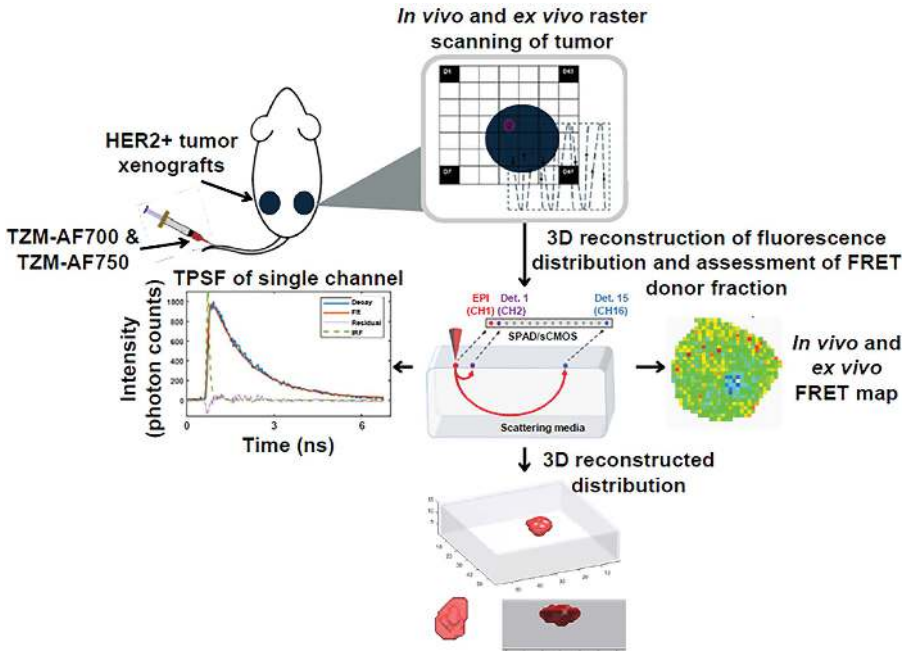


FIGURE 1.2 Schematic representation of the TD-MFMT imaging protocol.

two signals (at a 50:50 ratio) and focuses them on either camera using achromatic lenses. The optical filters used are FF01-780/12–25 (Semrock, NY), FF01-720/13–25 (Semrock, NY), and FF01-715/LP-25 (Semrock, NY).

1.3 Methods

1.3.1 Trastuzumab (TZM) Labeling

For MFLI and TD-MFMT FRET experiments, AF 700 donor (D) or AF750 acceptor (A) fluorophores were conjugated to TZM through monoreactive N-hydroxysuccinimide ester to lysine residues and incubated with 100 mM sodium bicarbonate, pH 8.3. Amicon Ultra-4-centrifugal filter units were used to remove unconjugated dye, followed by PBS washing. The protein concentration was adjusted to 1 mg/ml degree of labeling using a NanoDrop Lite spectrophotometer. We have determined that approximately two molecules of AF700 or AF750 dyes were conjugated per TZM molecule using the spectrophotometer DU 640.

1.3.2 Animal Imaging Experiments

Tumor xenografts were generated by injecting 10×10^6 HER2+ AU565 cells mixed with 1:1 Cultrex BME into the inguinal mammary fat pad of athymic nude mice. The tumors were allowed to grow for four weeks and monitored daily. Once the tumors were palpable, the mice were anesthetized and TZM FRET pair probes—TZM-AF 700(D) and TZM-AF 750 (A)—were injected through the retro-orbital route at the A:D ratio of 2:1 (A = 40 μg/ml and D = 20 μg/ml) in a volume range of 100–120 μl. MFLI was performed at 48 h post-injection in anesthetized mice bearing AU565 tumor xenografts. The small animal imaging system is equipped with an isoflurane anesthesia machine and warm stage maintained at 37.2 °C (Rodent WarmerX2, Stoelting, IL, USA). *In vivo* MFLI small animal imaging was performed under anesthesia and the vital signs of the mice were recorded using MouseOx Plus with a throat clamp. MFMT imaging was performed *ex vivo* in AU 565 tumors extracted from animals after sacrifice 48 h post-injection (p.i.).

1.3.3 MFLI FRET Imaging and Data Analysis

The FRET donor fraction was extracted using bi-exponential fitting via nonlinear least square (NLSF) based iterative fitting. FD% indicates the fraction of the population of TZM-AF 700 and TZM-AF750 molecules undergoing FRET within the region of interest (ROI) that is selected by tumor intensity thresholding. The FD% estimation analysis was calculated by fitting the fluorescence decays in each pixel of tumor ROI to the bi-exponential model:

$$I(t) = IRF(t) \otimes (\alpha_1 e^{\frac{-t}{\tau_1}} + \alpha_2 e^{\frac{-t}{\tau 2}}) + \varepsilon \quad (1)$$

$$A_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

$$A_2 = 1 - A_1$$

Here, I(t) represents observed fluorescence decay, IRF(t) is the instrument response function inherent to the system, α1 and α2 are the amplitude of the exponential terms; A1 and A2 are the FRET (FD%) and non-FRET donor fractions respectively, τ1 and τ2 are the quenched and unquenched donor lifetimes, t is time, ε is an additive Poissonian noise, and ⊗ represents convolution operation. The temporal point spread function (TPSF) at each pixel was fitted using the open-source software AlliGator (59), which uses the Levenberg-Marquardt nonlinear least square fitting algorithm to estimate the optimized bi-exponential model parameters. The fluorescence decays from the MFLI system were analyzed with the same fitting parameters and smoothing with Anscombe filtering. The violin plots, average values, and standard deviations can be reported (Figure 1.1G).

1.3.4 TD-MFMT Imaging and Data Analysis

The *ex vivo* MFMT imaging was performed immediately after tumor extraction in a Petri dish using a few drops of PBS to prevent drying. The tunable Ti-Sapphire laser was used as the excitation source (700 nm) and point scanning illumination was achieved using a galvo-mirror that raster scanned over a 6 mm × 6 mm area. A galvano-mirror was controlled via LabVIEW software to move the laser beam across the sample. The time-resolved data was collected by a customized SPAD array, which consists of a 32-channel linear array in which 16 channels are connected to the TCSPC (Figure 1.2). Each detector has a 50 μm diameter and a pitch size of 250 μm. TD and CW measurements are captured by raster scanning the field of view (FOV) under the control of LabVIEW. The FRET fraction at each scanning position is resolved by bi-exponentially fitting the fluorescence decay (TD measurement) using the NLSF method in AlliGator. The 3D fluorescence molecular distribution is fundamentally based on solving an inverse problem, which iteratively compares a Monte Carlo simulated forward light propagation model with the fluorescent emissions at different source-detector separations. The temporal resolution of the system is quantified by the full width at half maximum (FWHM) of the IRF and can be achieved to 120 ps. Additionally, the 3D spatial resolution is defined by the voxel size in the reconstruction and can be up to 100 μm.

1.3.5 Immunohistochemistry (IHC) Validation

After imaging, the tumors were fixed in 10% formalin followed by paraffin embedding and immunohistochemistry. The paraffinized tumors were subjected to sectioning to generate 5 μm tumor tissue sections. These sections were then deparaffinized, rehydrated, and subjected to epitope retrieval using 1 mM EDTA pH 8.0 for 30 min. Vectastain ABC Elite kit (Vector Labs, Burlingame, CA, # PK-610) was followed for IHC staining and Vector NovaRED (Vector) was used as a peroxidase substrate. Methyl green (Sigma Aldrich, # M 8884) was used for counterstaining. Hematoxylin and eosin (H&E) staining was used for morphological characterization.

The tumor tissue was stained for anti-HER2 (dilution 1:800) and rabbit monoclonal anti-TZM (dilution 1:100) primary antibodies. An Olympus BX40 microscope equipped with an Infinity 3 camera (Lumenera Inc., Ottawa, ON, Canada) was used for brightfield microscopy. Nanozoomer slide scanner (Hamamatsu Photonics) was used to acquire wide-field whole-tumor images.

1.4 Results

1.4.1 *In vivo* MFLI of Tumor Xenografts for Quantitation of HER2-TZM Binding in Live, Intact Mice

MFLI *in vivo* imaging provides a measure of time-resolved fluorescence signals that are used for the estimation of the fraction of the donor-labeled TZM undergoing FRET (FD%, amplitude of short lifetime component), thereby quantitating the amount of drug-target engagement. Figure 1.3A shows the FRET FD% maps of AU565 tumors at 48 h post-TZM injection. The quantitative pixel-by-pixel estimation of FD% levels and distribution show that the AU565 left tumor has a lower FRET signal compared to the right tumor, indicating higher TZM-HER2 binding in the right vs. left AU565 tumors (Figure 1.3B). After *in vivo* MFLI, the tumors were subjected to IHC validation of TZM and HER2 tumor distribution. The IHC data shows higher TZM levels in the right vs. left AU 565 tumors, confirming the MFLI FRET data (Figure 1.3C). The MFLI FRET measurements in whole living, intact animal experiments are validated by immuno-histological analysis, demonstrating the ability of MFLI FRET to measure drug-target engagement and intracellular drug delivery in a quantitative and longitudinal manner.

1.4.2 *Ex vivo* MFMT Imaging of AU565 Tumor to Visualize Intra-Tumor Heterogeneity of HER2-TZM Binding in Extracted Whole Tumor

We demonstrated the usage of the TD-MFMT system as a new imaging technique for quantifying the fluorescence lifetime as well as the volumetric distribution of fluorescent probes at millimeters depth (mesoscopic regime) with high resolution (hundreds micrometers) in scattering bio-tissues (22) (Figure 1.2). *Ex vivo* TD-MFMT can be used to monitor the HER2-TZM drug-target engagement via FLI FRET in whole tumors. Figure 1.4 shows that TD-MFMT displays the intra-tumor heterogeneity of HER2-TZM binding with higher resolution than MFLI (Figure 1.3). The FRET donor fraction obtained by TD-MFMT demonstrates correlation with immunohistological validation results (Figure 1.4). The areas indicated by solid black arrows show a similar TZM-HER2 binding pattern in the IHC and FD% maps. In contrast, high intensity areas appear to show an inverse correlation with high binding tumor regions. Interestingly, FRET signal distribution across the tumor appears to correlate strongly with HER2-TZM binding in tumor xenografts using anti-TZM IHC. However, the highly heterogenous FRET signal and IHC TZM distribution do not correlate with HER2 immunostaining, which displays a uniform distribution across the tumor. These results highlight TD-MFMT's capabilities for assessing intra-tumor heterogeneity of targeted drug delivery in preclinical applications.

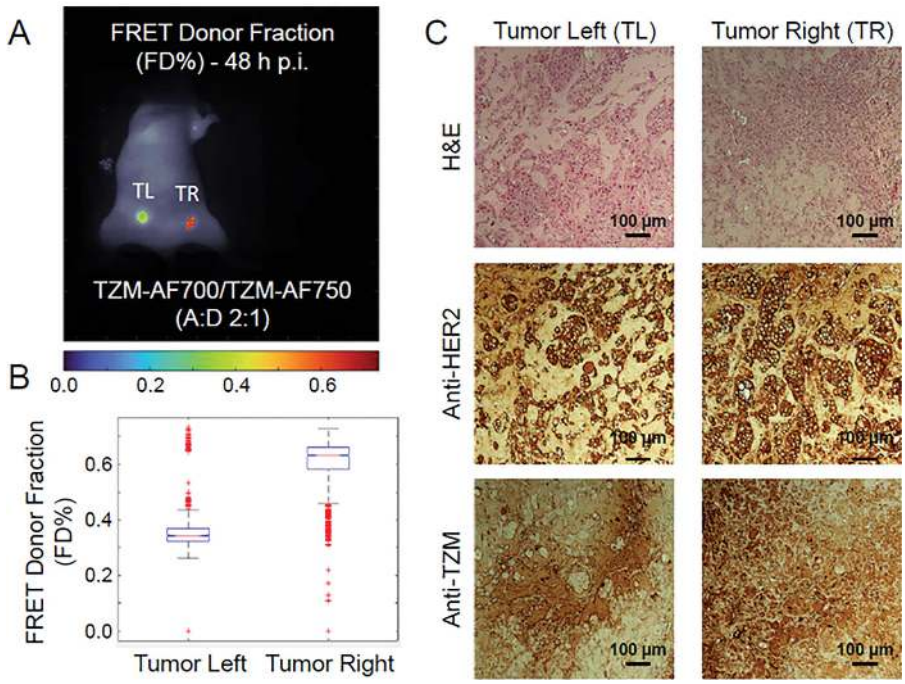


FIGURE 1.3 MFLI of HER2 tumor xenografts 48 h p.i. in live, intact nude mice. (A) FRET donor fraction tumor maps, (B) charts indicating the quantitation of FD%, and (C) IHC validation using H&E, HER2, and TZM immunostaining patterns.

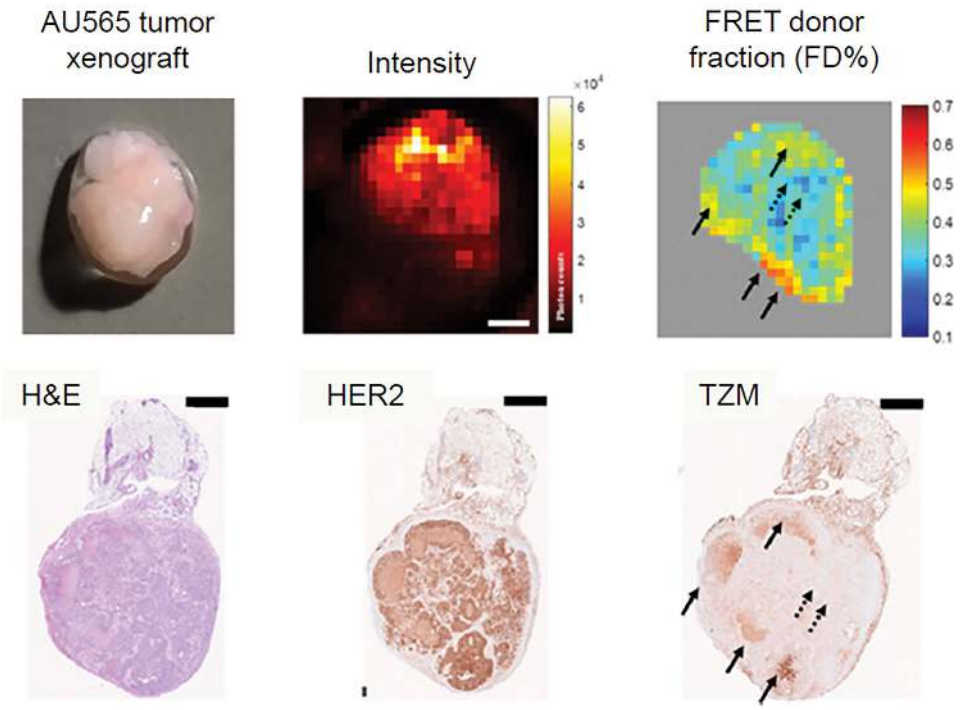


FIGURE 1.4 TD-MFMT *ex vivo* imaging of the dissected AU565 tumor xenograft. Top left: white light picture; middle: fluorescence intensity; and right: FRET donor fraction using MFMT imager. Bottom left: H&E, middle anti-HER2; and right: anti-TZM immunohistochemistry (IHC) results. TZM accumulation via IHC and FRET donor fraction are spatially correlated with solid arrows pointing at higher TZM-HER2 binding regions and dotted arrows indicating reduced TZM-HER2 regions.

1.5 Discussion

MFMT is a robust technology for reporting on the intra-tumor heterogeneity of antibody-receptor targeted delivery. The system is capable of *in vivo* and *ex vivo* imaging to acquire time-resolved (TD) and continuous-wave (CW) data. The TD-MFMT leverages both approaches to monitor drug-target engagement and create tomographic images of the sample (Figure 1.2). Testing the system on *ex vivo* samples qualitatively showed areas where the FRET fraction map created by the MFMT system correlated with the gold standard IHC of TZM, but not HER2 immunostaining results. TD-MFMT imaging is useful for acquiring maps of drug-target engagement immediately after tumor extraction without any manipulation. Nevertheless, there are areas of improvement in TD-MFMT that can be further investigated. First, several optical components, namely the scanning objective lens, are optimized for the visible light range. Upgrading the lens to make it more efficient in the NIR range will improve the collection efficiency of the system. This leads to the second shortcoming of the system, which is the time it takes to image the sample. A 2×2 mm^2 sample takes approximately 25 min per sample moving at 100 μm scanning points and an acquisition time of 1 sec. With improved optics, the collecting efficiency will increase, which will then notably decrease the acquisition time. Additionally, use of appropriate tumor xenograft models expressing NIR fluorescent proteins (NIR FPs) would allow the simultaneous assessment of drug-target engagement and tumor size. TD-MFMT also could be adequate to evaluate binding of NIR-labeled TZM or other antibody probes to tumor metastasis in various organs. The TD-MFMT system could also be used for measuring intra-tumor heterogeneity in clinical tumor tissue biopsies, which would provide direct information on the effectiveness of targeted therapy. However, any translation to the clinic would require FDA approval of NIR-labeled TZM drugs for injection prior to biopsy collection, which may be a difficult issue to overcome, or incubation of *ex vivo* tumor tissue samples with exogenously added NIR-labeled TZM.

Taken together MFLI or TD-MFMT measurements facilitate a highly sensitive, potentially multiplexed, FRET-based approach for quantifying drug-target engagement *in vitro*, *ex vivo*, and *in vivo*. The heterogeneity and complexity of the tumor microenvironment associated with the challenge of *in vivo* deep tissue noninvasive imaging demand constant innovation of instrumentation and computational methods. Herein, we show that TD-MFMT FRET measurements can determine the spatial distribution of drug-target engagement in HER2+ human breast tumor xenograft previously injected with TZM-AF 700 and TZM-AF 750 NIR FRET pair (Figure 1.4). MFMT *ex vivo* imaging clearly shows the intra-tumor heterogeneity of the TZM-HER2 binding across different regions of the tumor. MFLI and TD-MFMT FRET maps show a high correlation with IHC TZM immunostaining levels. The TD-MFMT higher resolution FRET tumor map suggests distinct levels of drug binding/distribution/penetration across the whole tumor. We demonstrate that MFMT is a promising

imaging modality to measure drug-target engagement across the whole tumor using FLI FRET imaging. MFMT is a novel preclinical imaging technology to dissect tumor microenvironment heterogeneities at near cellular scale resolution, including drug-receptor engagement, vasculature, and others.

1.6 Future Trends

Our goal is to expand the utility of MFLI and TD-MFMT FRET imaging to the clinical scenario, using either treated tumor biopsies *ex vivo* or live imaging in surgical settings. The major limitations of FLI are the availability of high-speed detectors and algorithms to compute the large data set for accurate evaluation of FLI FRET measurements. To achieve high sensitivity and contrast we are currently testing and validating the use of large time-gated solid-state CMOS SPAD cameras, such as SwissSPAD2 detectors, in comparison to the state-of-art ICCD camera. We have characterized the SwissSPAD2 detectors, which are sensitive enough to the level of individual photons, high signal-to-noise ratio, with unparalleled photon counting and time-resolved performances, and readout up to 97,700 frames per second. Implementation of artificial intelligence and deep learning models for FLI will enhance the performance of data processing and determination of lifetime-based parameters from the complex biodistribution of the fluorophore (60). We are also synergizing microlensing to increase detection sensitivity, as well as deep learning and/or artificial methods implemented at the edge. Moreover, site-specific labeled antibody methodologies will enable antibody labeling to occur in a more controlled and uniform manner compared to the stochastic NHS-ester labeled probes. The ability of SS2 to measure lifetimes and resolve small lifetime variations together with the integration of site-specific NIR antibody labeling methodologies will allow for the expansion of FLI FRET to the measurement of target engagement of a wide variety of clinically relevant antibodies and antibody-drug conjugates in preclinical tumor models as well as clinical tumor samples.

Acknowledgments

This work is supported by the National Institute of Health Grants R01 CA 237267 and R01 CA 271371. We thank the AMC histology core facility for the use of the Nanozoomer slide scanner and the immunostaining procedures. We thank the members of the Barroso and Intes labs for their helpful discussions and comments.

REFERENCES

1. Greenwalt I, Zaza N, Das S, Li BD. Precision medicine and targeted therapies in breast cancer. *Surg Oncol Clin N Am*. 2020 Jan;29(1):51–62.
2. Bose R, Ma CX. Breast cancer, HER2 mutations, and overcoming drug resistance. *N Engl J Med*. 2021 Sep 23;385(13):1241–3.
3. Menke-van der Houven van Oordt CW, McGeoch A, Bergstrom M, et al. Immuno-PET imaging to assess target engagement: experience from 89Zr-Anti-HER3 mAb (GSK2849330) in patients with solid tumors. *J Nucl Med*. 2019 Jul;60(7):902–9.
4. Henry KE, Dacek MM, Dilling TR, et al. A PET imaging strategy for interrogating target engagement and oncogene status in pancreatic cancer. *Clin Cancer Res*. 2019 Jan 1;25(1):166–76.
5. Pantel A, Makvandi M, Kim H, et al. PET imaging of PARP-inhibitor drug-target engagement predicts response in ovarian cancer in a pilot study. *J Nucl Med*. 2022 Jun 1;63(supplement 2):2486.
6. Yang M, Gao H, Sun X, et al. Multiplexed PET probes for imaging breast cancer early response to VEGF121/rGel treatment. *Mol Pharm*. 2011 Apr 4;8(2):621–8.
7. Rudkouskaya A, Faulkner DE, Sinsuebphon N, Intes X, Barroso M. Macroscopic fluorescence lifetime-based Förster resonance energy transfer imaging for quantitative ligand–receptor binding. In: Park K, editor. *Biomaterials for Cancer Therapeutics* (Second Edition). Woodhead Publishing; 2020. pp. 331–63 (Woodhead Publishing Series in Biomaterials).
8. Rudkouskaya A, Sinsuebphon N, Ochoa M, et al. Multiplexed non-invasive tumor imaging of glucose metabolism and receptor-ligand engagement using dark quencher FRET acceptor. *Theranostics*. 2020 Aug 15;10(22):10309–25.
9. Dmitriev RI, Intes X, Barroso MM. Luminescence lifetime imaging of three-dimensional biological objects. *J Cell Sci*. 2021 May 7;134(9):1–17.
10. Berezin MY, Achilefu S. Fluorescence lifetime measurements and biological imaging. *Chem Rev*. 2010 May 12;110(5):2641–84.
11. Pal R, Hom M, van den Berg NS, et al. First clinical results of fluorescence lifetime-enhanced tumor imaging using receptor targeted fluorescent probes. *Clin Cancer Res*. 2022 Jun 1;28(11):2373–84.
12. Ochoa M, Rudkouskaya A, Smith JT, Intes X, Barroso M. Macroscopic fluorescence lifetime imaging for monitoring of drug-target engagement. *Methods Mol Biol*. 2022;2394:837–56.
13. Kumar ATN, Carp SA, Yang J, Ross A, Medarova Z, Ran C. Fluorescence lifetime-based contrast enhancement of indocyanine green-labeled tumors. *J Biomed Opt*. 2017 Apr 1;22(4):40501.
14. Pal R, Lwin TM, Krishnamoorthy M, et al. Fluorescence lifetime of injected indocyanine green as a universal marker of solid tumours in patients. *Nat Biomed Eng*. 2023 Oct 16;7(12):1649–66.
15. Datta R, Heaster TM, Sharick JT, Gillette AA, Skala MC. Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications. *J Biomed Opt*. 2020 May;25(7):071203.
16. Sharman KK, Periasamy A, Ashworth H, Demas JN. Error analysis of the rapid lifetime determination method for double-exponential decays and new windowing schemes. *Anal Chem*. 1999 Mar 1;71(5):947–52.
17. Cheng H, Xie Y, Pellionisz P, et al. Dynamic optical contrast imaging (DOCI): system theory for rapid, wide-field, multispectral optical imaging using fluorescence lifetime contrast mechanism. In: Fei B, Linte CA, editors. *Medical Imaging 2019: Image-Guided Procedures, Robotic Interventions, and Modeling*. San Diego, United States: SPIE; 2019. p. 102.

18. Chen SJ, Sinsuebphon N, Rudkouskaya A, Barroso M, Intes X, Michalet X. In vitro and in vivo phasor analysis of stoichiometry and pharmacokinetics using short-lifetime near-infrared dyes and time-gated imaging. *J Biophoton*. 2019 Mar;12(3):e201800185.
19. Omer T, Intes X, Hahn J. Temporal data set reduction based on D-optimality for quantitative FLIM-FRET imaging. *PLoS One*. 2015 Dec 11;10(12):e0144421.
20. Lazzari-Dean JR, Gest AM, Miller EW. Optical estimation of absolute membrane potential using fluorescence lifetime imaging. *Elife*. 2019 Sep 23;8:e44522.
21. Levchenko SM, Pliss A, Peng X, Prasad PN, Qu J. Fluorescence lifetime imaging for studying DNA compaction and gene activities. *Light Sci Appl*. 2021 Nov 2;10(1):224.
22. Smith JT, Rudkouskaya A, Gao S, et al. In vitro and in vivo NIR fluorescence lifetime imaging with a time-gated SPAD camera. *Optica*. 2022 May 20;9(5):532–44.
23. Ardeshirpour Y, Sackett DL, Knutson JR, Gandjbakhche AH. Using in vivo fluorescence lifetime imaging to detect HER2-positive tumors. *EJNMMI Res*. 2018 Apr 4;8(1):26.
24. Kosaka N, Ogawa M, Sato N, Choyke PL, Kobayashi H. In vivo real-time, multicolor, quantum dot lymphatic imaging. *J Invest Dermatol*. 2009 Dec;129(12):2818–22.
25. Kosaka N, Mitsunaga M, Longmire MR, Choyke PL, Kobayashi H. Near infrared fluorescence-guided real-time endoscopic detection of peritoneal ovarian cancer nodules using intravenously injected indocyanine green. *Int J Cancer*. 2011 Oct 1;129(7):1671–7.
26. Sinsuebphon N, Rudkouskaya A, Barroso M, Intes X. Comparison of illumination geometry for lifetime-based measurements in whole-body preclinical imaging. *J Biophotonics*. 2018 Oct;11(10):e201800037.
27. Ozturk MS, Chen CW, Ji R, et al. Mesoscopic fluorescence molecular tomography for evaluating engineered tissues. *Ann Biomed Eng*. 2016 Mar;44(3):667–79.
28. Yang F, Ozturk MS, Zhao L, Cong W, Wang G, Intes X. High-resolution mesoscopic fluorescence molecular tomography based on compressive sensing. *IEEE Trans Biomed Eng*. 2015 Jan;62(1):248–55.
29. Venugopal V, Chen J, Barroso M, Intes X. Quantitative tomographic imaging of intermolecular FRET in small animals. *Biomed Opt Express*. 2012 Nov 8;3(12):3161–75.
30. Zhao L, Yang H, Cong W, Wang G, Intes X. Lp regularization for early gate fluorescence molecular tomography. *Opt Lett*. 2014 Jul 15;39(14):4156–9.
31. Rajoria S, Zhao L, Intes X, Barroso M. FLIM-FRET for cancer applications. *Curr Mol Imaging*. 2014;3(2):144–61.
32. Lutz MI, Schwaiger C, Hochreiter B, Kovacs GG, Schmid JA. Novel approach for accurate tissue-based protein colocalization and proximity microscopy. *Sci Rep*. 2017 Jun 1;7:2668.
33. Maus EA, Tan IC, Rasmussen JC, et al. Near-infrared fluorescence imaging of lymphatics in head and neck lymphedema. *Head Neck*. 2012 Mar;34(3):448–53.
34. Ogawa M, Kosaka N, Choyke PL, Kobayashi H. In vivo molecular imaging of cancer with a quenching near-infrared fluorescent probe using conjugates of monoclonal antibodies and indocyanine green. *Cancer Res*. 2009 Feb 12;69(4):1268–72.
35. Escobedo JO, Rusin O, Lim S, Strongin RM. NIR dyes for bioimaging applications. *Curr Opin Chem Biol*. 2010 Feb;14(1):64–70.
36. Alford R, Ogawa M, Hassan M, Gandjbakhche AH, Choyke PL, Kobayashi H. Fluorescence lifetime imaging of activatable target specific molecular probes. *Contrast Media Mol Imaging*. 2010;5(1):1–8.
37. Abe K, Zhao L, Periasamy A, Intes X, Barroso M. Non-invasive in vivo imaging of near infrared-labeled transferrin in breast cancer cells and tumors using fluorescence lifetime FRET. *PLoS One*. 2013 Nov 21;8(11):e80269.
38. Yuan N, Pandey V, Verma A, Williams JC, Intes X, Barroso M. Antibody-target binding quantification in living tumors using macroscopy fluorescence lifetime Forster resonance energy transfer imaging (MFLI FRET). In: *Visualizing and Quantifying Drug Distribution in Tissue VIII*. SPIE; 2024. pp. 17–20.
39. Masoud V, Pagès G. Targeted therapies in breast cancer: new challenges to fight against resistance. *World J Clin Oncol*. 2017 Apr 10;8(2):120–34.
40. Bluemn EG, Coleman IM, Lucas JM, et al. Androgen receptor pathway-independent prostate cancer is sustained through FGF signaling. *Cancer Cell*. 2017 Oct 9;32(4):474–89.e6.
41. Moreira J, Tobias A, O'Brien MP, Agulnik M. Targeted therapy in head and neck cancer: an update on current clinical developments in epidermal growth factor receptor-targeted therapy and immunotherapies. *Drugs*. 2017 May;77(8):843–57.
42. Sabnis AJ, Bivona TG. Principles of resistance to targeted cancer therapy: lessons from basic and translational cancer biology. *Trends Mol Med*. 2019 Mar;25(3):185–97.
43. Smith JT, Sinsuebphon N, Rudkouskaya A, Michalet X, Intes X, Barroso M. In vivo quantitative FRET small animal imaging: intensity versus lifetime-based FRET. *Biophys Rep*. 2023 May 9;3(2):100110.
44. Rudkouskaya A, Sinsuebphon N, Ward J, Tubbesing K, Intes X, Barroso M. Quantitative imaging of receptor-ligand engagement in intact live animals. *J Control Release*. 2018 Sep 28;286:451–9.
45. Vivekanandhan S, Knutson KL. Resistance to trastuzumab. *Cancers*. 2022 Oct 19;14(20):5115.
46. Verma A, Pandey V, Sherry C, et al. Fluorescence Lifetime Imaging for Quantification of Targeted Drug Delivery in Varying Tumor Microenvironments. *Adv Sci (Weinh)*. 2025; 12(3):e2403253. doi:10.1002/advs.202403253
47. Ochoa M, Rudkouskaya A, Yao R, Yan P, Barroso M, Intes X. High compression deep learning based single-pixel hyperspectral macroscopic fluorescence lifetime imaging in vivo. *Biomed Opt Express*. 2020 Oct 1;11(10):5401–24.
48. Smith JT, Yao R, Sinsuebphon N, et al. Fast fit-free analysis of fluorescence lifetime imaging via deep learning. *Proc Natl Acad Sci*. 2019 Nov 26;116(48):24019–30.
49. Rudkouskaya A, Smith JT, Intes X, Barroso M. Quantification of trastuzumab–HER2 engagement in vitro and in vivo. *Molecules*. 2020 Jan;25(24):5976.
50. Chen J, Venugopal V, Intes X. Monte Carlo based method for fluorescence tomographic imaging with lifetime multiplexing using time gates. *Biomed Opt Express*. 2011 Mar 14;2(4):871–86.

51. Yang H, Qian W, Yang L, Xie H, Jiang H. In vivo evaluation of a miniaturized fluorescence molecular tomography (FMT) endoscope for breast cancer detection using targeted nanoprobes. *Int J Mol Sci*. 2020 Dec 9;21(24):9389.
52. Shi J, Udayakumar TS, Wang Z, Dogan N, Pollack A, Yang Y. Optical molecular imaging-guided radiation therapy part 1: integrated x-ray and bioluminescence tomography. *Med Phys*. 2017 Sep;44(9):4786–94.
53. Yang F, Ozturk MS, Yao R, Intes X. Improving mesoscopic fluorescence molecular tomography through data reduction. *Biomed Opt Express*. 2017 Aug 1;8(8):3868–81.
54. Ozturk MS, Montero MG, Wang L, Chaible LM, Jechlinger M, Prevedel R. Intravital mesoscopic fluorescence molecular tomography allows non-invasive in vivo monitoring and quantification of breast cancer growth dynamics. *Commun Biol*. 2021 May 11;4(1):1–11.
55. Gao S, Li M, Smith JT, Intes X. Design and characterization of a time-domain optical tomography platform for mesoscopic lifetime imaging. *Biomed Opt Express*. 2022 Aug 10;13(9):4637–51.
56. Venugopal V, Intes X. Adaptive wide-field optical tomography. *J Biomed Opt*. 2013 Mar;18(3):036006.
57. Zhao L, Abe K, Barroso M, Intes X. Active wide-field illumination for high-throughput fluorescence lifetime imaging. *Opt Lett*. 2013 Oct 1;38(19):3976–9.
58. Zhao L, Abe K, Rajoria S, Pian Q, Barroso M, Intes X. Spatial light modulator based active wide-field illumination for ex vivo and in vivo quantitative NIR FRET imaging. *Biomed Opt Express*. 2014 Mar 1;5(3):944–60.
59. Chen SJ, Sinsuebphon N, Barroso M, Intes X, Michalet X. AlliGator: a phasor computational platform for fast in vivo lifetime analysis. In: *Optics in the Life Sciences Congress (2017)*, Paper OmTu2D2. Optica Publishing Group; 2017. p. OmTu2D.2.
60. Pandey V, Erbas I, Michalet X, et al. Deep learning-based temporal deconvolution for photon time-of-flight distribution retrieval. *Opt Lett*. 2024;49(22):6457–60. doi:10.1364/OL.533923

2

Combined Reflectance Confocal Microscopy–Optical Coherence Tomography for Skin Cancer Detection and Therapy Guidance

Nicusor Iftimia, John Grimble, Gopi Maguluri, Aditi Sahu, Miguel Cordova, and Milind Rajadhyaksha

2.1 Introduction

Unfortunately, despite long-term efforts to diagnose and treat cancer early, cancer continues to affect many people. Among many types of cancer, non-melanoma skin cancers (NMSCs) are among those with the highest incidence, with an estimated over 3.5 million cases of basal cell carcinoma (BCC) and 1.8 million cases of squamous cell carcinoma (SCC) diagnosed annually in the United States and over 1 million in other regions of the world (Europe, UK, Australia).[1,2] The annual rate is projected to increase by 3%–8% in the future.[3] Of these, 80%–90% of cases occur on the head and neck, and 65% on the face.[4] Early detection and accurate assessment of cancer margins can help to improve outcomes. Efforts are continuously being made to detect and treat cancers early, and thus to reduce their impact on people's lives.

Among various technologies for skin cancer diagnosis, optical imaging technologies seem to continue to play an important role in earlier cancer detection, as they can provide rapid, non-invasive, and accurate cancer detection at the point of care, and thus have great potential to improve global cancer care. Multiple clinical studies indicate that optical coherence tomography (OCT) and reflectance confocal microscopy (RCM) may play a very promising role in cancer diagnosis and therapy guidance.[5–12]

OCT is a fundamentally new type of optical imaging modality, which performs high-resolution, cross-sectional tomographic imaging of biologic systems by measuring the backscattered or back-reflected light.[13] OCT can provide discrimination between tissue microstructures based on optical coherence gating, which is based on low coherence interferometry (LCI). LCI measures the echo time delay and intensity of backscattered light by comparing it to light that has traveled a known reference path length and time delay. Measurements are performed using an optical interferometer, such as a Michelson or Mach Zehnder. In a Michelson interferometer (see Figure 2.1a), light from a source is directed onto a beam splitter and one of the beams is incident onto the sample to be imaged, while the second beam travels a reference path with a variable path length and time delay. The backscattered light from the sample is combined with reflected light from the reference arm and generates an interference pattern. The frequency of the interferogram is directly proportional with the optical pathlength mismatch between the two arms. The interference spectrum is detected by a photodetector and converted into an electrical signal, which is digitized and processed to extract the reflectivity profile of the sample. If the light source is purely monochromatic, an interference pattern is obtained all the time, while if it has a certain bandwidth D*l*, then an interference pattern is only obtained for a limited pathlength difference between the two arms of the interferometer, enabling sample sectioning with a resolution δ_z, defined by Equation 2.1:

$$\delta_z = 044 * (\lambda_0^2 / \Delta\lambda) \tag{2.1}$$

where λ_0 is the central wavelength of the light source.

By scanning the beam across the sample, a high-definition image of the sample can be obtained (see Figure 2.1b). Image resolutions of 1 to 15 μm can be achieved, which is one to two orders of magnitude higher than conventional ultrasound. The unique features of this technology enable a broad range of research and clinical applications.[14]

RCM can be used as a stand-alone or OCT complementary technology, as it provides submicron scale images of the biological samples. RCM is a non-invasive imaging technique that enables *in vivo* visualization of the superficial layers of the tissue in real time.[15] Resolution is almost comparable to conventional histology. RCM is mostly used for skin and mucosal layers of the tissue, allowing the clinician to perform a "virtual biopsy" of the tissue and obtain diagnostic clues while minimizing unnecessary biopsies.[16,17]

The illustration of the RCM principle is shown in Figure 2.2a. A diode laser is used as the source of coherent, monochromatic light. The emitted light is projected into the biological specimen after passing through a beam splitter and an objective lens, causing illumination of a small point within it, while also being scattered through a larger area of the sample. Tissue sectioning is performed by using a confocal aperture (pinhole) that allows for the detection of the photons coming from the focal plane of an imaging objective, while rejecting the photons coming from different depths. The pinhole selectively allows light from the focal point to pass through, hence confocal, and inhibits light from other tissue points or planes. By scanning the light across the sample using a pair of scanning mirrors, a high-resolution enface image of a specific plane within the examined specimen can be created (see Figure 2.2b and c).

In this chapter we present a customized RCM-OCT instrument developed at PSI, which combines both OCT and RCM

DOI: 10.1201/9781003449942-4

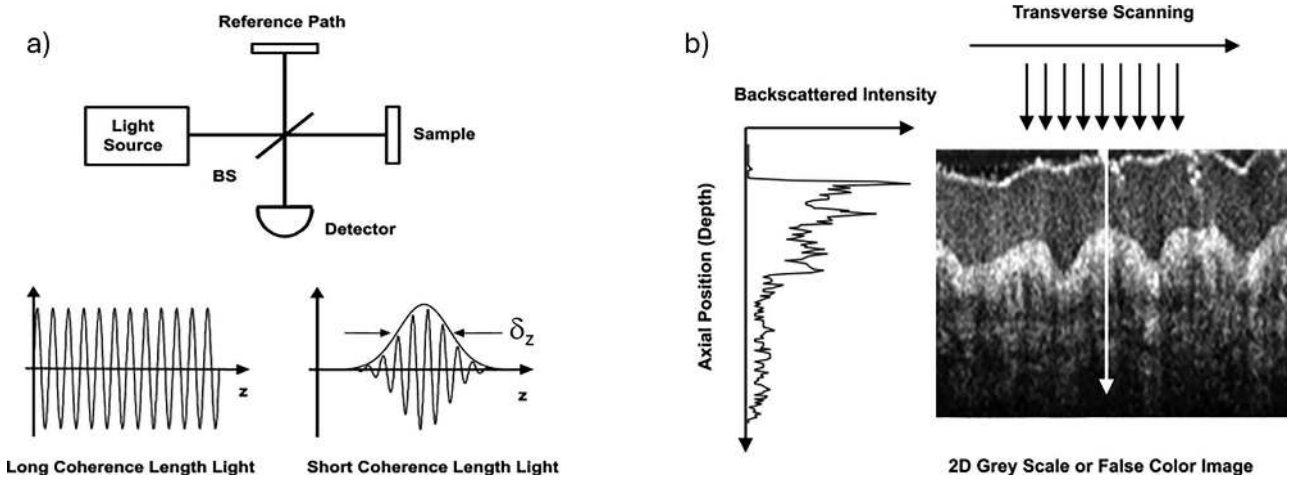


FIGURE 2.1 Schematic representation of LCI based on a Michelson setup (a) and example of OCT image generation by scanning the beam across the sample (b).

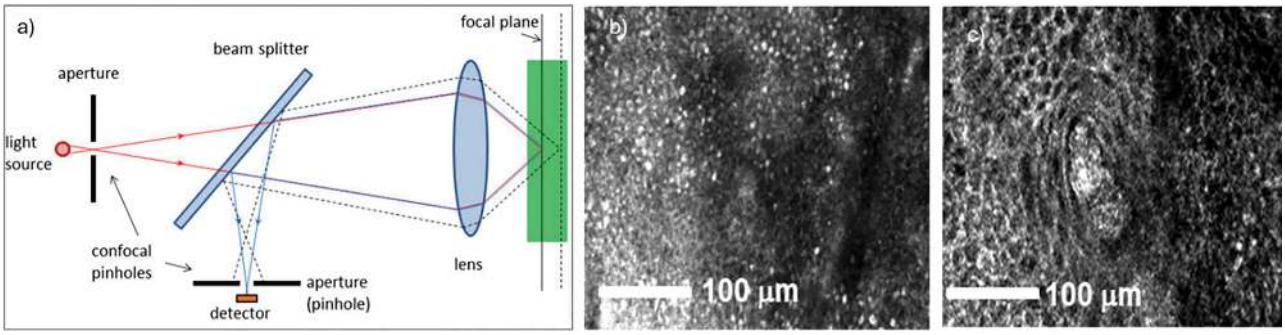


FIGURE 2.2 RCM principle (a) and examples of RCM images of the skin corneal layer (b) and epidermal layer (c).

within the same optical probe.[18] This instrument was developed with the goal of improving clinician ability to diagnose skin cancers, perform *in situ* biopsy, and guide various therapy modalities.

2.2 Materials

An RCM-OCT setup was designed for skin cancer diagnosis and therapy guidance. The goal of this development was to provide clinicians with a tool that will enable faster decisions on selecting the appropriate therapy for skin cancers, as well as for providing therapy guidance. While RCM provides submicron scale resolution and thus enables identification of morphological changes of the skin, with the downside of limited penetration depth, OCT imaging of the same lesion brings the benefit of resolving its depth of invasion. Although OCT and RCM can be used individually, their combined use within the same instrument provides more reliable diagnosis and therapy guidance of NMSCs. Specifically, their combined use provides better characterization of the lesion's margins, both in depth and laterally, as well as improved sensitivity and specificity, as previously demonstrated by our team.[19–22]

The schematic of the RCM-OCT instrumentation unit is shown in Figure 2.3A, while its implementation is shown in Figure 2.3B. The schematic of the handheld probe and its physical implementation are shown in Figure 2.3C, D, and E.

The OCT subsystem is based on a swept-source approach. It uses a fiber optic interferometer, with the handheld probe optics in the sample arm and an optical delay line in the reference arm. A micro-electro-mechanical systems (MEMS)-based swept source (Santec USA) with a 100 kHz A-scan rate, a bandwidth of 105 nm, and a central wavelength of 1310 nm is used for this system. The light is split in a 10/90 ratio and sent to the reference and sample arms of the interferometer, respectively. Two fiber circulators are used to maximize the return from both sample and reference arms of the interferometer. A balanced detector is used to minimize swept source intensity noise. Signal fringes are digitized at a rate of 500 MS/sec by a PCIe digitizer (AlazarTech, Model ATS 9352) and then processed in a GPU unit (NVidia, Model GTX 1060). A 4-analog output digitizer (National Instruments, Model PCIe 6323) is used to control the two axes of the MEMS mirror and the liquid lens. The liquid lens allows adjusting the focal plane of the OCT beam without affecting the position of the RCM focal plane.

The RCM subsystem uses a superluminescent diode as a light source with a central wavelength of 830 nm (Thorlabs, Model SLD830S-A10). The light is sent to the probe through a single mode/multimode (SM/MM) double clad coupler (Thorlabs, Model DC780SEB). The inner core is used to send the light to the probe, while the double clad is used to collect and transmit the light to an avalanche photodetector (Model APD410A, Thorlabs). The signal is digitized by an Alazar digitizer (AlazarTech, Model ATS9146) with 125 MS/s digitization speed. A 2-analog output digitizer (NI PCIe 6321) is used to control the amplitude of the resonant galvanometer and the slow-axis galvanometer to generate RCM images. The Z-focus motor is controlled through the USB line.

The instrumentation unit, approximately 2 × 2 × 2 feet in size, is placed on a wheeled cart, which also holds the monitor, the keyboard, and the mouse. A T-shape telescope was used to combine RCM and OCT within the same optical path (see lenses L1-L3 in Figure 2.3A). OCT and RCM beams are combined within the same optical path using a dichroic mirror before reaching the custom-designed imaging objective (NA = 0.8, Photongear, NY). This objective was designed to provide a working distance of 1 mm and a pupil size of 10 mm for the 800–1400 nm range. However, it was optimized to provide the best performance in the RCM mode at 830 nm. The design of this probe enables separate manipulation of the imaging objective NA in each mode (see US patent 9,655,521), and thus the preservation of the imaging capabilities of each of the two modes: submicron lateral resolution (~0.86 μm) and optical sectioning on the order of 2.5 μm in the RCM mode, with an imaging depth of ~300 μm, and ~10 μm lateral resolution with ~1.2 mm imaging depth in the OCT mode.

The OCT and RCM imaging beams (1.31 μm and 830 nm, respectively) are delivered from the instrumentation unit to the probe through fiber optic patch cords (SMF-28E fiber for OCT and double-clad DCF780S fiber for RCM). The light beams are then collimated using Thorlabs collimators (TC12APC-1310 for OCT and TC25APC-780 for RCM) and sent to two separate scanning engines. The OCT beam is scanned by a dual-axis MEMS mirror (Model A7B2–1, Mirrrorcle, CA), while the RCM beam is scanned by an engine consisting of a CRS12 kHz resonant scanner (Cambridge Technologies, MA), generating the line scan, and 7 mm aperture slow axis scanner

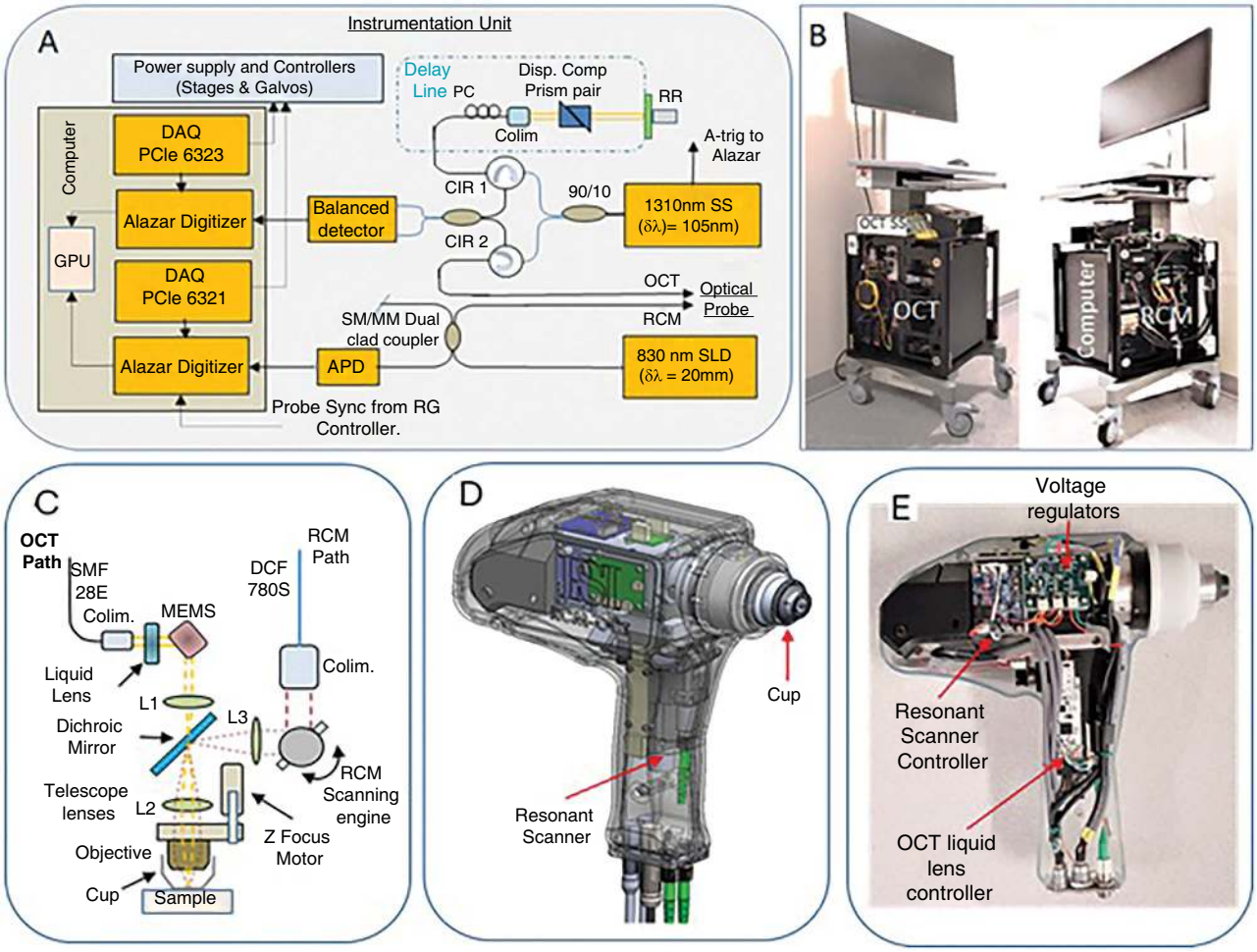


FIGURE 2.3 RCM-OCT unit: (A) schematic and (B) physical implementation. Handheld probe: (C) simplified schematic; (D) CAD design, (E) physical implementation. Legend: CIR—circulator; BS—beam splitter; PC—polarization controller; APD—avalanche photodetector; GPU—graphical processing unit; DAQ—data acquisition card; RR—retroreflector; SS—swept source; SLD—superluminescent diode [18].

(Model 6200H, Cambridge Technology, MA). The field of view (FOV) provided by this probe is ~2 mm × 2 mm in the OCT mode and 0.70 mm × 0.70 mm in the RCM mode. The objective was designed to provide optimal performance when a 0.6 mm thickness transparent window is placed between the objective and the sample to be imaged.

Imaging at different depths for RCM is enabled by a small stepper motor, which moves the position of objective lens relative to a stationary optically transparent cup-and-window, while the focus of OCT is handled by a liquid lens (Model D-A-39N1, Edmund Optics) over a range of ~ 1.5 mm. A Z-focus motor (Model DM0620, MicroMo) is used to control the focus with a 5-micron resolution over a 400-μm range. However, a faster depth sectioning approach will be developed within the very near future, as the motor somewhat limits the speed of acquiring 3D stacks. The cup-and-window is the only element of the probe that is in contact with the patient's skin that gently flattens and stabilizes the imaged area. Index matching oil (n = 1.5) is applied to the skin surface to optically interface to the window and eliminate potential back-reflections, while ultrasound gel (n = 1.34) is used inside the cup as the necessary immersion medium for the objective lens.

The controllers for the resonant scanner, MEMS scanner, Z-focus motor, and liquid lens are integrated within the probe. A voltage regulator is placed as well inside the probe to regulate the voltages necessary for all these controllers.

The probe is 7.5 inches in length, 8 inches in height, and 4 inches in width. It weighs less than 1.5 lb and therefore is relatively easier to operate than its predecessor. The all-fiber-optic configuration of this instrument allowed placing the laser illumination and detection modules into the instrumentation unit, making the probe lighter and smaller.

2.3 Methods

The initial evaluation of the RCM-OCT technology was carried out under the protocol approved by the institutional review board of Memorial Sloan Kettering Cancer Center. Written and verbal informed consent were obtained from each patient before imaging. Prior to treatment, patients underwent imaging with the combined RCM-OCT probe. Two types of lesions were examined: (1) previously biopsied, histopathologically confirmed basal cell carcinomas and (2) clinically or dermoscopically suggestive of BCC. Lesions were excluded if histopathologic examination was unavailable after imaging.

A total of 85 BCC lesions from 55 patients (27 female [49%] and 28 male [51%]) with a median age of 59 (21–90 years) were examined. Out of the 85 lesions, 25 were previously biopsied, while 60 were clinically or dermoscopically suggestive of BCC. Patient face was the most common imaged site (38 [45%]) followed by the trunk (35 [41%]).

Clinical and dermoscopic images of the lesions were captured with a handheld dermatoscope (VeosDS3; Canfield Scientific). The combined RCM-OCT imaging was performed using the handheld probe. Following a rigorously verified procedure,[23] we placed adhesive paper rings (3M) at the clinical margins for image navigation. Dermoscopic examination of the area within the paper ring was carried out with a handheld dermatoscope. The ring was divided into four hypothetical quadrants and imaging in dermoscopically suspicious areas was performed in each quadrant, following a clockwise approach. If dermoscopically suspicious areas were found in any of the quadrants, then BCC features were examined with the combined RCM-OCT probe. On encountering suspicious foci, we acquired RCM stacks and OCT raster images. Annotation of features on RCM and OCT was performed on the RCM stacks and OCT rasters immediately afterward. The entire procedure of imaging and image annotation took on average 20 to 30 minutes. The imaging results were examined again while the patient was in the room, and if inconclusive images were identified, the procedure was repeated in the culprit areas.

2.4 Results

Dermoscopy has shown typical signs of BCC: irregular vasculature and irregular shape skin pigmentation (see example in Figure 2.4a). Both OCT and RCM imaging (see Figure 2.4b, c, and e) simultaneously revealed the typical features of the BCCs. RCM showed cord-like structures with peripheral palisading admixed with a fibrotic stroma (see arrows in a representative image in Figure 2.4d), which is a common feature for superficial BCC.[21,22] The cross-sectional OCT images showed hypoechoic areas in the upper dermis along the dermal-epidermal junction (see arrows in the selected frame in Figure 2.4b out of the 512 collected frames), which are characteristic of superficial and early nodular BCCs. The enlarged and densely packaged nuclei (i.e., increased nuclear volume relative to cellular cytoplasm) in the tumor area scatter light weakly, giving rise to a hypoechoic area relative to the surrounding strongly scattering bright-appearing stroma in the OCT image (see the enface image in Figure 2.4c).

The diagnostic accuracy for the RCM-OCT examined lesions was evaluated in nonbiopsied clinically or dermoscopically suspicious lesions. RCM imaging correctly diagnosed BCCs in 47 of 48 lesions (98%) with high confidence (99%). OCT imaging identified all 48 lesions (100%) with high confidence (88%) for 42 of these lesions. Low-confidence RCM cases had cords or tumor islands without clear palisading and/or clefting. Low-confidence OCT cases manifested as very superficial or dark areas in the deeper dermis, which can be caused as well by the strong light absorption in large-size blood vessels. In all lesions identified with low confidence by either RCM or OCT, the complementary technology (OCT or RCM) helped increase confidence in the diagnosis. In summary, of the 85 investigated lesions, 60 (71%) were found to be positive for BCC on histopathologic examination after imaging (see example of histopathology correlation in Figure 2.4e). For previously biopsied lesions, only 12 of the 25 (48%) showed residual BCC features. The likely cause was that BCCs might have been entirely removed during biopsy. In nonbiopsied clinically or dermoscopically suspicious lesions, 48 of 60 (80%) were BCCs. Other diagnoses for the remaining lesions included three seborrheic keratosis, one sebaceous hyperplasia, two lichenoid keratosis, one squamous cell carcinoma, two scar, one dermal fibrosis and actinic changes, one acanthoma, and one cylindroma. These results are summarized in Table 2.1.

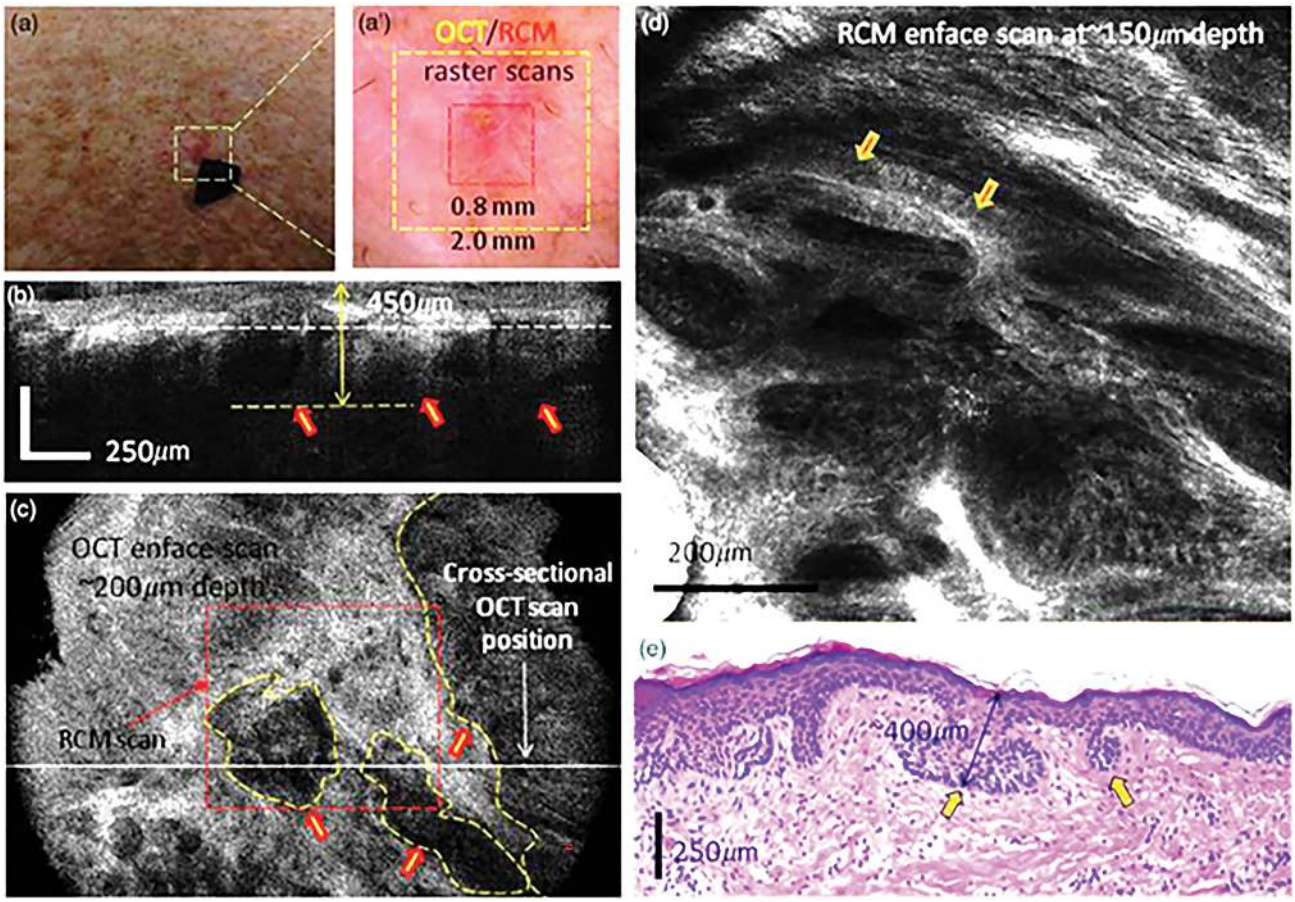


FIGURE 2.4 (a) Clinical image showing an erythematous macule on the back. (a′) Dermoscopy image showing shiny white lines and serpentine vessels. (b) Cross-sectional and (c) enface OCT images showing multiple hypoechoic areas (see arrows), suggestive of BCC. (d) Reflectance confocal microscopy showing cord-like structures with peripheral palisading (arrows) admixed with a fibrotic stroma, suggestive of BCC. (e) Histology image of the lesion (confirming the diagnosis of superficial BCC), showing multiple small tumor nests originating from the epidermis (hematoxylin and eosin, 4× magnification) [21].

TABLE 2.1

Summary of the RCM-OCT Imaging Results

			Predictive Value, % (95% CI)	
Modality	**Sensitivity, % (95% CI)**	**Specificity, % (95% CI)**	**Positive**	**Negative**
Biopsied lesions (n = 25)				
RCM	100 (73.5–100)	38.5 (13.9–68.4)	0.6 (36.1–80.9)	1 (47.8–100)
OCT	100 (73.5–100)	23.1 (5.0–53.8)	54.5 (32.2–75.6)	100 (29.2–100)
RCM-OCT	100 (73.5–100)	23.1 (5.0–53.8)	54.5 (32.2–75.6)	100 (29.2–100)
Nonbiopsied lesions (n = 60)				
RCM	97.9 (88.9–99.9)	75.0 (42.8–94.5)	94.0 (83.5–98.7)	90.0 (55.5–99.7)
OCT	100 (92.6–100)	58.3 (27.7–84.8)	90.6 (79.3–96.9)	100 (59.0–100)
RCM-OCT	100 (92.6–100)	75.0 (42.8–94.5)	94.1 (83.8–98.8)	100 (66.4–100)

2.5 Discussion

The results of this pilot study show that RCM and OCT imaging, when combined into a single probe, enable improved diagnosis of NMSC. Orthogonal OCT images help to localize suspicious structures, and the cellular resolution of enface RCM stacks identifies these structures. Furthermore, OCT raster scans provide volumetric information, which helps to display enface images and confirm suspicious features that mimic tumors or tumor-related features (e.g., vellus follicles and tumor nests).

Along with its diagnostic capability, OCT enabled the measurement of deep margins, simultaneously within

the same FOV as RCM, which highlights a key advantage of the combined probe in assessing tumor margins. High correlation between depth estimated by OCT and depth measured in histopathologic examination was observed for all BCCs, which is in concordance with findings in our published studies.[21] Minor differences in depth evaluation between OCT and histopathologic measurements was observed in shallow lesions. This could be attributed to possible misregistration between OCT and histological sections and tissue retraction and shrinkage effects.[22] A lower correlation was observed for deeper lesions (> 800 μm), which is likely the result of poorer image quality and decreased resolution and contrast of OCT when used for greater depths. In addition, because of the maximum depth limitation of OCT (1000 μm), greater differences from (and lower correlations with) histological measurements are to be expected. This finding is significant for the management of superficial and early nodular BCCs (approximately 0.2–0.5 mm deep), which can be treated with nonsurgical alternatives.

2.6 Future Trends

Underestimation of cancer margins leads to insufficient treatment and recurrence of the skin cancer. Therefore, image-guided therapy is highly desirable to precisely define cancer margins and guide nonsurgical therapy.

Nonsurgical treatments are preferred for superficial and nodular types of BCCs, which are shallow (depth ~200–500 μm) and less aggressive compared to the other deeper and more aggressive types (micronodular, infiltrative, sclerosing). The American Academy of Dermatology, American Society for Dermatologic Surgery, American Society for Mohs Surgery, and American College of Mohs Surgery recently issued "appropriate use" guidelines for these alternative nonsurgical therapies.[23,24] However, the main challenge in triaging cases suitable for nonsurgical treatments and in delivering effective treatment is the difficulty in accurately delineating cancer margins before treatment and confirming clearance afterwards. This is because the current diagnosing aiding approaches, including dermoscopy and ultrasound, do not provide the necessary information about the margins of these cancers and the status of the lesion after therapy.

Among the nonsurgical treatment modalities, radiotherapy (RT) and laser ablation are gradually becoming used more often for treating lesions on the exposed areas of the skin.[25–30] RT is also chosen when Mohs surgery cannot be used due to the very large depth of cancer invasion in the deeper dermis and beyond. Besides being highly effective, both laser ablation and RT come with the additional benefit that are not associated with scarring or disfigurement.

Unfortunately, the lack of adequate imaging tools that can be used in the clinics to assist physician-patient decision in treatment planning, for example biopsy vs. no biopsy, surgery vs. nonsurgical therapy, has negatively impacted the clinical spread of nonsurgical methods. For example, in the case of RT, without adequate planning there is typically additional healthy skin irradiated than necessary. Overtreatment, especially with RT, leads to side effects such as dermatitis and subcutaneous fibrosis, which can lead to poor cosmetic outcome, as well as secondary malignancies. Laser ablation, although very promising, has not gained large spread due to the absence of proper guidance techniques. Without guidance, laser therapy cannot always be effective since the laser fluence, and the number of irradiation passes need to be selected based on the accurate knowledge of the cancer margins. Therefore, laser ablation is not yet mentioned in the NCCN guidelines since more studies are needed to prove its effectiveness.

The proposed technology seems to provide the necessary data needed for nonsurgical therapy guidance. Therefore, we foresee its rapid clinical adoption.

REFERENCES

1. Deady S, Sharp L, Comber H. Increasing skin cancer incidence in young, affluent, urban populations: a challenge for prevention. *Br J Dermatol*. 2014 Mar 26. doi:10.1111/bjd.12988. [Epub ahead of print]
2. Donaldson MR, Coldiron BM. No end in sight: the skin cancer epidemic continues. *Semin Cutan Med Surg*. 2011;30(1):3–5.
3. Ciuciulete AR, Stepan AE, Andreiana BC, Simionescu CE. Non-melanoma skin cancer: statistical associations between clinical parameters. *Curr Health Sci J*. 2022 Jan–Mar;48(1):110–5. doi:10.12865/CHSJ.48.01.16. Epub 2022 Mar 31. PMID: 35911940; PMCID: PMC9289586.
4. Griffin LL, Ali FR, Lear JT. Non-melanoma skin cancer. *Clin Med*. 2016 Feb;16(1):62–5. doi:10.7861/clinmedicine.16-1-62. PMID: 26833519; PMCID: PMC4954336.
5. Hall ET, Fernandez-Lopez E, Silk AW, Dummer R, Bhatia S. Immunologic characteristics of nonmelanoma skin cancers: implications for immunotherapy. *Am Soc Clin Oncol Educ Book*. 2020;40:1–10. doi:10.1200/EDBK_278953. [DOI] [PubMed] [Google Scholar]
6. Avanaki MRN, Hojjatoleslami A, Sira M, Schofield JB, Jones C, Podoleanu AG. Investigation of basal cell carcinoma using dynamic focus optical coherence tomography. *Appl Opt*. 2013 Apr 1;52(10).
7. Chen CJ, Sierra H, Cordova M, Rajadhyaksha M. Confocal microscopy guided laser ablation for superficial and early nodular basal cell carcinoma. *JAMA Dermatol*. 2014 May 14. doi:10.1001/jamadermatol.2013.10225. [Epub ahead of print]
8. Gonzalez SG, Gill M, Halpern AC (Eds). *Reflectance Confocal Microscopy of Cutaneous Tumors—An Atlas with Clinical, Dermoscopic and Histological Correlations*. London: Informa Healthcare; 2008.
9. Hinz T, Ehler L-K, Hornung T, et al. Preoperative characterization of basal cell carcinoma comparing tumour thickness measurement by optical coherence tomography, 20-MHz ultrasound and histopathology. *Acta Derm Venereol*. 2012;92: 132–7.
10. Guitera P, Pellacani G, Longo C, Seidenari S, Avramidis M, Menzies SW. In vivo reflectance confocal microscopy enhances secondary evaluation of melanocytic lesions. *J Invest Dermatol*. 2009;129:131–8.
11. Guitera P, Menzies SW, Longo C, et al. In vivo confocal microscopy for diagnosis of melanoma and basal cell carcinoma using a two-step method: analysis of 710 consecutive clinically equivocal cases. *J Invest Dermatol*. 2012; 132(10):2386–94.

12. Razi S, Khan S, Truong TM, et al. Cutaneous squamous cell carcinoma: an up-to-date comprehensive review with a focus on contemporary optical imaging diagnostic modalities. *Dermato*. 2023;3:161–81. doi:10.3390/dermato 3020013.
13. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991 Nov 22;254(5035):1178–81. doi:10.1126/science.1957169. PMID: 1957169; PMCID: PMC 4638169.
14. Al-Mujaini A, Wali UK, Azeem S. Optical coherence tomography: clinical applications in medical practice. *Oman Med J*. 2013 Mar;28(2):86–91. doi:10.5001/omj.2013.24. PMID: 23599874; PMCID: PMC3628201.
15. Levine A, Markowitz O. Introduction to reflectance confocal microscopy and its use in clinical practice. *JAAD Case Rep*. 2018 Nov 10;4(10):1014–23. doi:10.1016/ j.jdcr.2018.09.019. PMID: 30456275; PMCID: PMC6232695.
16. Flores ES, Cordova M, Kose K, Phillips W, Nehal KS, Rajadhyaksha M. Feasibilty of intraoperative imaging during Mohs surgery with reflectance confocal microscopy. *Proceedings of SPIE Photonics West (BiOS—Biomedical Optics)*, 2014 (abstract, conference paper).
17. Pan ZY, Lin JR, Cheng TT, et al. In vivo reflectance confocal microscopy of Basal cell carcinoma: feasibility of preoperative mapping of cancer margins. *Dermatol Surg*. 2012;38(12):1945–50.
18. Maguluri G, Grimble J, Mujat M, Park J, Caron A, Iftimia N. Fiber-based hand-held RCM-OCT probe for noninvasive assessment of skin lesions and therapy guidance. *Transl Biophoton*. 2022;4. doi:10.1002/tbio.202200002.
19. Iftimia N, Ferguson RD, Mujat M, et al. Combined reflectance confocal microscopy/optical coherence tomography imaging for skin burn assessment. *Biomed Opt Express*. 2013;4(5):680–95.
20. Scimone MT, Krishnamurthy S, Maguluri G, et al. Assessment of breast cancer surgical margins with multimodal optical microscopy: a feasibility clinical study. *PLoS One*. 2021;16(2):e0245334. doi:10.1371/journal.pone. 0245334.
21. Iftimia N, Yélamos O, Chen CJ, et al. Handheld optical coherence tomography-reflectance confocal microscopy probe for detection of basal cell carcinoma and delineation of margins. *J Biomed Opt*. 2017 Jul 1;22(7):76006. doi:10.1117/1.JBO.22.7.076006. PMID: 28697233; PMCID: PMC5995139.
22. Sahu A, Yelamos O, Iftimia N, et al. Evaluation of a combined reflectance confocal microscopy–optical coherence tomography device for detection and depth assessment of basal cell carcinoma. *JAMA Dermatol*. 2018;154(10):1175–83.
23. Motley R, Kersey P, Lawrence C. Multiprofessional guidelines for the management of the patient with primary cutaneous squamous cell carcinoma. *Br J Dermatol*. 2002;146(1):18–25.
24. Bichakjian C, Armstrong A, Baum C, et al. Guidelines of care for the management of basal cell carcinoma. *J Am Acad Dermatol*. Online ahead of print. doi:10.1016/ j.jaad.2017.10.006. [DOI] [PubMed]
25. Amini S, Viera MH, Valins W, et al. Nonsurgical innovations in the treatment of nonmelanoma skin cancer. *J Clin Aesthet Dermatol*. 2010;3(6):20–34.
26. Choudhary S, Tang L, Elsaie ML, et al. Lasers in the treatment of nonmelanoma skin cancer. *Dermatol Surg*. 2021;37(4):409–25.
27. Iyer S, Friedli A, Bowes L, et al. Full face laser resurfacing: therapy and prophylaxis for actinic keratoses and nonmelanoma skin cancer. *Lasers Surg Med*. 2004;34(2):114–9.
28. Jiang SB, Levine VJ, Nehal KS, Baldassano M, Kamino H, Ashinoff RA. Er:YAG laser for the treatment of actinic keratoses. *Dermatol Surg*. 2000 May;26(5):437–40.
29. Konnikov N, Avram M, Jarell A, Tannous Z. Pulsed dye laser as a novel non-surgical treatment for basal cell carcinomas: response and follow up 12–21 months after treatment. *Lasers Surg Med*. 2011;43(2):72–8.
30. Lapidoth M, Adatto M, Halachmi S. Treatment of actinic keratoses and photodamage with non-contact fractional 1540-nm laser quasi-ablation: an ex vivo and clinical evaluation. *Lasers Med Sci*. 2013 Feb;28(2):537–42. doi:10.1007/ s10103-012-1103-6. Epub 2012 Apr 27.

3

In Vivo *Confocal Laser Endomicroscopy: An Imaging Biomarker for Risk Stratification of Precancerous Pancreatic Cystic Lesions*

Da Yeon Ryoo, Bryn Koehler, Vanisha Patel, and Somashekar G. Krishna

3.1 Introduction

3.1.1 Pancreatic Cystic Lesions (PCLs)

Pancreatic cancer is one of the deadliest forms of cancer known to humanity, accounting for 8.2% of all cancer deaths in the United States and with an estimated five-year survival rate of 11.5%.[1] The global incidence of pancreatic cancer is projected to increase as well, to 18.6 per 100,000 persons by the year 2050.[2] The high mortality of pancreatic cancer and its rapidly increasing incidence pose a significant public health burden. Therefore, it is imperative now, more than ever, to find efficient methods to screen, detect, and treat pancreatic adenocarcinoma. About 20% of pancreatic adenocarcinomas arise from pancreatic cystic lesions (PCLs), which offers an opportunity to screen for premalignant PCLs.[3] Considering that the incidence of PCLs in US populations is approximately 3% to 15%, the risk is significant.[4] However, not all types of PCLs have malignant potential, emphasizing the importance of developing methods that accurately identify their subtypes.

Pancreatic cysts are further classified as mucinous and non-mucinous PCLs. Of the two subtypes, mucinous PCLs have the potential to progress to pancreatic adenocarcinoma. Mucinous PCLs include intraductal papillary mucinous neoplasms (IPMNs) and mucinous cystic neoplasms (MCNs). MCNs represent about 20% of all resected PCLs and are thought to have a lower malignant potential, although it continues to be debated.[5] High-risk features of MCNs that indicate premalignancy include cyst size larger than 3 cm, presence of mural nodule, dilated pancreatic duct (PD) larger than 6 mm, and peripheral eggshell calcification.[6–8] IPMNs are the most common type of PCLs and are more frequently observed in elderly (≥ 60 years of age) individuals.[6] They are further classified as main duct, branch duct, and mixed IPMNs. They are characterized by intraductal dysplastic epithelium and goblet cells that produce mucin.[9] All IPMNs have malignant potential at baseline with at least a low-grade dysplasia, but the risk varies depending on cyst size, growth rate, the presence of mural nodules or solid component, and dilation of PD.[6]

Non-mucinous PCLs consist of serous cystadenoma (SCA), solid pseudopapillary neoplasm (SPN), and cystic neuroendocrine neoplasm (cystic-NET). Compared to mucinous PCLs, non-mucinous PCLs tend to show a lower risk of malignancy potential. SCAs, the most common benign pancreatic neoplasm, are slow growing and have less than 0.5% occurrence of transforming to malignant serous cystadenocarcinoma.[10–13] The reported cases of malignant serous cystadenocarcinomas were of sizes larger than 4 cm.[10] SPNs are rare, making up 0.3% to 2.6% of all pancreatic tumors, and are also known to be largely benign.[14] Although margin-negative surgical resection is thought to be curative in most patients, some tumors will show indications of malignancy or metastatic disease. However, even in the case of metastatic malignant disease, the overall survival remains high. Cystic-NETs comprise around 6.5% to 36.1% of all pancreatic neuroendocrine tumors and are thought to be less functional and less aggressive than solid NETs.[15]

Due to the difference in malignant potential between mucinous and non-mucinous PCLs, much research has been done on methods of differentiation between the two. However, the lack of a universal standardized method of PCL differentiation is repeatedly recognized.

PCL differentiation is subsequently followed by risk stratification, which guides more invasive testing and management. As with PCL differentiation, a universal method for PCL risk stratification has not been established, although there are many novel contenders. Despite clinical guidelines for pancreatic cyst diagnosis and management, cysts with relatively low malignant potential often get overtreated, leading to unnecessary testing, including radiological imaging, endoscopy procedures, and inadvertent pancreatectomies. The ability to accurately diagnose and risk stratify pancreatic cysts will decrease the overtreatment and offer proper triaging based on an accurate level of concern. The importance of developing universal methods of PCL identification and risk stratification is emphasized by the high morbidity and mortality rates of pancreatic surgeries, reaching approximately 20% to 40% and 2%, respectively.[9,16–18]

3.1.2 Pancreatic Cystic Lesion Imaging Modalities

Computed tomography (CT) scan, magnetic resonance imaging (MRI), and endoscopic ultrasound (EUS) are the most commonly utilized imaging methods for the evaluation of cyst morphology.[19] EUS is a merging of flexible endoscopy and ultrasound that was first developed in the 1980s to help improve visualization of the pancreatobiliary system.[20] Placing the ultrasound probe inside the body allows for visualization of structures deep to air-filled and bony structures that were not previously possible. Advancements in this technology

DOI: 10.1201/9781003449942-5

over the past four decades have allowed it to be a sophisticated tool in diagnosing and treating many gastrointestinal pathologies. The essential components of all echoendoscopes are the ultrasound transducer, a video chip for recording, a light source, and a water irrigation system. While transcutaneous ultrasound utilizes gel to minimize air interference with ultrasound transmission, echoendoscopes utilize small balloons around the transducers that can be inflated using water irrigation systems.[21]

Over time, newer fixed curved transducers were developed to allow endoscopists to obtain fine-needle aspiration (FNA) while visualizing both the cyst and the needle. These transducers produced a sector-shaped image parallel to the endoscope (Figure 3.1), allowing the needle to be advanced through the scope and visualized in the imaging field for biopsy.[20] Although there was some early resistance due to the perceived superiority of radial images of linear transducers, most EUS cases are now performed with fixed curved transducers.

EUS-FNA of cyst fluid can be used to measure and evaluate carcinoembryonic antigen (CEA), glucose, and cytology. Cyst fluid CEA and cytology are the two recommended modalities of pancreatic cyst differentiation, although with very low clinical evidence, and recently, the addition of glucose has improved standard evaluation.[22] Although widely adopted, cyst fluid analysis has variable sensitivity and specificity when it comes to differentiating PCLs, ranging from 58%–93% to 87%–98%, respectively.[23–27] Of the various components of cyst fluid analysis, cyst fluid glucose of less than 50 mg/dL had the highest sensitivity (~90%) for diagnosing mucinous cysts.[27,28] More recently, cyst fluid next-generation sequencing (NGS) analysis to detect molecular markers has risen to the horizon, surpassing the diagnostic accuracy of cyst fluid CEA and cytology for differentiation and detection of PCL subtypes.[29,30]

One of the most promising advances in EUS technology is the development of EUS-guided confocal laser endomicroscopy (EUS-CLE). EUS-CLE, first invented in 2004, enables *in vivo* histopathological evaluation of gastrointestinal and pancreatobiliary processes, allowing assessment of organ systems at subcellular level in their natural environment.[31] Confocal laser endomicroscopy images are obtained via low-power laser. The low-power laser illuminates the tissue of interest, and the reflected light is refocused through the pinhole onto the detection system, allowing a high-resolution image.[32] Compared to prior diagnostic modalities of PCL, the ability to directly visualize the cystic surface for diagnosis is deemed more advantageous compared to former indirect methods of diagnosis via analysis of cyst fluid content and imaging.[33] Additionally, the ability to diagnose PCLs without the invasiveness of cyst resection for histopathology analysis makes EUS-CLE more favorable.

Many clinical applications of CLE exist, taking advantage of its ability to offer high-resolution optical biopsy of the tissue in question. Some examples of CLE application in the clinical setting include colorectal cancer screening,[34] management of Barrett's esophagus,[35] detection of esophageal squamous cell carcinoma,[36] *in vivo* diagnosis of *H. pylori* in the stomach,[37] diagnosis of celiac disease, assessing severity of ulcerative colitis,[38,39] and differentiation of urothelial carcinoma of the bladder.[40] Additionally, bronchoscopic CLE has recently been discussed as a modality for detecting and managing intrathoracic diseases such as malignancy, interstitial lung disease, and obstructive lung disease.[41]

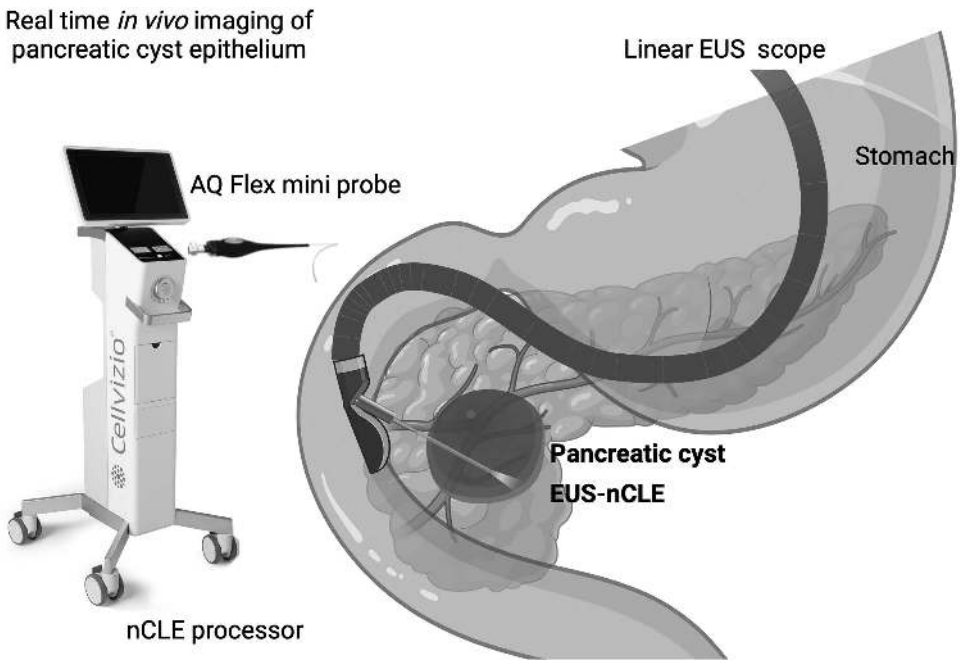


FIGURE 3.1 Endoscopic ultrasound with confocal endomicroscopy involves a linear echoendoscopy where FNA needle is actuated into the pancreatic cyst. The endomicroscopy processor, along with the AQ-FLEX CLE miniprobe, is utilized. The miniprobe is preloaded before entering the cyst with the needle and then advanced beyond the needle tip for imaging.

3.2 Materials

3.2.1 EUS-Guided Confocal Laser Endomicroscopy (EUS-CLE)

Cellvizio® 100 Series System, a probe-based and needle-based confocal laser endomicroscopy (pCLE/nCLE) system manufactured by Mauna Kea Technologies in Paris, France, was used.

The new platform is Next-Generation Cellvizio®, the real-time *in vivo* cellular imaging platform. Built with enhanced capabilities, the entirely new system uses pCLE/nCLE advanced imaging technology to deliver cellular visualization virtually anywhere in the human body.

3.2.2 Chemicals

A fluorescent contrast is necessary for CLE imaging. Intravenous fluorescein is the most commonly used contrast agent. Fluorescein (2.5 ml to 5 ml of 10% fluorescein sodium) is intravenously injected immediately prior to nCLE imaging of pancreatic cysts.

3.3 Methods

The mini probe to perform EUS-CLE (Figure 3.2) in the pancreas is the AQ-Flex probe (Cellvizio; Mauna Kea Technologies, Paris, France). AQ-Flex miniprobe is introduced through a 19-gauge FNA needle, and this technique is therefore termed needle-based CLE (nCLE).[42] To acquire images via EUS-nCLE, the PCL is identified via EUS, and intravenous fluorescein is administered. The PCL is then punctured with the needle, and the mini probe is actuated beyond the tip of the needle for endomicroscopic video acquisition. The probe is then unlocked and retracted, allowing for PCL fluid aspiration.

Key features associated with PCL subtypes were identified and validated by directly comparing *in vivo* EUS-nCLE image patterns with *ex vivo* probe-based CLE (pCLE) and surgical histopathology.[43] Figure 3.3 demonstrates the nCLE imaging features of common types of PCLs. IPMNs demonstrated "finger-like" papillary projections of variable length with overlying epithelium and an underlying vascular core. MCNs had single or layering epithelial bands of variable thickness without a papillary conformation. Cystic-NETs were characterized by dark clusters or trabeculae of cells separated by bright vascular spaces. SCAs had a classic "fern pattern" of vascularity where parallel vessels emanated from a central vessel.[43]

Over the past eight years, our research group has conducted studies demonstrating the promising results of utilizing nCLE to diagnose and risk stratify PCLs.

3.4 Results

3.4.1 Confocal Endomicroscopy: Diagnosis, Cost-Benefit Analysis, and Learning Techniques in the Management of Pancreatic Cysts

In a preliminary study using a novel nCLE classification system, we assessed accuracy, interobserver agreement (IOA), and interobserver reliability (IOR) in diagnosing 29 patients with PCLs. Observers reviewed high-yield nCLE videos from

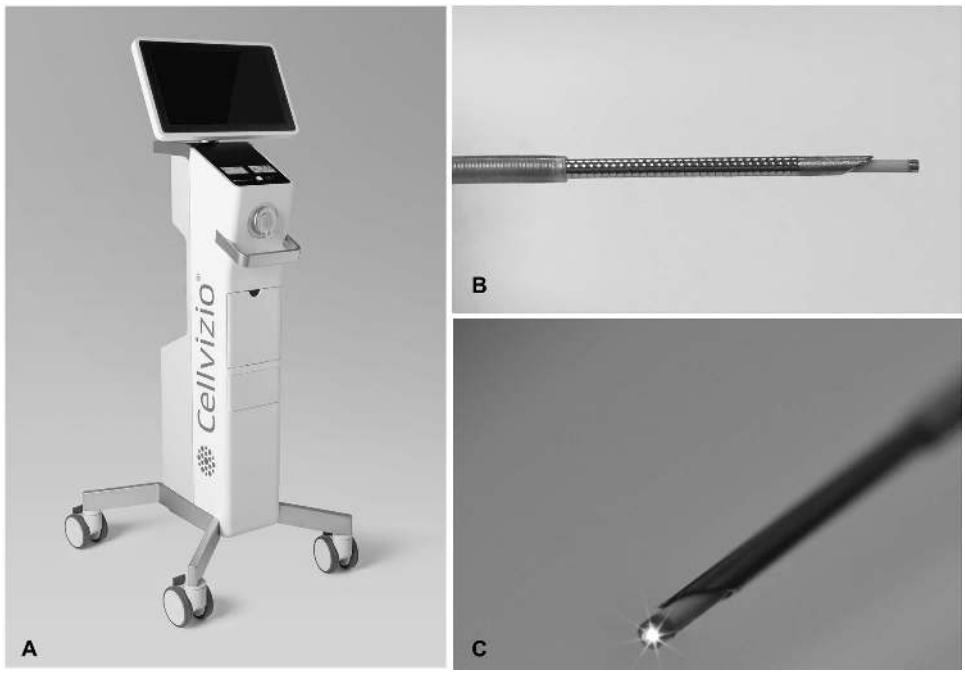


FIGURE 3.2 EUS-nCLE equipment: (A) processor, (B) AQ-FLEX CLE miniprobe positioned past the bevel of the needle, and (C) AQ-FLEX CLE miniprobe with laser endomicroscopy.

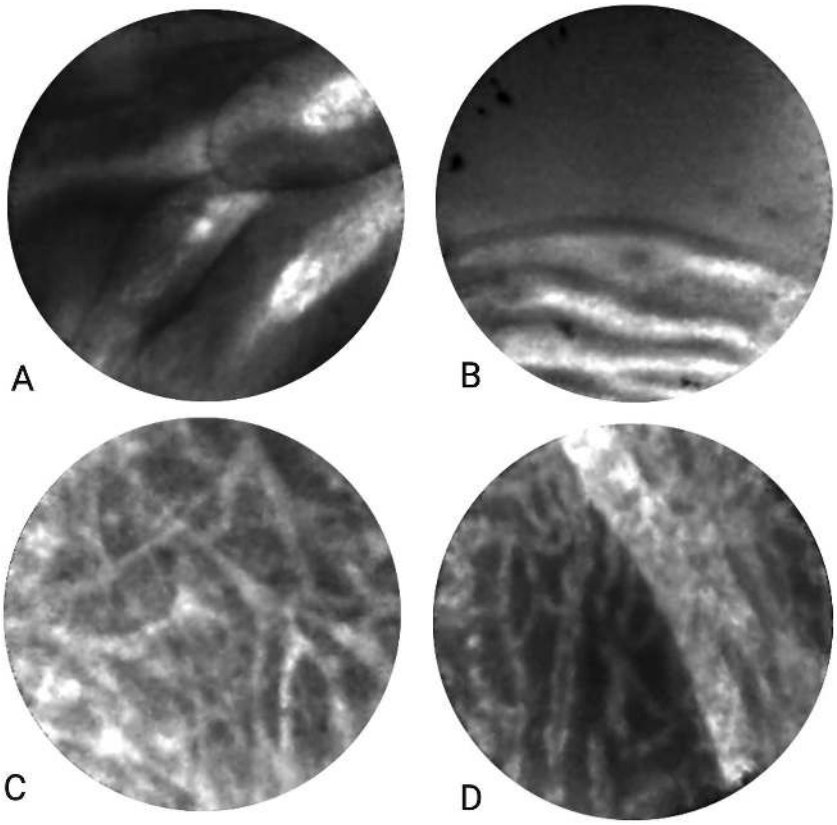


FIGURE 3.3 EUS-nCLE imaging features of common types of pancreatic cystic lesions.

each subject and achieved 95% sensitivity, 94% specificity, and 95% accuracy for differentiating PCLs, with near-perfect IOA (κ = 0.81) and IOR (κ = 0.86).[44] In a subsequent study involving a larger cohort of independent observers (n = 13) reviewing a large subset of subjects with PCLs (n = 76 subjects), there was an "almost perfect" IOA (κ = 0.82, 95% CI 0.77–0.87) and IOR (κ = 0.82, 95% CI 0.78–0.85) for differentiating mucinous from non-mucinous PCLs. The highest IOA was observed for SCA, followed by IPMNs. The individual accuracies of EUS-nCLE in diagnosing PCL subtypes were high for non-mucinous cysts (SCA: 98%; cystic-NET/SPN: 96%; pseudocyst: 96%) and slightly lower for mucinous lesions (IPMN: 86%; MCN: 84%).[45] In parallel to the aforementioned interobserver studies, we compared the diagnostic accuracy of EUS-nCLE to cyst fluid CEA and cytological analysis in differentiating mucinous from non-mucinous PCLs. nCLE demonstrated higher accuracy (98% sensitivity, 94% specificity, and 97% accuracy) compared to the reference standard of CEA and cytology analysis (74% sensitivity, 61% specificity, and 71% accuracy), p<001.[46]

While EUS-nCLE has proven to be a valuable and impactful diagnostic modality, its use is limited due to a lack of widespread expertise and structured training in developing this skill set. To address this, a sample training program was devised to assist endosonographers in accurately distinguishing between mucinous and non-mucinous PCLs. A database of 80 videos was curated for learners to review in a two-phase process. Following this, learners underwent competency assessments in interpreting EUS-nCLE videos. Competency was achieved with a median of 38 assessments. Post-training, learners successfully diagnosed mucinous and non-mucinous PCLs with high accuracy (89%) and confidence (83%), with a lasting effect observed after the eight-week washout period. Furthermore, the IOA was substantial (κ = 0.63) in this cohort of observers.[47]

Another obstacle in the widespread implementation of EUS-nCLE is the concern about the financial burden associated with the technology. The standard of care diagnosis for evaluation of PCL includes clinical history, PCL morphology (imaging studies and EUS), and cyst fluid analysis. A cost-benefit analysis of adding EUS-nCLE to the standard of care was conducted. In this cohort of 93 patients with a mean PCL diameter of 3.6 cm, there was a net saving of $442,438 (mean: $4,757.40/patient). Notably, adding EUS-nCLE to the standard of care prevented one unwarranted pancreatic surgery for every ten patients.[48]

3.4.2 Confocal Endomicroscopy for the Risk Stratification of IPMNs

Advanced neoplasia in IPMNs is progression to high-grade dysplasia and/or invasive adenocarcinoma (HGD-Ca). Our preliminary study focused on identifying novel endomicroscopy markers of histopathological progression in branch duct (BD) IPMNs. Blinded external observers reviewed high-yield nCLE videos from 26 patients with BD-IPMNs with reference histological diagnosis. The study found that the papillary

epithelial thickness (κ = 0.61) and darkness (κ = 0.55) demonstrated the highest IOA for detecting advanced neoplasia.[46,49] A combination of papillary thickness and darkness revealed the highest sensitivity (94%), specificity (100%), and area under the curve (AUC = 0.96%) for the detection of advanced neoplasia in IPMNs. This study demonstrated that parameters indicative of advanced neoplasia, such as cellular stratification, loss of polarity, and increased nuclear/cytoplasmic ratio, could be quantitatively measured in nCLE images.[49]

A potential barrier of incorporating EUS-nCLE for risk-stratification of IPMNs is the time and labor intensity of reviewing imaging to quantify papillary epithelial thickness and darkness. Hence, we developed an investigational convolutional neural network (CNN)-based artificial intelligence (AI) algorithm to detect high-grade dysplasia/adenocarcinoma in IPMNs using nCLE data.[50] The preliminary study used edited nCLE videos from 35 subjects with IPMNs, and surgical histopathology was available for comparison. Two CNN-AI algorithms were designed—one based on segmentation and measurement of papillary epithelial thickness and darkness and the other based on holistic nCLE feature extraction. For detecting HGD-Ca in IPMNs, the diagnostic performance of the CNN-AI algorithms was compared with the 2017 Fukuoka guidelines.[51] Both CNN-AI algorithms yielded higher sensitivity (CNN-AI 83% vs. Fukuoka guidelines 55.6%) and accuracy (CNN-AI 86% vs. Fukuoka guidelines 74%) with marginally lower specificity (CNN-AI 88% vs. Fukuoka guidelines 94%).[50,52] Developing such an algorithm could streamline the process of real-time image analysis, minimizing interobserver disagreement. This study demonstrated the potential of nCLE and AI algorithms in improving the diagnosis, risk stratification, and detection of HGD-Ca in BD-IPMNs.

3.5 Discussion

3.5.1 Advancements in Ongoing Research: Confocal Endomicroscopy and Artificial Intelligence Integration

The Confocal Laser Endomicroscopy as an IMaging Biomarker for the Diagnosis of Pancreatic Cystic Lesions (CLIMB) study (clinicaltrials.gov NCT 03492151) is an ongoing prospective multicenter study with 12 tertiary care centers. The aim of the CLIMB study is to assess the diagnostic accuracy of nCLE in conjunction with cyst fluid markers and imaging to distinguish mucinous from non-mucinous PCLs, diagnose specific subtypes, and detect IPMNs with advanced neoplasia.

3.5.2 Confocal Endomicroscopy for the Differentiation of Pancreatic Cysts

We pooled 194 patients with a definitive histopathological diagnosis from the INDEX single-center study from 2015–2018 (clinicaltrials.gov NCT 02516488) and the ongoing CLIMB multicenter study. We found 111 (57.2%) patients with mucinous PCL, and 87% had at least one worrisome feature based on the 2017 Fukuoka guidelines.[51] EUS-nCLE demonstrated a sensitivity, specificity, and accuracy of 98%, 95%, and 97%, respectively, for the detection of mucinous PCLs. In comparison, cyst fluid CEA, cytology, or fluid glucose reached a maximum accuracy of 80%. Overall, EUS-nCLE was effective in identifying specific PCL subtypes with accuracies of 88% for IPMNs (n = 78), 90% for MCNs (n = 30), 98% for SCAs (n = 40), and 94% for cystic-NET or SPN (n = 16).

3.5.3 Interobserver Agreement among Pathologists and Performance of Confocal Endomicroscopy in Risk-Stratification of IPMNs

Risk stratification of IPMNs with nCLE remains an area that requires continued improvement and research. We have conducted a series of projects assessing the performance of EUS-nCLE compared to the Fukuoka Criteria and IOA studies of pancreaticobiliary (PB) pathologists in diagnosing HGD and identifying IPMN subtypes.[51,53]

We compared the performance of the Fukuoka guidelines with preoperative EUS-nCLE in 65 patients who underwent surgical resection of IPMNs[51] from the INDEX (clinicaltrials.gov NCT 02516488) and CLIMB (clinicaltrials.gov NCT 0349215) study databases. Preoperative EUS-nCLE predicted HGD-Ca with a sensitivity, specificity, and accuracy of 68% (95% CI 48%–83%), 87% (95% CI 73%–94%), and 80% (95% CI 68%–88%), respectively. In comparison, the revised Fukuoka Criteria had a sensitivity, specificity, and accuracy 42% (95% CI 26%–61%), 95% (95% CI 83%–99%), and 74% (95% CI 62%–83%), respectively. Discrepancies in dysplasia classification were observed in 20% of cases, where nCLE indicated low-grade dysplasia (LGD) despite HGD-Ca on histopathology in eight cases (12.3%). Among these cases, five had HGD-Ca involvement in less than 25% of the BD-IPMN, as evaluated by two independent PB pathologists. Overall, nCLE augmented preoperative risk stratification of BD-IMPNs when compared to the reference standard Fukuoka Criteria. However, EUS-nCLE may be limited in accurately detecting localized HGD-Ca when the progression is confined to less than 25% of the cyst epithelium, particularly if it occurs beyond the reach of the EUS-FNA needle that houses the nCLE probe during *in vivo* examination.

There is limited data on IOA among PB pathologists in diagnosing HGD-Ca in resected BD-IPMNs. Hence, we assessed IOA between two blinded PB pathologists who reviewed 65 surgical specimens to differentiate dysplasia. There was substantial IOA between the two pathologists with a κ value of 0.71 (95% CI 0.54–0.88). The two PB pathologists disagreed in nine cases (14%). Therefore, a third PB pathologist reviewed the case and assisted with achieving consensus. Of note, in these cases, HGD-Ca was observed in less than 25% of the entire surface epithelium of BD-IPMNs. In total, 26 (40%) of BD-IPMNs were classified as HGD. The IOA for invasive adenocarcinoma was almost perfect with κ = 0.815 (95% CI 0.57–1.00). The IOA for the extent of HGD-Ca within a BD-IPMN was moderate (κ = 0.507, 95% CI 0.36–0.66). The takeaway from this analysis was that when HGD-Ca is focal (less than 25% of BD-IPMN), a second review may be helpful in identifying more subtle and localized areas of HGD within a cyst in clinical practice.

Accurate classification of the BD-IPMN subtype has prognostic relevance as non-gastric subtypes have malignant potential and risk of disease progression, recurrence, and metachronous IPMN development in the remaining pancreas.[54–58] We are also focusing efforts to identify preoperative morphologic characteristics associated with specific IPMN subtypes. Using data from the initial IOA study, one PB pathologist re-reviewed 65 cases to reclassify IPMNs by their specific subtypes—gastric, intestinal, and pancreatobiliary. Intraductal oncocytic papillary neoplasms (IOPN), a distinct tumor type from IPMN, was included in this analysis because both IOPN and IPMN present similarly as pancreatic cystic lesions due to their involvement of branch ducts. If disagreements occurred between initial and new subtype classifications, a second blinded PB pathologist reviewed the case and reached a consensus. We found 53 (81.5%) were of the gastric subtype and 12 (18.5%) were non-gastric (five intestinal, five pancreatobiliary, two oncocytic). Overall, 26 (40%) BD-IPMNs had HGD-Ca. The non-gastric subtypes were strongly associated with HGD-Ca (92%, $p< 0.001$). Pre-surgical nCLE characteristics were scrutinized in relation to the new classification of histopathological subtypes. The non-gastric BD-IMPNs more often had increased papillary epithelium thickness (≥ 50 microns; 67% vs. 23%, $p = 0.006$) and greater density (three or more papillae in one field; 58% vs. 21%, $p = 0.028$). By identifying these new possible criteria and pending further studies, EUS-nCLE can assist in the preoperative differentiation of IPMNs into gastric and non-gastric subtypes, which carries significant prognostic implications.

3.5.4 Functional Artificial Intelligence Algorithm for Risk-Stratification of IPMNs

Artificial intelligence can potentially revolutionize various aspects of patient management, including imaging biomarkers and early detection of cancers. We can harness the power of AI to enhance a clinician's ability to process fast-moving real-time nCLE imaging data efficiently. Our initial CNN-AI algorithm utilized human-edited high-yield nCLE videos (< 2 minutes) for training and testing. Our current research involves full-length, unedited nCLE videos ranging from seven to ten minutes. We created two new nCLE-AI algorithms that can process raw nCLE videos to identify high-yield frames capturing pathognomonic papillary features and another to risk-stratify BD-IPMNs by capturing key features for risk stratification (Figure 3.4). We compared our algorithms' performance to histopathology and the 2017 Fukuoka High-Risk Criteria.[51] We reviewed EUS-nCLE data from 66 patients, of which 41% had BD-IPMNs with HGD-Ca on pathology. The AI algorithm successfully detected HGD-Ca with a sensitivity of 78% (95% CI 59%–89%), specificity of 72% (95% CI 56%–84%), and diagnostic accuracy of 74% (95% CI 63%–83%). In comparison, the 2017 Fukuoka Criteria yielded a sensitivity of 44% (95% CI 28%–63%), specificity of 95% (95% CI 83%–99%), and comparable diagnostic accuracy of 74% (95% CI 63%–83%). The combination of EUS-nCLE AI and 2017 Fukuoka HR criteria improved the sensitivity, specificity, and accuracy to 85% (95% CI 68%–94%), 69% (95% CI 54%–81%), and 76% (95% CI 64%–85%), respectively. Our

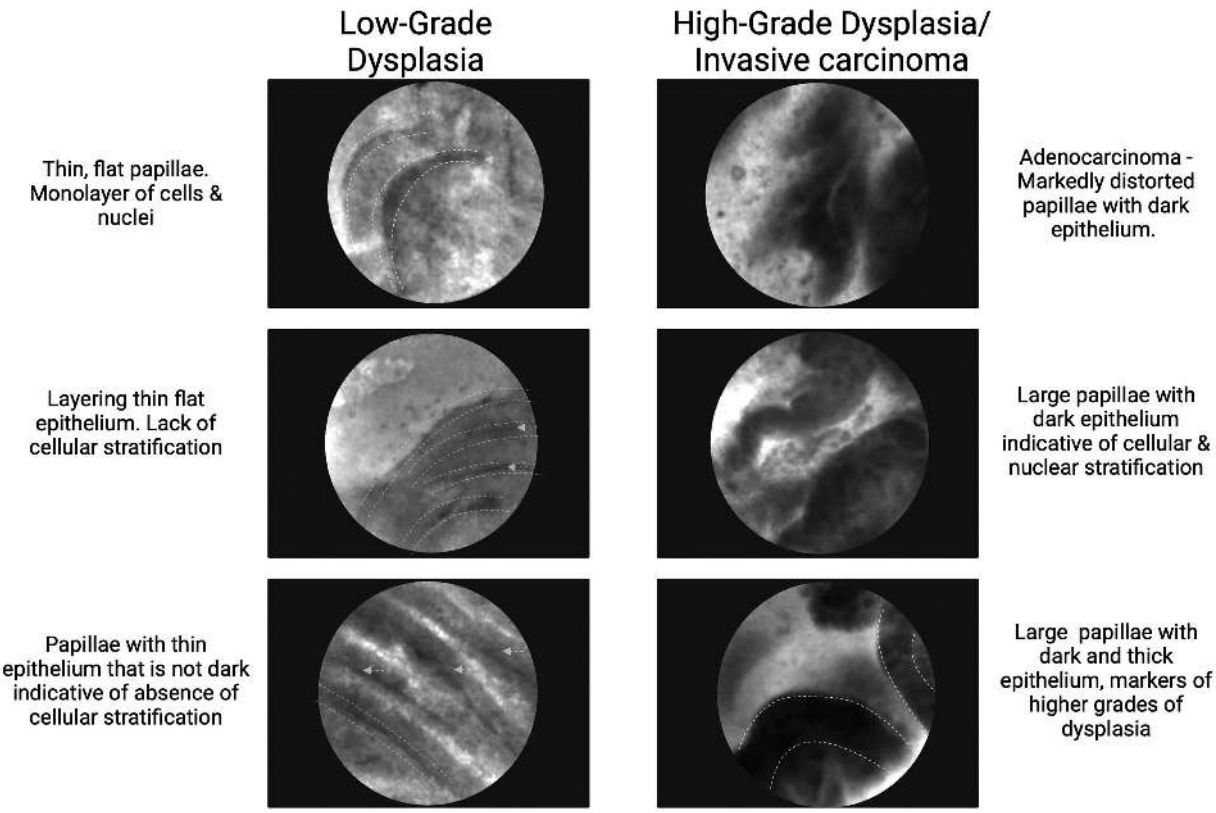


FIGURE 3.4 EUS-nCLE imaging features that differentiate IPMNs.

model compares favorably to established guidelines regarding diagnostic performance while remaining impartial and ensuring reproducibility by eliminating interobserver variability. With continued training and testing, the performance of these models will improve, and they can potentially be adapted to integrate key clinical data and information to fine-tune their performance. However, further studies are needed to validate these techniques.

3.6 Future Trends

Recent global developments in AI technology, the emergence of regenerative AI, and the enhanced technological hardware for processing large data will undoubtedly find application in medicine. While the current method of pancreatic cancer diagnosis and management has primarily relied on real-life human interpretation of images and laboratory findings, perhaps artificial intelligence could be applied to streamline the process of pancreatic cancer screening, detection, risk stratification, treatment, and management. Several AI algorithms have been developed, although it has yet to reach the point where it could be reliably applied to patient care.

Another area of future research in the preoperative risk stratification of IPMNs is the combination of EUS-nCLE with immune biomarkers that could allow for the detection or targeted treatment of advanced neoplasia. Antibodies and peptides are two examples of biomarkers conjugated to fluorescent agents to target various molecules and identify malignancy in CLE.[59,60] One study used monoclonal antibodies that targeted EGFR as molecular probes for CLE to improve the diagnosis of gastric cancer in an animal model and found the technique to be feasible.[61] Therefore, it is reasonable to consider that a similar approach could be undertaken in the preoperative risk stratification of PCLs and possibly in the targeted treatment of malignant PCLs.

The current 19-gauge needle used for nCLE also limits the surface area we can assess within a cystic lesion. We can evaluate roughly 40%–50% of the epithelial surface area in genu, body, and tail cysts, 30% of pancreatic head cysts, and 20% or less of uncinate process cysts. The needle is challenging to maneuver despite physical maneuvers (rotating torso and elevating arm) and the use of the large scope wheel. Since the surface area of the cyst epithelium sampled is limited, the ability to identify more of the cyst epithelium and particularly discrete areas of HGD-Ca within cysts is hindered. Decreasing the diameter of the mini probe will allow for housing the device in thinner EUS needles, permitting greater maneuverability and visualization of IPMNs. This is important since we have demonstrated that nCLE can potentially not detect HGD-Ca when involvement is less than 25% of the IPMN. Similarly, needles with tip deflection can also be useful in visualizing IPMN areas that are otherwise difficult to reach.

The development of EUS-nCLE has enhanced the accuracy of diagnosing specific types of PCLs and improved the risk stratification of IPMNs. We are conducting multicenter studies to substantiate this focus further and build an extensive repository of CLE imaging from PCLs. Currently, our focus is on generating, optimizing, and ultimately applying AI in real-time CLE-guided risk stratification of PCLs. This will improve precision management by reducing unwarranted pancreatic surgeries and enabling early detection of advanced neoplasia in precancerous pancreatic cysts.

REFERENCES

1. Institute NC. *Cancer Stat Facts: Pancreatic Cancer.* NIH National Cancer Institute.
2. Hu JX, Zhao CF, Chen WB, et al. Pancreatic cancer: a review of epidemiology, trend, and risk factors. *World J Gastroenterol*. 2021 Jul 21;27(27):4298–21. doi:10.3748/wjg.v27.i27.4298
3. Buerlein RCD, Shami VM. Management of pancreatic cysts and guidelines: what the gastroenterologist needs to know. *Ther Adv Gastrointest Endosc*. 2021 Jan–Dec;14:26317745211045769. doi:10.1177/26317745211045769
4. Canto MI, Hruban RH, Fishman EK, et al. Frequent detection of pancreatic lesions in asymptomatic high-risk individuals. *Gastroenterology*. 2012 Apr;142(4):796–804; quiz e14–5. doi:10.1053/j.gastro.2012.01.005
5. Nilsson LN, Keane MG, Shamali A, et al. Nature and management of pancreatic mucinous cystic neoplasm (MCN): a systematic review of the literature. *Pancreatology*. 2016;16(6):1028–36. doi:10.1016/j.pan.2016.09.011
6. Anand N, Sampath K, Wu BU. Cyst features and risk of malignancy in intraductal papillary mucinous neoplasms of the pancreas: a meta-analysis. *Clin Gastroenterol Hepatol*. 2013 Aug;11(8):913–21; quiz e59–60. doi:10.1016/j.cgh.2013.02.010
7. Yamao K, Yanagisawa A, Takahashi K, et al. Clinicopathological features and prognosis of mucinous cystic neoplasm with ovarian-type stroma: a multi-institutional study of the Japan pancreas society. *Pancreas*. 2011 Jan;40(1):67–71. doi:10.1097/MPA.0b013e3181f749d3
8. Goh BK, Tan YM, Chung YF, et al. A review of mucinous cystic neoplasms of the pancreas defined by ovarian-type stroma: clinicopathological features of 344 patients. *World J Surg*. 2006 Dec;30(12):2236–45. doi:10.1007/s00268-006-0126-1
9. Scheiman JM, Hwang JH, Moayyedi P. American gastroenterological association technical review on the diagnosis and management of asymptomatic neoplastic pancreatic cysts. *Gastroenterology*. 2015 Apr;148(4):824–48 e22. doi:10.1053/j.gastro.2015.01.014
10. Van Dyke TJ, Johlin FC, Bellizzi AM, Howe JR. Serous cystadenocarcinoma of the pancreas: clinical features and management of a rare tumor. *Dig Surg*. 2016;33(3):240–8. doi:10.1159/000444721
11. Slobodkin I, Luu AM, Höhn P, et al. Is surgery for serous cystic neoplasms of the pancreas still indicated? Sixteen years of experience at a high-volume center. *Pancreatology*. 2021 Aug;21(5):983–9. doi:10.1016/j.pan.2021.03.020
12. Jais B, Rebours V, Malleo G, et al. Serous cystic neoplasm of the pancreas: a multinational study of 2622 patients under the auspices of the International Association of Pancreatology and European Pancreatic Club (European Study Group on Cystic Tumors of the Pancreas). *Gut*. 2016 Feb;65(2):305–12. doi:10.1136/gutjnl-2015–309638
13. Assawasirisin C, Qadan M, Aimprasittichai S, et al. Pancreatic serous cystadenoma: a continuing diagnostic challenge. *Ann Surg*. 2024 Jan 17. doi:10.1097/SLA.0000000000006203

14. Lubezky N, Papoulas M, Lessing Y, et al. Solid pseudopapillary neoplasm of the pancreas: management and long-term outcome. *Eur J Surg Oncol*. 2017 Jun;43(6):1056–60. doi:10.1016/j.ejso.2017.02.001
15. Zhu JK, Wu D, Xu JW, et al. Cystic pancreatic neuroendocrine tumors: a distinctive subgroup with indolent biological behavior? A systematic review and meta-analysis. *Pancreatology*. 2019 Jul;19(5):738–50. doi:10.1016/j.pan.2019.05.462
16. Oh HC, Seo DW, Song TJ. Resolution of a septated pancreatic cyst by booster endoscopic ultrasonography-guided ablation. *J Dig Dis*. 2011 Dec;12(6):497–9. doi:10.1111/j.1751–2980.2011.00538.x
17. Allen PJ, D'Angelica M, Gonen M, et al. A selective approach to the resection of cystic lesions of the pancreas: results from 539 consecutive patients. *Ann Surg*. 2006;244(4):572.
18. Firkins SA, Hart PA, Porter K, et al. Incidence and risk factors for new-onset diabetes mellitus after surgical resection of pancreatic cystic lesions: a MarketScan study. *Pancreas*. 2022 May 1;51(5):427–34. doi:10.1097/MPA.0000000000002054
19. Ardeshna DR, Cao T, Rodgers B, et al. Recent advances in the diagnostic evaluation of pancreatic cystic lesions. *World J Gastroenterol*. 2022 Feb 14;28(6):624–34. doi:10.3748/wjg.v28.i6.624
20. Cazacu IM, Luzuriaga Chavez AA, Saftoiu A, Vilmann P, Bhutani MS. A quarter century of EUS-FNA: progress, milestones, and future directions. *Endosc Ultrasound*. 2018;7(3):141–60. doi:10.4103/eus.eus_19_18
21. Chung A, Kwan V. Endoscopic ultrasound: an overview of its role in current clinical practice. *Australas J Ultrasound Med*. 2009 May;12(2):21–9. doi:10.1002/j.2205–0140.2009.tb00050.x
22. Elta GH, Enestvedt BK, Sauer BG, Lennon AM. ACG clinical guideline: diagnosis and management of pancreatic cysts. *Am J Gastroenterol*. 2018 Apr;113(4):464–79. doi:10.1038/ajg.2018.14
23. Thornton GD, McPhail MJ, Nayagam S, Hewitt MJ, Vlavianos P, Monahan KJ. Endoscopic ultrasound guided fine needle aspiration for the diagnosis of pancreatic cystic neoplasms: a meta-analysis. *Pancreatology*. 2013;13(1):48–57. doi:10.1016/j.pan.2012.11.313
24. Wu QM, Guo YN, Xu YQ. Diagnostic performance of endoscopic ultrasound-guided fine-needle aspiration in pancreatic lesions. *Eur Rev Med Pharmacol Sci*. 2018 Mar;22(5):1397–401. doi:10.26355/eurrev_201803_14485
25. Nigam N, Rastogi A, Bhatia V, Sureka B, Jain P, Bihari C. EUS-guided FNA in diagnosing pancreatic lesions: strength and cytological spectrum. *J Cytol*. 2019;36(4):189–95. doi:10.4103/JOC.JOC_5_18
26. Muthusamy VR, Chandrasekhara V, Acosta RD, et al. The role of endoscopy in the diagnosis and treatment of cystic pancreatic neoplasms. *Gastrointest Endosc*. 2016 Jul;84(1):1–9. doi:10.1016/j.gie.2016.04.014
27. Ohtsuka T, Fernandez-Del Castillo C, Furukawa T, et al. International evidence-based Kyoto guidelines for the management of intraductal papillary mucinous neoplasm of the pancreas. *Pancreatology*. 2023 Dec 28. doi:10.1016/j.pan.2023.12.009
28. Pfluger MJ, Jamouss KT, Afghani E, et al. Predictive ability of pancreatic cyst fluid biomarkers: a systematic review and meta-analysis. *Pancreatology*. 2023 Nov;23(7):868–77. doi:10.1016/j.pan.2023.05.005
29. Paniccia A, Polanco PM, Boone BA, et al. Prospective, multi-institutional, real-time next-generation sequencing of pancreatic cyst fluid reveals diverse genomic alterations that improve the clinical management of pancreatic cysts. *Gastroenterology*. 2023 Jan;164(1):117–33.e7. doi:10.1053/j.gastro.2022.09.028
30. Springer S, Wang Y, Dal Molin M, et al. A combination of molecular markers and clinical features improve the classification of pancreatic cysts. *Gastroenterology*. 2015 Nov;149(6):1501–10. doi:10.1053/j.gastro.2015.07.041
31. Saghir SM, Dhindsa BS, Daid SGS, et al. Efficacy of EUS-guided needle-based confocal laser endomicroscopy in the diagnosis of pancreatic lesions: a systematic review and meta-analysis. *Endosc Ultrasound*. 2022;11(4):275–82. doi:10.4103/EUS-D-20-00122
32. Bhutani MS, Koduru P, Joshi V, et al. EUS-guided needle-based confocal laser endomicroscopy: a novel technique with emerging applications. *Gastroenterol Hepatol*. 2015 Apr;11(4):235–40.
33. Neumann H, Kiesslich R, Wallace MB, Neurath MF. Confocal laser endomicroscopy: technical advances and clinical applications. *Gastroenterology*. 2010 Aug;139(2):388–92, 392.e1–2. doi:10.1053/j.gastro.2010.06.029
34. Buchner AM, Shahid MW, Heckman MG, et al. Comparison of probe-based confocal laser endomicroscopy with virtual chromoendoscopy for classification of colon polyps. *Gastroenterology*. 2010 Mar;138(3):834–42. doi:10.1053/j.gastro.2009.10.053
35. Kiesslich R, Gossner L, Goetz M, et al. In vivo histology of Barrett's esophagus and associated neoplasia by confocal laser endomicroscopy. *Clin Gastroenterol Hepatol*. 2006 Aug;4(8):979–87. doi:10.1016/j.cgh.2006.05.010
36. Pech O, Rabenstein T, Manner H, et al. Confocal laser endomicroscopy for in vivo diagnosis of early squamous cell carcinoma in the esophagus. *Clin Gastroenterol Hepatol*. 2008 Jan;6(1):89–94. doi:10.1016/j.cgh.2007.10.013
37. Kiesslich R, Goetz M, Burg J, et al. Diagnosing Helicobacter pylori in vivo by confocal laser endoscopy. *Gastroenterology*. 2005 Jun;128(7):2119–23. doi:10.1053/j.gastro.2004.12.035
38. Günther U, Daum S, Heller F, et al. Diagnostic value of confocal endomicroscopy in celiac disease. *Endoscopy*. 2010 Mar;42(3):197–202. doi:10.1055/s-0029-1243937
39. Kiesslich R, Goetz M, Lammersdorf K, et al. Chromoscopy-guided endomicroscopy increases the diagnostic yield of intraepithelial neoplasia in ulcerative colitis. *Gastroenterology*. 2007 Mar;132(3):874–82. doi:10.1053/j.gastro.2007.01.048
40. Breda A, Territo A, Guttilla A, et al. Correlation between confocal laser endomicroscopy (Cellvizio). *Eur Urol Focus*. 2018 Dec;4(6):954–9. doi:10.1016/j.euf.2017.05.008
41. Kramer T, Wijsman PC, Kalverda KA, Bonta PI, Annema JT. Advances in bronchoscopic optical coherence tomography and confocal laser endomicroscopy in pulmonary diseases. *Curr Opin Pulm Med*. 2023 Jan 01;29(1):11–20. doi:10.1097/MCP.0000000000000929

42. Karia K, Kahaleh M. A review of probe-based confocal laser endomicroscopy for pancreaticobiliary disease. *Clin Endosc*. 2016 Sep;49(5):462–6. doi:10.5946/ce.2016.086
43. Krishna SG, Modi RM, Kamboj AK, et al. *In vivo* and *ex vivo* confocal endomicroscopy of pancreatic cystic lesions: a prospective study. *World J Gastroenterol*. 2017;23(18):3338–48.
44. Krishna SG, Brugge WR, Dewitt JM, et al. Needle-based confocal laser endomicroscopy for the diagnosis of pancreatic cystic lesions: an international external interobserver and intraobserver study (with videos). *Gastrointest Endosc*. 2017 Oct;86(4):644–54.e2. doi:10.1016/j.gie.2017.03.002
45. Machicado JD, Napoleon B, Lennon AM, et al. Accuracy and agreement of a large panel of endosonographers for endomicroscopy-guided virtual biopsy of pancreatic cystic lesions. *Pancreatology*. 2022 Nov;22(7):994–1002. doi:10.1016/j.pan.2022.08.012
46. Krishna SG, Hart PA, Malli A, et al. Endoscopic ultrasound-guided confocal laser endomicroscopy increases accuracy of differentiation of pancreatic cystic lesions. *Clin Gastroenterol Hepatol*. 2020 Feb;18(2):432–40 e6. doi:10.1016/j.cgh.2019.06.010
47. Machicado JD, Napoleon B, Akshintala V, et al. Structured training program on confocal laser endomicroscopy for pancreatic cystic lesions: a multicenter prospective study among early-career endosonographers (with video). *Gastrointest Endosc*. 2023 Dec;98(6):953–64. doi:10.1016/j.gie.2023.07.019
48. Luthra A, Hart PA, Papachristou GI, et al. Cost-benefit analysis and resource implications of endoscopic ultrasound-guided confocal endomicroscopy in pancreatic cysts. *Techn Innov Gastrointest Endosc*. 2022;24(1):35–44.
49. Krishna SG, Hart PA, DeWitt JM, et al. EUS-guided confocal laser endomicroscopy: prediction of dysplasia in intraductal papillary mucinous neoplasms (with video). *Gastrointest Endosc*. Mar 2020;91(3):551–63.e5. doi:10.1016/j.gie.2019.09.014
50. Jiang J, Chao WL, Cao T, et al. Improving pancreatic cyst management: artificial intelligence-powered prediction of advanced neoplasms through endoscopic ultrasound-guided confocal endomicroscopy. *Biomimetics*. 2023 Oct 19;8(6). doi:10.3390/biomimetics8060496
51. Tanaka M, Fernandez-Del Castillo C, Kamisawa T, et al. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. *Pancreatology*. 2017 Sep–Oct;17(5):738–53. doi:10.1016/j.pan.2017.07.007
52. Machicado JD, Chao WL, Carlyn DE, et al. High performance in risk stratification of intraductal papillary mucinous neoplasms by confocal laser endomicroscopy image analysis with convolutional neural networks (with video). *Gastrointest Endosc*. 2021 Jan 16. doi:10.1016/j.gie.2020.12.054
53. Longnecker DS, Adsay NV, Fernandez-del Castillo C, et al. Histopathological diagnosis of pancreatic intraepithelial neoplasia and intraductal papillary-mucinous neoplasms: interobserver agreement. *Pancreas*. 2005 Nov;31(4):344–9. doi:10.1097/01.mpa.0000186245.35716.18
54. Kwon JE, Jang KT, Ryu Y, et al. Subtype of intraductal papillary mucinous neoplasm of the pancreas is important to the development of metachronous high-risk lesions after pancreatectomy. *Ann Hepatobiliary Pancreat Surg*. 2019 Nov;23(4):365–71. doi:10.14701/ahbps.2019.23.4.365
55. Furukawa T, Hatori T, Fujita I, et al. Prognostic relevance of morphological types of intraductal papillary mucinous neoplasms of the pancreas. *Gut*. 2011 Apr;60(4):509–16. doi:10.1136/gut.2010.210567
56. Ideno N, Ohtsuka T, Kono H, et al. Intraductal papillary mucinous neoplasms of the pancreas with distinct pancreatic ductal adenocarcinomas are frequently of gastric subtype. *Ann Surg*. 2013 Jul;258(1):141–51. doi:10.1097/SLA.0b013e31828cd008
57. Mino-Kenudson M, Fernández-del Castillo C, Baba Y, et al. Prognosis of invasive intraductal papillary mucinous neoplasm depends on histological and precursor epithelial subtypes. *Gut*. 2011 Dec;60(12):1712–20. doi:10.1136/gut.2010.232272
58. Miyasaka Y, Ohtsuka T, Tamura K, et al. Predictive factors for the metachronous development of high-risk lesions in the remnant pancreas after partial pancreatectomy for intraductal papillary mucinous neoplasm. *Ann Surg*. 2016 Jun;263(6):1180–7. doi:10.1097/SLA.0000000000001368
59. Hsiung PL, Hardy J, Friedland S, et al. Detection of colonic dysplasia in vivo using a targeted heptapeptide and confocal microendoscopy. *Nat Med*. 2008 Apr;14(4):454–8. doi:10.1038/nm1692
60. Sturm MB, Joshi BP, Lu S, et al. Targeted imaging of esophageal neoplasia with a fluorescently labeled peptide: first-in-human results. *Sci Transl Med*. 2013 May 08;5(184):184ra61. doi:10.1126/scitranslmed.3004733
61. Hoetker MS, Kiesslich R, Diken M, et al. Molecular in vivo imaging of gastric cancer in a human-murine xenograft model: targeting epidermal growth factor receptor. *Gastrointest Endosc*. 2012 Sep;76(3):612–20. doi:10.1016/j.gie.2012.05.013

4

Portable Confocal Microscopy for Aiding Diagnosis and Treatment of Skin Cancer

Momoka Sugimura, Jingwei Zhao, Kenneth Marcelino, Yongjun Kim, Rafael Romero, and Dongkyun Kang

4.1 Introduction

Skin cancer remains a grave health concern. It is the most common type of cancer in the United States [1]. Including both melanoma and non-melanoma skin cancers, the incidence of skin cancer surpasses that of all other cancers, and the incidence rate continues to increase. Annually, the United States records approximately 5.4 million new cases of basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) [2]. These non-melanoma skin cancers are not generally life threatening but still cause significant morbidities and psychological and financial burdens. Conversely, melanoma had 82,054 new cases and 11,906 deaths in the United States in 2019 [3]. The financial burden of skin cancer treatment in the United States exceeds $8 billion annually, ranking skin cancer as the fifth most expensive cancer covered by Medicare [4].

This issue is further complicated for those living in rural areas, who face substantial barriers to accessing dermatology care. Rural residents experience higher rates of skin cancer, partly due to the increased exposure to UV radiation from outdoor work, such as farming [5]. Seeking treatment often means traveling to dermatology clinics in urban areas, which brings about additional financial and logistical challenges such as transportation costs and long delays for appointments. A study between 2014 and 2018 found that individuals in rural Michigan had a nearly 50% higher risk of developing and dying from melanoma compared to urban residents [6]. Similarly, people experiencing homelessness have up to a 25% increased risk of malignant or premalignant skin growth, largely due to the increased sun exposure and lack of healthcare access: as many as 76% of people experiencing homelessness never underwent a skin exam [7]. These situations underscore the significant disparities in skin cancer care adversely affecting underserved populations.

The current standard of care for diagnosing skin cancer includes dermoscopy and skin biopsy. Dermoscopy is a noninvasive imaging tool that provides magnified images of skin lesions. Dermoscopy's ability to reveal detailed skin lesion patterns makes it invaluable for identifying potentially malignant skin lesions. However, its effectiveness is somewhat dependent on the clinician's expertise, with sensitivity and specificity varying widely from 32% to 90%, making it less reliable for lesions with equivocal characteristics [8].

When a suspicious lesion is identified with dermoscopy, the lesion is biopsied for further histologic examination. This procedure involves the removal of skin tissue, which is then fixed, embedded, sectioned, stained, and examined under a microscope. There are several types of skin biopsies, including shave biopsy (removing a small section of the top layer of skin), punch biopsy (removing a circular section down to the subcutaneous layer), and excisional biopsy (removing an entire lump or area of the skin lesion). The choice of biopsy type depends on the suspected condition. The procedure typically involves local anesthesia and may require stitches, depending on the depth and size of the tissue removed. While skin biopsies are necessary for accurate diagnosis and treatment of skin cancer, the invasive nature of biopsies often leads to scarring and potential discomfort. Moreover, if the biopsy reveals malignancy, it often requires additional visits for the complete removal of the malignant tissue. One challenge in the current diagnostic process with dermoscopy is ensuring the complete removal of cancerous lesions, as dermoscopy does not provide sufficient resolution to map out the accurate margin of cancerous lesions.

Over the past two decades, reflectance confocal microscopy (RCM) has emerged as a promising diagnostic tool for addressing some of the challenges in the standard methods for diagnosing skin cancers. RCM is a noninvasive imaging technique that provides real-time, cellular-resolution images of the epidermis and superficial dermis. Integrating RCM with the standard dermatology workflow could significantly enhance the accuracy of skin cancer detection. RCM particularly excels at identifying skin lesions that are difficult for dermoscopy to assess, increasing the specificity to around 90%, while maintaining high sensitivity [8]. This advancement helps minimize unnecessary biopsies of benign lesions. Furthermore, decreasing the number of unnecessary biopsies of benign lesions with RCM can also reduce the financial burden on both patients and the healthcare system. In an analysis of 619 lesions examined via RCM, 326 (50.2%) were determined to be benign [9]. The overall expenses associated with RCM diagnostics and follow-up processes amounted to $42,037, which was significantly less than the $95,905 that biopsies of the 326 benign lesions would have cost [9]. RCM also can be beneficial in guiding excisional procedures. Before excising a lesion, surgeons can use RCM to delineate the margins of a tumor more accurately. This ensures the complete removal of tumors through excision and can lead to fewer surgical sessions and a lower risk of recurrence [10]. Beyond surgical applications, RCM proves valuable in monitoring the effectiveness of noninvasive treatments, offering a

DOI: 10.1201/9781003449942-6

noninvasive method to assess treatment outcomes and guiding subsequent patient care decisions [10].

Figure 4.1 illustrates the schematic of RCM. Light from a source (e.g., near-infrared laser) passes through a beam expander and is reflected by a beam splitter. A beam scanning mechanism is used to steer the field angle of the light on an objective lens. The objective lens focuses the light onto a tight spot on the sample. Certain cellular structures in the skin reflect the incident light. For instance, cell membranes and cytoplasm of keratinocytes in the epidermis reflect the incident light more strongly than the nuclei, providing contrast for visualizing a honeycomb pattern of keratinocytes. The reflected light from the sample is collected by the objective lens, de-scanned by the scanning mechanism, and passes through the beam splitter. After the beam splitter, the light is focused by a focusing lens onto a pinhole. The pinhole effectively rejects the light originating from out-of-focus regions of the sample and therefore mainly lets in-focus light pass through, which is detected by a photodetector. The commercially available implementations of RCM are Vivascope 1500 and 3000 (Caliber ID, Rochester, NY), with the 3000 device having the benefit of being handheld for increased portability [11]. The Vivascope RCM devices provide a lateral resolution of 0.5–1.0 μm and an axial resolution of 2–5 μm. Vivascope 1500 has a field of view (FOV) of 500 μm × 500 μm, while Vivascope 3000 has a FOV of 750 μm × 750 μm. RCM images are acquired at the speed of 6–9 fps. Mosaicking is usually done to visualize the entire skin lesion up to the area of 8 mm × 8 mm in Vivascope 1500.

Recognizing these benefits of RCM, the Centers for Medicare and Medicaid Services granted reimbursement codes for the acquisition and interpretation of RCM images [8]. This is a significant step towards making RCM more accessible to a wide range of healthcare providers beyond a few academic and research hospitals. However, the widespread utilization of RCM has been hampered partly by the high device cost.

Another potential advancement that could enhance the utilization of RCM is increasing the imaging speed. With the FDA-cleared RCM devices, the procedural time required to capture RCM images for a single skin lesion is not negligible, comparable to the time necessary for preparing, explaining, and conducting a biopsy procedure [12]. This is partly because the RCM imaging speed is limited to 6–9 frames/sec (fps) [11], and the acquisition of numerous RCM images prolongs the procedural time. Additionally, patient movement during imaging can introduce motion artifacts, compromising the image quality. High imaging speed can help mitigate these issues, ensuring the acquisition of clear images even when minor patient movements exist.

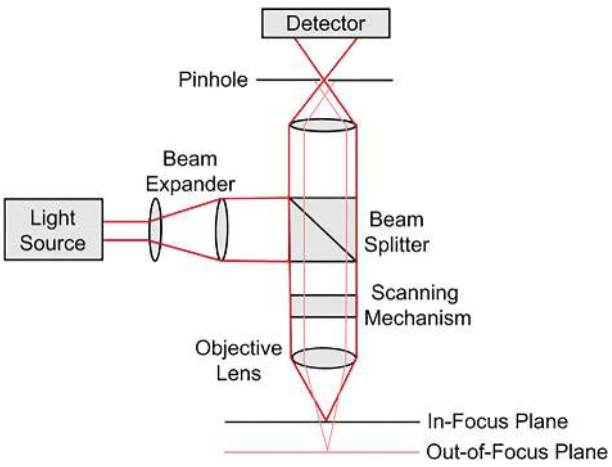


FIGURE 4.1 Schematic of reflectance confocal microscopy (RCM).

In this chapter, we will discuss the development of a new RCM technique termed portable confocal microscopy (PCM). In PCM, expensive optoelectrical components needed for RCM imaging such as galvo scanners and high-bandwidth detectors are replaced by inexpensive optoelectrical components such as diffraction gratings and consumer-grade CMOS sensors [13, 14]. We will first review the working principle of RCM and PCM and explain how PCM can be built with a low material cost. We will then review the standard methods and materials for acquiring PCM images of skin lesions. Finally, we will present representative PCM images of common skin lesions and discuss the limitations and future directions.

4.2 Materials

4.2.1 Device Components

An LED (part number = SFH 4780S; central wavelength = 820 nm; 40%-of-maximum bandwidth = 40 nm; emission area = 0.75 mm × 0.75 mm; viewing angle = 20°; maximum optical power = 660 mW) was purchased from Osram. A fold mirror (part number = 43-878), and a multi-element camera lens (part number = 59-781, f = 50 mm) were purchased from Edmund Optics. An aspherical collimator (part number = 354240-B; f = 8 mm), a cylindrical lens (part number = LJ1874L1-B; f = 7.7 mm), three achromatic doublets (part number ACN127-030-B; f = 30 mm), and another fold mirror (part number = PFSQ05-03-M01) were purchased from Thorlabs. A water-immersion objective lens (part number = MRD07420; magnification = 40; NA = 0.8) was purchased from Nikon. Illumination and detection slits (part number = 2-25-3+M-0.5; width = 25 μm; length = 3 mm) were purchased from National Apertures. A transmission grating (part number = PING-Sample-012; groove density = 1,765 lpmm) was purchased from Ibsen Photonics. A CMOS sensor (part number = acA1300-200 μm; 1280 × 1024 pixels; pixel size = 4.8 μm) manufactured by Basler was purchased from Graftek Imaging.

Most of the optical elements were mounted in custom, 3D-printed mechanical holders (printed by Form 3, Formlabs, and uPrint SE Plus; Stratasys). A miniature, three-axis translation mount (part number = CXYZ05A) used to align the detection slit was purchased from Thorlabs.

4.2.2 Materials for Testing and Imaging

A positive USAF resolution target (part number = 58–198; Minimum = Group 0, Element 1; Maximum = Group 9, Element 3) was purchased from Edmund Optics. An index-matching gel manufactured from Alcon was purchased from Amazon. A skin adapter was 3D printed.

4.3 Methods

4.3.1 Portable Confocal Microscopy

Portable confocal microscopy utilizes a scanning technique termed spectrally encoded imaging, where spectral components of light are used to encode a spatial coordinate of the sample. The schematic of a PCM device is shown in Figure 4.2. Light from a near-infrared LED is focused onto the illumination slit by a condenser lens and cylindrical lens. The line illumination from the illumination slit is then collimated by a collimation lens and is diffracted by a grating. The diffracted light is focused on the sample by an objective lens. Each line of the sample is illuminated by a distinctive wavelength. The light scattered back from the sample is collected by the same objective lens. Dispersion by the first grating is reversed by the second grating, making light from the sample focused on the detection slit regardless of the illumination wavelength. Light after the detection slit is dispersed by the third grating to generate two-dimensional confocal images on the CMOS sensor. The confocal images are then transferred to a laptop or other mobile computing platforms via a USB cable.

The PCM alignment process combines alignment of each sub-components and alignment of the overall confocal conjugation. The process begins with aligning the camera lens coupled with a CMOS sensor using a bandpass filter to achieve a sharp image of distant objects at the design wavelength. The illumination unit, including the LED, condenser lens, cylindrical lens, and illumination slit, is mounted on a fixed 3D-printed base plate. The sharpness of the illumination slit is then optimized by adjusting the illumination collimation lens (collimation lens 1). Subsequently, components like gratings, mirrors, and an objective lens are passively mounted into a custom, 3D-printed holder using tight fits. The detection collimation lens (collimation lens 2) is aligned by using a laser diode to illuminate the detection slit from the tip of the objective lens. Adjustments to the detection collimation lens are made by finding the smallest width of the slit image. Finally, a focusing lens is placed in front of the detection slit. The lateral positions of the detection slit and the axial position of the focusing lens are adjusted, along with the axial position of the USAF resolution target around the focal plane of the objective lens. These three components are adjusted until the smallest resolvable patterns are visualized and the smallest axial response is achieved.

The PCM imaging performance is tested by imaging the USAF resolution target and measuring the width of line spread function. The axial response curve is measured by acquiring a sequence of PCM images while translating the USAF resolution target. PCM images are acquired by custom software developed on the LabVIEW (manufacturer = National Instruments) platform.

The PCM devices were developed with a material cost of $4,000–$5,000. The off-the-shelf objective lens cost was the biggest contributor to the material cost, which can be reduced with the development of a custom objective lens using fewer lens elements. The dimensions of the device were 22 cm × 17.5 cm × 10 cm, and the weight was about 1 kg. The lateral and axial resolution was 1.6 μm and 6.0 μm, respectively, which was comparable to the resolution of the FDA-cleared RCM devices. The illumination power on the skin was 0.16 mW, significantly lower than the maximum permissible light exposure per the ANSI standard. The imaging speed was 20 fps.

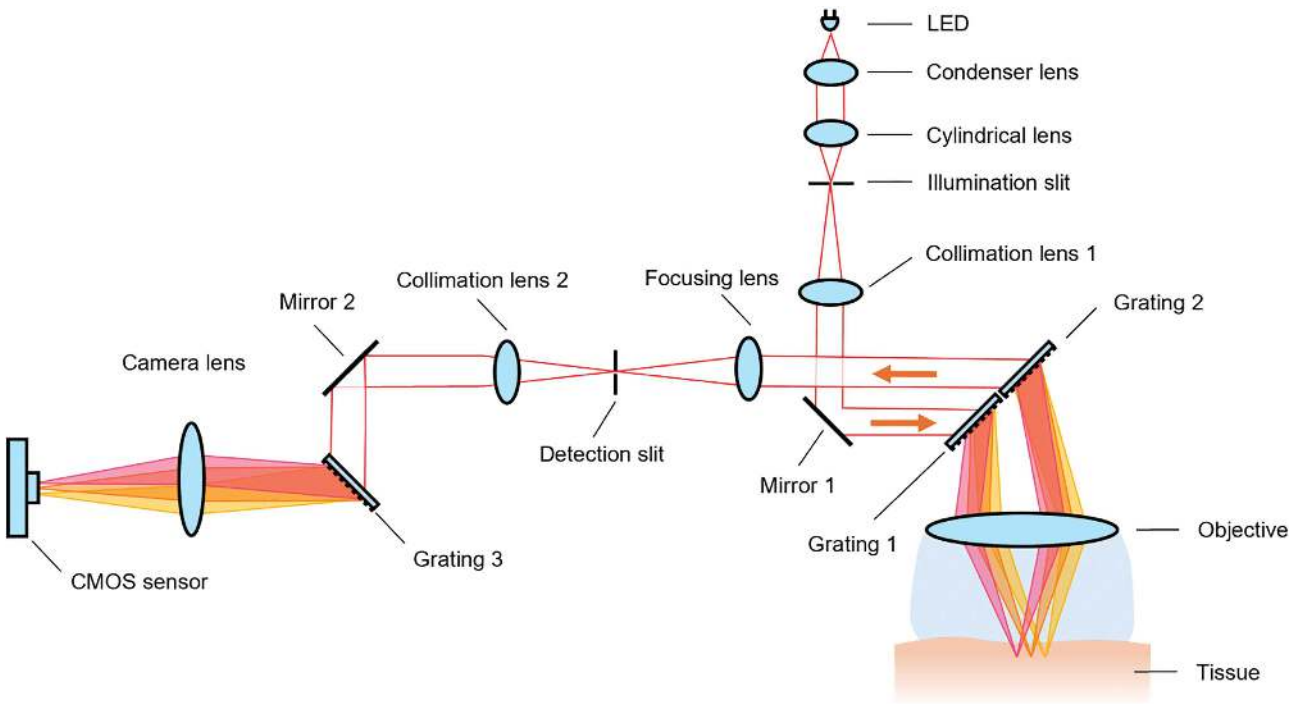


FIGURE 4.2 Schematic of portable confocal microscopy (PCM).

PCM offers several key benefits that make it promising for widespread utilization. First, PCM's use of an inexpensive near-infrared LED for illumination and diffraction gratings for confocal imaging significantly reduces the device cost and complexity compared to standard RCM devices that use lasers and scanning mirrors. Moreover, the PCM design obviates the need for mechanically moving parts, contributing to a more compact device size. PCM operates on a rechargeable battery that powers the LED, and a single USB cable needs to be connected to a laptop for operation. The imaging speed of PCM is no longer constrained by mechanical scanning limits but by the imaging speed of the CMOS sensor (typically higher than 200 fps) and the signal-to-noise ratio.

4.3.2 Skin Imaging Procedure

A typical procedure for imaging skin lesions with PCM is illustrated in Figure 4.3. First, double-sided adhesive ring tape is attached to the skin with the opening of the adhesive ring tape centered on the skin lesion of interest. A skin adapter, composed of a 3D-printed holder and a magnetic washer, is attached to the adhesive ring tape. Mineral oil (refractive index = 1.47) or liquid with a similar refractive index is applied on the skin lesion to fill in the ridges of the skin and index matching. A transparent plastic disc (material = fluorinated ethylene propylene, FEP, refractive index = 1.34) is placed over the mineral oil and attached to the double-sided adhesive ring tape to flatten the surface being imaged. An index-matching gel is applied to the FEP disc, which makes contact with the water-immersion objective lens. The procedure concludes with the placement of the distal end of the PCM device on the metallic washer. The distal end of the PCM device has multiple miniature magnets, which establish a stable connection between the PCM device and the skin.

Alternatively, the PCM device can be operated without making a stable connection with the skin but rather maneuvered by free hands. This operational mode provides a quick method for imaging skin lesions, albeit a less stable connection between the device and the skin. In the free-hand operation mode, a double-sided adhesive ring tape is attached to the distal end of the PCM device. An index-matching gel is applied on the distal end of the objective lens. An FEP disc is attached to the double-sided adhesive ring tape, while uniformly spreading the index-matching gel. The procedure is completed by applying mineral oil directly on the skin and gently placing the PCM device on the skin lesion.

After selecting the skin lesion for examination, the device is powered on by turning on the battery and LED. Real-time PCM images are transferred to a laptop via a USB cable and displayed on the laptop screen. To enhance the visibility of the lesion, the exposure time and/or gain is adjusted as needed. Increasing the gain while maintaining the exposure time increases brightness without reducing the imaging speed. The PCM device is laterally translated as needed to ensure comprehensive coverage of the skin lesion. At a region of interest, axial scanning of the objective lens is initiated to acquire PCM images from multiple imaging depths. Upon completing the imaging session, the PCM device is powered off by turning off

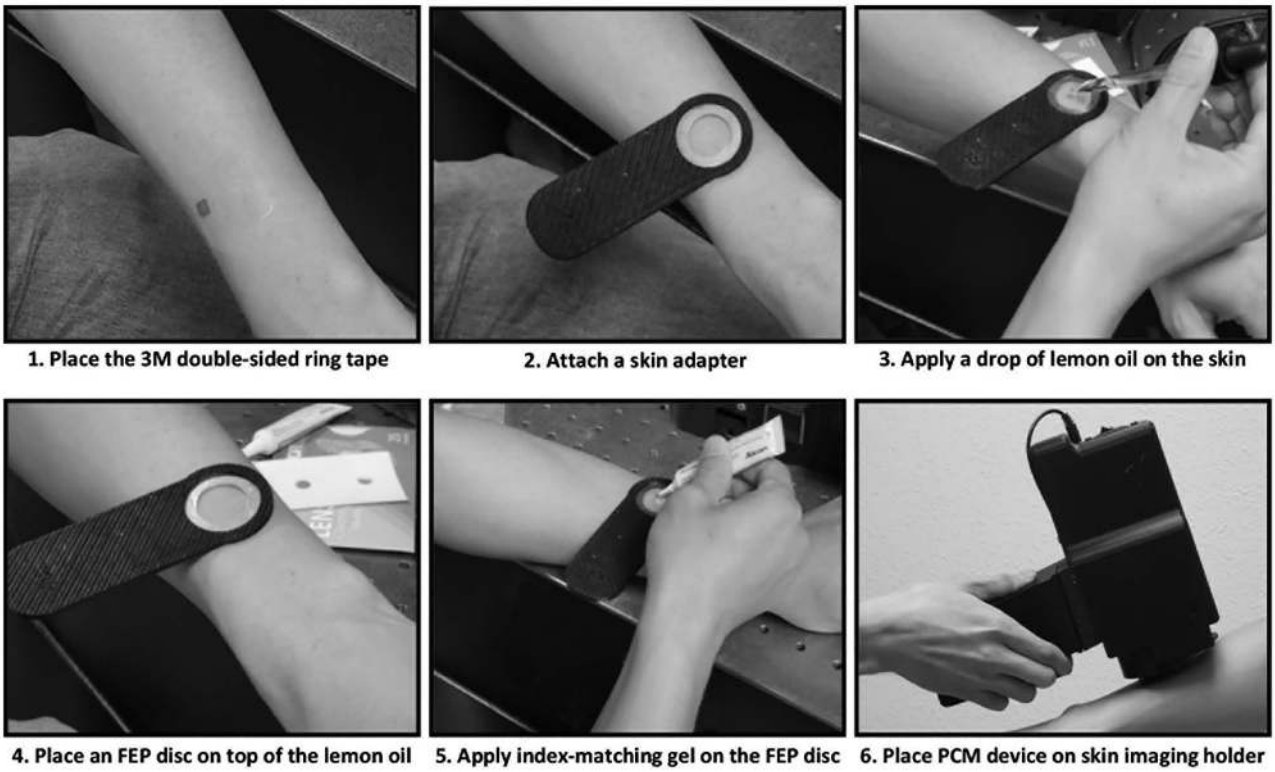


FIGURE 4.3 PCM imaging procedure.

the LED and battery, and the USB cable is disconnected from the laptop. The objective lens is cleaned with an alcohol swab, similar to the cleaning of RCM and dermoscopy devices.

4.4 Results

We evaluated the performance of PCM in comparison to the commercial RCM device (Vivascope 1500) [15]. When imaging basal cell carcinoma, PCM visualized well-known features of BCC such as tumor islands (Figure 4.4A), convoluted blood vessels (Figure 4.4B), and dark silhouette (Figure 4.4C) in a similar manner to RCM images obtained from the corresponding skin lesions (Figure 4.4D–F). PCM was also found promising in visualizing characteristic features of benign lesions, including keratinocyte disarray in actinic keratosis (Figure 4.4G), curved and tubular vessels in seborrheic keratosis (Figure 4.4H), and edged dermal papillae in melanocytic nevi (Figure 4.4I). These features were similar to those visualized in corresponding RCM images (Figure 4.4J–L).

PCM has several advantages over the standard RCM devices. The portability of the device makes it easy to image lesions at anatomically hidden locations and facilitates the transport of the device between exam rooms. The PCM imaging speed is higher than the speed of the standard RCM devices, which could facilitate imaging of dynamic events such as blood flow. Imaging of blood flow was previously shown as beneficial in monitoring patients with graft-versus-host disease [16].

4.5 Discussion

PCM is a low-cost implementation of an FDA-cleared RCM technology. By using diffraction gratings, broadband LEDs, and slit apertures to obviate the need for beam-scanning devices and their associated electronics, PCM can reduce the cost and size of RCM devices. PCM prototypes were built with a low material cost ($4,000–5,000) and are lightweight (1 kg). The imaging speed is 20 fps, which surpasses that of the standard RCM devices.

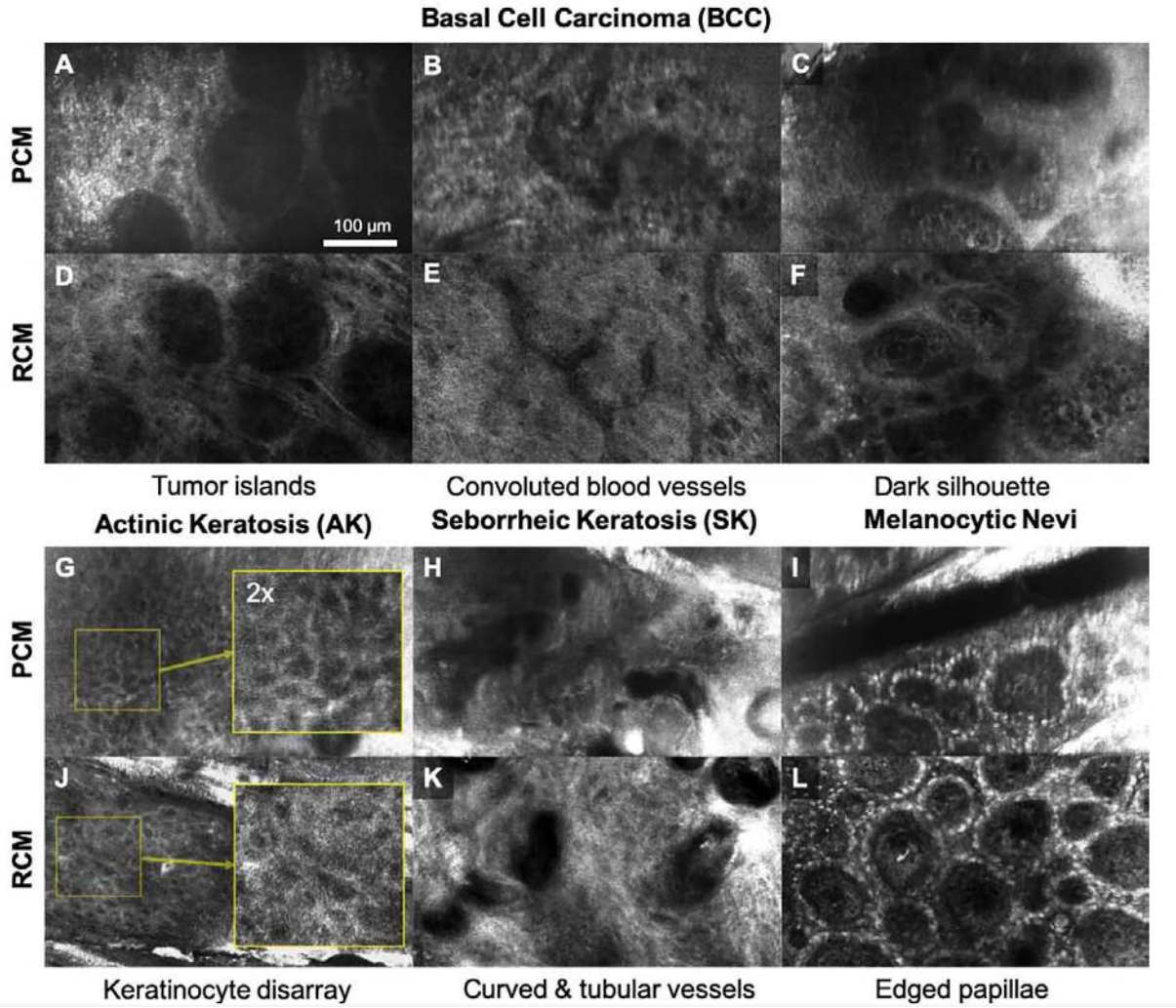


FIGURE 4.4 PCM and RCM images of skin lesions commonly encountered in dermatology care. A–C, G–I: PCM images; D–F, J–L: RCM images. Figures reprinted with permission from Reference 15.

Recently, an enhanced version of PCM utilizing a superluminescent diode and rotating diffuser was developed to increase imaging capabilities [17]. This version achieved a remarkably fast imaging speed of human skin at 107 fps due to the higher power of the superluminescent diode. This PCM device also achieved the high lateral resolution of 1.1 μm and axial resolution of 2.8 μm while minimizing the speckle artifacts.

A pilot imaging study showed that PCM could visualize key morphologic features associated with malignant and benign lesions in real time. The visualized morphologic features were comparable to those visualized in corresponding RCM images. All these results indicate that PCM is promising for aiding the diagnosis and treatment of skin cancer in a wide range of healthcare settings (e.g., rural clinics, mobile clinics). For example, PCM can enable the same-visit diagnosis and treatment of the most common skin cancer, BCC, in a mobile clinic visiting rural communities. This can potentially reduce the physical and financial burdens on patients living in rural communities.

4.6 Future Directions

While promising, the current PCM prototypes have several limitations. First, the use of the slit aperture degrades the confocal optical sectioning capability. This can potentially reduce the image contrast when imaging deeper regions of the tissue. The preliminary imaging results appear to indicate that some of the key morphologic features relevant to the diagnosis of malignant and benign lesions are readily examined in PCM. However, further studies will be needed to have a thorough understanding of which morphologic features are visible and which are not in PCM images when compared to RCM images.

Second, the cost and size of the PCM device will need to be reduced further to be adopted in a wide range of healthcare settings. Currently, the most expensive item in the PCM device is the off-the-shelf objective lens, which uses dozens of optical elements to perform well for a large range of wavelengths. A custom objective lens with a similar resolution could be developed that uses only a few optical elements by optimizing the lens for a small range of wavelengths used for PCM [18, 19]. The use of a reduced number of optical elements could potentially reduce the cost and size of the PCM device further.

Additionally, the small FOV of PCM can pose challenges in locating suspicious regions for cancer detection. Integrating a wide-field imaging modality with PCM can enable reliable identification of regions of interest [20]. Multiple PCM images can be stitched together to generate a large-area PCM image as was demonstrated with the handheld RCM device [21].

Finally, the interpretation of PCM images requires training, as the image contrast differs from that of hematoxylin and eosin (H&E)-stained histopathological images. Recently, a deep learning-based algorithm was developed to transfer RCM images into pseudo-H&E images [22], and the same algorithm can be adapted to transfer PCM images into pseudo-H&E images, which could facilitate the analysis of image features by less experienced healthcare professionals.

Acknowledgments

PCM technology development was supported by the National Institutes of Health (R01EY033349, R41CA261214, R21TW011222, R21TW010221). The University of Arizona has a technology-licensing agreement with ArgosMD on the PCM technology. DK has the rights to receive royalties because of this licensing agreement. DK serves as a scientific advisor to ArgosMD.

REFERENCES

1. N. Hasan, A. Nadaf, M. Imran, U. Jiba, A. Sheikh, W. H. Almalki, S. S. Almujri, Y. H. Mohammed, P. Kesharwani, and F. J. Ahmad, "Skin cancer: understanding the journey of transformation from conventional to advanced treatment approaches," *Molecular Cancer* 22, 168 (2023).
2. H. W. Rogers, M. A. Weinstock, S. R. Feldman, and B. M. Coldiron, "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012," *JAMA Dermatology* 151, 1081–1086 (2015).
3. P. Aggarwal, P. Knabel, and A. B. Fleischer Jr, "United States burden of melanoma and non-melanoma skin cancer from 1990 to 2019," *Journal of the American Academy of Dermatology* 85, 388–395 (2021).
4. E. Linos, K. A. Katz, and G. A. Colditz, "Skin cancer—the importance of prevention," *JAMA Internal Medicine* 176, 1435–1436 (2016).
5. C. Smit-Kroner, and S. Brumby, "Farmers sun exposure, skin protection and public health campaigns: an Australian perspective," *Preventive Medicine Reports* 2, 602–607 (2015).
6. R. A. Shellenberger, T. M. Johnson, F. Fayyaz, B. Swamy, J. Albright, and A. C. Geller, "Disparities in melanoma incidence and mortality in rural versus urban Michigan," *Cancer Reports* 6, e1713 (2023).
7. M. Adly, T. E. Woo, D. Traboulsi, D. Klassen, and J. Hardin, "Understanding dermatologic concerns among persons experiencing homelessness: a scoping review and discussion for improved delivery of care," *Journal of Cutaneous Medicine and Surgery* 25, 616–626 (2021).
8. M. Rajadhyaksha, A. Marghoob, A. Rossi, A. C. Halpern, and K. S. Nehal, "Reflectance confocal microscopy of skin in vivo: from bench to bedside," *Lasers in Surgery and Medicine* 49, 7–19 (2017).
9. S. Muzumdar, R. Wu, M. J. Rothe, and J. M. Grant-Kels, "Reflectance confocal microscopy decreases the cost of skin lesion diagnosis: a single institution retrospective chart review," *Journal of the American Academy of Dermatology* 86, 209–211 (2022).
10. J. M. Guilera, "The role of reflectance confocal microscopy in clinical trials for tumor monitoring," *Dermatologic Clinics* 34, 519–526 (2016).
11. "VivaScope system: non-invasive cellular imaging of the skin," https://www.caliberid.com/vivascope-system.html.
12. N. Shahriari, J. M. Grant-Kels, H. Rabinovitz, M. Oliviero, and A. Scope, "Reflectance confocal microscopy: principles, basic terminology, clinical indications, limitations, and practical considerations," *Journal of the American Academy of Dermatology* 84, 1–14 (2021).

13. E. E. Freeman, A. Semeere, H. Osman, G. Peterson, M. Rajadhyaksha, S. Gonzalez, J. N. Martin, R. R. Anderson, G. J. Tearney, and D. Kang, "Smartphone confocal microscopy for imaging cellular structures in human skin in vivo," *Biomedical Optics Express* 9, 1906–1915 (2018).
14. C. Gong, D. B. Stratton, C. N. Curiel-Lewandrowski, and D. Kang, "Speckle-free, near-infrared portable confocal microscope," *Applied Optics* 59, G41–G46 (2020).
15. C. Curiel-Lewandrowski, D. B. Stratton, C. Gong, and D. Kang, "Preliminary imaging of skin lesions with near-infrared, portable, confocal microscopy," *Journal of the American Academy of Dermatology* 85, 1624–1625 (2021).
16. I. Saknite, J. R. Patrinely, Z. Zhao, H. Chen, A. Beeghly-Fadiel, T. K. Kim, M. Jagasia, M. Byrne, and E. R. Tkaczyk, "Association of leukocyte adhesion and rolling in skin with patient outcomes after hematopoietic cell transplantation using noninvasive reflectance confocal videomicroscopy," *JAMA Dermatology* 158, 661–669 (2022).
17. M. Sugimura, K. Marcelino, R. Romero, J. Zhao, Y. Kim, A. Nessaee, K. Kim, D. Stratton, C. Curiel-Lewandrowski, and J. Garfinkel, "High-speed reflectance confocal microscopy using speckle modulation," *Biomedical Optics Express* 15, 4877–4890 (2024).
18. D. Kang, R. W. Carruth, M. Kim, S. C. Schlachter, M. Shishkov, K. Woods, N. Tabatabaei, T. Wu, and G. J. Tearney, "Endoscopic probe optics for spectrally encoded confocal microscopy," *Biomedical Optics Express* 4, 1925–1936 (2013).
19. J. Zhao, Y. Kim, M. Sugimura, K. Marcelino, R. Romero, B. Liang, M. Khan, E. Yang, and D. Kang, "Miniature Objective Lens for Scattering-based Light Sheet Microscopy," in *Optica Biophotonics Congress: Biomedical Optics 2024 (Translational, Microscopy, OCT, OTS, BRAIN)* (Optica Publishing Group, Fort Lauderdale, Florida, 2024), p. MS1A.5.
20. D. L. Dickensheets, S. Kreitinger, G. Peterson, M. Heger, and M. Rajadhyaksha, "Wide-field imaging combined with confocal microscopy using a miniature f/5 camera integrated within a high NA objective lens," *Optics Letters* 42, 1241–1244 (2017).
21. K. Kose, M. Cordova, M. Duffy, E. S. Flores, D. H. Brooks, and M. Rajadhyaksha, "Video-mosaicing of reflectance confocal images for examination of extended areas of skin in vivo," *British Journal of Dermatology* 171, 1239–1241 (2014).
22. J. Li, J. Garfinkel, X. Zhang, D. Wu, Y. Zhang, K. De Haan, H. Wang, T. Liu, B. Bai, and Y. Rivenson, "Biopsy-free in vivo virtual histology of skin using deep learning," *Light: Science & Applications* 10, 233 (2021).

5

Quantitative Fluorescence Imaging, Light-Triggering, and Monitoring of Chemodrug Release from Liposomes In Vivo by a *Mesoscopic-Scale Theranostic Endoscope*

Ulas Sunar, Rasel Ahmmed, Eli Kluiszo, and Semra Aygun-Sunar

5.1 Introduction

A major challenge in the management of advanced ovarian cancer is the presence of disseminated microscopic metastases (micromets) within the intraperitoneal cavity.[1,2] Despite surgery and adjuvant chemotherapy, as many as 60% of patients can show occult disseminated disease, with only a ~43% survival rate.[3–7] Surgical resection is often incomplete because it is challenging to detect micromets. Systemic chemotherapy is often ineffective and has toxic side effects. Conventional imaging techniques are less sensitive than reassessment surgeries and are ineffective in detecting these micromets.[8,9] Traditional imaging techniques, such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), and ultrasound, demonstrate less sensitive detection than reassessment surgeries. Despite surgery and adjuvant chemotherapy, as many as 60% of patients can show occult disseminated disease.[1,5,9] Moreover, systemic chemo can induce toxic side effects.[2,4,7,10] Thus, there exists an unmet need for both sensitive imaging technologies for improved detection and therapeutic approaches that selectively treat ovarian cancer micromets while avoiding healthy tissue. Liposomal nanocarriers have enhanced the biodistribution and efficacy of anti-cancer drugs such as doxorubicin (Dox).[11–13] Liposomal Dox is routinely used in recurrent ovarian carcinoma resistant to first-line platinum-based chemotherapy agents.[14] However, physiological barriers in release kinetics and insufficient tumor uptake ratios lead to sub-optimal biodistribution.[7,15] Thus, an imaging-guided drug delivery approach is needed to improve detection and treatment while reducing toxicity.

Porphyrin-based compounds (photosensitizers) are used for photodynamic therapy (PDT) and imaging because they strongly fluoresce and preferentially accumulate ~2-3 fold in malignant cells and have demonstrated some success in detecting ovarian micromets with qualitative imaging approaches.[7,11,16–20] Since micromets remain on the epithelial surface, PDT is an attractive approach for fulfilling the adjunctive therapy role to surgery. However, high normal tissue toxicity has been a major obstacle for the clinical translation of PDT for treating ovarian miromets. Seminal works by Hasan[4,7,21–24] and Mordon[1,25–28] groups have shown epidermal growth factor receptor (EGFR) and folate-targeted approaches reducing the peritoneal tissue toxicity of PDT by using semi-quantitative imaging approaches in animal models, though further work is needed to demonstrate clinical translation feasibility of imaging and treatment efficacy of these approaches under quantitative imaging guidance.

Our approach for the treatment is called chemophototherapy (CPT),[10,13,29–37] which involves the imaging and controlled release of nano-constructs: porphyrin-phospholipid doped (PoPD) liposomes, triggered directly by near-infrared (NIR) light (Fig.5.1a). Our Porphyrin conjugate (PoP) is based on HPPH, which has been through over 10 human phototherapy clinical trials. Additionally, NIR light is a clinically applicable stimulus that can permeabilize PoPD liposomes on demand with excellent temporal and spatial control. By restricting drug delivery to a specified region, injected doses, and therefore systemic side effects, can be reduced. These PoPD liposomes have negligible actuation in the "off state" and minimal interference with healthy tissues when triggered in micromets. PoPD-Dox triggered by NIR light holds great potential to only release Dox at the desired site and time. This approach is similar to photodynamic therapy (PDT) for recurrent gynecologic malignancies.[2,11,28] The treatment light can be delivered via wide-field illumination or a laparoscopic approach for releasing PoPD-Dox at tumor sites within the intraperitoneal cavity.[2,28,38]

To rationally develop image-guided detection and optimized and targeted drug delivery, knowledge about drug concentrations at and near the target cells is essential. Our approach capitalizes on the high sensitivity and specificity of fluorescence contrast, the fact that both Dox and Porphyrin are highly and differentially fluorescent,[39] and the high resolution of endoscopic imaging. However, accurate quantification of drug concentrations in tumors remains an ongoing challenge because raw fluorescence signal is affected by tumor absorption and scattering properties, which confound the true fluorescence contrast.[40,41] Spatial frequency domain imaging (SFDI) projects spatially modulated light (structured illumination) onto tissue (e.g. micromets) and measures the tissue's spatially modulated transfer function, which is related to optical absorption and scattering. SFDI allows fast quantitative imaging of tissue optical (absorption and scattering) parameters as well as fluorescence concentrations of exogenously administrated drugs. Recently, we expanded the SFDI technique to include endoscopic applications for

DOI: 10.1201/9781003449942-7

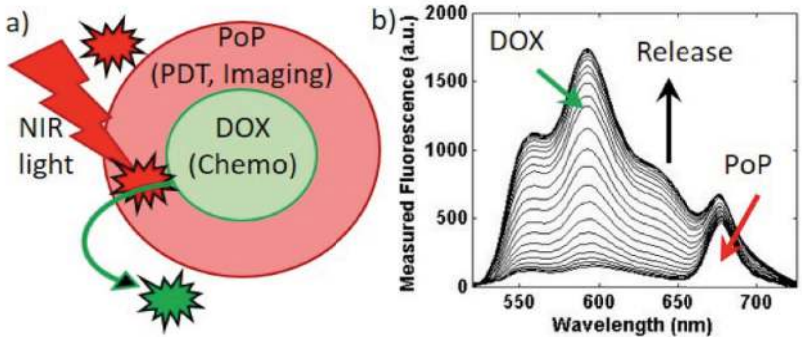


FIGURE 5.1 Liposomal formulation of "nano-balloon" for imaging and treatment. (a) PoP red (or Dox) fluorescence can be used to localize the cancer cell. NIR light triggers the drug release (doxorubicin, or Dox), where Dox is in the middle of the "nano-balloon." When NIR light activates the PoP (HPPH-porphyrin), it releases the Dox. (b) The combined fluorescence spectra from Dox and PoP indicate that fluorescence increases during the release, which allows image-guided drug delivery approach.

quantitative fluorescence imaging of drug concentration.[33,42,43] Our dual-channel endoscope approach uses two image fibers for patterned illumination and detection in both reflectance and fluorescence imaging modes to quantify optical and fluorescence parameters, respectively. Our approach uses SFDI-quantified optical parameters to provide *a priori* information on absorption and scattering at the excitation and emission wavelengths (λ_x, λ_m) to correct the raw fluorescence signal and obtain absolute fluorescence concentration for improved visualization of micromets.[17] Moreover, we have demonstrated that quantifying concentrations of porphyrins and Dox is possible. We use fluorescence of the PoP liposome's outer Porphyrin layer (PoP (HPPH), shown in red, Fig. 5.1a), which, before release, is the main imaging marker, to quantify initial drug concentration. Co-localized imaging of Dox and Porphyrin fluorescence utilizes two fluorescence contrasts. Initial Porphyrin fluorescence will allow quantification of the maximum Dox available for release. Once Dox is released at the precise locations of structured NIR light (Fig. 5.1b), Dox fluorescence is the imaging marker used to quantify locally released Dox. Knowledge of tissue optical properties allows us to rapidly quantify drug fluorescence concentration in target tissues for near-real-time feedback. As the imaging is non-contact, wide-field, and fast, it determines the pharmacokinetics of drug release in near real-time.

5.2 Materials

5.2.1 Doxorubicin (Dox) Encapsulated in PoP-Liposomes

We demonstrated potent tumor ablation with CPT using PoP liposome formulations of doxorubicin, termed "long-circulating Dox in PoP liposomes" (LC-Dox-PoP), described by Lovell and Sunar,[29,33,42,44–46] has long-circulation in blood and is highly stable in storage, and independently confirmed by the NCI Nanocharacterization Lab program. Both PoP and entrapped Dox are stable in circulation in mice, and laser treatment results in potent tumor ablation. LC-Dox-PoP can potently ablate tumors when NIR light is applied. LC-Dox-PoP was prepared through ethanol injection, extrusion, and active drug loading, as outlined previously.[34] The liposomes were composed of 53% DSPC (distearoylphosphatidylcholine), 5% DSPE-PEG-2K (distearoylphosphatidylethanolamine-polyethylene glycol), 40% cholesterol, and 2% Pyro-lipid. These components were selected for their biocompatibility and effectiveness in drug delivery. To synthesize the PoP liposomes, the lipids were first dissolved in heated ethanol. This lipid solution was then combined with ammonium sulfate to facilitate the formation of liposomes. The mixture underwent extrusion through membranes of various pore sizes to ensure uniformity in liposome size. Following extrusion, a buffer exchange was performed using dialysis in HEPES buffer containing 10% sucrose to remove any unencapsulated materials and prepare the liposomes for drug loading. Dox (LC Labs) was then loaded into the liposomes through an active loading process. The liposome mixture was incubated at 60°C for one hour, allowing Dox to be encapsulated efficiently. The amount of Dox encapsulated in the liposomes was quantified using fluorescence measurements, ensuring precise control over drug dosage. 480 nm excitation and 590 nm emission were used to detect Dox in a fluorescence plate reader (TECAN Safire) PoP-liposome self-assembly status and elution position could also be tracked without interference by using 420 nm excitation and 670 nm emission for the PoP.

5.2.2 Mouse Model of Ovarian Cancer

To study the effectiveness of the Dox-PoP liposomes in vivo, a mouse model of ovarian cancer was established using the human ovarian adenocarcinoma cell line NIH:OVCAR-3.[43] Briefly, these cells were obtained from the American Type Culture Collection (ATCC) and were specifically chosen for their relevance to human ovarian cancer. Nine female athymic nude mice were injected subcutaneously with 1×10^7 OVCAR-3 cells. This method ensured that the tumors would develop in a controlled and reproducible manner. Once the tumors reached a diameter of approximately 3 mm, the mice were administered intravenous injections of LC-Dox-PoP liposomes, with each mouse receiving 50 μg/mL of LC-Dox-PoP, totaling 5 mg/kg.

Imaging was performed one hour after the liposome injection to assess the distribution and concentration of Dox within the tumors. Following the imaging, the tumors were exposed to near-infrared (NIR) treatment light at 660 nm with an intensity of 300 mW/cm² and a beam diameter of 1.6 cm. This light exposure triggered the release of Dox from the PoP liposomes directly at the tumor site. Fluorescence measurements were taken every 10 s during the treatment to monitor the release dynamics of Dox in real time. In total, seven mice received LC-Dox-PoP + treatment light, one mouse received the injection of LC-Dox-PoP, but no treatment light, and one mouse received treatment light without any injected LC-Dox-PoP. After the treatment, the mice were euthanized, and tissue samples were collected from the tumors, surrounding muscle, and contralateral muscle tissues.

5.3 Methods

The quantification of the absolute fluorescence concentration in vivo requires two different imaging modes (**reflectance** and **fluorescence imaging**) to be done concurrently or sequentially with fast data acquisition. Reflectance imaging allows quantification of optical parameters, which in turn are used to compensate light attenuation at both excitation and emission wavelengths to correct for measured fluorescence to obtain absolute fluorescence concentration. Below we start with the SFDI reflectance imaging on how to obtain the measurements and quantify optical parameters. Then, we moved to fluorescence imaging, and then the approach for correction for quantifying the fluorescence concentrations of the fluorophores.

5.3.1 Optical Property Quantification Using SFDI Reflectance Imaging

5.3.1.1 Model (Diffuse Light Propagation in the Spatial Frequency Domain)

Near-infrared (NIR) light propagation in living tissue is dominated by light scattering. Thus, living tissue acts like milk or clouds in the sky, and is not transparent like water. While photons diffuse through turbid media, some may get absorbed or scattered. Figure 5.2a illustrates photons from a point source propagating in a scattering dominant turbid media. Scattering phenomena cause deviations in the path of propagation, while optical absorption leads to some photons lost in the medium. Optical absorption and reduced scattering are denoted as μ_a and μ_s', respectively and attribute to the attenuation of light. Optical absorption, μ_a, is formally defined as the product of molar concentration $C[mol/L]$ and molar extinction coefficient $\epsilon\left[cm^{-1}/M\right]$, $\mu_a = C\epsilon$, that is a function of wavelength and mainly dependent on oxygenated and deoxygenated hemoglobin in the NIR spectrum. The other dominating factor in tissue is scattering, μ_s', which is caused by a phenomenon known as Mie scattering. Light attenuated by both absorption and scattering effects and accurate quantification of optical parameters of absorption and scattering remains a challenge in diffuse optical imaging that involves thick and deep living tissues. This section will describe the spatial frequency domain approach in the photon diffusion domain that involves two-dimensional sensors like charged coupled device (CCD) cameras that are highly suitable for non-contact wide-field imaging for surgery guidance and therapy monitoring in near-real-time.

Spatial frequency domain imaging (SFDI) in reflectance mode is a wide-field imaging technique that quantifies the optical absorption (μ_a), and scattering (μ_s') properties. It has advantages over point spectroscopic techniques because it can provide local maps (images) of these optical contrasts. The signals we wish to detect with SFDI are modeled as a spatially modulated photon density wave in the spatial frequency domain (Fig. 5.2b and c). In the context of light diffusion where scattering, μ_s', is many orders of magnitude greater than absorption, μ_a, ($\mu_s' \gg \mu_a$), the light diffusion in a homogeneous medium can be described by a photon diffusion equation:

$$\nabla^2\varphi - \mu_{eff}^2\varphi = -3\mu_{tr}q. \tag{5.1}$$

Here, φ is the light fluence rate related to measured power per unit area, q is the source illumination, μ_{eff} is the effective attenuation, related to the product of absorption and scattering

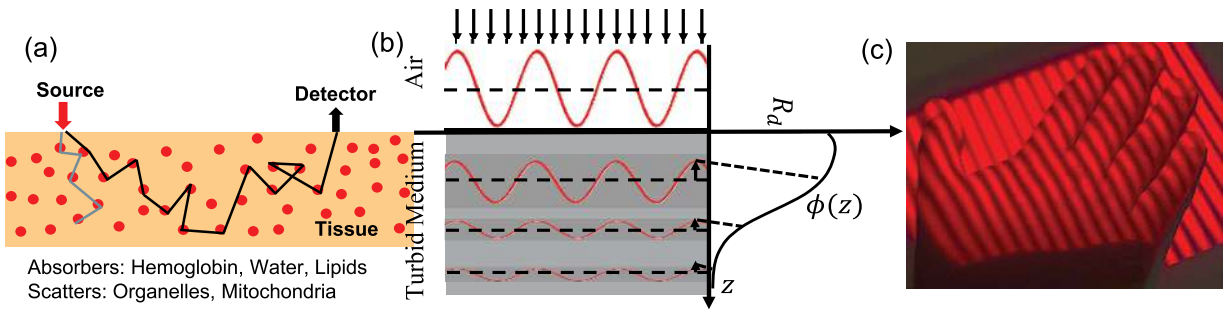


FIGURE 5.2 (a) Microscopic picture of light propagation (scattering and absorption) in living tissue. Photons have multiple scattering (black line) and some get absorbed (gray line) before being collected by a detector. (b) Macroscopic picture of spatially modulated light as a wave. The measured reflected signal, reflectance Rd, is also modulated at the same frequency as the light source, with measured signal attenuates with respect to distance from the source. This modulated signal response allows quantification of both optical absorption and scattering parameters. (c) An example of spatially modulated light measurement of the hand, is red light with a sinusoidal pattern illuminates the hand, and reflected light is measured with a camera detector.

parameters, $\mu_{eff} = \sqrt{3\mu_s^{'}\mu_a}$, and $\mu_{tr} = \left(\mu_a + \mu_s^{'}\right)$ is the total attenuation of light contributed by both absorption and scattering. In the spatial frequency domain, we spatially (here sinusoidal illumination) modulate the light source to project a light pattern that is normal to the to the plane of incidence:

$$q = q_0(z)cos(k_x x + \alpha), \tag{5.2}$$

where q_0 is the power at depth z, the spatial frequency is $k_x = 2\pi f_x$ and α represents phase shift offset in radians. If the media's response to the source illumination is proportional to the input intensity, the sinusoidal pattern reflecting from the media will have the same spatial frequencies. Therefore, we can model the measured light intensity or fluence rate (φ) reflecting from the media with the same frequency as the source, q:

$$\varphi = \varphi_0(z)cos\left(k_x x + \alpha\right). \tag{5.3}$$

Substituting Equations 5.2 and 5.3 into 1 yields a one-dimensional photon diffusion equation for the fluence rate as a function of depth z, which is the Helmholtz-type differential equation:

$$\frac{d^2}{dz^2}\varphi_0(z) - \mu_{eff}^{'2}\varphi_0(z) = -3\mu_{tr}q_0(z), \tag{5.4}$$

This diffusion equation (Equation 5.4) implies that the solution for fluence rate can be modeled as an exponentially-damping signal with respect to depth z with a decay rate of $\mu_{eff}^{'}$, which is defined as the effective attenuation,

$$\mu_{eff}^{'} = \sqrt{\mu_{eff}^2 + k_x^2} = 1/\delta_{eff}^{'} \tag{5.5}$$

Equation 5 indicates that the decay rate depends on the usual effective attenuation ($\mu_{eff}^{'}$) for the continuous wave (CW), DC term case when there is no modulation ($k_x, f_x = 0$), with an additional term of spatial frequency modulation, k_x^2. From this, we can also infer that spatial modulation allows depth sectioning (or depth sensitivity), with the effective penetration depth being defined as $\delta_{eff}^{'} = 1/\mu_{eff}^{'}$. Thus, the effective attenuation increases with spatial frequency and higher frequencies are more attenuated (tissue acts as a low-pass filter). Thus, higher frequencies are more sensitive to superficial tissue, while lower frequencies probe deeper tissue. Since absorption length ($1/\mu_a \sim 100mm$) is much larger than and scattering length ($1/\mu_s^{'} \sim 1mm$) higher frequencies are more sensitive to scattering (photons definitely scatter at the illuminated superficial region). The diffuse reflectance (the measured reflected light), $R_d(k)$, can be obtained as:

$$R_d(k) = \frac{3Aa'}{\left(\mu_{eff}^{'}/\mu_{tr} + 1\right)\left(\mu_{eff}^{'}/\mu_{tr} + 3A\right)}, \tag{5.6}$$

where $a' = \mu_s^{'}/\mu_{tr}$ is the reduced albedo and $A = \left(1 - R_{eff}\right)/2\left(1 + R_{eff}\right)$ is the proportionality constant. Here, n is the refractive index of tissue and $R_{eff} \approx 0.0636n + 0.668 + \left(0.710/n\right) - \left(1.440/n^2\right)$ is the effective reflection coefficient. Thus, one can simulate photon diffuse reflectance by using Eq. (5.6) with respect to different spatial frequencies for different cases of optical parameters by using this simple and fast analytical solution via MATLAB. The optical properties, μ_a and $\mu_s^{'}$, can be recovered by fitting this model to the experimental data. Towards this aim, in the next subsection, we briefly discuss how to obtain the experimental data.

5.3.1.2 Projection, Demodulation, and Detection of Spatially Modulated Light

The measured reflected intensity, I, has two components, DC (f=0), and AC, $I = I_{AC} + I_{DC}$, where I_{DC} is the DC offset, and only the SC component is modulated

$$I_{AC} = M_{AC}\left(x, f_x\right) \cdot \cos\left(2\pi f_x x + \alpha\right) \tag{5.7}$$

where M_{AC} is the amplitude of the reflected modulated intensity at frequency f_x. During the measurements, three equally spaced phase shifted sinusoidal intensity patterns are projected as shown in Figure 5.3.

To recover the reflectance, we first recover the modulated amplitude (M_{AC}) in Eq. (5.8) by using the demodulation approach, where we illuminate the sample with the sinusoidal patterns at three different phase offsets (0, 2π/3, 4π/3) at the same spatial frequency (Fig. 5.3), and obtain AC and DC components (M_{DC}, M_{AC}) by:

$$M_{AC}\left(x_i, f_x\right) = \frac{\sqrt{2}}{3}\sqrt{\left[I_1\left(x_i\right) - I_2\left(x_i\right)\right]^2 + \left[I_2\left(x_i\right) - I_3\left(x_i\right)\right]^2 + \left[I_3\left(x_i\right) - I_1\left(x_i\right)\right]^2}, \tag{5.8}$$

$$M_{DC}\left(x_i, f_x = 0\right) = \frac{1}{3}\left[I_1\left(x_i\right) + I_2\left(x_i\right) + I_3\left(x_i\right)\right] \tag{5.9}$$

where I_1, I_2, I_3 represent the reflected modulated intensity images at three different phase offsets, like the ones shown in Figure 5.3. Experimental modulation amplitude, $M_{AC}\left(x_i, f_x\right)$ depends on the source intensity, the modulation transfer function (MTF) of the imaging system, and the tissue diffuse reflectance, R_d, which is the MTF of the tissue sample:

$$M_{AC}\left(x_i, f_x\right) = I_0 \cdot MTF\left(x_i, f_x\right) \cdot R_d\left(x_i, f_x\right) \tag{5.10}$$

Since we are only interested in tissue response, R_d, one can obtain it by normalizing the measurements to a reference phantom with known optical properties, where intensity and instrument response is assumed to be the same for both unknown and reference measurements. Thus, unknown diffuse reflectance, R_d, which contains the tissue response (optical properties) can be written as:

$$R_d\left(x_i, f_x\right) = \frac{M_{AC}\left(x_i, f_x\right)}{M_{AC,ref}\left(x_i, f_x\right)} \cdot R_{d,ref}\left(f_x\right), \tag{5.11}$$

where M_{AC} is the measured modulated amplitude of the unknown tissue, $M_{AC,ref}$ is the measured modulated amplitude

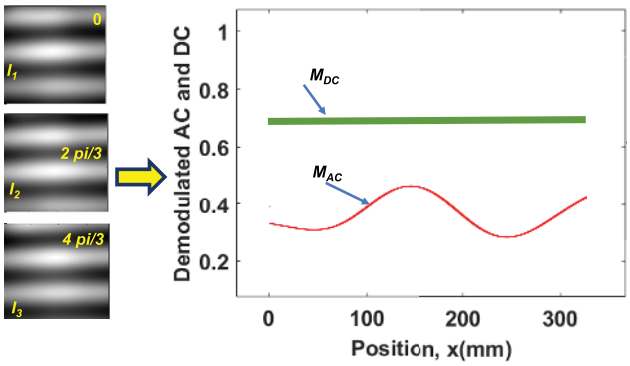


FIGURE 5.3 SFDI projections (*Intensity, I*) on a sample captured with a camera. For a single spatial frequency, f_x, there are three intensity components (I_1, I_2, I_3) for three phases (0, 2pi/3, 4pi/3) are projected on the sample. Then the signal is demodulated to obtain M_{DC} and M_{AC} components by using Equations (5.8 & 5.9). 1D signals (with respect to the spatial position (x)) of M_{DC} and M_{AC} components are plotted for simplicity and clarification purposes

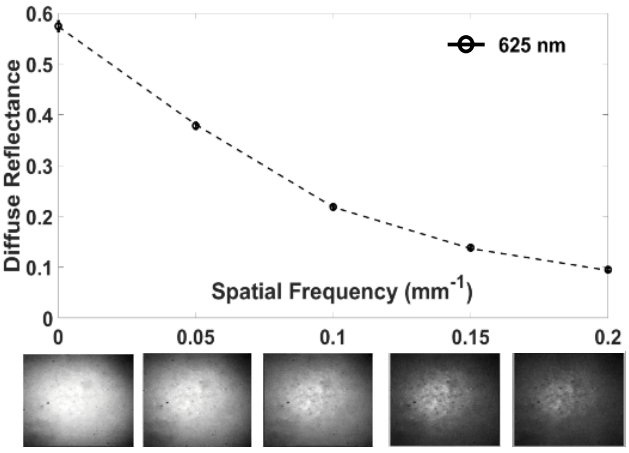


FIGURE 5.4 Diffuse reflectance of a tissue-mimicking phantom. The characteristic exponential decay curve as a function of spatial frequency is seen here. The optical properties are approximated by calibrating against and fitting the diffuse reflectance of a known reference.

of the known reference, and $R_{d,ref}$ is the diffuse reflectance of the known reference phantom. $R_{d,ref}$ can be predicted with assumed reference optical parameters. Thus, this normalization allows for eliminating absolute intensity and instrument response measurements. An example of R_d curve measured from a human ovarian tissue-mimicking phantom is shown in Figure 5.4.

Note that each pixel in a reflectance image at spatial frequency, f_x, corresponds to its own point in the reflectance curve. For example, a set of 100 × 100 reflectance images for x spatial frequencies would have 10,000 reflectance curves.

5.3.1.3 Quantification of Optical Parameters

One of the main advantages of SFDI is the sensitivity to both scattering and absorption by projecting different spatial frequencies. This can be seen clearly in Figure 5.5a, where R_d curve is plotted with respect to spatial frequencies by changing the optical parameters at a single pixel. For example, Figure 5.5b clearly shows that R_d values do not change at thigh frequencies for the cases of different absorption coefficients. This indicates that R_d at high frequencies is not sensitive to the absorption parameter (μ_a) but only sensitive to the scattering parameter (μ_s').

Similarly, Figure 5.5a indicates that scattering parameters is sensitive to both high and low spatial frequencies. Thus, this spatial frequency-dependent sensitivity allows for accurate quantification of both μ_a and μ_s', by performing measurements at different spatial frequencies with a combination of low and high frequencies. There is a trade-off between the accuracy of optical property quantification and data acquisition time, thus in most experimental settings one typically choose fewer spatial frequencies to reduce the acquisition time, for example, three spatial frequencies at DC, $low-frequency\,mm^{-1}$, and $high-frequency\,mm^{-1}$ (e.g., 0.1 mm^{-1}) components can be used to obtain accurate quantification of both μ_a and μ_s'. For example, in our endoscopic system, we used 4 spatial frequencies from 0 to 0.2 mm^{-1}. Then, optical absorption (μ_a) and scattering (μ_s') parameter maps can be obtained by fitting

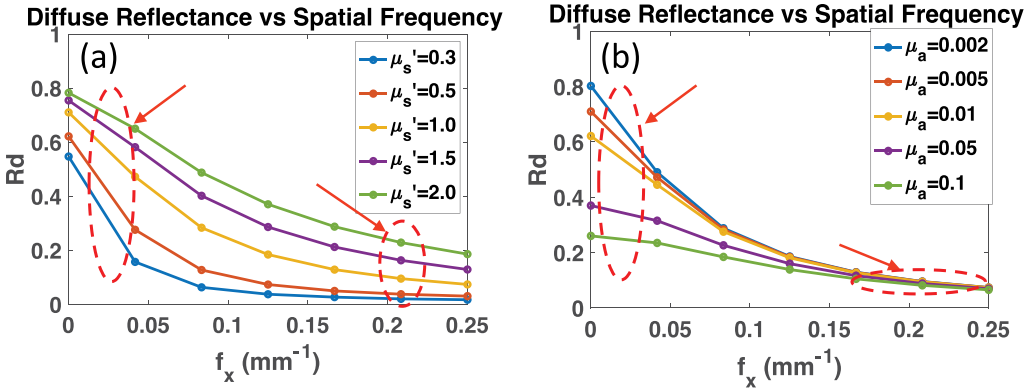


FIGURE 5.5 Relationship between spatial frequency and sensitivity to optical properties. (a) When μ_a is kept constant changes in μ_s' generate contrast visible all over the f_x. (b) When μ_s' is kept constant changes in μ_a generate contrast at low f_x

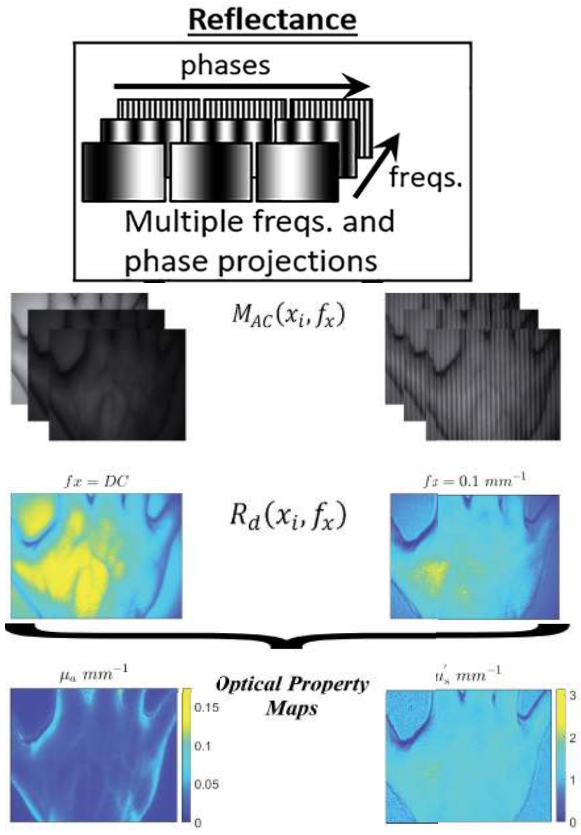


FIGURE 5.6 Flow chart for quantifying optical properties. In this example, there are two spatial frequencies that are demodulated and fit to obtain the optical property maps. Data provided from openSFDI.org.

spatial frequency domain reflectance ($R_d(f)$) data for each pixel with a frequency-domain Monte Carlo or diffusion model using the lsqcurvefit nonlinear fitting algorithm on MATLAB[47]. A flow chart of optical property map quantification is shown in Figure 5.6.

5.3.2 Quantitative Fluorescence Imaging (qFI) for Assessing Fluorescence Concentration

As discussed previously, the liposomal formulation of "nanoballoon" for imaging and treatment has two components (PoP based on HPPH-porphyrin) and Doxorubicin (Dox) that both have distinct excitation (**x**) and emission (**m**) spectrums. The quantitative fluorescence imaging of each component can allow local drug concentration *in vivo*. In this section, we describe briefly the quantification of the fluorescence concentration applicable to any fluorophore. Fluorescence generation and detection in tissue is a three-step process, which involves excitation source propagation in absorbing and scattering media, then fluorophore absorption and fluorescence generation, and then generated fluorescence propagation to the detector, again in absorbing and scattering media. The excitation (**x**) light source propagation is strongly dependent on the optical properties of the medium at the excitation wavelength (λ_x), as denoted by $\mu_{a,x}$ and $\mu_{s,x}^{'}$, while fluorescence emission (**m**) propagation is strongly dependent on the optical properties of the medium at the emission wavelength (λ_m), as denoted by $\mu_{a,m}$ and $\mu_{s,m}^{'}$ (Fig 5.7a). The SFDI technique quantifies optical properties at both excitation (**x**) and emission (**m**) wavelengths, thus modeling the light attenuation at both wavelengths accurately, which leads to the quantification of drug fluorescence concentration accurately.

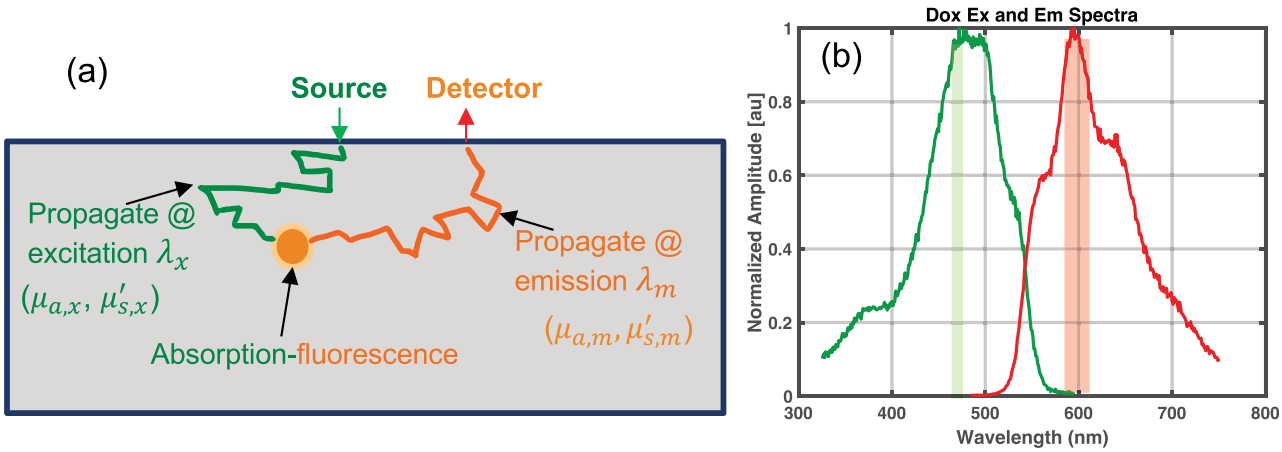


FIGURE 5.7 Fluorescence generation and propagation is a three-step process, which involves excitation source propagation, then fluorophore absorption and fluorescence generation, and generated fluorescence propagation to the detector. The excitation (**x**) source propagation is strongly dependent on the optical properties of the medium at the excitation wavelength (λ_x), while fluorescence propagation is strongly dependent on the optical properties of the medium at the emission (**m**) wavelength (λ_m). (b) Normalized Dox excitation and emission spectra. Shaded areas indicated excitation filter ~480±10 nm and 590±30 nm emission filter

Figure 5.7b shows the excitation and emission spectra of Dox fluorophore. We usually excite the fluorophore at λ_x = ~480nm using a light emitting diode (LED) or laser source with 10nm cleanup filter in front of each light source. On the detector side we use two filters changeable with a filter wheel. When the optical density (OD) filter is on, diffuse reflectance measurement is acquired at both excitation and emission wavelengths (480 nm, 590 nm), called here as $R_d(\lambda_x)$ or simply R_x, and $R_d(\lambda_m)$ or simply R_m, and then we illuminate the fluorophore at the emission band (590 nm), and we obtain the filter position is switched to the fluorescence filter (590±30 nm) and raw (uncorrected) diffuse fluorescence emission is measured, called as $F_{raw}(\lambda_m)$ or F_{raw}.

Based on the three-step process of fluorescence generation and detection, it is clear to note that the measured fluorescence (F_{raw}) is affected by the tissue optical properties at both excitation and emission wavelengths and attenuated by the light diffusion (light attenuation) through the thick samples like *in vivo* cases. There are several models available to obtain quantitative fluorescence concentration information. The simplest approaches are based on empirical corrections such as by taking the ratio of raw fluorescence to reflectance at the excitation wavelength (F_{raw}/R_x) to normalize or cancel out excitation related variations on the sample. More rigorous approaches involve compensation of attenuation at both excitation and emission wavelengths. Then, the corrected fluorescence is as:[48–53]

$$qFI = F_{corr} = F_{raw} / X(\lambda_x,\lambda_m) \tag{5.12}$$

where F_{raw} is the measured raw signal, $X(\lambda_x,\lambda_m)$ is the correction factor and F_{corr} is the corrected fluorescence by compensating the optical attenuation due to absorption, scattering and light propagation at both excitation and emission wavelengths (λ_x,λ_m). A simple expression for $X(\lambda_x,\lambda_m)$ is derived previously:[48–53]

$$X(\lambda_x,\lambda_m) = (1-R_x) * R_m / \mu_{a,x} \tag{5.13}$$

Alternatively, we have been using the model developed by Gardner *et al.*[33,43,54] because this model suits well to our SFDI experimental setup. The model assumes planar illumination, and it accounts for the presence of tissue absorption (μ_a) and scattering (μ_s') parameters based on reflectance geometry and the Monte Carlo Model, very similar to the SFDI approach.[33,42,43,54] As shown in the previous sections, SFDI can provide μ_a and μ_s' at both excitation and emission wavelengths as input to Gardner's model that uses these to correct for the intrinsic or attenuation corrected fluorescence with the correction factor, $X(\lambda_x,\lambda_m)$, as:

$$X(\lambda_x,\lambda_m) = (Term1 / Term2) - (Term3 / Term4) \tag{5.14}$$

where, $X(\lambda_x,\lambda_m)$, is the fluorescence attenuation correction factor,[17,54,55] is a bit more complicated expression having four components: $Term1 = C_1(\lambda_x)C_3(\lambda_m)$, $Term3 = C_2(\lambda_x)C_3(\lambda_m)$, $Term2 = K_1(\lambda_x)/\delta(\lambda_x) + K_3(\lambda_m)/\delta(\lambda_m)$, $Term4 : K_2(\lambda_x)/\delta(\lambda_x) + K_3(\lambda_m)/\delta(\lambda_m)$. Here, the parameters C_1, K_1, C_2, and K_2 are functions of the diffuse reflectance at the excitation wavelength, and two parameters C_3, and K_3, are functions of reflectance at the emission wavelengths, $\delta = 1/\sqrt{3\mu_a\mu_s'}$ is the effective penetration depth at each wavelength.

Figure 5.8 clarifies the correction approach. We prepared tissue-mimicking reflectance and fluorescence phantoms made by varying the optical properties and titrating the Dox concentrations. Phantoms were prepared using Intralipid 20% (Fresenius Kabi) for scattering and India Ink (Higgins) for

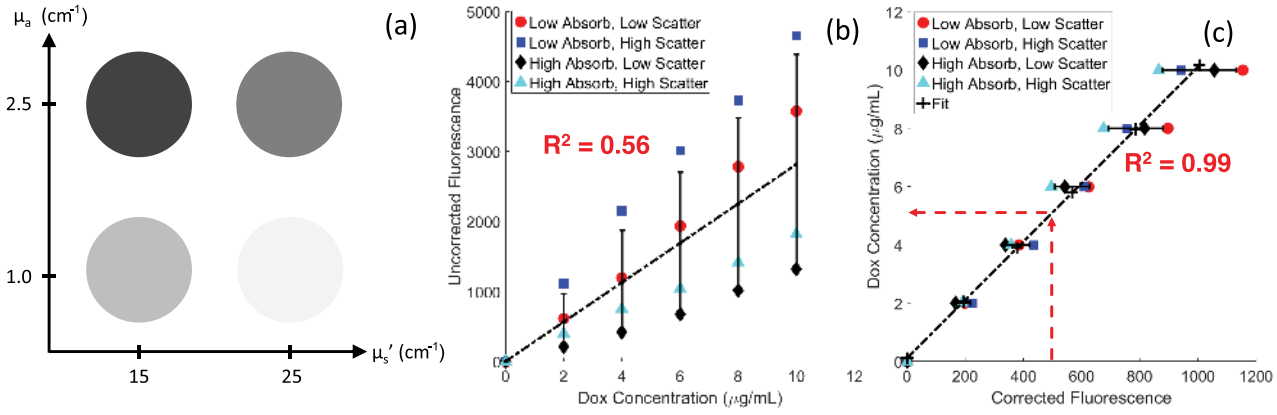


FIGURE 5.8 Assessing Dox concentration with optical attenuation correction. (a) A schematic of the measured phantom optical properties (b) Uncorrected (raw) Dox fluorescence vs Dox concentration, showing the large variation with respect to background optical property differences (R^2=0.56). (c) Attenuation corrected Dox fluorescence vs Dox concentration, indicating a strong linear relationship with respect to Dox concentration (R^2=0.99). Red arrow indicates for an arbitrary future measurement of the corrected fluorescence with a value of ~500 will map into absolute concentration of ~5 μg/mL.

absorption. Four phantoms were prepared with different combinations of optical properties (Fig. 5.8a): μ_a = 1.0 and 2.5 cm^{-1}, μ_s' = 15 and 25 cm^{-1} at 490 nm. For the calibration of Dox concentration, a stock solution of 1.0 mg/mL free-Dox was used. After baseline reflectance and fluorescence measurements of each phantom, increasing volumes of free-Dox were added to each (0, 200, 400, 600, 800, 1000 µL) and fluorescence measurements were acquired at each addition. Reconstructed raw fluorescence values with respect to concentration were used as a calibration curve to obtain absolute Dox concentrations. The raw fluorescence signal showed substantial variations with respect to optical parameters (Fig. 5.8b), but attenuation corrected fluorescence showed substantially less variation with respect to Dox concentration (Fig. 5.8c). The uncorrected raw fluorescence signals had a poor correlation with the Dox concentration (**R^2 = 0.56**), which improved substantially with the corrected fluorescence signal (**R^2 = 0.99**). By using the calibration factor, transition to Dox concentrations is established so that we can obtain the absolute fluorescence concentration in tissue with a linear calibration curve that maps the corrected fluorescence to concentration for each fluorophore. For example, in Figure 5.8c, the red arrow indicates that an arbitrary future measurement of the corrected fluorescence with a value of ~500 will map into an absolute concentration of ~5µg/mL.

5.3.3 Custom Endoscopic SFDI (eSFDI) for Imaging and Drug Release Setup

The schematic diagram of the custom-built bench-top endoscopic spatial frequency domain imaging (eSFDI) instrument featuring a dual-channel fiber endoscope design is shown in Figure 5.9a,b while Figure 5.9c represents the internal arrangement of the eSFDI system, as detailed previously [43]. This dual-channel strategy allowed two separate image fibers for light projection and imaging. The objective lenses positioned at the ends of the fiberscopes indicate the areas where the fields of view (FOVs) for both illumination and detection intersect. Two compact light-emitting diodes (LEDs), with central wavelengths of ~480 nm and ~590 nm were used for measuring Dox excitation/fluorescence (480/590 nm, Mightex, CA) in the LC-Dox-PoP nanoconstruct (shown in Figure 5.1). The 480 nm LED was filtered at 480 nm ± 20 nm to examine Dox's optical characteristics at its excitation peak, while a 590 nm ± 20 nm band-pass filter was utilized to quantify fluorescence properties at Dox's emission wavelength. The light was directed through a compact Digital Micromirror Device (DMD) module (LightCrafter 4500, Texas Instruments), where it was spatially modulated to the desired frequency and phase patterns. The sinusoidal patterns were projected by the illumination channel and captured by the other imaging-fiber channel that was coupled to the EMCCD camera (1004 x 1002 pixels, Andor, Ireland).

As the schematic of the instrumentation in Figure 5.9b shows, there was an achromatic lens of f=200mm (Thorlabs) placed in between 480/590 LED light (Mightex, CA) and a Digital Micromirror Device (DMD) light crafter 4500 (Texas Instruments) for enhancing the pattern intensity and making the wide field more homogeneous. The wide field illumination by the LED source passed through the achromatic lens and sinusoidal pattern reflected by the DMD and coupled to another achromatic lens of f=50mm (AC254-050-A-ML, Thorlabs) to collimate the light to the 2mm diameter with 13,000 elements of image fiber (Asahi Kasei). The field of view (FOV) of the customized objective lens was 70°. Coupler 1 and coupler 2 are reciprocal to each other in both image fibers which are a combination of several lenses such as 2 achromatic lenses of f=-100mm (ACN254-100-A-ML, Thorlabs), 1 each achromatic lens of f=60, 100 and 200mm (AC254-060/100/200-A-ML, Thorlabs) is shown in the down part of Figure 5.9b. The customized objective lens (diameter: 2mm, FOV: 70°) is placed at the distal end of both image fibers to enable the same projection and reflectance. The cross polarizer (Edmund Optics #29-489) is built in front of the endoscope, to reduce specular reflection. The focused patterns are relayed onto the surface through the image fiber, and reflected light is collected with the second fiber and coupled to the EMCCD camera (1004 × 1002 pixels, Luca, Andor, Belfast, Ireland). On the detector side, a 40x objective lens (Edmund optics) is placed in a way that the image fiber to camera distance is 100mm to couple

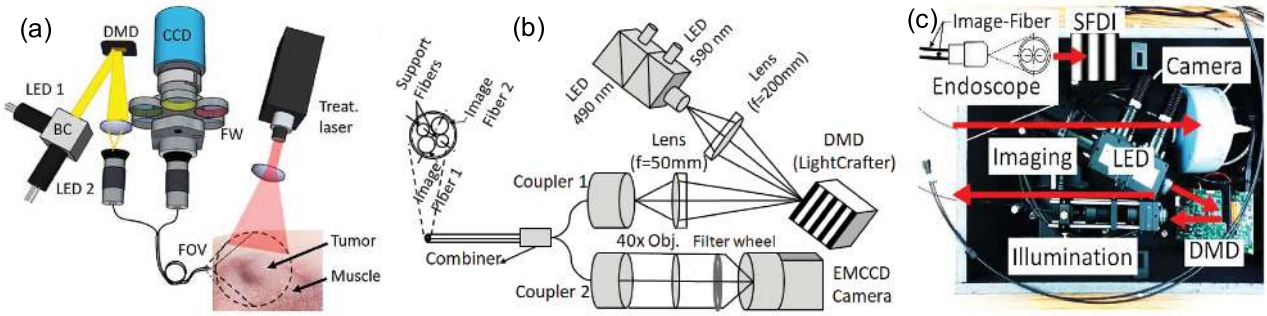


FIGURE 5.9 A dual-mode, dual-channel endoscope that enables integrated imaging and treatment strategies on a single platform. **(a)** Overview of the system functioning in two imaging modes (reflectance and fluorescence) to generate corrected fluorescence maps for determining the actual Dox concentration. **(b)** Schematic diagram of the eSFDI device. **(c)** Image of the compact and mobile version of the device, showcasing the individual components

the object. There was also a filter wheel (Thorlabs) through to control the desired wavelength

The camera was focused on the same field of view (FOV) as the projected pattern (1.6 cm x 1.6 cm). The acquisition duration was 0.1 s during reflectance mode, leading to a total SFDI acquisition time of about 2.4 s to capture images at both wavelengths. In fluorescence mode, a single 2s acquisition was conducted with the excitation light activated, followed by 2 s dark measurements, resulting in a total acquisition time of 4 s. Cross-polarizers were integrated into the front of the endoscope's distal end to eliminate specular reflections. For in vivo measurements, a 2 mm diameter image fiber with 13,000 elements (Asahi Kasei) was utilized to enhance photon counts. This endoscope is based on mesoscopic diffuse optical imaging, having approximately 100 μm of resolution and a few millimeters of penetration depth[43]. For the release experiments, a 660 nm laser with a fluence rate of ~300 mW/cm² served as the treatment light. The entire system was automated by a custom MATLAB software program.

5.4 Results

5.4.1 Quantitative Fluorescence Imaging (qFI) for Detection and Delineation of Tumor Cells *In Vitro*

Figure 5.10 demonstrates that accurate, quantitative fluorescence imaging (*qFI*) by SFDI improves visualization for image-guided detection and therapy planning. We made two wells (yellow arrows, Figure 5.10a) in ground pork to hold high-scattering solutions containing porphyrins like PoP (HPPH), which we translated in clinical trials of head and neck PDT. Each well contained the same fluorophore concentration and scattering, but the lower well also contained blood to increase optical absorption, mimicking the effect of higher tumor attenuation due to vasculature or surgery bleeding effects on measured fluorescence *in vivo*. The extra added blood mixed with porphyrin slightly leaked out, mimicking micrometastasis (micromets) or escaping residual malignant cells. Figure 5.10b shows the SFDI reflectance image (with spatial frequency, $f = 0$, DC term), where the lower well appears darker due to higher absorption. Blood absorption and scattering attenuated the raw fluorescence (F_{raw}), so the lower well appears to have less fluorescence signal (F_{raw}) and unclear contrast for a clinician during the surgery or treatment planning (Figure 5.10c). After correction for optical absorption and scattering, the corrected fluorescence signal (F_{corr}, *qFI*) in both wells is similar (even leaks representing "micromets, residual malignant cells" become visible, purple arrows), in alignment with the actual values of equal fluorophore concentrations (Figure 5.10d). Thus, these results indicate that (1) *qFI* improves visualization/detection with corrected fluorophore contrast mapping, which can be utilized for improved treatment/surgery planning, and (2) corrected fluorescence contrast allows assessing absolute fluorophore concentration (mg/mL), which can be used for optimizing PDT and CPT drug dosimetry for complete cell kill, as the complete cell kill is highly dependent on the drug concentration at the local site.

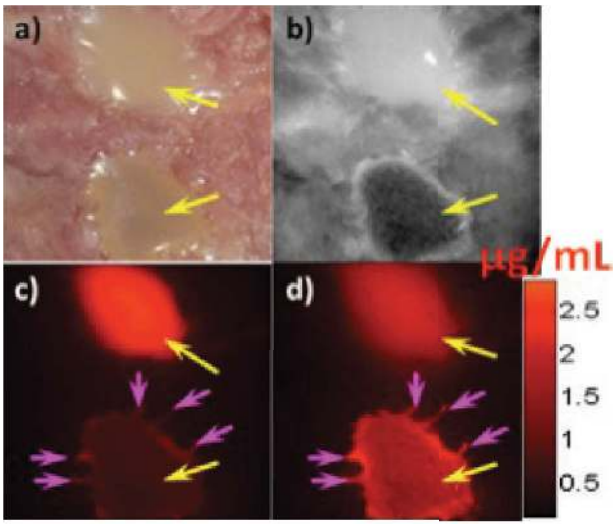


FIGURE 5.10 Optical parameters affect fluorescence contrast *in vivo*, and SFDI-derived quantitative fluorescence imaging (qFI) improves contrast. (a) Color, (b) reflectance (counts), (c) raw fluorescence, F_{raw} (counts), and (d) corrected fluorescence (F_{corr}) image (concentration), mg/mL.

5.4.2 qFI for Detection and Delineation of Tumors *In Vivo*

While we mainly focused on the endoscopic SFDI for intraperitoneal cavity imaging, the same approach for wide-field imaging applications for surgical guidance can be implemented.[5,6,39–41,56–58] We investigated the surgical guidance and theranostic potential of 5-aminolaevulinic acid photodynamic therapy (ALA-PDT) by employing a custom SFDI system to quantify the absolute concentration of the photosensitizer PpIX in nonmelanoma skin cancer models.[17] SFDI can serve as an effective diagnostic tool by accurately quantifying the absolute fluorescence concentration of PpIX for accurately delineating the tumor boundaries (detection), overcoming limitations of spectroscopic measurements that neglect absorption and scattering effects.[17,55] The wide-field SFDI makes it suitable for surgical guidance during tumor removal procedures marked by fluorescent markers.[5,6,39–41,56–58] Effective ALA-PDT requires sufficient accumulation of PpIX at the tumor site, yet uptake can vary significantly among patients and tumors.[2,33,39,41,59–70] Figure 5.11 shows representative images of a squamous cell carcinoma (SCC) tumor in a SCID mouse 1h after it-administration of ALA. Figure 5.11a shows the white-light image to indicate the tumor and injection location. The uncorrected fluorescence signal (Fig.5.11b) did not show tumor contrast (1152 ± 72 a.u. vs. 1411 ± 86 a.u.) in the tumor and surrounding normal respectively), but the absolute PpIX concentration image demonstrated a clear tumor contrast (Fig.5.11c) compared to normal surrounding tissue (1.10 ± 0.12 μg/mL vs. 0.76 ± 0.06 μg/mL respectively). This is possibly since

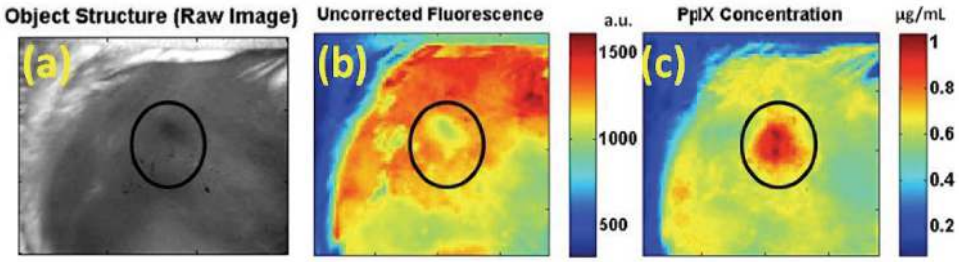


FIGURE 5.11 SFDI improves tumor contrast. Representative images of a SCC tumor 1h post ALA injection. **(a)** White light structural image showing the tumor area and ALA injection site, reflectance (counts) (**b**) The uncorrected fluorescence image [a.u.] did not show localized contrast. **(c)** The corrected fluorescence image (absolute PpIX fluorescence concentration) showed a high contrast between the tumor and surrounding area compared to the uncorrected image[17]

SCCs are well vascularized and have high microvessel density within the tumor compared to BCCs, which have even less microvessel density compared to the stroma. Light absorption (due to blood) and scattering (cell structure) attenuate raw fluorescence, this SCC tumor absorption parameter (μ_a) was ~61% and the scattering parameter μ_s' was ~12% higher than surrounding normal tissue at the excitation wavelength, and ~75% and ~12% higher at the emission wavelength, thus correcting the fluorescence for the light attenuation in living tissue was crucial to achieve this improved contrast. Overall, these results indicate that **1)** corrected fluorescence is related to fluorescence concentration, and **2)** quantitative fluorescence imaging improves detection. We demonstrated that accurate, quantitative fluorescence imaging by SFDI provides accurate and quantitative contrast and thereby improves the detection and delineation of tumors.

5.4.3 Quantitative Assessment of Temporal and Spatial Drug (Dox) Release Kinetics

Figure 5.12 shows Dox release kinetics in tissue phantoms, from our recent publications.[36,42,43] Uncorrected Dox fluorescence (a.u) (Figure 5.12a) shows a large difference in the final values for Dox fluorescence signals. This is most likely because the raw fluorescence signal is strongly affected by the optical parameters (absorption). However, correction for the effects of background optical parameters reveals that both phantoms did release similar concentrations of Dox (Figure 5.12b). There is a significant time delay for full release in the high-absorbing phantom (~4 min [blue] versus 8 min [red]), likely because the treatment light is more attenuated due to higher background absorption. As expected, the release kinetic characteristics are strongly dependent on the delivered light fluence rate and energy (light dosimetry). The administered light dose and drug concentration were the same, but additional treatment light was needed for full release. Note also that *quantified* Dox release in both phantoms was similar within ~9 min of treatment in contrast to unquantified fluorescence kinetics.

Next, we investigated the release kinetics in mice. Figure 5.12c shows three representative releases in tumor vs. two control mice (Control-1 corresponds to no drug injected, but treatment light applied, while Control-2 corresponds to drug injected, but no treatment light applied). Figures 5.12d, e, and f display representative images of whitelight and pre- and post-treatment maps of the quantified Dox fluorescence concentration at tumor sites. For all treated mice (N = 9), the average quantified Dox concentration was 1.0 ± 0.2 mg/mL before release and 8.0 ± 3.0 mg/mL at final release, indicating enhanced drug delivery at the target tumor sites. Notably, there were variations in the time-dependent release kinetics, particularly with the duration of light treatment. For instance, Mouse-1 (M1) exhibits slower release kinetics compared to Mouse-3 (M3), indicating the heterogeneity of release kinetics in individual tumors. To confirm our eSFDI quantification findings, we conducted *ex vivo* measurements using samples obtained from the imaged areas of *in vivo* drug release in mice. As depicted in Figure 5.12g, tumor tissue from mice treated with both drug and light exhibited an average Dox concentration of 5.6 ± 0.8, while peripheral muscle had 2.5 ± 0.7 mg/g, and contralateral muscle tissue that did not receive light had 1.3 ± 0.2 mg/g. Figure 5.12h shows the Pearson correlation coefficient between the average Dox levels quantified with eSFDI and those quantified through *ex vivo* analysis of 0.77, with a p-value of 0.0011, indicating a clear linear relationship between the two methods.

5.5 Discussion

5.5.1 Summary of Results

This research utilizes the endoscopic spatial frequency domain imaging (eSFDI) approach for real-time, precise drug delivery in the treatment of ovarian cancer. By integrating quantitative fluorescence imaging with spatial frequency domain imaging, this dual-channel, flexible fiber endoscope accurately determines both the optical and fluorescence properties of tissues. It allows for the correction of drug fluorescence based on tissue optical properties, ensuring highly accurate measurements of doxorubicin (Dox) concentrations in vivo. This precise quantification is pivotal for adjusting therapeutic interventions

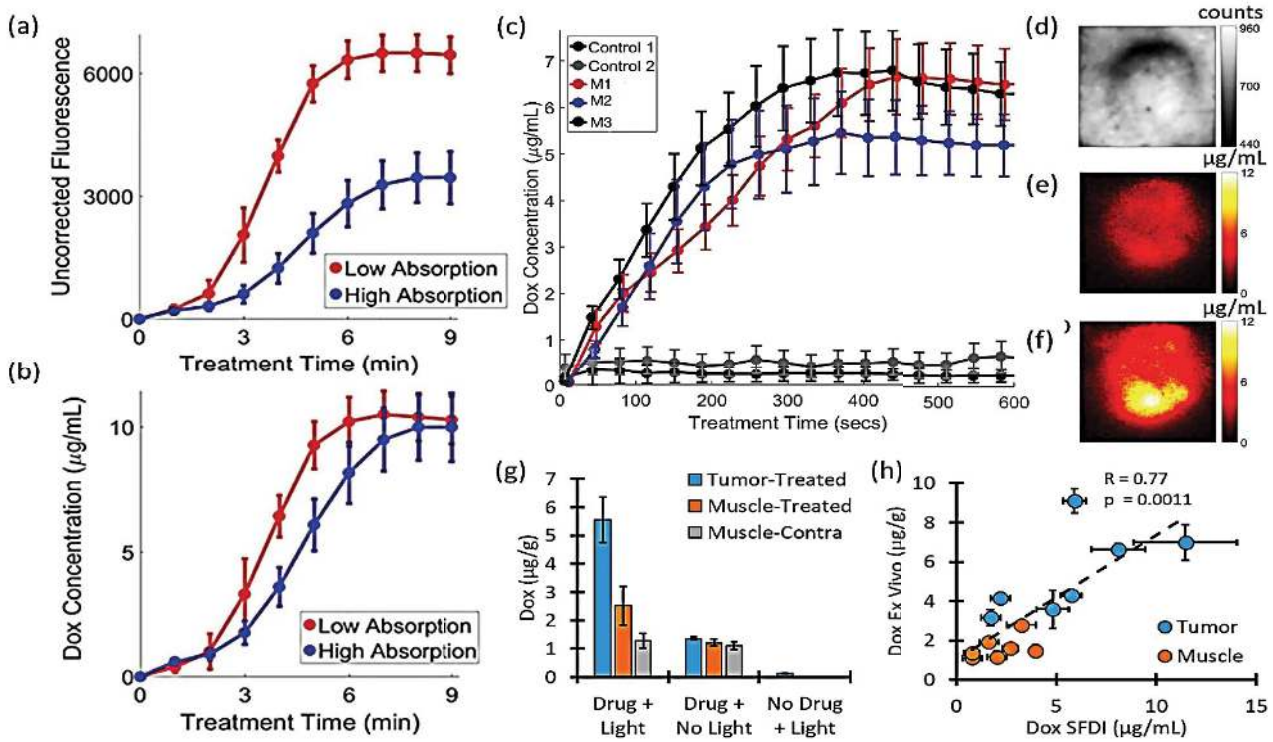


FIGURE 5.12 (a) Dox release in uncorrected raw fluorescence changes in phantoms with different optical properties. (b) Dox release as characterized by corrected fluorescence concentration changes. (c) Representative Dox release curves from five mice, indicate high release in targeted release, and *in vivo* Dox release kinetics show heterogeneity. (d) A representative white light structural image related to Mouse 3 (M3), shows the Dox release in the tumor. (e, f) Quantified Dox fluorescence concentration map 1-minute and 10-minute post-treatment light, respectively. (g) Comparison of Dox concentration with different treatment conditions at different regions of the tissue. (h) The Pearson correlation coefficient between the average Dox via eSFDI to the average Dox via *ex vivo* analysis.

to optimize drug delivery specifically to tumor sites, thereby enhancing treatment efficacy while minimizing systemic toxicity

5.5.2 Advancements in Imaging Technology and Personalized Medicine

The eSFDI system has demonstrated its capability to non-invasively quantify Dox concentrations in both phantoms and animal models. Our results support the importance of quantifying released Dox concentrations to accurately assess drug release and optimize treatment for individual tumors. Dox release kinetics revealed inter- and intra-tumor heterogeneity, highlighting the necessity for individualized light dosimetry and illumination patterns to achieve complete drug release while avoiding excessive illumination and drug release that could lead to normal tissue damage and toxicity. This novel system holds significant promise for advancing ovarian cancer treatment by enabling precise control of localized drug delivery. Real-time monitoring and adjustment of chemotherapy at the tumor site represent a significant advancement in personalized medicine, potentially increasing treatment efficacy while reducing adverse effects. Furthermore, by accurately delineating tumor boundaries, quantitative imaging modalities aid in treatment planning, tumor delineation, and monitoring of therapeutic efficacy. The integration of eSFDI with emerging diagnostic imaging with targeted therapy approaches hold great potential in managing disseminated cancer. Nanoparticle-based imaging agents and image-guided drug delivery systems represent innovative avenues for precise cancer imaging and therapy monitoring.

SFDI can operate as a diagnostic aid without the use of photosensitizers and chemo drugs, especially in preventative exams or direct treatment guidance, as physiological correlations can be made directly from intrinsic optical property measurements.[61,62,64,71,72] For example, in patients with actinic skin damage and nonmelanoma skin cancer, SFDI was used to obtain quantitative scattering and hemoglobin concentration contrasts that were indicative of the lesion stage.[57,61–64,73] The same approach can be implemented for oral and gastrointestinal cancers with endoscopic and laparoscopic approaches, Thus, eSFDI can provide early biomarkers of neoplasia and can be used by clinicians for frequent, non-invasive monitoring of at-risk individuals. In the referenced PDT studies, light delivery was applied for superficial lesions. However, for deeper and thicker tumors, alternative modes of light delivery can be considered. Interstitial fiber optics, endoscopes, or 3D fluorescence imaging would also provide a more accurate approach for reducing or eliminating partial volume effects.[18,74]

5.6 Future Trends and Broader Applications

Future adaptations might include drug targeting such as folate-targeting for ovarian cancer, as well as the integration of laparoscopic imaging approaches with a larger field of view, and snapshot acquisition methods to enhance the speed and precision of imaging, facilitating quicker clinical decisions during interventions.[56,75–78] The current analysis requires post-processing that takes about 5 minutes depending on the pixel number and binning. Implementing a GPU-based fast Monte Carlo model would allow mapping of the released drug concentrations in near real-time.[79] Fast mapping of drug concentrations would provide a time window for image-based feedback to optimize drug delivery. Furthermore, exploring higher resolution image fibers and imaging rods-based laparoscopic approaches could optimize the system's performance, ensuring more reliable and robust data for clinical use. The integration of advanced imaging technologies, such as deep learning (DL) algorithms for image analysis, will further improve the accuracy and efficiency of the eSFDI system and real-time decision-making. Artificial intelligence (AI) could assist in automating the interpretation of complex imaging data, providing clinicians with actionable insights faster and more accurately. This integration will be crucial in personalizing treatment plans, predicting patient responses to specific therapies, and optimizing the timing and dosage of drug delivery, allowing for customized treatment plans that maximize efficacy and minimize toxicity.

Beyond ovarian cancer, the eSFDI system can be adapted for other types of cancer and medical conditions requiring precise drug delivery and monitoring. Its versatility and noninvasive nature make it an attractive option for various gastrointestinal, brain, and lung diseases, potentially revolutionizing the approach to targeted therapy and real-time monitoring.

Acknowledgments

The authors acknowledge the funding support from NIH/NCI R01 (5R01CA243164) and NIBIB R21 (EB01914701). We also thank Dr. Luigi Belcastro and Farzad Azizizade for their feedback and edits.

REFERENCES

1. Azais H Tardivel M Kerdraon O Stallivieri A Frochot C Betrouni N Collinet P Mordon S., S. C. Assessment of the specificity of a new folate-targeted photosensitizer for peritoneal metastasis of epithelial ovarian cancer to enable intraperitoneal photodynamic therapy. A preclinical study. *Photodiagnosis Photodyn Ther* **13**, 130–138 (2016).
2. Chan JK Cuccia D Pham H Kimel S Gu M Hammer-Wilson MJ Liaw LHL Osann K DiSaia PJ Berns M Tromberg B Tadir Y., M. B. J. Laparoscopic photodynamic diagnosis of ovarian cancer using 5-aminolevulinic acid in a rat model. *Gynecol Oncol* **87**, 64–70 (2002).
3. Pectasides D, K. H. F. A. B. N. B. V. Z. J. A. A. Correlation of abdominal computed tomography scanning and second-look operation findings in ovarian cancer patients. *Am J Clin Oncol Cancer Clin Trials* **14**, 457–462 (1991).
4. Rizvi I Evans CL Abu-Yousif AO Muzikansky A Pogue BW Finkelstein D Hasan T., C. J. P. Synergistic enhancement of carboplatin efficacy with photodynamic therapy in a three-dimensional model for micrometastatic ovarian cancer. *Cancer Res* **70**, 9319–9328 (2010).
5. Longmire M Ogawa M Choyke PL Kobayashi H., K. N. Multicolor in vivo targeted imaging to guide real-time surgery of HER2-positive micrometastases in a two-tumor coincident model of ovarian cancer. *Cancer Sci* **100**, 1099–1104 (2009).
6. Van Dam GM Crane LMA Harlaar NJ Pleijhuis RG Kelder W Sarantopoulos A De Jong JS Arts HJG Van Der Zee AGJ Bart J Low PS Ntziachristos V., T. G. Intraoperative tumor-specific fluorescence imaging in ovarian cancer by folate receptor-α targeting: First in-human results. *Nat Med* **17**, 1315–1319 (2011).
7. Springa BQ, A.-Y. A. P. A. R. I. Z. X. M. Z. A. S. S. R. M. L. G. R. E. S. O. E. H. T. Selective treatment and monitoring of disseminated cancer micrometastases in vivo using dual-function, activatable immunoconjugates. *Proc Natl Acad Sci* **111**, 933–942 (2014).
8. Rose PG Miraldi F Abdul-Karim FW., F. P. Positive emission tomography for evaluating a complete clinical response in patients with ovarian or peritoneal carcinoma: Correlation with second-look laparotomy. *Gynecol Oncol* **82**, 17–21 (2001).
9. Sugiyama T Komai K Nishimura H Yakushiji M Nishimura H., N. T. Comparison of CA 125 assays with abdominopelvic computed tomography and transvaginal ultrasound in monitoring of ovarian cancer. *Int J Gynecol Obstet* **54**, 251–256 (1996).
10. Luo D Miranda D Lovell JF., C. K. A. Chemophototherapy: An Emerging Treatment Option for Solid Tumors. *Adv Sci*. **4**, (2017).
11. Staropoli N Botta C Fiorillo L Grimaldi A Lama S Caraglia M Salvino A Tassone P Tagliaferri P., C. D. Pegylated liposomal doxorubicin in the management of ovarian cancer: A systematic review and metaanalysis of randomized trials. *Cancer Biol Ther* **15**, 707–720 (2014).
12. Allen P. R., T. M. & C. Liposomal drug delivery systems: from concept to clinical applications. *Adv. Drug Deliv. Rev.* **65**, 36–48 (2013).
13. Luo Carter K. A. & Lovell J. F., D. Nanomedicine and Nanobiotechnology. in *Nanomedical engineering: shaping future nanomedicines* vol. 7 169–188 (Wiley Interdisciplinary Reviews, 2015).
14. Gordon, A. N. et al. Recurrent epithelial ovarian carcinoma: a randomized phase III study of pegylated liposomal doxorubicin versus topotecan. *J Clin Oncol* **19**, 3312–3322 (2001).
15. Li L., S. D. & H. Pharmacokinetics and biodistribution of nanoparticles. *Mol Pharm* **5**, 496–504 (2008).
16. Boussios S Katsanos K Pavlidis N., P. G. Systemic treatment-induced gastrointestinal toxicity: Incidence, clinical presentation and management. *Ann. Gastroenterol.* **25**, 106–118 (2012).

17. Sunar U Morgan J Zeitouni N Henderson BW., R. D. J. Quantification of PpIX concentration in basal cell carcinoma and squamous cell carcinoma models using spatial frequency domain imaging. *Biomed Opt Express* **4**, (2013).
18. Ozturk MS Sunar U Intes X., R. D. Mesoscopic Fluorescence Tomography of a Photosensitizer (HPPH) 3D Biodistribution in Skin Cancer. *Acad Radiol.* **21**, 271–280 (2014).
19. Mo W Sunar U., R. D. Imaging a photodynamic therapy photosensitizer in vivo with a time-gated fluorescence tomography system. *J Biomed Opt*. **17**, (2012).
20. Mallidi, S. *et al*. Beyond the Barriers of Light Penetration: Strategies, Perspectives and Possibilities for Photodynamic Therapy. *Theranostics* **6**, 2458–2487 (2016).
21. Wang S Scholzen T Zhang Z Vogel A Hasan T Rahmanzadeh R., H. G. A light-controlled switch after dual targeting of proliferating tumor cells via the membrane receptor EGFR and the nuclear protein Ki-67. *Sci Rep* **6**, (2016).
22. Spring BQ Zheng LZ Mai Z Watanabe R Sherwood ME Schoenfeld DA Pogue BW Pereira SP Villa E Hasan T., B. S. R. A photoactivable multi-inhibitor nanoliposome for tumour control and simultaneous inhibition of treatment escape pathways. *Nat Nanotechnol* **11**, 378–387 (2016).
23. Celli JP Blanden AR Massodi I Glidden MD Pogue BW Hasan T., R. I. An imaging-based platform for high-content, quantitative evaluation of therapeutic response in 3D tumour models. *Sci Rep* **4**, (2014).
24. Zhong W Rizvi I Mai Z Spring BQ Yun SH Hasan T., C. J. P. In vivo high-resolution fluorescence microendoscopy for ovarian cancer detection and treatment monitoring. *Br J Cancer* **101**, 2015–2022 (2009).
25. Sheng C Hasan T Pogue BW., H. P. J. Photobleaching-based dosimetry predicts deposited dose in ALA-PpIX PDT of rodent esophagus. *Photochem Photobiol.* **83**, 738–748 (2007).
26. Sheng Pogue B. W. Wang E. Hutchins J. E. & Hoopes P. J., C. Assessment of photosensitizer dosimetry and tissue damage assay for photodynamic therapy in advanced-stage tumors. *Photochem. Photobiol*. **79**, 520–525 (2004).
27. Henderson BW, D. T. J. How does photodynamic therapy work? *Photochem Photobiol*. **55**, 145–157 (1992).
28. Azais H Bonner C Kerdraon O Tardivel M Jetpisbayeva G Frochot C Betrouni N Collinet P Mordon S., Q. G. Fischer 344 Rat: A Preclinical Model for Epithelial Ovarian Cancer Folate-Targeted Therapy. *Int J Gynecol Cancer* **25**, 1194–1200 (2015).
29. Luo D. Razi A. Geng J. Shao S. Giraldo D. Sunar U. Ortega J. and Lovell J. F., C. K. A. Doxorubicin encapsulated in stealth liposomes conferred with light-triggered drug release. *Biomaterials* **75**, 193–202 (2016).
30. Luo, D. et al. Porphyrin-phospholipid liposomes with tunable leakiness. *Journal of Controlled Release* **220**, 484–494 (2015).
31. Luo, D. et al. Rapid Light-Triggered Drug Release in Liposomes Containing Small Amounts of Unsaturated and Porphyrin–Phospholipids. *Small* **12**, 3039–3047 (2016).
32. Carter KA, W. S. G. J. L. D. S. S. L. JF. Metal Chelation Modulates Phototherapeutic Properties of Mitoxantrone-Loaded Porphyrin-Phospholipid Liposomes. *Mol Pharm*. **13**, 420–427 (2016).
33. Kress J. Carter K. A. Luo D. Shao S. Lele S. Lovell J. F. and Sunar U., R. D. J. Quantitative imaging of light-triggered doxorubicin release. *Biomed. Opt. Express* **6**, (2015).
34. Carter, K. A. et al. Porphyrin–phospholipid liposomes permeabilized by near-infrared light. *Nat. Commun*. **5**, (2014).
35. Carter KA, L. D. R. A. G. J. S. S. O. J. L. JF. Sphingomyelin liposomes containing porphyrin-phospholipid for irinotecan chemophototherapy. *Theranostics* **6**, 2329–2336 (216AD).
36. Lovell JF, J. C. H. E. J. H. K. C. R. J. C. W. C. W. W. L. V., Z. G. Porphysome nanovesicles generated by porphyrin bilayers for use as multimodal biophotonic contrast agents. *Nat Mater* **10**, 324–332 (2011).
37. Miranda D, L. JF. Mechanisms of light-induced liposome permeabilization. *Bioeng Transl Med.* **1**, 267–276 (2016).
38. Major, A. L. et al. In vivo fluorescence detection of ovarian cancer in the NuTu-19 epithelial ovarian cancer animal model using 5-aminolevulinic acid (ALA). *Gynecol Oncol* 122–132 (1997) doi:https://doi.org/10.1006/gyno.1996.4502.
39. Bogaards A Trachtenberg J Wilson BC Lilge L., S. H. J. In vivo quantification of fluorescent molecular markers in real-time by ratio imaging for diagnostic screening and image-guided surgery. *Lasers Surg Med* **39**, 605–613 (2007).
40. Yang B Tunnell JW., S. M. Attenuation-corrected fluorescence extraction for image-guided surgery in spatial frequency domain. *J Biomed Opt* **18**, (2013).
41. Yang VX Herman P Wilson BC., M. P. J. A multispectral fluorescence imaging system: design and initial clinical tests in intra-operative Photofrin-photodynamic therapy of brain tumors. *Lasers Surg Med* **32**, 224–232 (2003).
42. Sunar U. Rohrbach D. J. Carter K. A. Luo D. Shao S. Lele S. and Lovell J. F, K. J. Light-triggered doxorubicin release quantified by spatial frequency domain imaging and diffuse optical spectroscopy. *Optics InfoBase Conference Papers* Preprint at https://opg.optica.org/abstract.cfm?uri=ots-2016-JW4A.3 (2016).
43. Kress J Carter KA Luo D Poon C Aygun-Sunar S Shao S Lele S Lovell JF Sunar U., R. D. J. A dual-channel endoscope for quantitative imaging, monitoring, and triggering of doxorubicin release from liposomes in living mice. *Sci Rep*. **7**, (2017).
44. Rohrbach Carter K. A. Luo D. Shao S. Aygun-Sunar S. Lovell J. F. & Sunar U., D. J. Fluence Rate-Dependent Kinetics of Light-Triggered Liposomal Doxorubicin Assessed by Quantitative Fluorescence-Based Endoscopic Probe. *Int J Mol Sci* **26**, (2025).
45. Ghosh S, C. U. S. U. L. JF. Chemophototherapeutic Ablation of Doxorubicin-Resistant Human Ovarian Tumor Cells 1. *Photochem Photobiol*. **99**, 844–849 (2023).
46. Baran T. H., T. M. & F. Recovery of intrinsic fluorescence from single-point interstitial measurements for quantification of doxorubicin concentration. *Lasers Surg. Med*. **45**, 542–550 (2013).
47. Cuccia DJ Durkin AJ Ayers FR Tromberg BJ., B. F. Quantitation and mapping of tissue optical properties using modulated imaging. *J Biomed Opt* **14**, (2009).
48. C., A. K. K. M. M. Y. & W. B. Quantification of in vivo fluorescence decoupled from the effects of tissue optical properties using fiber-optic spectroscopy measurements. *J. Biomed. Opt*. **15**, (2010).

49. Valdés, P. A. *et al*. Quantitative fluorescence in intracranial tumor: Implications for ALA-induced PpIX as an intraoperative biomarker - Clinical article. *J Neurosurg* **115**, 11–17 (2011).
50. Valdes, P. A., Angelo, J. P., Choi, H. S. & Gioux, S. qF-SSOP: real-time optical property corrected fluorescence imaging. *Biomed. Opt. Express* **8**, 3597–3605 (2017).
51. Valdes, P. A. *et al*. System and methods for wide-field quantitative fluorescence imaging during neurosurgery. *Opt. Lett*. **38**, 2786–2788 (2013).
52. Saager, R. B., Durkin, A. J., Cuccia, D. J., Saggese, S. & Kelly, K. M. Quantitative fluorescence imaging of protoporphyrin IX through determination of tissue optical properties in the spatial frequency domain. *J Biomed Opt* **16**, 126013 (2011).
53. Yang, B., Sharma, M. & Tunnell, J. W. Attenuation-corrected fluorescence extraction for image-guided surgery in spatial frequency domain. *J Biomed Opt* **18**, 80503 (2013).
54. Gardner CM Welch AJ., J. S. L. Fluorescence spectroscopy of tissue: recovery of intrinsic fluorescence from measured fluorescence. *Appl Opt* **35**, (1996).
55. Saager RB, C. D. S. S. K. K. D. AJ. Quantitative fluorescence imaging of protoporphyrin IX through determination of tissue optical properties in the spatial frequency domain. *J Biomed Opt* **16**, 126013 (2011).
56. Valdes, P. A., Angelo, J. P., Choi, H. S. & Gioux, S. qF-SSOP: real-time optical property corrected fluorescence imaging. *Biomed. Opt. Express* **8**, 3597–3605 (2017).
57. Zeitouni, N. C., Rohrbach, D. J., Aksahin, M. & Sunar, U. Preoperative Ultrasound and Photoacoustic Imaging of Nonmelanoma Skin Cancers. *Dermatologic Surgery* **41**, (2015).
58. Gioux, S. *et al*. Three-dimensional surface profile intensity correction for spatially modulated imaging. *J Biomed Opt* **14**, 034045 (2009).
59. Sunar U, R. D. M. J. Z. N. Noninvasive imaging of absolute PpIX concentration distribution in nonmelanoma skin tumors at pre-PDT. in *Proc. SPIE 8568, Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XXII, 85680C* (SPIE 8568, San Francisco, California, United States, 2013). doi:https://doi.org/10.1117/12.2005078.
60. Angelo JP, C. S. O. M. S. U. G. S. I. X. Review of structured light in diffuse optical imaging. *J Biomed Opt*. **24**, 1–20 (2018).
61. Rohrbach DJ Huihui J Saager R Keymel K Paquette A Morgan J Zeitouni N Sunar U., M. D. Preoperative mapping of nonmelanoma skin cancer using spatial frequency domain and ultrasound imaging. *Acad Radiol*. **21**, (2014).
62. Travers JB, P. C. B. T. R. B. B. C. R. D. B. S. T. J. R. M. D. H. K. K. C. L. B. S. R. C. M. E. G. S. H. H. R. A. D. M. M. N. C. E. S. U. Quantifying skin photodamage with spatial frequency domain imaging: Statistical results. *Biomed Opt Express* **10**, 4676–4683 (2019).
63. Poon, C. *et al*. Mesoscopic imaging of actinic skin damage using spatial frequency domain imaging. in *Biophotonics Congress: Biomedical Optics Congress 2018 (Microscopy/Translational/Brain/OTS)* CF3B.5 (Optica Publishing Group, 2018). doi:10.1364/TRANSLATIONAL.2018.CF3B.5.
64. Travers, J. B. *et al*. Noninvasive mesoscopic imaging of actinic skin damage using spatial frequency domain imaging. *Biomed Opt Express* **8 6**, 3045–3052 (2017).
65. Wagnieres GA, S. W. W. B. In vivo fluorescence spectroscopy and imaging for oncological applications. *Photochem Photobiol*. **68**, 603–632 (1998).
66. Lilge, L., Molpus, K., Hasan, T. & Wilson, B. C. Light Dosimetry for Intraperitoneal Photodynamic Therapy in a Murine Xenograft Model of Human Epithelial Ovarian Carcinoma. in *Photochemistry and Photobiology* vol. 68 281–288 (American Society for Photobiology, 1998).
67. Pogue, B. W., Lilge, L., Patterson, M. S., Wilson, B. C. & Hasan, T. Absorbed photodynamic dose from pulsed versus continuous wave light examined with tissue-simulating dosimeters. *Appl. Opt*. **36**, 7257–7269 (1997).
68. Moriyama, E. H., Kim, A., Bogaards, A., Lilge, L. & Wilson, B. C. A ratiometric fluorescence imaging system for surgical guidance. *Advances in Optical Technologies* (2008) doi:10.1155/2008/532368.
69. Wilson, N. A., Abu-Shumays, R., Lieberman, K., Akeson, M. & Dunbar, W. B. Feedback control of a DNA molecule tethered in a nanopore to repeatedly probe DNA-binding enzymes. in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* 5745–5748 (2008). doi:10.1109/IEMBS.2008.4650519.
70. Wilson, BC. & Patterson, MS. The physics, biophysics and technology of photodynamic therapy. *Phys Med Biol* **53**, (2008).
71. Abdelrehim T, S. U. Intraperitoneal Cavity Surface Reconstruction for Ovarian Cancer Micrometastasis by SFDI Laparoscopy. in *Biophotonics Congress: Biomedical Optics 2022 (Translational, Microscopy, OCT, OTS, BRAIN)* (Optica Publishing Group, Fort Lauderdale, Florida, United States, 2022).
72. Poon, C. *et al*. Early assessment of burn severity in human tissue ex vivo with multi-wavelength spatial frequency domain imaging. *Toxicology in Vitro* **52**, 251–254 (2018).
73. Sunar, U., Rohrbach, D. J., Morgan, J. & Zeitouni, N. Imaging nonmelanoma skin cancers with combined ultrasound-photoacoustic microscopy. in *Optical Methods for Tumor Treatment and Detection: Mechanisms and Techniques in Photodynamic Therapy XXII* (eds. Kessel, D. H. & Hasan, T.) vol. 8568 856816 (SPIE, 2013).
74. Sunar, U., Mo, W. & Rohrbach, D. PDT induced changes assessed by time-gated fluorescence tomography in a head and neck tumor model. in *Biomedical Optics and 3-D Imaging* BSu5A.9 (Optica Publishing Group, 2012). doi:10.1364/BIOMED.2012.BSu5A.9.
75. Angelo JP Gioux S., van de G. M. Real-time endoscopic optical properties imaging. *Biomed Opt Express*. **8**, 5113–5126 (2017).
76. van de Giessen, M., Angelo, J. P. & Gioux, S. Real-time, profile-corrected single snapshot imaging of optical properties. *Biomed. Opt. Express* **6**, 4051–4062 (2015).

77. Angelo J. Lee B. T. Bigio I. J. and Gioux S., V. C. R. Ultrafast optical property map generation using lookup tables. *J. Biomed. Opt*. **21**, (2016).
78. Applegate M Angelo J Austin W Tabassum S Aguénounon E Tilbury K Saager R Gioux S Roblyer D, K. K. OpenSFDI: an open-source guide for constructing a spatial frequency domain imaging system. *J Biomed Opt*. **25**, 1–13 (2020).
79. Fang, Q. & Boas, D. A. Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units. *Opt. Express* **17**, 20178–20190 (2009).

6

Fluorescence Imitating Brightfield Imaging (FIBI): A Simple Slide-Free Microscopy Approach

Nathan Anderson, Farzad Fereidouni, and Richard Levenson

6.1 Introduction

Histopathology, the study of tissue specimens under a microscope, reached near-contemporary performance in the late 1800s. By the 1890s, histology slides were being prepared with today's methodology, namely formalin fixation and paraffin embedding (FFPE), followed by sectioning and staining with hematoxylin and eosin (H&E) (1). This reliable technique facilitated a revolution in diagnostic medicine, and to this day, H&E-stained slides remain the gold standard for evaluation of tissue pathology.

While H&E-based methods remain at the center of tissue diagnostics, they nevertheless have significant drawbacks that have been tolerated until recently. The modern histology lab requires a suite of specialized equipment and a team of skilled technicians, resulting in high operational costs and difficulty in deploying labs to lower-resource environments. Perhaps the most troublesome aspect of FFPE slides is that they are inherently slow to prepare; even in well-equipped laboratories, *stat* tissue processing takes hours to perform (2), and nearly one in four biopsies are associated with turnaround times of three days or longer (3). Rapid histology, at point of care, could dramatically improve the velocity of patient care. Furthermore, pathology stands out as a field of medicine that has resisted modernization and adoption of digital formats; only in 2017 did the FDA approve the first digital whole slide scanner (4). Physical histology slides represent an analog platform in a world embracing digitization; migration to digital can facilitate remote diagnostics as well as computational image analysis.

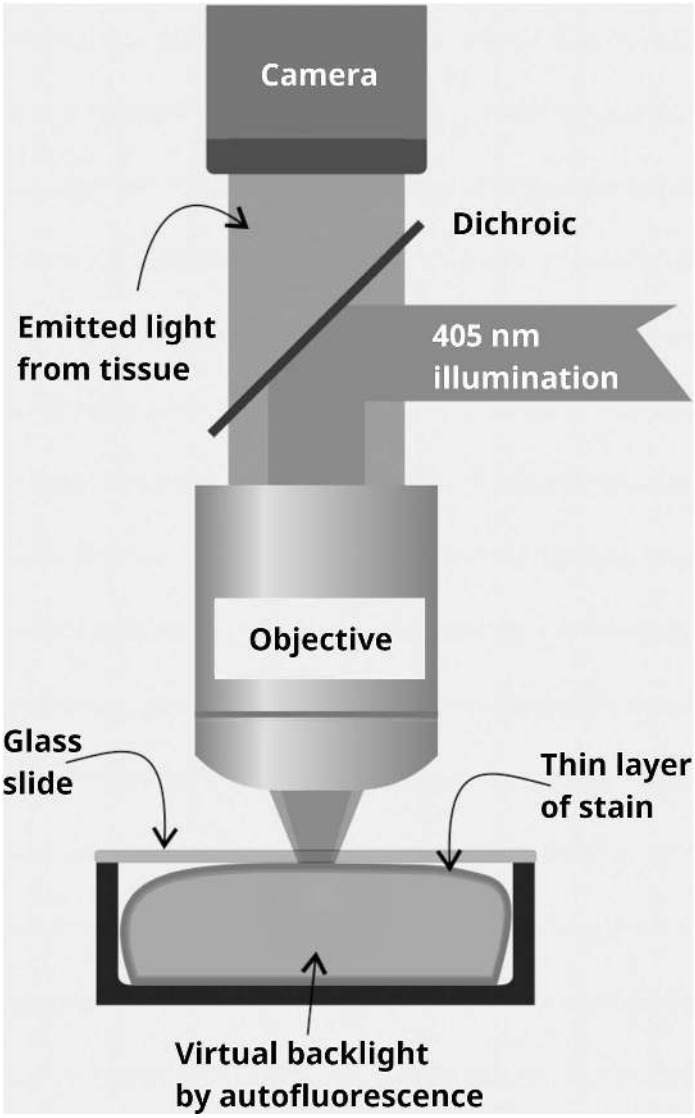
A particular drawback of standard FFPE tissue processing is its unsuitability for intraoperative use, including tumor margin detection. Intraoperative histology deployed for such surgical guidance often involves frozen sections prepared using cryotomes (5). These tools rapidly cut frozen tissue samples into thin sections without the need for prior fixation and paraffin embedding, allowing the specimens to be evaluated, typically within 10–30 minutes (6). However, technical issues render the method inferior to FFPE in most qualities other than speed. Tissues analyzed this way are prone to freezing artifacts (7), and improper handling may result in significant tissue damage (8). The process may consume much of the specimen and thus may require subsequent excision of additional tissue for downstream analytics. Moreover, some specimens are not well suited for frozen sections at all; breast is a good example, as the fatty tissue does not freeze or cut well.

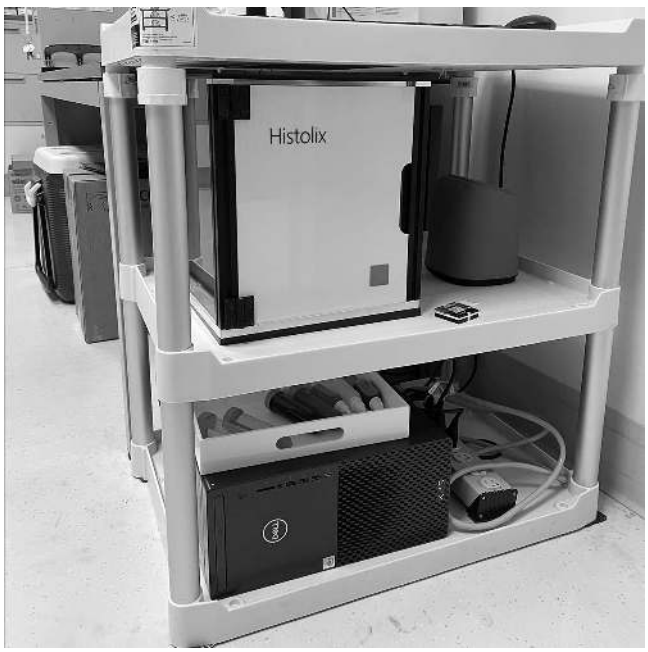
A series of technologies are now rapidly emerging under the rubric of "slide-free microscopy." These methods, though varied in approach, can rapidly generate diagnostic-quality histology while circumventing the pitfalls of H&E slides and frozen sections. Technologies in this space include structured illumination, conventional reflectance and fluorescence confocal microscopy, multi-photon imaging, spectrally encoded confocal microscopy, stimulated Raman scattering microscopy, light-sheet microscopy, and optical coherence tomography, among others (9).

Here we describe a technology termed "fluorescence imitating brightfield imaging" or FIBI. FIBI relies on a simple phenomenon: excitation light can generate autofluorescence in thick (< 100 μm) tissue specimens. In standard fluorescence microscopy, efforts are made to minimize tissue autofluorescence, as it can interfere with detection of desired signals. In contrast, FIBI involves excitation of specimens using a 405-nm LED deployed in an epifluorescence optical light path. The excitation light penetrates around a millimeter into tissue to generate diffuse broadband autofluorescence (10), some of which returns through the surface of the tissue that had been briefly stained prior to imaging. This transillumination light bearing surface-based contrast is then captured by a conventional objective lens and color camera. Staining is essential; without it, no useful structures are detected. Fortunately, in this application, standard histological stains such as hematoxylin and eosin can be used. The resulting FIBI images then closely recapitulate the appearance of conventional H&E slides under a brightfield microscope.

This workflow allows for diagnostic quality (11), direct-to-digital histological images to be generated in less than five minutes from the time tissue is received. By eliminating the need for extensive tissue processing, histology can be made available for interpretation almost immediately, at reduced cost and effort. Importantly, we have shown that prior FIBI imaging has no impact on subsequent FFPE processing or downstream molecular analyses (manuscript in preparation), indicating that FIBI imaging can be inserted into current clinical workflows without detriment. Finally, the advantages of this approach are not confined to increased speed and reduced cost: images themselves can display novel properties that can enrich the content.

DOI: 10.1201/9781003449942-8





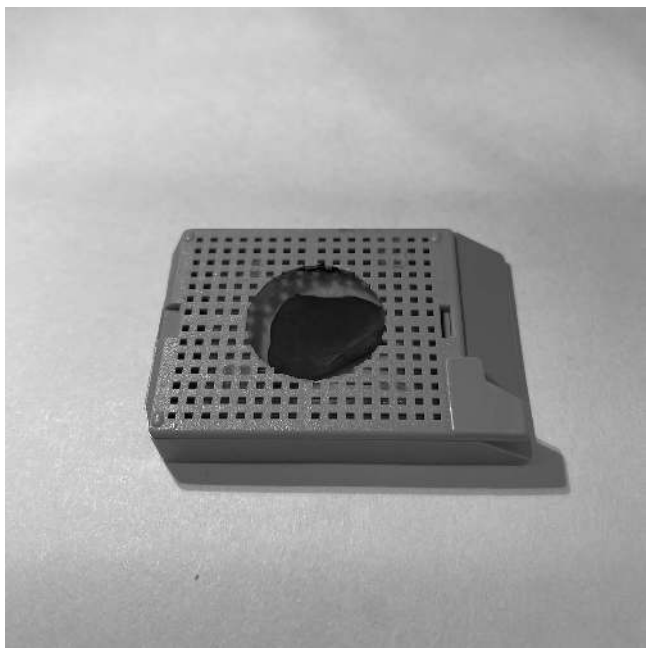


FIGURE 6.1 Schematic of FIBI mechanism (a), a functional hardware assembly (b) on a cart, and a windowed cassette containing a stained specimen (c) for imaging.

6.2 Materials

6.2.1 Optical Design and Components

Advantages of the FIBI hardware design include simplicity, speed, modularity, small form factor, and minimal use of custom parts. A functional assembly is described here (Figure 6.1a), but alternative optomechanical components may be substituted to accommodate larger sample sizes or achieve higher scan speeds. Epifluorescence excitation is generated by a 405 nm light-emitting diode (LZ1–00UB00, ams Osram), though other wavelengths may be used. The light is directed through a broadband dichroic (Di03-R405-t1–25x36, Semrock) before being focused onto the sample by a 10X (Nikon, Plan Apo, 0.3 NA) or 20X (Nikon, Plan Apo, 0.75 NA) objective lens. Scattered and reflected excitation light is rejected by a long-pass emission filter (FF01–430/LP-25, Semrock), and images are captured with a scientific-grade CCD color camera (MD091CU-SY, Ximea), though CMOS cameras also work well. Mounted samples are fixed to a motorized stage with an XY travel range of 25 mm × 50 mm (X-LSM025B and X-LSM025A, Zaber Technologies) as well as a Z-positioner (XVSR20A, Zaber Technologies) to facilitate imaging at multiple focal planes as needed to achieve extended depth of focus for capturing irregular tissue surfaces. A simple webcam is used to capture a preview image of the sample to aid in scan configuration.

6.3 Methods

6.3.1 Specimen Preparation

FIBI imaging can be performed on fresh (unfixed) tissue specimens and samples that have been chemically fixed, previously frozen, or even recovered from FFPE blocks via deparaffinization. Compared to fixed tissue, freshly resected specimens can pose challenges: they exhibit lower levels of autofluorescence (12, 13), requiring longer camera exposure times, and they can be deformable and difficult to slice cleanly (14). Rapid fixation techniques such as microwave irradiation can increase tissue firmness, making specimens easier to handle and cut.

Fresh or fixed samples are typically hand-sliced with a surgical blade to generate a flat surface to be scanned. Importantly, the specimen should be at least 100 μm thick (unpublished data), but there is no upper limit for specimen thickness, as excitation and imaging take place on the same surface.

After slicing, samples are submerged and agitated in a bath of filtered 0.5 mg/ml Mayer's hematoxylin (H3136, MilliporeSigma) for 30 seconds before being briefly rinsed in a bath of deionized water or phosphate buffered saline (PBS). The process is then repeated with a solution of 2.5 mg/ml eosin (E4382, MilliporeSigma) in acidic ethanol. The staining process is flexible and may be adjusted; for example, stains and washing agents may be applied to the tissue surface by hand with transfer pipettes. Already prepared solutions are commercially available (STE0150/HXMMHLT, StatLab) and can be used directly out of the bottle.

After staining, the specimen is placed in a sample holder: a small tissue sample can be mounted in a standard histology cassette (for example, Leica IP ActivFlo Routine I Cassette) with an imaging window cut in the lid and a coverslip positioned just below (Figure 6.1c). Following FIBI imaging, the windowed lid can be replaced with an unmodified one, and the cassette can be directly submitted for paraffin embedding and subsequent standard histology if desired. For specimens too large to fit in a standard cassette, a simple sample holder can be machined or 3D printed to allow gentle compression of tissues against a glass or plastic window. To maintain sample hydration during imaging, a few drops of PBS can be placed on the coverslip before positioning the specimen. Standard histology foam can be positioned between the specimen and the backplate to facilitate gentle compression. The specimen is then mounted on the microscope's XY stage for imaging.

6.3.2 Software and Scanning

At present, FIBI microscopes are controlled by a custom software interface developed in the .NET environment. An intuitive click-and-drag interface allows the user to reposition the stages for real-time navigation and viewing of the specimen. Scans are configured by selecting regions of interest on the preview image. Acquisition of individual fields of view (approximately 1 mm^2 at 10X) takes place with exposure times that typically range from 5 to 50 ms for fixed tissues, though higher values may be needed for unfixed specimens. In scans utilizing extended depth of focus (EDOF), each field of view is captured at various focal planes and the best in-focus regions are automatically selected and combined into a single image using the methodology first described in Goldsmith, 2000 (15). After acquisition and EDOF processing, the resulting images are stitched into a single .svs file using a stitching library from Microvisioneer (Esslingen am Neckar, Germany). Post-processing occurs in parallel with stitching and can include adjustments such as flatfield correction, histogram matching (16), and dehazing.

6.4 Results

FIBI employs an epifluorescence light path to deliver excitation light onto thick (i.e., > 100 μm) tissues whose surfaces have been stained, typically with hematoxylin and eosin. The excitation light generates diffuse autofluorescence within the specimen, and some of the resulting light returns through the stained surface of the tissue to be captured with a color camera. Combining this "virtual backlight" with absorbing stains effectively mimics many aspects of traditional bright-field H&E microscopy. Examples are shown in Figure 6.2.

Although most of the contrast in FIBI images can be attributed to absorbance, other fluorescence phenomena take place that produce additional visual content. For example, autofluorescence also occurs at the surface of the specimen, and strong signals generated by components such as collagen and elastin reveal structures that might otherwise only be visible with special stains (17, 18) (Figure 6.3a). The color profile of FIBI images can be modified at acquisition by adjusting gains of the individual camera color channels. Post-capture color corrections (automatic or manual) can also be employed to emphasize certain features or generate images that more closely resemble standard H&E.

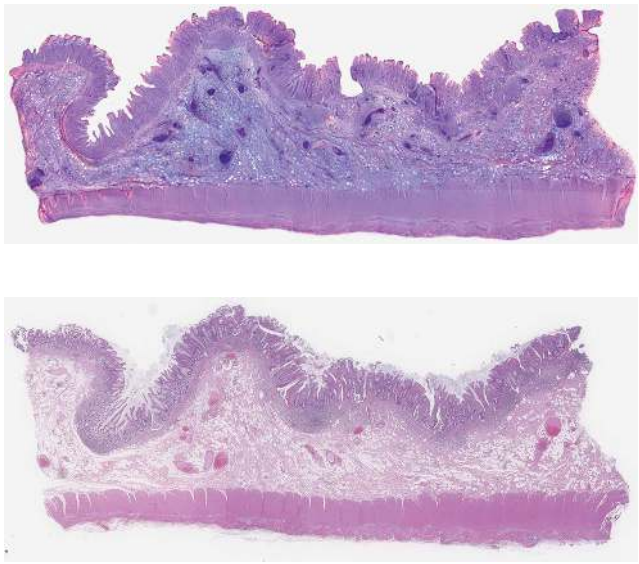


FIGURE 6.2 Specimen of human colon with mild ulcerative colitis. FIBI scan (a), and the corresponding H&E slide subsequently prepared from the same specimen (b) after FIBI.

6.4.1 FIBI Image Characteristics

Beyond displaying additional colors, FIBI images can be enriched by physical content arising from the fact that specimens are thick, rather than being cut into 5 μm slices (Figure 6.3c). Moreover, they are imaged before (rather than after) being exposed to xylene, so adipose cell contents are not dissolved away (19) (Figure 6.3e).

Traditional histology slides feature additional artifacts stemming from standard tissue processing methods (19). Because harsh tissue processing steps are not used for FIBI imaging, resulting scans are expected to better represent histology as it might appear *in situ* and can reveal structural components not readily seen on standard slides.

6.4.2 Validation

In a recent validation study (11), 100 tissue specimens representing 24 tissues were imaged using FIBI before being submitted for routine histological processing. Resulting H&E slides (which were nearly identical to their FIBI counterparts) were scanned to generate digital whole slide images. A panel of pathologists was asked to make a diagnosis on each image, and results tabulated. In this study, there was a 97% clinical concordance between FIBI diagnoses and reference diagnoses.

6.5 Discussion and Conclusion

6.5.1 Versatility and Accessibility

FIBI can be used in a wide range of applications and environments. The hardware was designed to be easy to adopt: the system itself is roughly the size of a standard microwave, enabling it to fit in typical surgery, radiology, and pathology workspaces, and it can be readily shipped worldwide. The mechanical and optical components are inexpensive compared to other slide-free approaches; the images are acquired using simple LEDs (rather than lasers), a standard color camera,

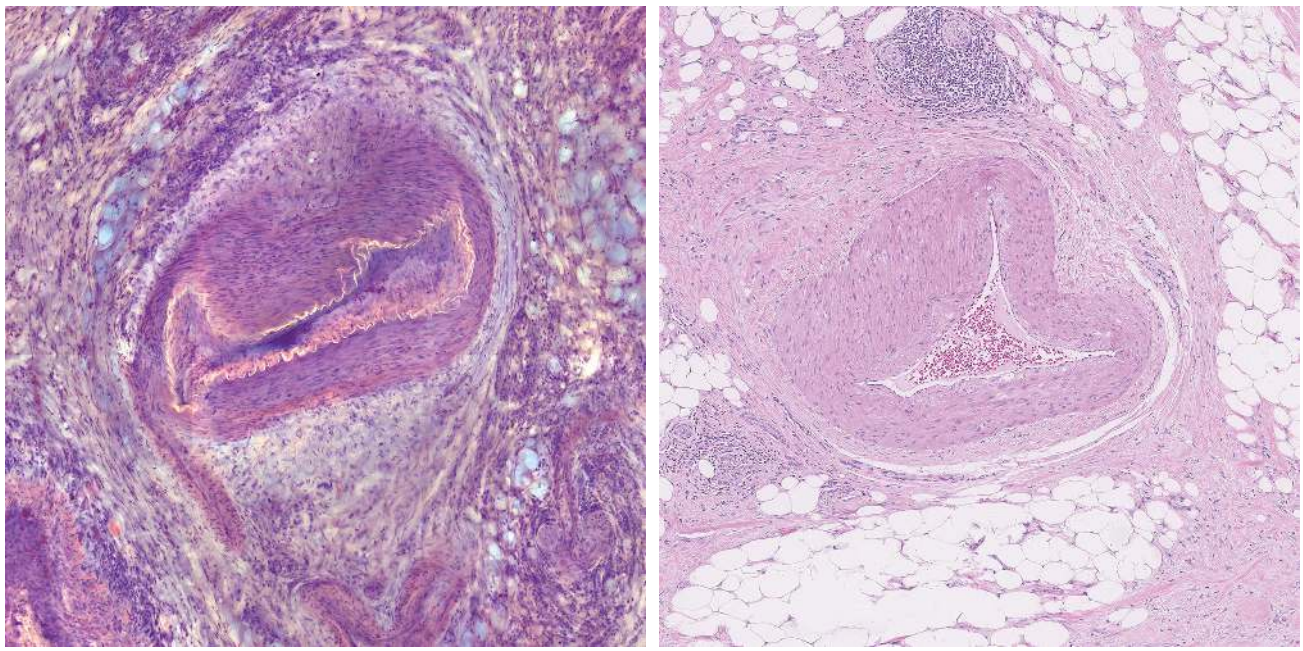


FIGURE 6.3 Novel features of FIBI histology (left) compared to conventional H&E histology (right). Elastin layers in a vessel (a) are more evident than in the corresponding H&E slide (b). Small bowel villi demonstrate structural continuity afforded by imaging unsectioned tissue (c), corresponding H&E (d). FIBI image of intact adipose tissue with easily detectable capillaries (e) compared to regions containing adipocytes whose fat deposits have been dissolved during processing (f).

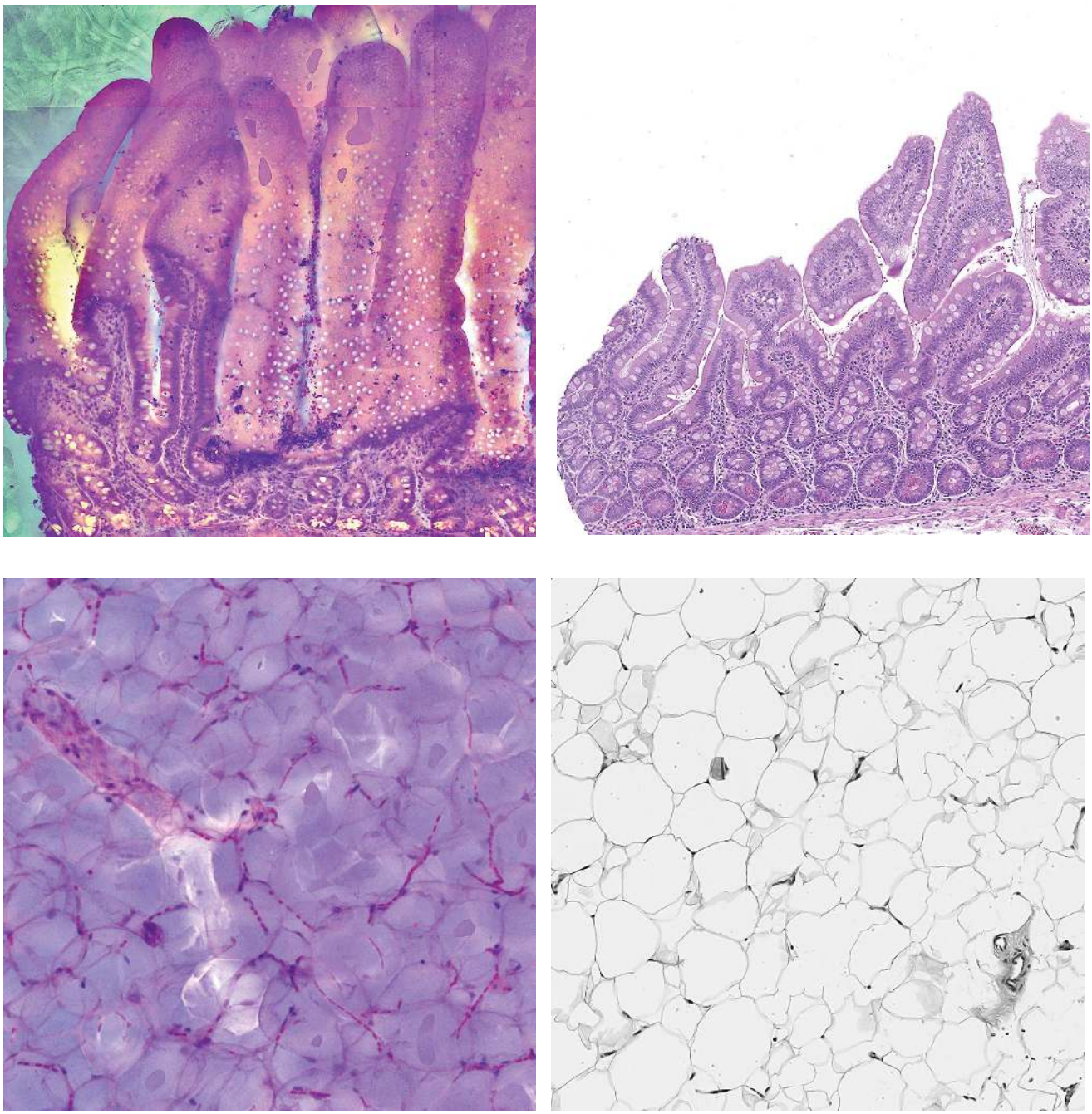


FIGURE 6.3 (continued)

straightforward sample stages, and typical microscope lenses. In addition to managing specimens of approximately the size that fits on standard H&E slides, it is straightforward to increase the sample size that can be accommodated. This can be accomplished just by enlarging the stage and sample holding hardware—the rest of the microscope system can remain unchanged. Preparing samples and configuring scans require little more than a surgical blade and standard histology stains; less technical ability is required than for standard histological tissue processing. Because the process is simple, fast, and can be largely automated, FIBI has the potential to help alleviate the widespread shortage of skilled histotechnologists (20, 21). FIBI's various advantages suggest many possible use cases, such as the following.

6.5.2 Applications: Surgical Guidance and Intraoperative Margin Detection

FIBI's compatibility with fresh tissue specimens makes it a suitable candidate to replace frozen sections for intraoperative

guidance, and it would be especially valuable for situations in which frozen sectioning is either technically challenging or not available at all. The resulting images can be technically superior to what is often generated by the freezing and cutting steps involved in current methods, and unlike frozen sectioning, it does not consume valuable tissue specimens. A specific application is evaluating surgical margins, helping surgeons to optimally manage tumor resections while avoiding unnecessary damage to vital neighboring structures.

6.5.3 Applications: Rapid Biopsy Evaluation

It can be important to assess tissue specimens, often acquired by core-needle biopsy, for adequacy—to ensure that the target lesion was in fact sampled and that the amount recovered is sufficient to support downstream tests for which the biopsy was obtained. This rapid on-site evaluation (ROSE) is typically performed using cytology-based methods, which do not provide any tissue structure information. By contrast, such specimens can be immediately and non-destructively imaged via FIBI to provide both adequacy readings and high-quality, potentially diagnostic histology. The specimens can then be processed via standard H&E workflow (as demonstrated in Figure 6.2) or immediately submitted for molecular studies.

6.5.4 Applications: Research

FIBI can produce histology-scale images that can present new structural insights because of its combination of novel color contrast and spatial continuity of small structures, as described previously. While these might be visible with alternative imaging methods (confocal, etc.), FIBI's ease of use and low cost can make such features available in many research settings.

6.5.5 Applications: Deployment to Low-Resource Countries

We believe that FIBI can change how tissue evaluation is performed at well-equipped medical facilities, but its true impact may lie in low-resource countries (LRCs). Across Africa, for example, cancer incidence is rising dramatically due in part to increasing life expectancies, and cancer death rates in LRCs already far exceed those in the US (22, 23). Lack of diagnostic histology services and pathologists contribute to disease impact, as early diagnosis is challenging (24–26). In regions with limited healthcare infrastructure, FIBI could fill a need for standard histopathology services thanks to its low cost and portability. And as images are direct-to-digital, they can easily be shared with pathologists worldwide for remote interpretation, addressing the significant shortage of pathologists in many LRCs.

6.5.6 Challenges

Slide-free microscopy methods such as FIBI offer powerful advantages over conventional histology methods, but as usual, such benefits come with tradeoffs. Many of FIBI's ideal use cases involve imaging of fresh tissues, which exhibit lower levels of autofluorescence than fixed specimens, potentially affecting scan speed. Scan durations are of little concern for biopsy-sized specimens (wherein scanning a fresh specimen takes roughly two minutes), but this may pose a challenge if larger scans (e.g., organ whole-mounts) are desired. Future iterations of FIBI may include higher LED intensity or alternative light sources, and faster hardware components such as stage and camera.

Compared to fixed tissue specimens, fresh tissue also exhibits a lower propensity for nuclear staining due to reduced membrane permeability (27). This may be overcome by prolonging the nuclear staining step, though further investigation is needed to determine ideal durations and identify if/how different tissue types behave in this regard.

Prolonged or repeated exposure of the stained tissue to FIBI's excitation light can result in bleaching of the eosin stain (Figure 6.4a). This is typically not a problem for automated scans (as compared to manual tissue viewing). If necessary, a quick restain with eosin can restore contrast.

Because FIBI is used on specimens far thicker than conventional H&E slides, signals from subsurface cells may be captured, causing images to display greater cellularity than expected in standard histology (Figure 6.4). Actual results depend on the magnification and numerical aperture (NA) of the collecting lens, and in most cases have no impact on the value of the resulting FIBI images. However, too many overlapping cells can make it hard to appreciate nuclear and subnuclear morphology and can affect overall cellularity estimates. One can defer to a subsequent H&E slide if necessary to evaluate such cytological details. We are currently working on alternative optical arrangements and accompanying software tools to recover high-resolution (thinner sampling) signals. Stay tuned.

Finally, as with any change in technology and workflow, it will be non-trivial to introduce FIBI-based tissue imaging into the standard workflow and staffing needs of existing pathology providers. When and where tissue reads occur, whether practitioners must be on-site and available at time of imaging, who will be responsible for performing and reading the FIBI scans, and of course, how capital costs and reimbursement issues will be managed will all have to be worked out.

6.5.7 Conclusion

In addition to describing the phenomena that make FIBI possible, this chapter discusses methods for constructing a functional hardware assembly and preparing tissue for scanning. Furthermore, we outline the novel histological features in FIBI images as well as various use cases for FIBI in both clinical and research settings. We are encouraged by FIBI's early success in validation studies, but it is well understood that there is still room for refining methods and exploring new applications. Luckily, FIBI possesses a favorable combination of traits (simplicity, low cost, and the ability to provide enriched histology) that facilitates collaboration. We hope that future studies with new partners will help guide process development and further validation, eventually culminating in FDA approval for maximum impact worldwide.

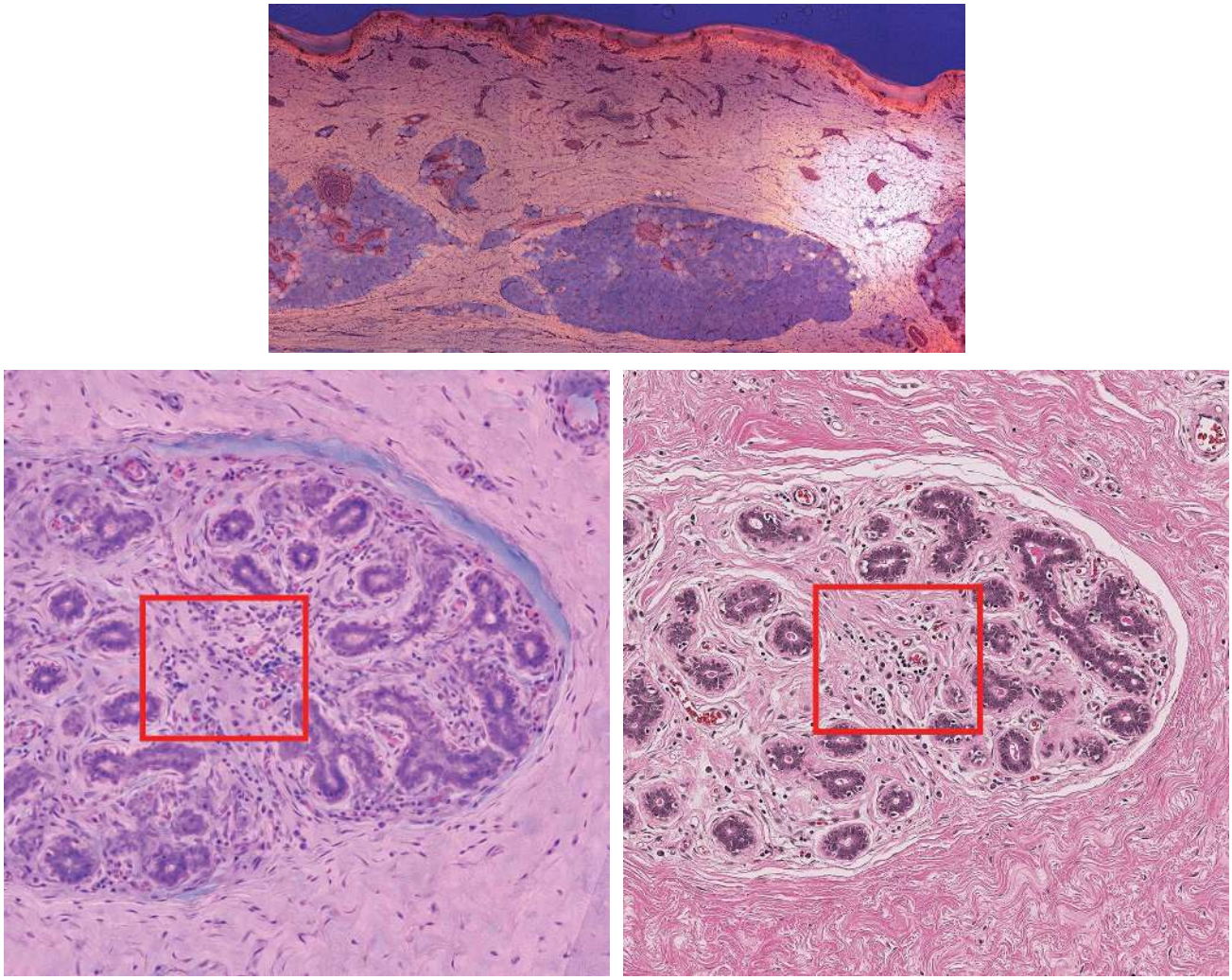


FIGURE 6.4 Tradeoffs of FIBI imaging. Dermal tissue (pig foot) demonstrates bleaching in a region of prolonged (intentional) application of incident FIBI excitation light (a). Benign breast has higher cellularity in FIBI (b), where the indicated region contains 136 visible cells. The corresponding H&E (c) has 118 cells in approximately the same region.

6.6 Future Trends

This chapter discusses the core principles behind FIBI and how the technology generates images that functionally resemble H&E slides. However, the technology is still in its infancy, and improvements to the hardware, software, and tissue preparation protocols are ongoing. Future iterations of FIBI should facilitate even more utility, faster scans, and even higher image quality.

6.6.1 Hardware

Current versions of FIBI employ a single excitation light delivered via an epifluorescent light path. In most cases, this configuration generates enough autofluorescence that high-quality images can be captured using reasonable exposure times. However, there do exist cases where image quality is poor because of inherently low induction or transmission of autofluorescence (often due to necrosis or excessive blood, which absorbs light along both the excitation and the emission pathways). For tissues in which autofluorescence is poor, supplementary white light can be delivered onto the tissue surface with crossed polarizers positioned to block specular reflections. Early experiments are promising and offer the potential to better image tissues that have been challenging thus far.

6.6.2 Tissue Preparation

Prior to FIBI imaging, specimens are stained with hematoxylin and eosin; this is advantageous as these stains are familiar to histotechnologists and pathologists alike. However, we are examining alternate staining options in an effort to improve image content and minimize the impact of any possible photobleaching. In addition to alternate staining protocols, other treatments related to specimen preparation are being explored. As discussed earlier, fresh tissues can be more challenging to image than fixed specimens, as they exhibit lower levels of

autofluorescence, lower susceptibility to nuclear staining, and lack physical rigidity. Various treatments have shown promise in reducing or eliminating these challenges: brief (< 1 minute) formalin exposure can stiffen small specimens such as needle biopsies, surfactants can be used to permeabilize cell membranes for improved nuclear staining, and hydrogen peroxide improves fresh tissue imaging by reducing adverse light absorption by red blood cells (28). Furthermore, our team is investigating various methods of rapid fixation, including alternative chemical fixatives as well as microwave irradiation (29).

Future work is aimed at developing methods that can accomplish rapid immunostaining (immunohistochemistry and/or immunofluorescence) for point-of-care molecular phenotyping.

6.6.3 Automated Image Processing and Analysis

A given FIBI image is likely to be richer with information than a similar H&E slide, as the FIBI image will have more physical depth and color gamut. However, to many pathologists and computational diagnostic platforms alike, an image that deviates from the norm is one that poses challenges to the interpreter. It would be favorable if FIBI images could be (optionally) processed to maximize similarity with the H&E datasets that have been used to train the world's pathologists and computational diagnostic platforms.

At present, FIBI scans automatically undergo enhancements such as flatfield correction and dehazing, though this will soon be expanded to include color balancing to generate a color profile more akin to H&E slides. Custom tools will allow the original color profile (with a wide range of hues) to be overlaid on the more familiar color-balanced scan to display a hybrid image demonstrating advantages from both color palettes.

AI-based tools offer considerable potential for enhancement and conversion of FIBI images. Preliminary work has demonstrated a cycleGAN image conversion platform capable of translating FIBI images into "virtual H&E slides," which may be indistinguishable from the true tissue equivalent (30). However, similar solutions should acknowledge the possibility that by converting a FIBI image into something more akin to an H&E slide, potentially FIBI-specific diagnostic contrast may be eliminated.

AI tools for histological diagnosis already exist (31, 32), but these are generally trained on large datasets of H&E slides paired with diagnostic information. Similarly, large datasets will not be available during the early days of FIBI, and therefore, we are working to develop transfer-learning approaches that can take advantage of previous AI models in FIBI-image automated diagnostics.

REFERENCES

1. Hussein IH, Raad M. Once upon a microscopic slide: the story of histology. *J Cytol Histol* [Internet]. 2015 [cited 2024 Mar 28];06(06). Available from: https://www.omicsonline.org/open-access/once-upon-a-microscopic-slide-the-story-of-histology-2157-7099-1000377.php?aid=62840
2. Jali P, Donoghue M, Gadiwan M. A rapid manual processing technique for resource-limited small laboratories. *J Oral Maxillofac Pathol*. 2015;19(3):306.
3. Patel S, Smith JB, Kurbatova E, Guarner J. Factors that impact turnaround time of surgical pathology specimens in an academic institution. *Hum Pathol*. 2012 Sep;43(9):1501–5.
4. US Food and Drug Administration. FDA. 2020 [cited 2024 Mar 28]. *FDA allows marketing of first whole slide imaging system for digital pathology.* Available from: https://www.fda.gov/news-events/press-announcements/fda-allows-marketing-first-whole-slide-imaging-system-digital-pathology
5. Ferreiro JA, Myers JL, Bostwick DG. Accuracy of frozen section diagnosis in surgical pathology: review of a 1-year experience with 24,880 cases at mayo clinic rochester. *Mayo Clin Proc*. 1995 Dec;70(12):1137–41.
6. Jaafar H. Intra-operative frozen section consultation: concepts, applications and limitations. *Malays J Med Sci*. 2006 Jan;13(1):4–12.
7. Desciak EB, Maloney ME. Artifacts in frozen section preparation. *Dermatol Surg*. 2000 May;26(5):500–4.
8. Dey P. *Basic and Advanced Laboratory Techniques in Histopathology and Cytology* [Internet]. Singapore: Springer Singapore; 2018 [cited 2024 Mar 28]. Available from: http://link.springer.com/10.1007/978-981-10-8252-8
9. Liu Y, Levenson RM, Jenkins MW. Slide over: advances in slide-free optical microscopy as drivers of diagnostic pathology. *Am J Pathol*. 2022 Feb;192(2):180–94.
10. Lichtman JW, Conchello JA. Fluorescence microscopy. *Nat Methods*. 2005 Dec;2(12):910–9.
11. Borowsky AD, Levenson RM, Gown AM, Morningstar T, Fleury TA, Henderson G, et al. A pilot validation study comparing FIBI, a slide-free imaging method, with standard FFPE H&E tissue section histology for primary surgical pathology diagnosis [Internet]. *Pathology*; 2022 Mar [cited 2023 Jan 10]. Available from: http://medrxiv.org/lookup/doi/10.1101/2022.03.10.22272226
12. Collins T. *Autofluorescence: Causes and Cures* [Internet]. Wright Cell Imaging Facility, Toronto Western Research Institute, University Health Network; 2006. Available from: https://hwpi.harvard.edu/files/iccb/files/autofluorescence.pdf?m=1465309329
13. Wang X, Xie Y, Huang M, Yao L, Wang Y, Fei Y, et al. Effect of fixation and mounting on fluorescence lifetime of cellular autofluorescence. *IEEE J Sel Top Quantum Electron*. 2019 Jan;25(1):1–6.
14. Suvarna SK, Layton C, Bancroft JD, editors. *Bancroft's Theory and Practice of Histological Techniques*. Eighth edition. Amsterdam: Elsevier; 2019. 557 p.
15. Goldsmith NT. Deep focus; A digital image processing technique to produce improved focal depth in light microscopy. *Image Anal Stereol*. 2000 Nov 1;19(3):163.
16. Jia W, Zhang H, He X, Wu Q. A comparison on histogram based image matching methods. In: *2006 IEEE International Conference on Video and Signal Based Surveillance* [Internet]. Sydney, Australia: IEEE; 2006 [cited 2024 Mar 28]. p. 97–97. Available from: http://ieeexplore.ieee.org/document/4020756/
17. Koenig K, Schneckenburger H. Laser-induced autofluorescence for medical diagnosis. *J Fluoresc*. 1994 Mar;4(1):17–40.
18. Puchtler H, Waldrop FS, Valentine LS. Fluorescence microscopic distinction between elastin and collagen. *Histochemie*. 1973;35(1):17–30.

19. Taqi S, Sami S, Sami L, Zaki S. A review of artifacts in histopathology. *J Oral Maxillofac Pathol*. 2018;22(2):279.
20. Buesa RJ. Histology aging workforce and what to do about it. *Ann Diagn Pathol*. 2009 Jun;13(3):176–84.
21. Fimbres AM. The next generation of histologists. *J Histotechnol*. 2010 Sep;33(3):119–22.
22. Kohler RE, Gopal S, Miller AR, Lee CN, Reeve BB, Weiner BJ, et al. A framework for improving early detection of breast cancer in sub-Saharan Africa: a qualitative study of help-seeking behaviors among Malawian women. *Patient Educ Couns*. 2017 Jan;100(1):167–73.
23. Parkin DM, Bray F, Ferlay J, Jemal A. Cancer in Africa 2012. *Cancer Epidemiol Biomarkers Prev.* 2014 Jun 2;23(6):953–66.
24. Stalsberg H, Awuah B, Ibarra JA, Nsiah-Asare A. Re-establishing a surgical pathology service in Kumasi, Ghana: case report and discussion of barriers and key elements of a successful collaboration between low- and high-resource countries. *Cancer.* 2008 Oct 15;113(8 Suppl): 2338–46.
25. Yadav H, Shah D, Sayed S, Horton S, Schroeder LF. Availability of essential diagnostics in ten low-income and middle-income countries: results from national health facility surveys. *Lancet Glob Health*. 2021 Nov; 9(11):e1553–60.
26. El Saghir NS, Adebamowo CA, Anderson BO, Carlson RW, Bird PA, Corbex M, et al. Breast cancer management in low resource countries (LRCs): consensus statement from the Breast Health Global Initiative. *The Breast*. 2011 Apr;20:S3–11.
27. Cheng R, Zhang F, Li M, Wo X, Su YW, Wang W. Influence of fixation and permeabilization on the mass density of single cells: a surface plasmon resonance imaging study. *Front Chem*. 2019 Aug 23;7:588.
28. Hopp MT, Schmalohr BF, Kühl T, Detzel MS, Wißbrock A, Imhof D. Heme determination and quantification methods and their suitability for practical applications and everyday use. *Anal Chem*. 2020 Jul 21;92(14):9429–40.
29. Katoh K. Microwave-assisted tissue preparation for rapid fixation, decalcification, antigen retrieval, cryosectioning, and immunostaining. *Int J Cell Biol*. 2016;2016:1–9.
30. Engel TN, Abraham TM, Morningstar T, Fung MA, Rangchi A, Kiuru M, et al. Pilot study of fluorescence imitating brightfield imaging for rapid, slide-free dermatopathology. *J Cutan Pathol*. 2022;49(12):1060–6.
31. Da Silva LM, Pereira EM, Salles PG, Godrich R, Ceballos R, Kunz JD, et al. Independent real-world application of a clinical-grade automated prostate cancer detection system. *J Pathol*. 2021 Jun;254(2):147–58.
32. Shafi S, Parwani AV. Artificial intelligence in diagnostic pathology. *Diagn Pathol*. 2023 Oct 3;18(1):109.

7

Fluorescently Labeled Tyrosine Kinase Inhibitors for Intracellular Protein Target Imaging

Allison Solanki, Jesse Korber, Summer L. Gibbs, and Lei G. Wang

7.1 Introduction

Cancer remains a leading cause of death worldwide, with an estimated 20 million new cancer diagnoses and 10 million cancer-related deaths annually, despite significant advancements in diagnostic technologies and treatment regimens (1). Notably, metastatic disease accounts for 90% of cancer-related mortality (2), where cancer early detection and screening can substantially increase patient survival rates. However, existing cancer screening technologies (e.g., colonoscopy, mammography, pap smear, etc.) are often cumbersome, inaccessible, or expensive, decreasing patient compliance and limiting accessibility in low-resource settings. In contrast, liquid biopsies (i.e., blood, urine, saliva) have emerged as promising tools for noninvasive routine monitoring, aiding early cancer detection and facilitating routine treatment response monitoring (3–5). Newly discovered neoplastic cells found in peripheral blood offer an exciting opportunity for cancer detection with much improved sensitivity, permitting a significant advancement in noninvasive, blood-based assays over other current approaches (6–8). We envision that the development of molecularly targeted fluorescent contrast agents for the identification of neoplastic cells in peripheral blood samples will transform patient cancer screening, with the potential to make a profound impact on patient care.

Epidermal growth factor receptor (EGFR) was identified as an oncogenic driver gene and a receptor tyrosine kinase (TK) over three decades ago (9–14). Since this discovery, aberrant expression or activity of EGFR has been identified as an essential biological factor linked to a variety of human epithelial cancers, including non-small cell lung carcinoma (NSCLC), head and neck squamous cell carcinoma (HNSCC), and pancreatic ductal adenocarcinoma (PDAC) (15). The development of EGFR tyrosine kinase inhibitors (TKIs) has revolutionized treatment outcomes for patients harboring these genetic abnormalities (16). Notably, there has been an unprecedented focus on the clinical use of TKIs over the last decade, with eight drugs receiving FDA approval and many more advancing to clinical trials with marked success (16–18). First-generation EGFR inhibitors (e.g., Erlotinib) feature an anilinoquinazoline core and inhibit EGFR protein activity through reversible competitive binding with ATP. To enhance the efficacy of these inhibitors, second-generation TKIs (e.g., Afatinib, PD-168393) have been developed. These inhibitors maintain the quinazoline core but include a thiol-reactive acrylamide warhead for irreversible binding, leading to increased potency and cytotoxicity, especially in cancers with EGFR mutations (16, 17, 19). These structural designs have not only transformed low-affinity reversible TKIs into highly potent compounds targeting drug-resistant mutants of EGFR, but also provided an excellent framework for fluorescence imaging of EGFR-positive cancer cells using these TKIs as the targeting motif for molecular-specific contrast agent design.

In this chapter, we describe a novel imaging toolbox consisting of fluorescently labeled standard-of-care (SOC) drugs for the identification and localization of intracellular EGFR in live cells. As a proof of concept, we synthesized and validated three fluorescently labeled small-molecule TKIs: Erlotinib-tetramethylrhodamine isothiocyanate (Erl-TRITC), Afatinib-TRITC (Afa-TRITC), and PD168393-TRITC (PD-TRITC). Previously reported synthetic pathways generating the TKIs were modified to incorporate the TRITC fluorophore conjugation (20, 21). Comprehensive *in vitro* studies, including binding affinity assessment for purified EGFR protein, and live cell staining in cancer cell lines with varied EGFR expression levels demonstrated a direct correlation between the fluorescence intensity of the internalized, fluorophore-labeled TKIs and the quantified EGFR expression levels. These findings underscore the potential of fluorophore-labeled TKIs as an innovative and promising strategy. They are poised to establish a strong foundation for the development of blood-based assays for future cancer early detection and screening.

7.2 Materials

7.2.1 Chemicals

All chemicals and solvents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Hampton, NH) and used without purification. Fluorophores were purchased from Thermo Fisher Scientific (Waltham, MA) and used without purification.

7.2.2 Cell Culture

The human osteosarcoma cell line, U2OS, was a generous gift from Dr. Xiaolin Nan at Oregon Health and Science University (OHSU, Portland, OR). The human epidermoid carcinoma cancer cell line, A431, was generously provided by Dr. Kimberley Samkoe at Dartmouth College (Hanover, NH).

DOI: 10.1201/9781003449942-9

The human lung adenocarcinoma cell line, HCC827, was purchased from ATCC (Manassas, VA). The cell lines were maintained at passage numbers < 25 throughout experimentation and regularly tested for mycoplasma contamination. The cell lines were expanded in their respective growth media (A-431, U2OS: DMEM [Thermo Fisher Scientific] + 10% fetal bovine serum [FBS, VWR Scientific, Radnor, PA] + 1% penicillin/streptomycin/glutamine [Thermo Fisher Scientific]; HCC827: RPMI 1640 [Thermo Fisher Scientific] + 10% FBS + 1% penicillin/streptomycin/glutamine) and stored in a 37 °C/5% CO_2 incubator.

7.3 Methods

7.3.1 Structure Confirmation and Characterization

All synthesized molecules were identified using tandem liquid chromatography mass spectroscopy (LCMS) analysis (Agilent 1260 liquid chromatography system with a 6224 time of flight [TOF] mass spectrometer, Agilent, Santa Clara, CA) using a 2.7-micron C18 column (4.6 × 50 mm, Agilent). Crude product was purified by flash column chromatography with silica gel or by high-performance liquid chromatography (HPLC, Agilent 1260 system with a preparative pump). All final products were confirmed using nuclear magnetic resonance (NMR) spectroscopy (Bruker Avance II+ 400 MHz or Avance III 600 MHz, Bruker, Billerica, MA).

7.3.2 Synthesis of Fluorescently Labeled EGFR TKIs

N'(2-cyano-4-nitrophenyl)-*N*,*N*-dimethylformimidamide (2): Compound 1 (1.09 g, 6.68 mmol) was dissolved in dioxane (25 mL). *N*,*N*-dimethylformamide dimethyl acetal (0.98 mL, 7.4 mmol) was added to the solution. The reaction mixture was refluxed for 2 h, after which the solvent was removed under reduced pressure, resulting in compound 2 as a yellow solid, which was used in the next step without further purification. ESI+ calculated for $C_{10}H_{11}N_4O_2\left[M+H\right]^+$: 219.09 m/z, measured: 219.09 m/z.

N-(3-chloro-4-fluorophenyl)-6-nitroquinazolin-4-amine (6): Compound 2 (1.0 g, 4.59 mmol) and compound 3 (0.80 g, 5.50 mmol) were added to acetic acid (15 mL). The reaction mixture was heated at 100 °C for 3 h. After cooling to room temperature, the crystalized reaction mixture was filtered and washed with ether to yield compound 6 as a yellow solid. ESI+ calculated for $C_{14}H_9ClFN_4O_2\left[M+H\right]^+$: 319.04 m/z , measured: 319.05 m/z.

N-(3-ethynylphenyl)-6-nitroquinazolin-4-amine (7): Compound 2 (1.0 g, 4.59 mmol) and compound 4 (0.64 g, 5.50 mmol) were added to acetic acid (15 mL). The reaction mixture was heated at 100 °C for 3 h. After cooling to room temperature, the crystalized reaction mixture was filtered and washed with ether to yield compound 7 as a yellow solid. ESI+ calculated for $C_{16}H_{11}N_4O_2\left[M+H\right]^+$: 291.09 m/z, measured: 291.09 m/z.

N-(3-bromophenyl)-6-nitroquinazolin-4-amine (8): Compound 2 (1.0 g, 4.59 mmol) and compound 5 (0.94 g, 5.50 mmol) were added to acetic acid (15 mL). The reaction mixture was heated at 100 °C for 3 h. After cooling to room temperature, the crystalized reaction mixture was filtered and washed with ether to yield compound 8 as a yellow solid. ESI+ calculated for $C_{14}H_{10}BrN_4O_2\left[M+H\right]^+$: 345.00 (50%) and 347.00 (50%) m/z, measured: 345.01 (50%) and 347.01 (50%).

N^4-(3-ethynylphenyl)quinazoline-4,6-diamine (9): Compound 6 (0.55 g, 1.74 mmol) and Tin(II) chloride dihydrate (1.55 g, 6.87 mmol) were added to ethyl acetate (17 mL). The reaction mixture was heated to reflux for 2 h. After cooling to room temperature, the pH of the solution was adjusted to 9–10 using sodium hydroxide (NaOH). The product was extracted using ethyl acetate and water. The organic layers were collected and combined then washed with brine and water. The solvent was removed under reduced pressure, and column chromatography was performed in dichloromethane:methanol (95:5) to afford compound 9. ESI+ calculated for $C_{16}H_{13}N_4$ $[M+H]^+$: 261.11 m/z, measured: 261.11 m/z.

N^4-(3-chloro-4-fluorophenyl)quinazoline-4,6-diamine (10): Compound 7 (0.5 g, 1.72 mmol) and Tin(II) chloride dihydrate (1.55 g, 6.87 mmol) were added to ethyl acetate (17 mL). The reaction mixture was heated to reflux for 2 h. After cooling to room temperature, the pH of the solution was adjusted to 9–10 using NaOH. The product was extracted using ethyl acetate and water. The organic layers were collected and combined then washed with brine and water. The solvent was removed under reduced pressure. Column chromatography was performed in dichloromethane:methanol (95:5) to afford compound 10. ESI+ calculated for $C_{14}H_{11}ClFN_4\left[M+H\right]^+$: 289.06 m/z, measured: 289.06 m/z.

N^4-(3-bromophenyl)quinazoline-4,6-diamine (11): Compound 8 (0.6 g, 1.74 mmol) and Tin(II) chloride dihydrate (1.55 g, 6.87 mmol) were added to ethyl acetate (17 mL). The reaction mixture was heated to reflux for 2 h. After cooling to room temperature, the pH of the solution was adjusted to 9–10 using NaOH. The product was extracted using ethyl acetate and water. The organic layers were collected and combined then washed with brine and water. The solvent was removed under reduced pressure, and the resulting solid was purified by column chromatography dichloromethane:methanol (95:5) to afford compound 11. ESI+ calculated for $C_{14}H_{12}BrN_4\left[M+H\right]^+$: 315.02 (50%) and 317.02 (50%) m/z, measured: 315.02 (50%) and 317.02 (50%) m/z.

2–(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)-4–(3-(4-((3-ethynylphenyl)amino)quinazolin-6-yl)thioureido)benzoate (Erl-TRITC): Compound 9 (24.4 mg, 0.94 mmol) and tetramethylrhodamine isothiocyanate (TRITC, 50 mg, 0.113 mmol) were dissolved in anhydrous dimethylformamide (2 mL). The solution was stirred at room temperature overnight under N_2 protected from light. The crude product was purified by column chromatography DCM:MeOH (98:2) to give the pure product Erl-TRITC. ^{1}H NMR (DMSO-d_6, 600 MHz): δ_H 9.83 (s, 1H), δ_H 8.62 (s, 1H), δ_H 8.47 (s, 1H), δ_H 8.37 (s, 1H), δ_H 8.28 (s, 1H), δ_H 8.08 (s, 1H), δ_H 7.95–7.91 (m, 3H), δ_H 7.77 (d, 1H, J = 8.8 Hz), δ_H 7.40 (t, 1H, J = 8.0 Hz), δ_H 7.22 (d, 1H, J = 7.5 Hz), δ_H 7.13 (d, 1H, J = 8.3Hz), δ_H 6.50 (m, 7H), δ_H 4.20 (s, 1H), δ_H 2.94 (s, 12H).

4–(3-(4-((3-chloro-4-fluorophenyl)amino)quinazolin-6-yl)thioureido)-2–(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)benzoate (Afa-TRITC): Compound 10 (22.9 mg, 0.079 mmol) and TRITC (40.6 mg, 0.091 mmol) were dissolved in anhydrous dimethylformamide (3 mL). The solution was stirred at room temperature overnight under nitrogen (N_2) protected from light. The crude product was purified by dry loading and performing column chromatography in DCM:MeOH (93:7) to yield the pure product Afa-TRITC. ^{1}H NMR ($CDCl_3$: CD_3OD, 8:2, 600 MHz): δ_H 9.45 (d, 1H, J = 7.2 Hz), δ_H 8.96 (d, 1H, J = 8.3 Hz), δ_H 8.57 (s, 1H), δ_H 8.32 (s, 1H), δ_H 8.21 (s, 1H), δ_H 7.84 (s, 1H), δ_H 7.67 (s, 1H), δ_H 7.52 (d, 1H, J = 8.7 Hz), δ_H 7.25 (m, 3H), δ_H 7.17 (d, 2H, J = 9.4 Hz), δ_H 6.98 (d, 1H, J = 8.3 Hz), δ_H 6.66 (t, 1H, J = 8.7 Hz), δ_H 6.55 (d, 2H, J = 9.4 Hz), δ_H 5.81 (s, 2H), δ_H 2.88 (s, 12H).

4–(3-(4-((3-bromophenyl)amino)quinazolin-6-yl)thioureido)-2–(6-(dimethylamino)-3-(dimethyliminio)-3*H*-xanthen-9-yl)benzoate (PD-TRITC): Compound 11 (26.8 mg, 0.085mmol) and TRITC (45.4 mg, 0.102 mmol) were dissolved in anhydrous dimethylformamide (2 mL). The solution was stirred at room temperature overnight under nitrogen (N_2) protected from light. LCMS analysis indicated the presence of TRITC, an additional 4.5 mg of compound 11 was added and the reaction mixture was stirred overnight. The crude product was purified by column chromatography dichloromethane :methanol:acetonitrile (93:7:1) to afford PD-TRITC. ^{1}H NMR ($CDCl_3$: CD_3OD, 9:1, 400 MHz): δ_H 9.47 (m, 1H), δ_H 8.98 (d, 1H, J = 8.1 Hz), δ_H 8.63 (s, 1H), δ_H 8.35 (s, 1H), δ_H 8.25 (s, 1H), δ_H 7.87 (s, 1H), δ_H 7.71 (m, 1H), δ_H 7.55 (d, 1H, J = 8.9 Hz),δ_H 7.27 (m, 3H), δ_H 7.20 (d, 2H, J = 9.3 Hz), δ_H 7.01 (d, 1H, J = 8.6 Hz), δ_H 6.87 (d, 1H, J = 7.6 Hz), δ_H 6.80 (t, 1H, J = 7.9 Hz), δ_H 6.57 (d, 2H, J = 8.9 Hz), δ_H 5.84 (s, 2H), δ_H 2.91 (s, 12H).

7.3.3 Photophysical Property Characterization

All three fluorescently labeled TKIs (i.e., Erl-TRITC, PD-TRITC, Pd-TRITC) were solubilized in phosphate buffered saline (PBS, pH 7.4) containing 10% DMSO at a concentration of 10 mM. Absorption and fluorescence emission spectra were collected using a SpectraMax M5 spectrometer with a microplate reader (Molecular Devices, Sunnyvale, CA). All absorption spectra were reference corrected. Extinction coefficients for all screened imaging agents were calculated using Beer's Law plots of absorbance versus concentration. Relative quantum yields were calculated using tetramethylrhodamine (TMR) as the reference standard (22). All data reported were the average of multiple measurements.

7.3.4 Binding Affinity Characterization

In vitro binding affinity was evaluated for the fluorescently labeled EGFR TKIs as well as the parent EGFR TKIs. 75 µL of 1 mM samples of Erlotinib, Erl-TRITC, Afatinib, Afa-TRITC, PD-168393, and PD-TRITC were sent to Thermo Fisher Scientific for analysis using their LanthaScreen® Eu Kinase Binding Assay. This assay analyzed purified EGFR kinase (5 nM, ErbB1 d746–750, Thermo Fisher Scientific) bound to a 4X Alexa Fluor® conjugate Tracer 199 (5 nM; K_D 3.2 nM) detected by addition of the Europium (Eu)-labeled antibody Eu-anti-GST (2 nM) prepared in Kinase Buffer. The compounds were screened with a final DMSO concentration of < 1% per well, in a 10-point titration curve over a three-fold serial dilution, starting at a concentration of 10 µM. In addition to the compounds of interest, a 0% displacement control without inhibitor, a 100% displacement control containing the highest concentration of the known inhibitor used in that assay, and a known inhibitor (SB202190 standard curve, 10-point titration, IC_{50} = 165 nM) were loaded onto the plate. Testing compounds were incubated for 60 min at room temperature, and subsequent signal was read on a fluorescence plate reader. Percentage EGFR displacement was reported for each compound by Thermo Fisher Scientific.

7.3.5 MTS Assay Inhibition

Cells were trypsinized, counted, and plated in 96-well glass bottom cell culture plates (Cellvis, Mountain View, CA) at a density of 5×10^4 cells/mL (100 mL/well). Cells were allowed to adhere and equilibrate overnight. Media was then replaced with increasing concentrations of parent drug or TRITC-labeled drug (0, 0.0001, 0.001, 0.01, 0.1, 1.0, 10 mM) and incubated at 37 °C/5% CO_2. After 72 h, 20 µL of the MTS CellTiter 96® AQueous One Solution Cell Proliferation Assay Reagent (Promega, Madison, WI) were added to each well. The cells were incubated for 3 h at 37 °C/5% CO_2. Plates were read for absorbance at 490 nm (SpectraMax M5, Molecular Devices, Sunnyvale, CA). Wells containing only media were used as a control for background absorbance signal and cells permeabilized with 0.1% Triton-X served as a negative control. GraphPad Prism (La Jolla, CA) was used to analyze dose-response curves.

7.3.6 EGFR Receptor Quantification

To quantify EGFR receptor values, cells were trypsinized, counted, and resuspended in phenol-red free (PRF) media. Cells were divided into 5×10^6 aliquots and incubated with 5 mg/mL cetuximab (Eli Lilly, Indianapolis, IN) directly conjugated to Alexa Fluor (AF)488 (Thermo Fisher Scientific, 1:3.4 antibody:fluorophore conjugation ratio) for 30 min at 37 °C/5% CO_2. Cells were then washed 2 × 5 min with PRF media. The cells were pelleted and then resuspended in fresh PBS prior to analysis on an LSR Fortessa flow cytometer (Becton, Dickinson and Company, Franklin Lakes, NJ). The flow cytometer was configured with a 488–1 (525/50) Cy2 channel filter to detect AF488. A minimum of 1×10^5 cells were counted for each sample. To quantify EGFR receptor number, Quantum™ Simply Cellular external standard beads (Bangs Laboratories, Inc., Fishers, IN) were used.

7.3.7 Live-Cell Imaging Studies

Cells were plated in 96-well glass bottom cell culture plate (Cellvis) at a density of 1×10^5 cells/mL (100 mL/well) in PRF media and allowed to adhere and equilibrate overnight

at 37 °C/5% CO_2. Approximately 16 h after plating, cells were stained with AF488-conjugated cetuximab at a final concentration of 5 mg/mL in PRF media for 45 min at 37 °C/5% CO_2. The cells were rinsed 3x in fresh D-PBS (Thermo Fisher Scientific), followed by a 1 h incubation with the respective EGFR TKI-TRITC reagents at 37 °C/5% CO_2. Reagent media was then removed and cells were incubated with Hoechst 33342 nuclear staining solution (Thermo Fisher Scientific) at a final concentration of 1 mg/mL for 5 min. Wells were then immediately replaced with fresh PRF media and imaged on a Zeiss AxioObserver outfitted with a Yokogawa CSU-X1 spinning disk confocal unit with a 20x 0.8 PlanApo objective using the following exposure times as well as excitation (x) and emission (m) filter combinations for image collection: UV—1500 ms, 392.5/12.5x, 445/35m; AF488—2500 ms, 470/40x, 525/50m; AF555—3000 ms, 550/25x, 605/70m. Following image collection, the 96-well plate was removed from the stage and the media was replaced with fresh PRF media and immediately reimaged using equivalent settings. This process was repeated for a total of three media replacements. Cell viability was maintained throughout the imaging process via a temperature- and humidity-controlled stage.

7.3.8 Phospho-EGFR Imaging Studies

Immediately after the completion of live-cell imaging studies, the cells were fixed with 4% paraformaldehyde (PFA, Electron Microscopy Sciences, Hatfield, PA), permeabilized with 0.3% Triton-X for 5 min and washed 3 × 5 min in PBS. A 30-min blocking step with 3% bovine serum albumin (BSA, bioWORLD, Dublin, OH) was directly followed by primary monoclonal phospho-EGFR antibody application (Tyr1068) (D7A5) (Cell Signaling Technology, Danvers, MA) at a 1:100 dilution that was visualized with a donkey anti-rabbit secondary antibody (Jackson ImmunoResearch, West Grove, PA) conjugated to AF647 (Thermo Fisher Scientific) at a final staining concentration of 2 mg/mL. The cells were washed and post-fixed with 4% PFA for 10 min, prior to imaging on the Zeiss AxioObserver/Yokogawa spinning disk confocal microscope, with the addition of the following filter combination: AF647—3000 ms, 640/30x, 690/50m. Mean cellular fluorescence intensity was quantified on individual regions of interest (ROIs) using ImageJ. A minimum $n = 3$ ROIs were used for EGFR TKI-TRITC fluorescence signal analysis.

7.4 Results

7.4.1 Design and Synthesis of Fluorophore-TKI Conjugates

Three fluorescently labeled EGFR TKIs, Erlotinib, Afatinib, and PD168393 imaging agents, were synthesized, where fluorophore labeling positions were selected based on the protein crystal structure of TKI bound to EGFR (23). The synthesis of these imaging agents was completed in four steps (Figure 7.1). In brief, the commercially available starting material, 2-amino-5-nitrobenzonitrile (compound 1) was treated with *N,N*-dimethylformamide dimethylacetal in dioxane at reflux, affording intermediate 2 in quantitative yield.

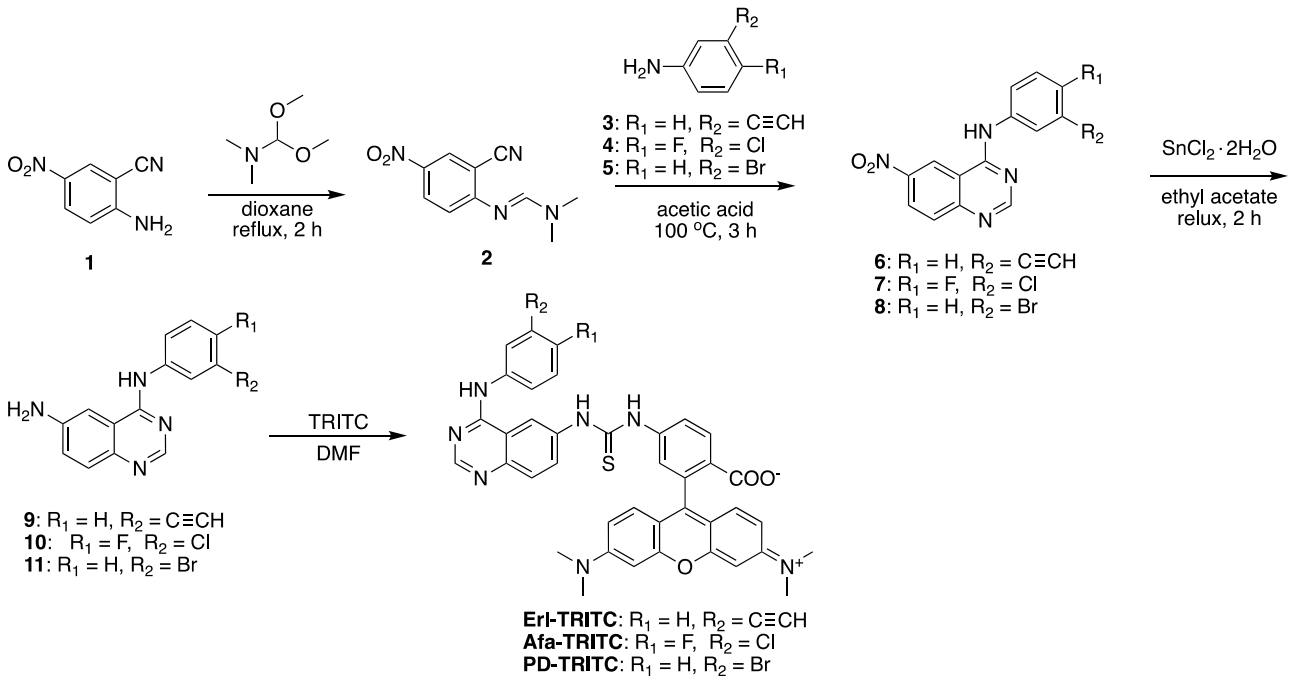


FIGURE 7.1 *Synthesis of fluorescently labeled TKIs for EGFR-specific labeling.* Synthetic scheme of the fluorescently labeled EGFR TKIs, including those modified from Erlotinib, Afatinib, and PD168393, resulted in three imaging agents, Erl-TRITC, Afa-TRITC, and PD-TRITC. The linker position for each imaging agent was carefully selected based on the co-crystal structure of these TKIs with the EGFR tyrosine kinase domain, guiding optimal design for drug labeling.

Subsequent electrophilic cyclization reactions followed by Dimroth rearrangement reaction with respective aniline substrates (compounds 3, 4, and 5), produced the anilinoquinazoline intermediates. Products from these first two steps were sufficiently pure and used in the next step without further purification. After the nitro reduction reaction, primary amine products 9, 10, and 11 were generated, and then conjugated with TRITC to yield the final fluorophore-labeled EGFR TKIs, Erl-TRITC, Afa-TRITC, and PD-TRITC as dark red solids.

7.4.2 Photophysical Properties of Fluorophore-TKI Conjugates

Photophysical property characterization was completed for all three imaging agents and unconjugated parent fluorophore (i.e., TMR) to evaluate how TKI conjugation impacted optical performance (Figure 7.2a). The absorption and emission spectra were collected in phosphate buffered saline solution containing 10% DMSO. Both Erl-TRITC and PD-TRITC exhibited absorption and emission maximum wavelength with a red shift of < 10 nm compared to their parent fluorophore, TMR. Surprisingly, Afa-TRITC showed an unexpected blue shift of 27 nm in absorption with unchanged emission. Despite these changes, the Stokes shifts of Erl-TRITC and PD-TRITC were comparable to that of TMR, with Afa-TRITC displaying approximately double the Stokes shift of TMR. The quantum yield of these three imaging agents were similar, ranging from 0.6% to 1.7%, while the extinction coefficient value of Afa-TRITC ($\varepsilon_{\text{PBS buffer}} = 67{,}200\ M^{-1}cm^{-1}$) was much higher than Erl-TRITC ($\varepsilon_{\text{PBS buffer}} = 24{,}700\ M^{-1}cm^{-1}$) and PD-TRITC ($\varepsilon_{\text{PBS buffer}} = 17{,}900\ M^{-1}cm^{-1}$). These changes are largely due to the TKI conjugation altering the spirolactone open(fluorescent)/close(non-fluorescent) equilibria, which are well-studied properties of the rhodamine-based fluorophores (24–26). The overall brightness of all three imaging agents were similar, with values ranging from 0.3 to 0.4 (Figure 7.2b).

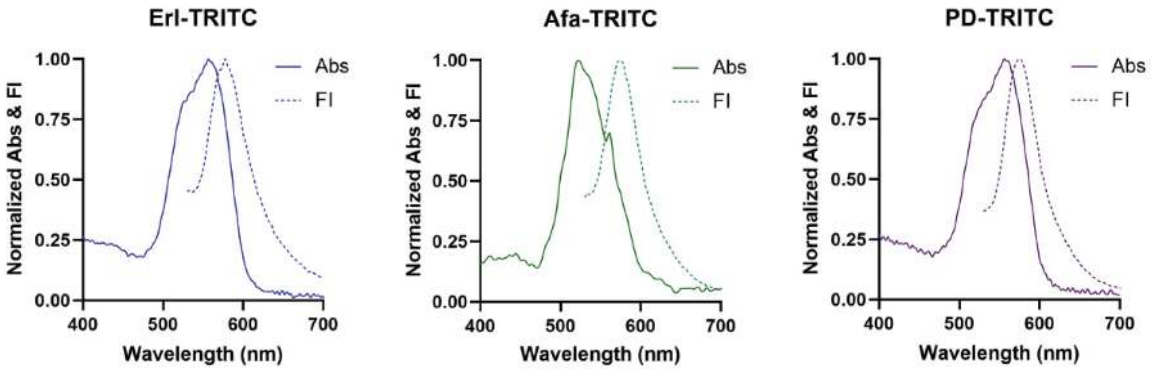
7.4.3 Assessment of EGFR Protein Binding Affinity and Cytotoxicity

To assess if fluorophore conjugation to the TKIs affected protein-drug interactions, we quantified the binding affinities of the imaging agents on purified EGFR protein. Only a small decrease in binding affinity was observed for the fluorescently labeled TKIs compared to their respective parent EGFR TKIs, suggesting that the conjugation did not significantly impair the intended binding of the TKIs (Figure 7.3a). To further confirm that the TRITC-modified imaging agents behaved in a similar manner to the parent drug, cytotoxicity tests to obtain IC_{50} for the imaging agents, unmodified TKIs, and the parent fluorophore were completed on a canonical EGFR mutation positive cell line (HCC827). Quantified IC_{50} values ranged from 380 to 570 nM for the three imaging agents, which were higher than their unlabeled TKI counterparts with IC_{50} values ranging from 18 to 90 nM. However, they retained a substantially higher affinity for the protein target than did the parent fluorophore control (i.e., TMR), suggesting the binding affinity of these imaging agents was primarily derived from the drug component rather than the fluorophore (Figure 7.3b).

7.4.4 Live-Cell Imaging of EGFR Protein with Fluorophore-TKI Conjugates

To demonstrate the feasibility and compatibility of our fluorophore-TKI conjugates for live-cell imaging, we

(a)



(b)

Compound	MW (g/mol)	λ_{abs} (nm)	ε ($M^{-1}cm^{-1}$)	$FWHM_{abs}$ (nm)	λ_{em} (nm)	Stokes Shift (nm)	$FWHM_{em}$ (nm)	Φ (%)	Brightness ($M^{-1}cm^{-1}$)
Erl-TRITC	703.82	556	24,700	81	577	21	68	1.2	0.30
Afa-TRITC	732.23	522	67,200	71	573	51	59	0.6	0.40
PD-TRITC	758.70	556	17,900	81	576	20	56	1.7	0.30

FIGURE 7.2 *Photophysical properties of TKI-TRITC derivatives.* (a) Normalized absorption (solid line) and fluorescence emission (dashed line) spectra of Erl-TRITC (left), Afa-TRITC (middle), and PD-TRITC (right). (b) Tabulated spectral properties of three TKI-TRITC derivatives. Maxima extinction coefficient (ε), quantum yield (Ø), and brightness were measured in PBS containing 10% of DMSO.

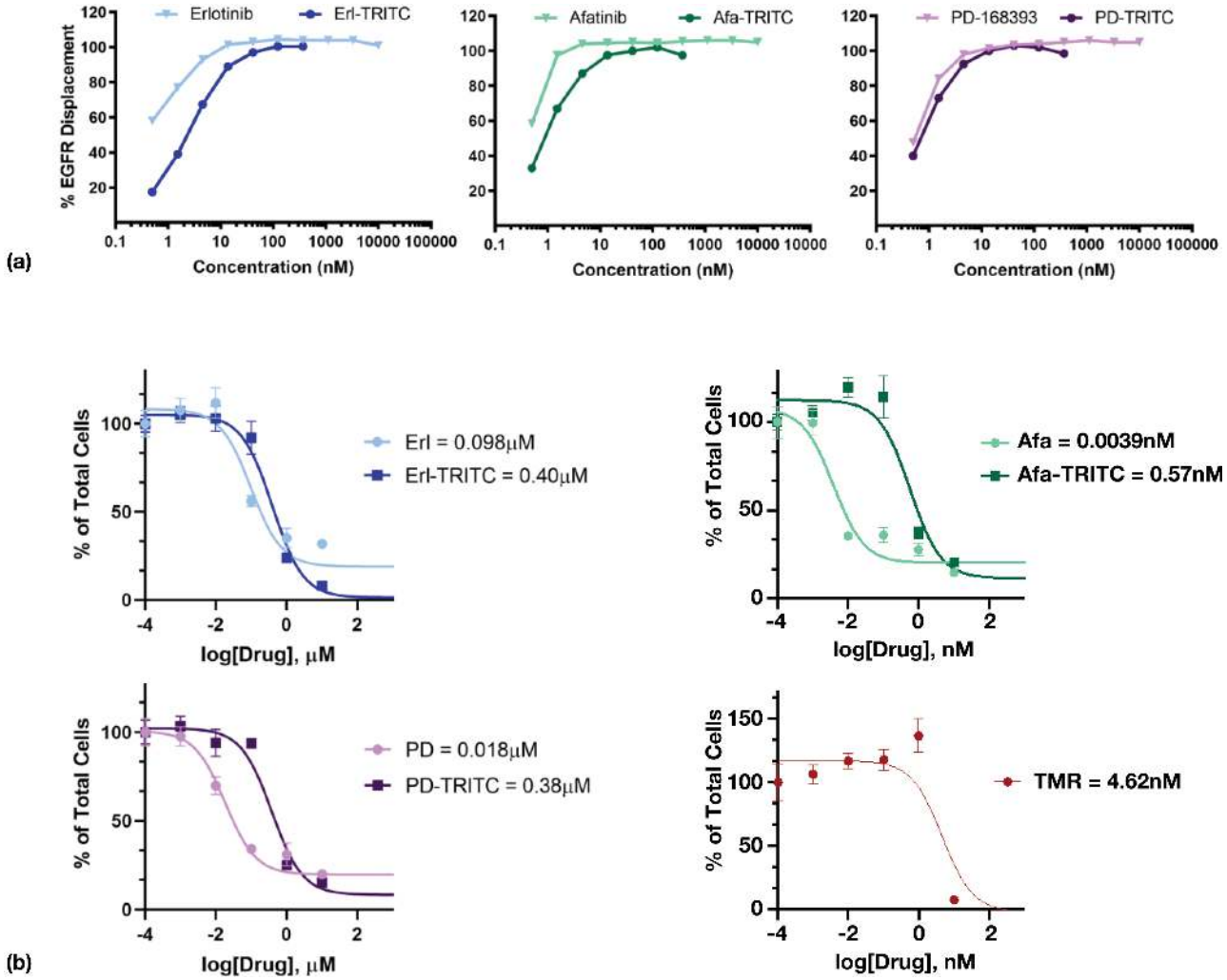


FIGURE 7.3 *EGFR protein binding affinity and cytotoxicity evaluation of the TKI-TRITC derivatives.* (a) Binding affinity assessment on purified EGFR protein, where the percentage of displacement by the parent drug (i.e., Erlotinib, Afatinib, or PD-168393) was compared to the fluorophore-labeled version (i.e., Erl-TRITC, Afa-TRITC, or PD-TRITC) respectively. All three imaging agents showed a minor decrease in binding affinity compared to unlabeled parent TKIs. (b) Cytotoxicity was evaluated using an MTS assay to compare Erlotinib versus Erl-TRITC, Afatinib versus Afa-TRITC, PD-168393 versus PD-TRITC, and unmodified TRITC (i.e., TMR) in HCC827 cells. Data presented as mean ± SD from $n = 3$ independent experiments. Cellular comparison of IC_{50} values (listed in the key for each tested agent) suggested the TKI-TRITC imaging agents retained the targeting capabilities of the parent TKIs.

conducted *in vitro* live-cell imaging studies. Three cell lines with increasing endogenous EGFR expression levels, as confirmed by antibody-receptor quantification, were utilized: A431 > HCC827 > U2OS (Figure 7.4a). Initial live-cell imaging studies showed cellular fluorescence intensities were consistent with EGFR receptor number for all three imaging agents, where the A431 cell line showed the highest uptake, followed by HCC827 cells, while U2OS had a low signal (Figure 7.4b). Moreover, cell co-staining studies with EGFR antibody cetuximab directly conjugated to AF488 ($EGFR_{AB}$ AF488) confirmed the correlation between TKI-TRITC signal and EGFR expression (Figure 7.4c). Finally, co-staining for EGFR and phospho-EGFR (pEGFR) illustrated that the TKI-TRITC derivatives localized in patterns consistent with known drug binding sites and EGFR activity (Figure 7.4d), validating the utility of these imaging agents for the cancer cell identification and quantification through intracellular imaging of EGFR in live cells under physiological conditions.

7.5 Discussion

In this chapter, we described a novel imaging approach for future cancer detection and screening, by utilizing fluorescently labeled SOC drugs for intracellular protein target localization in live cells. Traditional imaging techniques, such as immunohistochemistry and immunofluorescence-based approach, use antibodies that are bulky, expensive, and have long pharmacokinetic profiles. Additionally, these antibody-based approaches require cell destructive fixation and permeation for intracellular target imaging and quantification. In

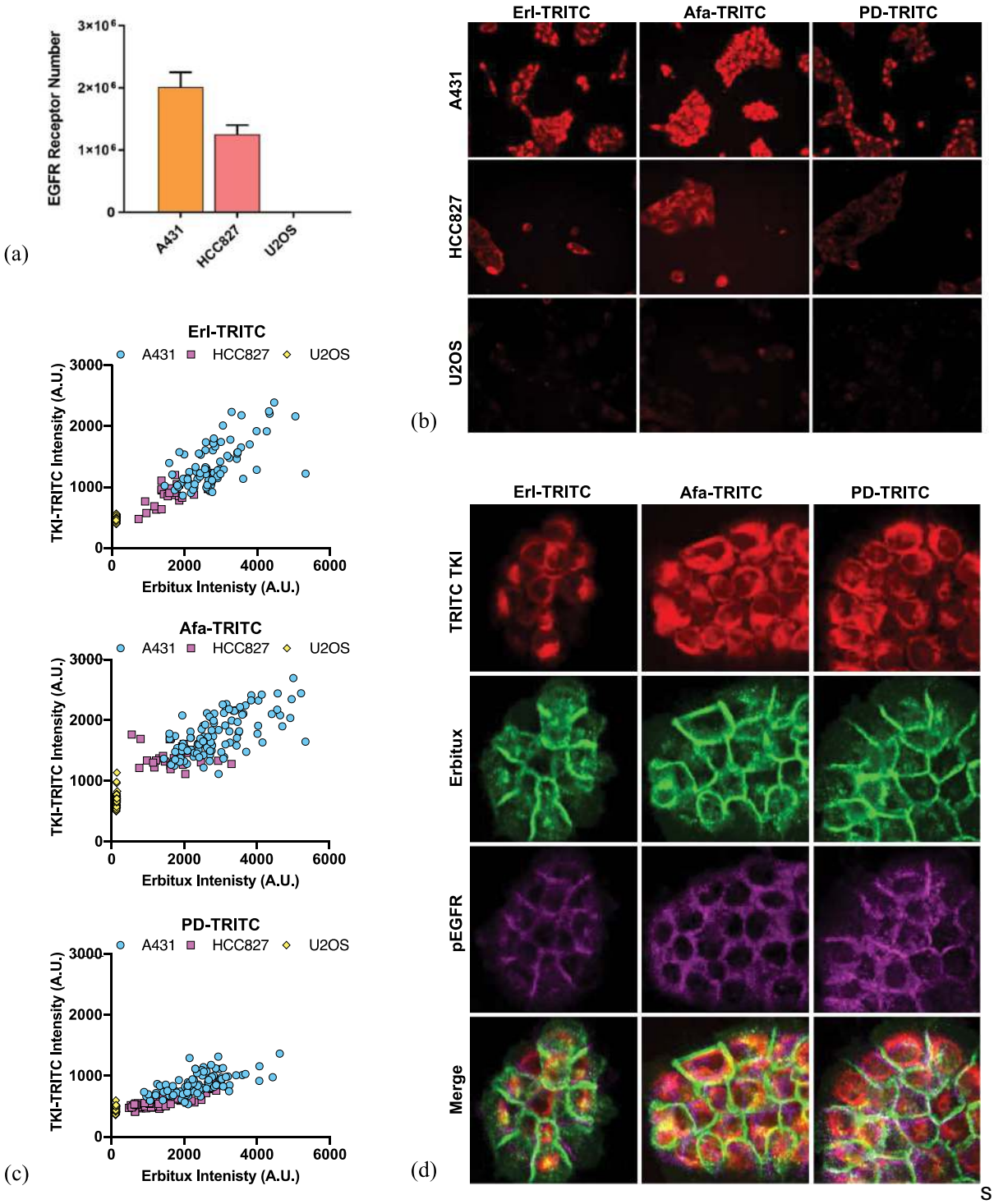


FIGURE 7.4 *Quantitative assessment of fluorophore-TKI conjugates uptake and distribution in live cells.* (a) Comparison of EGFR expression levels in A431, HCC827, and U2OS cells via flow cytometry. (b) Live-cell staining comparison of the three cell lines with each of the TKI-TRITC derivatives: Erl-TRITC (left column), Afa-TRITC (middle column), and PD-TRITC (right column). (c) Scatterplots show a strong correlation between TKI-TRITC intensities and $EGFR_{AB}$ signals across all three imaging agents in A431 (blue circle), HCC827 (magenta square), and U2OS (yellow diamond). (d) Representative fluorescence images of intracellular localization for TKI-TRITC derivatives (red), $EGFR_{AB}$-AF488 staining (green), and phospho-EGFR staining (magenta). Merged images for all three stains are shown in the bottom row.

contrast, small-molecule-based probes offer several advantages, including low cost, high stability, easy modification, and ready cell permeability (21, 25, 27–29). These features make small-molecule-based imaging agents particularly suitable for early detection and screening via imaging the cancer-specific biomarkers exhibited in both the extracellular and the intracellular spaces of cancer cells.

Our proof-of-concept study involved the design, synthesis, and validation of three fluorescently labeled EGFR TKIs, including Erl-TRITC, Afa-TRITC, and PD-TRITC. The photophysical characterization of these imaging agents indicated an impact on their optical properties post-conjugation. Nonetheless, these imaging agents are compatible with many standard commercial fluorescence imagers. Findings from protein binding studies confirmed that the modifications minimally altered the binding affinity of the TKIs after fluorophore labeling and maintained their intended binding capabilities for protein target imaging and quantification. Importantly, even though the IC_{50} values of the fluorescently labeled TKIs were higher compared to their parent drugs, they were still highly effective at EGFR-targeted binding and imaging. Further *in vitro* live-cell imaging studies conducted across cell lines with varied EGFR expression levels confirmed the feasibility of using these fluorescent TKIs for live cell imaging under physiological conditions. Importantly, quantitative assessments revealed a strong correlation between TKI-TRITC intensities and EGFR expression, and co-staining experiments validated the TKIs' specific intracellular localization, mirroring that of intracellular EGFR and pEGFR, substantiating the specificity and utility of these probes for the detection and imaging of EGFR-positive cancer cells.

In summary, these results highlight the potential of fluorescently labeled TKIs as innovative tools for cancer-specific imaging in live cells that are compatible with liquid biopsies. Such advancement opens new opportunities to transform cancer early detection and facilitate routine monitoring of treatment response.

7.6 Future Trends

We envision this imaging toolbox to serve as imaging probes to efficiently detect cancer early. Future work will involve extending this synthetic pipeline to create a library of small molecule TKIs with optimized fluorophores and linkers to improve target binding sensitivity and specificity. A particular avenue involves the incorporation of an enzyme-recognizing warhead as the linker between the fluorophore and the TKI. This approach will enable covalent labeling of cancer cells with retained fluorescence signal while background cells (i.e., non-cancer cells) demonstrate low to no signal, thereby enhancing contrast and improving diagnostic accuracy and sensitivity. Additionally, the development of activatable probes represents another significant avenue for cancer screening and detection using TKI-based imaging agents. Functionally, these probes can be designed to remain non-fluorescent, providing a black background until they are activated by cancer-associated protein biomarkers. Such framework would drastically reduce background signal, improving the detection sensitivity range and enhancing the precision of cancer diagnosis. Finally, this study suggests that it is feasible to extend this technology to other molecular inhibitors as the targeting moiety for subsequent fluorescence labeling, cancer detection, and screening. By integrating these novel imaging probes into the analysis of liquid biopsy samples, we anticipate a transformative impact on the landscape of cancer early detection and screening platforms.

7.7 Acknowledgement

This work was supported by funding from the Collins Medical Trust (Wang) and the OHSU Foundation (Gibbs).

REFERENCES

1. Organization WH. *Cancer*. 2020.
2. Christofori G. New signals from the invasive front. *Nature*. 2006;441(7092):444–50.
3. Alix-Panabières C, Pantel K. Liquid biopsy: from discovery to clinical application. *Cancer Discovery*. 2021;11(4):858–73.
4. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic—implementation issues and future challenges. *Nature Reviews Clinical Oncology*. 2021;18(5):297–312.
5. Nikanjam M, Kato S, Kurzrock R. Liquid biopsy: current technology and clinical applications. *Journal of Hematology & Oncology*. 2022;15(1):131.
6. Gast CE, Silk AD, Zarour L, Riegler L, Burkhart JG, Gustafson KT, et al. Cell fusion potentiates tumor heterogeneity and reveals circulating hybrid cells that correlate with stage and survival. *Science Advances*. 2018;4(9):eaat7828.
7. Powell AE, Anderson EC, Davies PS, Silk AD, Pelz C, Impey S, et al. Fusion between Intestinal epithelial cells and macrophages in a cancer context results in nuclear reprogramming. *Cancer Research*. 2011;71(4):1497–505.
8. Rizvi AZ, Swain JR, Davies PS, Bailey AS, Decker AD, Willenbring H, et al. Bone marrow-derived cells fuse with normal and transformed intestinal stem cells. *Proceedings of the National Academy of Sciences of the United States of America*. 2006;103(16):6321–5.
9. Downward J, Yarden Y, Mayes E, Scrace G, Totty N, Stockwell P, et al. Close similarity of epidermal growth factor receptor and v-erb-B oncogene protein sequences. *Nature*. 1984;307(5951):521–7.
10. Eckhart W, Hutchinson MA, Hunter T. An activity phosphorylating tyrosine in polyoma T antigen immunoprecipitates. *Cell*. 1979;18(4):925–33.
11. Hunter T, Sefton BM. Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proceedings of the National Academy of Sciences of the United States of America*. 1980;77(3):1311–5.
12. Kawamoto T, Sato JD, Le A, Polikoff J, Sato GH, Mendelsohn J. Growth stimulation of A431 cells by epidermal growth factor: identification of high-affinity receptors for epidermal growth factor by an anti-receptor monoclonal antibody. *Proceedings of the National Academy of Sciences of the United States of America*. 1983;80(5):1337–41.

13. Riedel H, Dull TJ, Schlessinger J, Ullrich A. A chimaeric receptor allows insulin to stimulate tyrosine kinase activity of epidermal growth factor receptor. *Nature*. 1986;324(6092):68–70.
14. Schlessinger J. Signal transduction by allosteric receptor oligomerization. *Trends in Biochemical Sciences*. 1988;13(11):443–7.
15. Brand TM, Iida M, Wheeler DL. Molecular mechanisms of resistance to the EGFR monoclonal antibody cetuximab. *Cancer Biol Ther*. 2011;11(9):777–92.
16. Cohen P, Cross D, Jänne PA. Kinase drug discovery 20 years after imatinib: progress and future directions. *Nature Reviews Drug Discovery*. 2021;20(7):551–69.
17. Abdeldayem A, Raouf YS, Constantinescu SN, Moriggl R, Gunning PT. Advances in covalent kinase inhibitors. *Chemical Society Reviews*. 2020;49(9):2617–87.
18. Singh J. The ascension of targeted covalent inhibitors. *Journal of Medicinal Chemistry*. 2022 Apr 28;65(8):5886–5901. doi: 10.1021/acs.jmedchem.1c02134. Epub 2022 Apr 19.
19. Boike L, Henning NJ, Nomura DK. Advances in covalent drug discovery. *Nature Reviews Drug Discovery*. 2022 Dec;21(12):881–898. doi: 10.1038/s41573-022-00542-z. Epub 2022 Aug 25. PMID: 36008483; PMCID: PMC9403961.
20. Marvania B, Lee P-C, Chaniyara R, Dong H, Suman S, Kakadiya R, et al. Design, synthesis and antitumor evaluation of phenyl N-mustard-quinazoline conjugates. *Bioorganic & Medicinal Chemistry*. 2011;19(6):1987–98.
21. Wang LG, Montaño AR, Combs JR, McMahon NP, Solanki A, Gomes MM, et al. OregonFluor enables quantitative intracellular paired agent imaging to assess drug target availability in live cells and tissues. *Nature Chemistry*. 2023;15(5):729–39.
22. Johnson I. Review: Fluorescent probes for living cells. *The Histochemical Journal*. 1998;30(3):123–40.
23. Johnson LN. Protein kinase inhibitors: contributions from structure to clinical compounds. *Quarterly Reviews of Biophysics*. 2009;42(1):1–40.
24. Beija M, Afonso CAM, Martinho JMG. Synthesis and applications of Rhodamine derivatives as fluorescent probes. *Chemical Society Reviews*. 2009;38(8): 2410–33.
25. Dai M, Yang YJ, Sarkar S, Ahn KH. Strategies to convert organic fluorophores into red/near-infrared emitting analogues and their utilization in bioimaging probes. *Chemical Society Reviews*. 2023;52(18):6344–58.
26. Wang L, Sibrian-Vazquez M, Escobedo JO, Wang J, Moore RG, Strongin RM. Spiroguanidine rhodamines as fluorogenic probes for lysophosphatidic acid. *Chemical Communications*. 2015;51(9):1697–700.
27. Wang L, Sibrian-Vazquez M, Escobedo JO, Wang J, Moore RG, Strongin RM. Spiroguanidine rhodamines as fluorogenic probes for lysophosphatidic acid. *Chemical Communications (Cambridge, England)*. 2015;51(9):1697–700.
28. Wang L, Barth CW, Sibrian-Vazquez M, Escobedo JO, Lowry M, Muschler J, et al. Far-red and near-infrared seminaphthofluorophores for targeted pancreatic cancer imaging. *ACS Omega*. 2017;2(1):154–63.
29. Wang LG, Munhenzva I, Sibrian-Vazquez M, Escobedo JO, Kitts CH, Fronczek FR, et al. Altering fundamental trends in the emission of xanthene dyes. *The Journal of Organic Chemistry*. 2019;84(5):2585–95.

8

Real-Time Cancer Metabolism Detection Using a Nanocoil Integrated Hyperpolarized Micromagnetic Resonance Spectrometer

Thasin Peyear, Roberta Pigliapochi, and Kayvan R. Keshari

8.1 Introduction

Metabolic reprogramming has been implicated in a variety of diseases, including cancer [1]. A hallmark trait of cancer is the "Warburg effect" [2], which promotes glucose uptake and preferentially produces lactate, even in the presence of oxygen [3]. This trait enables cancer cells to meet their metabolic requirements, grow uncontrollably [3], and avoid immune detection in the tumor microenvironment [4]. Studying the glycolytic rate, also known as "flux," may provide insight into the metabolic function of cancer cells and how they can respond to changes in their environment in real time, such as nutrient availability or drug use [4, 5]. Therefore, the real-time metabolic profile characterization is required to further understand their function. The growing interest in metabolic dysfunction has prompted the development of tools for metabolic research, including hyperpolarized (HP) nuclear magnetic resonance (NMR) spectroscopy, which is used to study the kinetics of metabolic processes [6–9].

The development of nuclear spin hyperpolarization in combination with NMR spectroscopy enables real-time metabolic studies in living cells. Dynamic nuclear polarization (DNP) is the most frequent method for hyperpolarization, and it may increase the NMR signal by 10^4 times [10], enabling the study of time-resolved conversion rates of hyperpolarized substrates [11]. [1-^{13}C]Pyruvate is a commonly used probe for studying the Warburg effect because of its quick cellular absorption and importance in metabolism [12, 13] and because it is easily hyperpolarized; [1-^{13}C]pyruvate has a relatively strong and long-lasting hyperpolarized signal with respect to its polarization (%) and T_1 relaxation [14–18]. Traditional DNP-NMR is a challenging method for studying cells due to low sensitivity and requiring a large number of cells (~10^7) per experiment [6]. As a result, we developed and improved upon a novel platform for measuring metabolic flux in living cells in real time without interfering with chemical or biological processes [6, 7, 9].

This chapter introduces a rapid and sensitive NMR platform that is ideal for small-scale metabolic flux analysis (Figure 8.1). This iteration builds on the previously designed microcoil platform by reducing the volume and number of cells needed to measure metabolic flux in multiple mass-limited cell suspension samples, resulting in a nanocoil-based hyperpolarized micromagnetic resonance spectroscopy (HMRS) platform [8]. The nanocoil-based HMRS has a dual-tuned nanocoil circuit that can process small amounts of biological material, allowing for sensitive ^{13}C and ^{1}H molecule detection with only a small number of cells (~10^4). The HMRS system is 10^3 times more sensitive than conventional DNP-NMR methods for determining glycolytic flux, improves experimental efficiency, and allows for more repetitions and high throughput while using fewer samples [6, 8]. This system works with a wide variety of cells and is especially useful when working with limited samples, such as primary cells or rare cell types.

8.2 Materials

8.2.1 Nanocoil

- Sylgard™ 184 silicone encapsulant, clear, 0.5 kg kit (Dow Corning).
  - Elastomer base
  - Curing agent
- 44 AWG (American wire gauge) magnet wire (Remington).
- 27 AWG Needle (Becton, Dickinson and Company).
- Vacuum desiccator.
- Gorilla super glue (or any all-purpose glue).
- Any small plastic disposable container such as a hexagonal weighing dish or 5 mL disposable polystyrene beaker.
- Oven.
- Razor blade.
- PTFE tube with an outer diameter of 0.6 mm and an inner diameter of 0.3 mm (Sigma-Aldrich).

8.2.2 Nanocoil Circuit

- Custom printed circuit board (or any non-magnetic soldering board).
- Non-magnetic variable and static capacitors, and SMA connector.
- Soldering iron and soldering wire.
- Non-magnetic SMA connectors (Digikey).
- Network analyzer (Keysight).

DOI: 10.1201/9781003449942-10

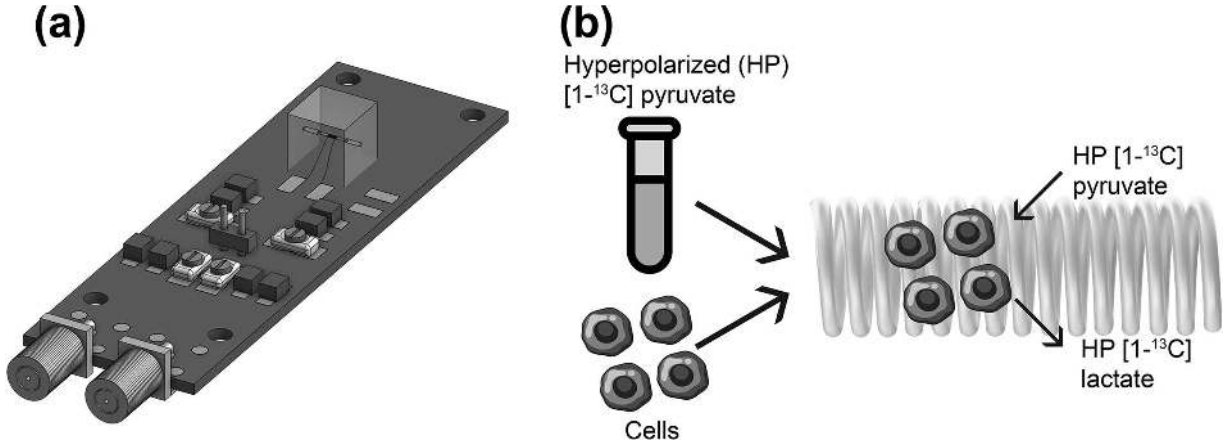


FIGURE 8.1 The nanocoil-based HMRS platform. (a) An illustration of the nanocoil-based HMRS platform, with the nanocoil embedded in the transparent block. (b) In the platform cells are mixed with hyperpolarized [1-^{13}C]pyruvate and placed in the nanocoil, where we monitor the conversion of hyperpolarized [1-^{13}C]pyruvate to hyperpolarized [1-^{13}C]lactate. Adapted from reference [8].

8.2.3 Hyperpolarization

- [1-^{13}C]Pyruvic acid (Sigma-Aldrich).
- AH-111501 radical (GE).
- SPINlab polarizer (GE).
- 10 N sodium hydroxide solution (Fisher Scientific).
- Deuterium oxide (D_2O) (Sigma-Aldrich).
- Trizma hydrochloride solution (TRIS) (Sigma-Aldrich).
- Ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich).
- Deuterium chloride (Sigma-Aldrich).

8.2.4 Hyperpolarized Cell Experiment

- MOLM13 and NOMO1 (DSMZ) acute myeloid leukemia cell lines.
- RPMI-1640 (Thermo Fisher) medium containing 10% fetal bovine serum (Thermo Fisher) and 1% penicillin-streptomycin, or the appropriate medium for your favorite cell lines.
- ABT199 (Cayman Chemical).
- Incubator set to 37 °C and 5% CO_2.
- Centrifuge.
- MRI scanner (preclinical or clinical version).
- 1 T Spinsolve spectrometer (Magritek) or any NMR instrument.

8.3 Methods

8.3.1 Nanocoil Fabrication

1. The 44 AWG magnet wire is wound around the 27 AWG needle 15 times. To achieve optimal performance, the winding should be as tight as possible (Figure 8.2a). This can be accomplished by twisting the wire after it has been coiled and/or applying a small amount of glue to the magnet wire (not between the wire and the needle). This will result in an 80 nL nanocoil.
2. To embed the coil in the polydimethylsiloxane (PDMS) block, mix the two components of the silicone elastomer kit, elastomer base and curing agent, at a 10:1 ratio, respectively.
3. Place the PDMS solution in the vacuum desiccator for about 30–45 min to remove all the bubbles produced by mixing.
4. Place the nanocoil in a disposable plastic container with support on the ends of the needle to suspend the nanocoil.
5. Add the PDMS solution into the container until the nanocoil is completely submerged, then vacuum for another 15–20 minutes to remove any remaining bubbles.
6. Place the container in an oven at 55 °C overnight to cure/harden the mixture.
   a. Instead of using support to suspend the nanocoil, cure a layer of PDMS, then place the nanocoil on top and embed it with a second layer of PDMS.
   b. Make sure there are no bubbles near or around the nanocoil because they will disrupt magnetic field homogeneity (Figure 8.2c).
7. Cut the PDMS to the appropriate size (Figure 8.2b).
8. Slowly remove the 27 AWG needle, revealing the microfluidic channel that goes through the nanocoil.
9. Insert tubing into both sides of the nanocoil microfluidic channels allowing for easy loading.
   a. Incorporate a reservoir and a microfluidic device for easy loading. Alternatively, insulin syringes may also be used to load samples.

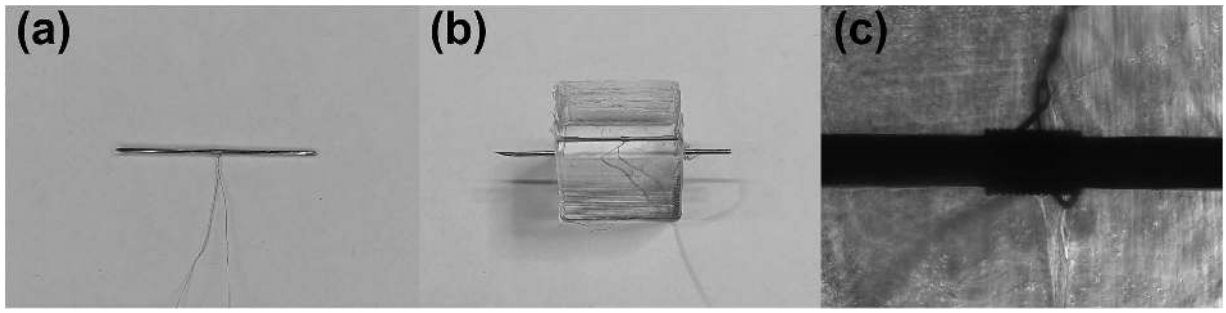


FIGURE 8.2 Fabrication of the nanocoil in the PDMS block. (a) The 44 AWG magnet wire is wound 15 times around the 27 AWG needle, the wire is then twisted five times to maintain the winding. (b) The PDMS block is cut to the appropriate size after being cured overnight at 55 °C. (c) 4X microscopy demonstrates that the windings of the nanocoil are maintained and there are no bubbles in the PDMS block.

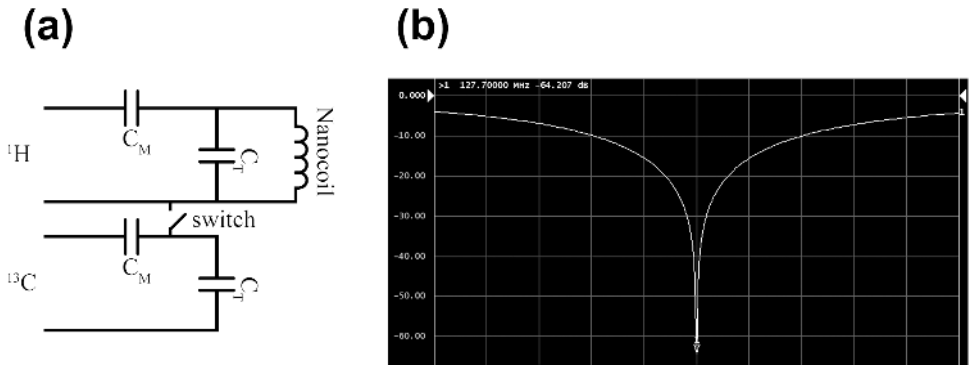


FIGURE 8.3 Design of the nanocoil resonance circuit. (a) Resonance circuit schematic where C_M is the net capacitance of matching capacitors and C_T is the net capacitance of tuning capacitors. The circuit also contains a switch that can be set to either ^{1}H or ^{13}C. (b) The network analyzer shows that the ^{1}H is tuned and matched to 127.7 MHz (^{1}H resonance frequency at 3 T).

8.3.2 Nanocoil Circuit

1. We have printed a custom printed circuit board (PCB) following the resonance circuit in Figures 8.1a and 8.3a.
   a. Alternatively, a nonmagnetic soldering board can be used.
2. The nanocoil is the inductor of the circuit diagram and can be soldered onto the PCB board (Figure 8.3a).
3. A combination of static and variable capacitors are soldered in parallel to the tune (C_T) and match (C_M) on the PCB's ^{1}H side. The network analyzer is used to tune to the ^{1}H resonance frequency of the 3 T MRI (127.7 MHz).
4. The SMA connector is added to the end of the ^{1}H circuit.
   a. For connecting to the MRI.
5. A switch is added to connect the circuit, and an additional combination of static and variable capacitors are soldered in parallel to the tune (C_T) and match (C_M) on the ^{13}C side of the PCB. The network analyzer is used to tune to the ^{13}C resonance frequency of the 3 T MRI (32.1 MHz).
   a. When the switch connects the two circuits, only ^{13}C spectroscopy is allowed.
6. The SMA connector is added to the end of the ^{13}C circuit.

8.3.3 Hyperpolarized NMR Experiment with Cells

1. SPINlab polarizer is used to prepare our sample using DNP. This method uses a fluid path to hyperpolarize the sample before mixing it with the dissolution buffer.
2. To prepare the hyperpolarized pyruvate sample, dissolve AH-111501 radical in neat [1-^{13}C]pyruvate and thoroughly mix it to a final concentration of 15 mM radical.
3. To prepare the dissolution buffer, 100 mM TRIS and 1 mM EDTA are dissolved in D_2O and the pH is reduced to physiological conditions using deuterium chloride.
4. Load 20 µL of pyruvate sample and 6.5 mL of dissolution buffer into the fluid path of the SPINlab polarizer.
5. Load the fluid path and hyperpolarize the sample at 5 T in 0.8 K for at least 1.5 hours with microwave irradiation at 139.88 GHz (the optimal frequency for our polarizer), and heat the dissolution buffer to 130 °C.
6. We expected a dissolution of about 3 mL out, resulting in a hyperpolarized 100 mM pyruvate solution. The pH of the dissolution is adjusted to around 7.4 by adding ~25.3 µL sodium hydroxide.

7. Prepare the cells to the appropriate concentration and incubate them at 37 °C for 5 min before the experiment.
   a. Here we resuspended $2.5 \cdot 10^7$ cells/mL in a volume of 100 µL, equivalent to 2000 cells per 80 nL.
8. Prepare the nanocoil PCB by connecting it to the MRI and placing it in the magnetic field isocenter.
9. Load the nanocoil with a sample set of cells (around the same cell density), and then adjust the wobble and shim accordingly in the ^{1}H channel.
   a. Wash the nanocoil thoroughly with water to remove the sample cells and any residue.
10. Switch to the ^{13}C channel and prepare your ^{13}C-NMR signal acquisition protocol.
11. When the cells are ready, start the dissolution process and sync a timer with the time of dissolution.
    a. Dissolution buffer is rapidly mixed with the hyperpolarized pyruvate, flushed through the fluid path and into a sample cup outside of the polarizer, and neutralized with sodium hydroxide.
12. A small aliquot of hyperpolarized pyruvate is quickly taken to the MRI and mixed with the cell suspension, resulting in an expected final concentration of 5 mM pyruvate.
13. The cell pyruvate mixture is then loaded into the nanocoil and moved to the isocenter of the MRI, followed by the ^{13}C-NMR signal acquisition.
    a. The pulse sequence consisted of a series of 45° RF pulses with a one-second delay that lasted for 2–3 min.
    b. Once the signal decays push in more cells into the nanocoil to increase the number of repetitions
    c. We obtained nine repeats from a single sample by inserting "fresh" cells into the nanocoil nine times (Figure 8.4a).
14. To ensure proper hyperpolarization, the amount of polarization (%) and T_1 relaxation should be measured using a 1 T Spinsolve spectrometer and 0.6 mL of dissolution 240 s after the dissolution.
15. Verify the pyruvate concentration using NMR after the experiment, and the hyperpolarized signal completely decays.
16. The cell viability should be consistent throughout the experiments and optionally the cells can be recycled for more experiments such as PCR, western blots, and scRNA seq.

8.3.4 HP Spectral Data Analysis

1. The spectra are loaded into the MestReNova (Mestrelab Research) software for analysis.
2. The stack of obtained free induction decay (FID) was subjected to Fourier transformation, phase correction, and 3 Hz exponential apodization (Figure 8.4a).
3. In this experiment, we have three peaks that originate from hyperpolarized pyruvate, pyruvate hydrate formed by the hydration of pyruvate, and lactate formed in the cells (Figure 8.4a).
4. The pyruvate to lactate conversion rate (k_{PL}/s) is calculated by integrating the area of each peak and the lactate signal's ratio to the total carbon signal (pyruvate + pyruvate hydrate + lactate).
5. The lactate production rate, or "flux," is calculated by multiplying the k_{PL} by the pyruvate concentration and normalizing for the number of cells in the nanocoil (Figure 8.4c).
6. With this platform, you can test multiple cell lines (Figure 8.4c) under different conditions and treatments with the least amount of sample possible without interfering with chemical or biological processes.

8.4 Results

8.4.1 Quality of the Nanocoil

We can use the prior method to make a nanocoil, however this requires a high level of precision, making it one of the most important and difficult steps. The NMR signal is dependent on the homogeneity of the magnetic field. This can be affected by a variety of factors, including how the magnet wire is wound around the 27 AWG needle, the number of knots or bends in the copper wire, and whether removing the needle affected the nanocoil's integrity. An easy way to test the quality of the nanocoil is to test it with a few NMR standards. Once we fabricate the nanocoil and connect it to a circuit board, we tune and match to the Larmor frequencies of ^{1}H and ^{13}C at 3 T (at the system's field strength). We then connect the nanocoil to the MRI and evaluate its quality by filling it with water and acquiring ^{1}H NMR spectra. We should look for a single water peak (without splitting) at 4.7 ppm with a full width at half maximum (FWHM) of 30 Hz or less (the narrower the better); we also expect to see a broad PDMS peak near 0 ppm. We can also perform first-order shimming to improve the FWHM of water. After testing the water sample, we can test the nanocoil with other standards in the ^{1}H or ^{13}C channel to observe splitting. For example, using [^{13}C,^{15}N]urea, we were able to resolve the splitting pattern from ^{13}C–^{15}N coupling [8].

8.4.2 Quantifying Metabolic Flux in Cells

The nanocoil can detect hyperpolarized NMR signal and characterize metabolic flux while only using 2000 cells per 80 nL. To show this capability we characterized the metabolic flux of [1-^{13}C]pyruvate to [1-^{13}C]lactate in two acute myeloid leukemia cell lines, MOLM13 and NOMO1, which have previously shown to have distinct k_{PL}s [19]. Each of the cell lines were mixed with hyperpolarized [1-^{13}C]pyruvate and characterized with the nanocoil as described in the aforementioned methods. As shown in Figure 8.5a, we observed a relative difference between the two cell lines; additionally, we observed an increase in metabolic flux when we incubated MOLM13 cells with ABT199 (Figure 8.5b), which is known to increase k_{PL} [6, 8].

8.5 Discussion

This chapter demonstrates how to construct the 80 nL nanocoil and its sensitivity and specificity in characterizing metabolic

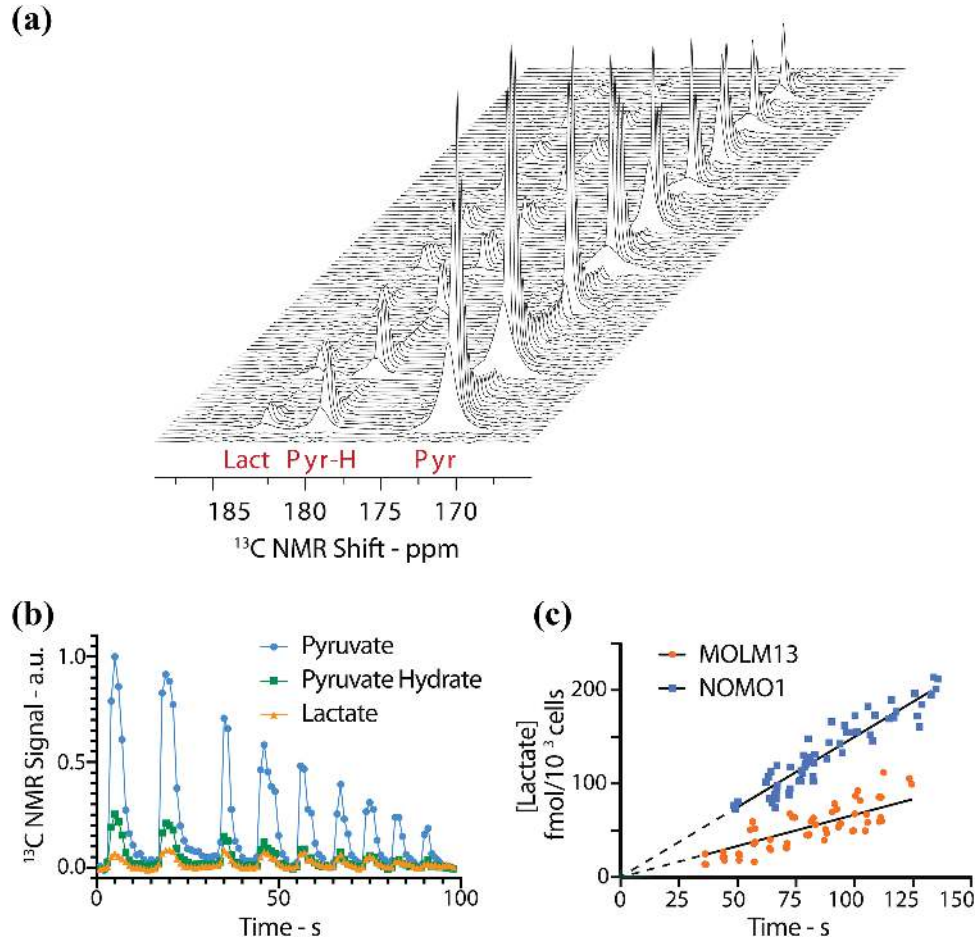


FIGURE 8.4 Analysis of metabolic flux with the nanocoil-based HMRS. (a) NMR spectra of hyperpolarized metabolites with hyperpolarized [1-^{13}C] pyruvate and approximately 2000 MOLM13 cells. (b) Peak-area plots for pyruvate, pyruvate hydrate, and lactate show when data was collected on the cell and when the nanocoil was reloaded with "fresh" cells. They also show that the signal decays over time due to T_1. (c) Comparison of the metabolic flux between MOLM13s and the more glycolytic NOMO1s. The solid line shows a linear fit of the lactate to total carbon ratio over time, while the dotted line represents extrapolation to zero when the cell suspension and hyperpolarized pyruvate were mixed. Adapted from reference [8].

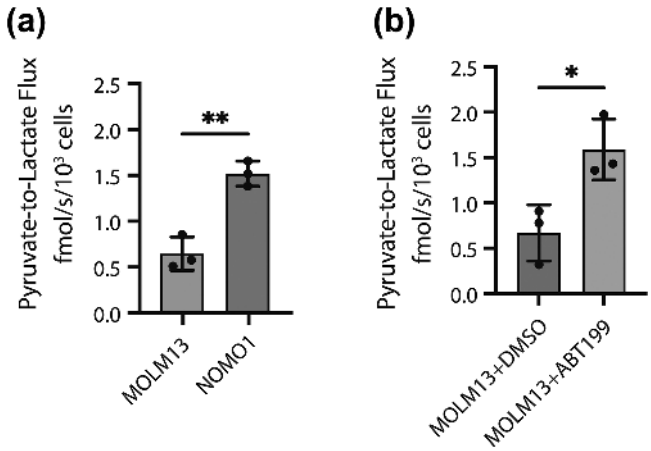


FIGURE 8.5 The metabolic flux of MOLM13 and NOMO1 cells and change in flux after drug treatment. (a) When comparing the metabolic flux of MOLM13 (orange) and NOMO1 (blue) cell lines, the NOMO1 cell lines have significantly higher k_{PL}/s ($p = 0.0027$). (b) After 24 hours of treatment with 50 nM ABT199, MOLM13's k_{PL}/s significantly increased when compared to MOLM13 treated with 0.1% DMSO ($p = 0.025$). Adapted from reference [8].

flux with a small sample size. When coupled with hyperpolarization, the nanocoil can detect metabolites produced in the fmol range with only 2000 cells, making this system orders of magnitude more sensitive than previous iterations [8]. The nanocoil has several advantages over other methods, including the ability to measure changes in metabolism in real time, being nondestructive and not interfering with any biological processes, and being compatible with a variety of probes, allowing for the measurement of the kinetics of various metabolic processes. With the integration of microfluidics, the nanocoil allows for high-throughput analysis of metabolic flux of a wide range of cell lines under various conditions.

8.6 Future Trends

Future development is required to streamline this protocol. This may include using a bioreactor to regulate temperature, oxygen/CO_2, and nutrient supply in order to create more ideal conditions for cell growth and testing. We may couple the nanocoil with other methods to "recycle" the cells and use them for downstream analysis with different techniques such as scRNA-seq or flow cytometry. To increase the system's throughput, we can incorporate several nanocoils on a single PCB and characterize several different cells/conditions at the same time using previously developed methods [19]. Because of its versatility, the nanocoil can be tailored to meet the requirements of a wide range of applications.

REFERENCES

1. Pavlova, Natalya N. and Craig B. Thompson, The emerging hallmarks of cancer metabolism. *Cell Metabolism*, 2016. 23(1): p. 27–47.
2. Warburg, O., F. Wind, and E. Negelein, The metabolism of tumors in the body. *The Journal of General Physiology*, 1927. 8(6): p. 519–30.
3. Vander Heiden, M.G., L.C. Cantley, and C.B. Thompson, Understanding the warburg effect: the metabolic requirements of cell proliferation. *Science*, 2009. 324(5930): p. 1029–1033.
4. Chang, C.-H., et al., Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell*, 2015. 162(6): p. 1229–1241.
5. Rodrigues, T.B., et al., Magnetic resonance imaging of tumor glycolysis using hyperpolarized 13C-labeled glucose. *Nature Medicine*, 2014. 20(1): p. 93–7.
6. Jeong, S., et al., Real-time quantitative analysis of metabolic flux in live cells using a hyperpolarized micromagnetic resonance spectrometer. *Science Advances*, 2017. 3(6): p. e1700341.
7. Lees, H., et al., Multi-sample measurement of hyperpolarized pyruvate-to-lactate flux in melanoma cells. *NMR in Biomedicine*, 2021. 34(3): p. e4447.
8. Pigliapochi, R., et al., Hyperpolarized nano-NMR platform for quantification of mass limited samples. *Analytical Chemistry*, 2024. 96(30): p. 12225–12229.
9. Jeong, S. and K.R. Keshari, Hyperpolarized micro-NMR platform for sensitive analysis of in vitro metabolic flux in living cells. *Methods in Molecular Biology,* 2022. 2393: p. 561–569.
10. Ardenkjaer-Larsen, J.H., et al., Increase in signal-to-noise ratio of > 10,000 times in liquid-state NMR. *Proceedings of the National Academy of Sciences of the United States of America*, 2003. 100(18): p. 10158–63.
11. Comment, A. and M.E. Merritt, Hyperpolarized magnetic resonance as a sensitive detector of metabolic function. *Biochemistry*, 2014. 53(47): p. 7333–7357.
12. Hu, S., et al., 13C-pyruvate imaging reveals alterations in glycolysis that precede c-Myc-induced tumor formation and regression. *Cell Metabolism*, 2011. 14(1): p. 131–142.
13. Keshari, K.R., et al., Hyperpolarized 13C-pyruvate magnetic resonance reveals rapid lactate export in metastatic renal cell carcinomas. *Cancer Research*, 2013. 73(2): p. 529–538.
14. Di Gialleonardo, V., et al., Multinuclear NMR and MRI reveal an early metabolic response to mTOR inhibition in sarcoma. *Cancer Research*, 2017. 77(11): p. 3113–3120.
15. Dong, Y., et al., Hyperpolarized MRI visualizes warburg effects and predicts treatment response to mTOR inhibitors in patient-derived ccRCC xenograft models. *Cancer Research*, 2019. 79(1): p. 242–250.
16. Miloushev, V.Z., et al., Metabolic imaging of the human brain with hyperpolarized 13C pyruvate demonstrates 13C lactate production in brain tumor patients. *Cancer Research*, 2018. 78(14): p. 3755–3760.
17. Granlund, K.L., et al., Hyperpolarized MRI of human prostate cancer reveals increased lactate with tumor grade driven by monocarboxylate transporter 1. *Cell Metabolism*, 2020. 31(1): p. 105–114.e3.
18. Salamanca-Cardona, L., et al., In vivo imaging of glutamine metabolism to the oncometabolite 2-hydroxyglutarate in IDH1/2 mutant tumors. *Cell Metabolism*, 2017. 26(6): p. 830–841.e3.
19. Lee, G., et al., Micro-slab coil design for hyperpolarized metabolic flux analysis in multiple samples. *Bioengineering (Basel)*, 2022. 10(1).

9

Detection of TMEM Doorways and Their Activity Required for Metastasis

George S. Karagiannis, Robert J. Eddy, Joseph A. Sparano, Jesus Anampa Mesias, Thomas E. Rohan, Craig Branch, John S. Condeelis, David Entenberg, and Maja H. Oktay

9.1 Introduction

The most common cause of mortality from breast cancer is distant metastasis, a process involving dissemination of cancer cells from the primary tumor to secondary sites (e.g., brain, bone, liver). Recent findings indicate that metastatic spread may continue from secondary to tertiary metastatic foci, thereby causing an exponential increase in metastatic burden and contributing to patient demise.[1–3]

About 30% of breast cancer patients will progress to metastatic disease, but it is difficult to distinguish between those likely to manifest metastases and those in whom the disease will never progress. Therefore, efforts are being made to develop tools to accurately identify the patients who will develop metastatic disease above and beyond classical clinicopathologic features.

Human breast cancer is an exceptionally heterogeneous disease. The simple classification of disease type includes determining the cell of cancer origin (ductal or lobular), the extent of cancer spread within the breast parenchyma (*in situ* or invasive), and the expression of cell signaling molecules that can regulate cell proliferation and estrogen dependence (estrogen, progesterone receptor) or HER2/Neu signaling (HER2/Neu protein overexpression or gene amplification).[4] Beyond this classification, the pathological evaluation of excised tumor also specifies tumor size, tumor grade (typically following the Nottingham classification[5]), histologic type for the *in situ* component (e.g., cribriform, solid, papillary) as well as for the invasive component (e.g., mucinous, tubular, medullary), the status of resection margins after surgery (clear or involved by tumor), the presence or absence of lympho-vascular invasion, and the number of lymph nodes involved by the tumor.[4] In addition to these pathological characteristics, the clinical management of breast cancer also includes consideration of host factors such as menopausal status, comorbidities, and other factors such as body-mass index,[6] social determinants of health,[7] and geographic ancestry,[8] which may play a role in disease progression and response to treatment.

The clinicopathological factors listed previously frequently do not provide enough information to stratify patients by risk of locoregional recurrence and metastasis in distant organs. In response to this need, several molecular tests have been developed and are currently in clinical use. For example, Oncotype DX,[9,10] Mammaprint,[11] and PAM50[12] provide prognostic information regarding distant recurrence risk in ER-positive, HER2-negative breast cancer. The Oncotype DX test generates a recurrence score (RS) based on the expression of 21 genes (16 test genes and five reference genes) in breast cancer tissues.[10] The test genes are associated with proliferation (Ki67, STK15, Survivin, CCNB1, MYBL2), HER2 signaling (GRB7, HER2), estrogen signaling (ER, PGR, BCL2, SCUBE2), invasion (MMP11, CTSL2), metabolism (GSTM1), macrophage infiltration (CD68), and chaperon regulator (BAG1).[10] The RS score ranges from 0 to 100, with lower scores indicating a lower risk of recurrence. Two prospective studies, including TAILORx in node-negative and RxPONDER in 1–3 positive axillary nodes, demonstrated that RS can also be used to spare the use of adjuvant chemotherapy in about 70% of women who might otherwise have been recommended adjuvant chemotherapy based on clinicopathologic features alone.[9,13–15]

A common feature of these tests is that they all directly or indirectly measure cancer cell proliferation and estrogen-dependence, but to varying degrees.[11,12,16,17] However, the metastatic cascade requires more than just cancer's proliferative capacity and estrogen dependence. As such, these tests have shown variable performance in their ability to predict risk of metastasis.[17]

In breast cancer, our group has investigated the tumor microenvironment and identified portals on blood vessels through which cancer cells intravasate and disseminate to distant sites (Figure 9.1A).[18] These portals are called tumor microenvironment of metastasis (TMEM) doorways.[17,19,20] Each TMEM doorway is composed of a cancer cell expressing the actin regulatory protein, Mena, a perivascular macrophage, and an underlying endothelial cell, all in direct physical contact.[17,19,20]

We showed that the density of TMEM doorways (TMEM doorway score) is prognostic of distant metastasis in patients with estrogen receptor (ER^+) $HER2^-$ breast cancer.[21,22] In addition, we found that the assembly and activity of TMEM doorways can be increased in residual breast cancer of certain patients following treatment with neoadjuvant chemotherapy.[23] Importantly, such an increase in TMEM doorway score and activity is abrogated in animal models if chemotherapy is combined with a small molecule inhibitor of TIE2,[23] a key signaling pathway of perivascular macrophages at TMEM doorways, which inhibits TMEM doorway activity and cancer cell dissemination by TMEM doorways.[24] Collectively, these

DOI: 10.1201/9781003449942-11

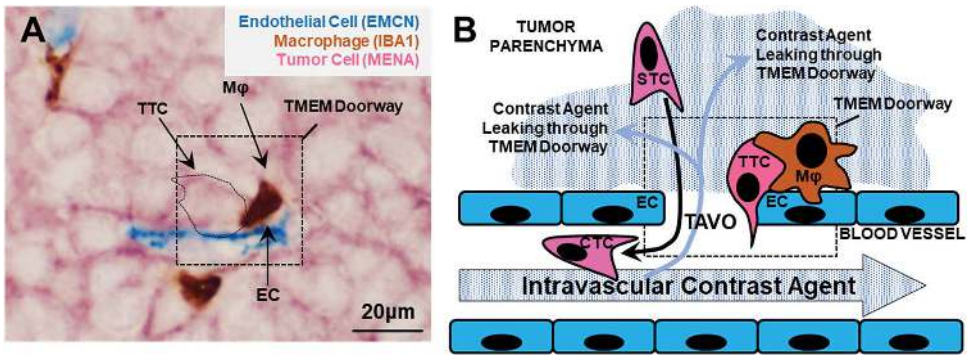


FIGURE 9.1 *TMEM doorway assembly and activity, upon which "TMEM Doorway Score" and "TDAM Score" are based.* TMEM doorways (dotted squares) are assembled via the physical proximity of an endothelial cell (EC), a perivascular macrophage (Mφ), and a TMEM doorway-related tumor cell (TTC) expressing Mena. **(A)** An image of TMEM doorway visualized using our standardized triple immunohistochemistry protocol for assessing TMEM doorway score. **(B)** Upon activation, TMEM doorways dismantle endothelial cell integrity, causing TMEM-associated vascular opening (TAVO). TAVO is taken advantage of by "streaming" tumor cells (STC) for intravasation (black arrow) and can be visualized using TMEM-doorway dextran or MRI: A TMR-dextran or gadolinium-based contrast agent (GBCA) is injected, and it circulates through the vasculature (including tumor vasculature). Upon reaching active TMEM doorways, contrast agent leakage into the tumor parenchyma (blue arrows) occurs due to dissolution of endothelial integrity and is quantified as "TMEM-Doorway-Activity-MRI" (TDAM). Note the color-matching of individual TMEM-doorway associated cells in the IHC (A) and the infographic (B).

observations indicate therapeutic opportunities, based on selective inhibition of the TMEM doorway-dependent cancer cell dissemination machinery.[25–27]

Although TMEM doorway score (number of TMEM doorways per tissue area) is prognostic for metastatic outcome, this approach cannot determine if TMEM doorways are indeed active (i.e., in a biological state that allows prometastatic tumor cells to escape into the peripheral circulation). To overcome this obstacle, we have developed complementary methods to allow for the specific assessment of TMEM doorway-associated vascular opening (TAVO) events that are involved in tumor cell intravasation into the peripheral circulation (Figure 9.1B). In mice, TAVO events can be quantified by measuring the leakage of intravenously injected high molecular weight dextran from vessels into the tumor parenchyma through open TMEM doorways.[28] Because such probes cannot be introduced into humans, an analogous approach using a gadolinium-based contrast agent (GBCA) has been developed for patients, thus leading to the development of a noninvasive MRI assay for direct TAVO assessment.[28] TAVO assessment by MRI is detailed more in the discussion.

In summary, although cancer cell dissemination is a critical step in the metastatic cascade, it is not inherently captured in current clinically available diagnostic/prognostic assays, leading to suboptimal assessment of metastatic risk.[17] It is thus important to incorporate dissemination markers in the clinical care of breast cancer patients. Doing so would have the added benefit of giving clinical researchers tools to assess anti-dissemination therapeutics more directly. In this chapter, we discuss how TMEM doorway biology, as assessed by the traditional TMEM doorway score and the accompanying TMEM doorway activity assays, may be exploited to determine cancer dissemination risk, both at baseline and in response to treatment, which may advance patient care, particularly in advanced stages of disease.

9.2 Materials

9.2.1 Chemicals

10% Neutral Buffered Formalin (9400–1) was purchased from Epredia, Fisher Scientific, Waltham, MA. Antigen Unmasking Buffer, pH 9.0 (H-3301), ImmEdge Pen (H-4000), Bloxall Blocking Solution (SP-6000), Blue Blocking Serum, VECTASTAIN ABC-AP Kit (AK-5001), Substrate Working Solution from VECTOR Blue AP Substrate Kit (SK-5300), ImmPACT VIP Substrate, Peroxidase (HRP) (SK-4605), and VectaMount Express Mounting Medium (H-5700) were purchased from Vector Laboratories, Inc., Newark, CA. Dako Wash Buffer (S3006), Dako Dual Endogenous Enzyme Block (S2003), Dako anti-CD68 mouse monoclonal antibody, clone PG-M1 (M0876), Dako Antibody Diluent (S0809), Dako Liquid DAB + Substrate Chromogen System (K3468), Dako Envision Flex antibody diluent (K8006), Dako EnVision FLEX+ Mouse Linker (K8021), and Dako EnVision+ System-HRP labelled polymer (K4001) were purchased from Agilent Technologies, Inc., Santa Clara, CA. Anti-CD31 rabbit polyclonal antibody (ab182981) was purchased from Abcam, Cambridge, UK. Anti-Mena mouse monoclonal antibody, clone A351F7D9 (MAB2635), 155 kDa tetramethylrhodamine isothiocyanate-dextran (T1287), fetal bovine serum (FBS) (F2442), and fish skin gelatin (G7765) was purchased from MilliporeSigma, Burlington, MA. Fast Green Substitute for Light Green working solution (S232b) was purchased from Poly Scientific R&D Corp, Bay Shore, NY. Anti-Iba-1 rabbit polyclonal antibody (019-19741) was purchased from Fujifilm Wako Pure Chemicals Corp, Osaka Japan. Anti-endomucin rat polyclonal antibody, V.7C7 (sc-65495), was purchased from Santa Cruz Biotechnology, Dallas, TX. Alkaline Phosphatase conjugated AffiniPure Donkey anti-rat IgG (712-055-150) was purchased from Jackson ImmunoResearch Labs, Inc., West

Grove, PA. Bovine serum albumin (BSA) (BP1600-100) was purchased from Fisher Scientific, Hampton, NH. Tween-20 (TWEEN201) was purchased from MP Biologicals, Irving, CA. Anti-tetramethylrhodamine (TRITC) rabbit polyclonal antibody (A6397), goat anti-rat IgG Alexa Fluor 647 (A-21247) and goat anti-rabbit IgG Alexa 488 (A-11034), DAPI (62248), and ProLong Gold antifade mountant (P36934) was purchased from Thermo Fisher, Waltham, MA.

9.2.2 Equipment

HistoCore Multicut microtome was purchased from Leica, Nussloch, GmbH. Tissue-Tek Accu-Edge high profile microtome blades (63068-HP) were purchased from Sakura Finetech USA, Torrence, CA. Tissue floatation bath was purchased from Boekel Scientific, Feasterville, PA. Superfrost Plus Microscope Slides, 25 × 75 × 1 mm, were purchased from Fisher Scientific, Pittsburg, PA. Microtome Isotemp convection oven (151030510) was purchased from Thermo Fisher, Waltham, MA. Slide staining station (2935M7) was purchased from Medicus Health, Kentwood, MI. Electric food steamer, 800W (DZG-D80A1), was purchased from Secura. Dako AutostainerLink 48 was purchased from Agilent Technologies, Inc., Santa Clara, CA. Digital whole slide scanner PANNORAMIC 250 Flash III was purchased from 3DHistech, Budapest, Hungary.

9.3 Methods

9.3.1 Methodology for TMEM Doorway Scoring

9.3.1.1 Selection of Tissue for TMEM Doorway Score Analysis

The starting material for measuring TMEM doorway score is formalin-fixed paraffin-embedded (FFPE) tissue samples from either human patients with, or mouse models of, breast cancer. Typically, each tumor is fixed with 20x volume of 10% neutral buffered formalin for 48–72 hours at room temperature. After removing the fixative, the tumor sample is divided into several embedding cassettes (3 × 2.5 × 0.4 cm), stored in 70% ethanol, processed into FFPR tissue blocks, cut into 5 µm thick tissue sections on a microtome, and placed onto 25 × 75 × 1 mm standard microscope slides according to standard FFPE protocols used in any histopathology laboratory.[29] These sections are then stained with hematoxylin and eosin (H&E)[30] and evaluated by a pathologist, who chooses the best blocks for subsequent TMEM doorway staining. For both human and mouse studies, it is important to use slides that contain at least 5 mm of contiguous tumor area with less than 10% covered by necrosis, inflammation, tissue folds, tears, or other tissue-embedding artifacts.

9.3.1.2 TMEM Doorway Staining Protocol

Due to species-dependent differences, antibody availability and optimizations, and tumor antigen-processing procedures, distinct protocols have been developed and validated for clinical (human) and preclinical (mouse) studies.

9.3.1.2.1 Human Breast Cancer Protocol

Formalin-fixed paraffin-embedded tumor sections (5 µm) are placed in a convection oven for 15 min at 58 °C until paraffin is melted. Slides are dewaxed in xylene, two times for 10 min each, rehydrated in 100% ethanol for 2 min, 95% ethanol for 2 min, 70% ethanol for 2 min, and 50% ethanol for 2 min, and placed in distilled H_2O. Slides are antigen retrieved in 95 °C Antigen Unmasking Buffer, pH 9.0, placed in a steamer for 20 min and allowed to cool to 22 °C for 30 min. Slides are rinsed in Dako Wash Buffer for 5 min and the tumor area is traced using an ImmEdge Pen. To identify macrophages, slides are blocked with Dako Dual Endogenous Enzyme Block for 10 min, rinsed for 2 min, and incubated with Dako anti-CD68 mouse monoclonal antibody, clone PG-M1 diluted in Dako Antibody Diluent at 1:100 dilution for 30 min. Slides are rinsed and incubated with Dako Liquid DAB + Substrate Chromogen System for 4 min, rinsed and blocked in Bloxall Blocking Solution for 10 min to quench endogenous AP activity, and then blocked with Blue Blocking Serum, VECTASTAIN ABC-AP Kit for 20 min. To identify endothelial cells, slides are incubated with anti-CD31 rabbit polyclonal antibody at 1:500 dilution in Blue Blocking Serum for 30 min. Slides are rinsed and incubated with VECTASTAIN ABC-AP Kit, AP-Rabbit IgG secondary antibody for 30 min. Slides are rinsed and incubated for 10 min with freshly prepared Substrate Working Solution from VECTOR Blue AP Substrate Kit, rinsed, and blocked in Dako Dual Endogenous Enzyme Block for 10 min. To identify tumor cells, slides are rinsed and incubated with anti-Mena mouse monoclonal antibody, clone A351F7D9, diluted at 1:500 in Dako Envision Flex antibody diluent for 1 hr, and rinsed and incubated with Dako EnVision FLEX+ Mouse Linker for 10 min to amplify the signal. Slides are rinsed and incubated with Dako EnVision+ System-HRP labelled polymer for 20 min, rinsed, and incubated with freshly prepared ImmPACT VIP Substrate, Peroxidase (HRP) for 20 min. Slides are rinsed and briefly counterstained with Fast Green Substitute for Light Green working solution for 2 sec, washed two times with distilled H_2O, and mounted with VectaMount Express Mounting Medium and allowed to cure for 16 hr prior to scanning.

9.3.1.2.2 Mouse Mammary Cancer Protocol

Slides are dewaxed as described for human breast cancer protocol (Section 9.3.1.2.1). To identify macrophages, slides are blocked with Dako Dual Endogenous Enzyme Block for 10 min and incubated with anti-Iba-1 rabbit polyclonal antibody diluted 1:1000 in Dako Antibody Diluent for 30 min, and rinsed and incubated with ImmPRESS-HRP Goat anti-rabbit IgG polymer reagent for 30 min. Slides are rinsed and incubated with Dako Liquid DAB + Substrate Chromagen System for 2 min, and rinsed and blocked with Dual Endogenous Enzyme Block (Agilent) for 10 min. To identify endothelial cells, slides are incubated with anti-endomucin rat polyclonal antibody, V.7C7 at 1:300 dilution in Dako Antibody Diluent for 1 hr. Slides are rinsed and incubated with Alkaline Phosphatase conjugated AffiniPure Donkey anti-rat IgG at 1:150 dilution in Dako Antibody Diluent for 1 hr. Slides are rinsed and incubated for 10 min with freshly prepared Substrate Working Solution from VECTOR Blue AP Substrate Kit, rinsed and blocked in Dako Dual Endogenous Enzyme Block for 10 min.

After rinsing, tumor cells are identified by incubation with anti-Mena antibody and are counterstained and mounted as described in Section 9.3.1.2.1 for human breast cancer tissue.

9.3.1.3 TMEM Doorway Scanning

Upon completion of staining, each slide is evaluated by a pathologist to confirm the quality of the stain, and if deemed appropriate, the slides are imaged on an automated digital whole slide scanner (Pannoramic P250 Flash III, 3DHistech, Budapest, Hungary) using a 20x 0.8NA objective lens, and high-speed, 4-megapixel, color CMOS camera (pixel size = 0.24 μm). Overview image intensity thresholding is used to automatically identify tissue regions for high resolution scanning. The instrument is set to automatically focus and stitch the individual fields of view (FOVs) into single pyramidal composite image files. Further analyses and quantifications are performed either on the original glass slides by trained pathologists or on their digital representations, if intended for automated analysis and digital pathology.

9.3.1.4 TMEM Doorway Quantification

9.3.1.4.1 Manual Method

For manual scoring, pathologists evaluate the entire slide under the microscope or using the digital whole slide scan. Once the areas of the highest vascularity are identified, images of ten high-power fields of view (438 × 330 μm) are captured and used for scoring. Images are loaded into Adobe Photoshop and assessment of TMEM doorway scores is performed using the "circle" tool to create ~50 μm circles around TMEM doorways, avoiding overlap between adjacent TMEM doorways. Each FOV is limited to a maximum of 20 TMEM doorways. After TMEM doorways are marked, the "marked" images are saved as separate files. The total number of TMEM doorways for each image is tabulated, and the scores from all ten images are then summed to give a final TMEM score for each patient sample, expressed as the total number of TMEM doorways in ten high-power fields.[21]

9.3.1.4.2 Automated Method

Image analysis is conducted using a custom application in Visiopharm's VIS (Visiopharm Integrator System) software using their VisiomorphDP module, focusing exclusively on predefined invasive tumor regions, as determined by pathologists. We have tested a number of image analysis packages for this purpose and found that VIS the best suited software package. We were not successful in performing the type of morphometric analysis listed later with Fiji. However, since the time of our testing, QuPath, an application specifically designed for digital whole slide scan analysis, has been released and has advanced significantly. Though not tested for this application, we conjecture that the process could be accomplished using QuPath.[31] The process involves generating "regions of interest" (ROIs), requiring about one to two minutes of a pathologist's time per case. The analysis workflow is described in detail in our original publication[32] and is concluded in five main steps:

a. *Preprocessing*: Images captured using the slide scanner, as indicated earlier, are split into "Feature Channels" to enhance signal contrast for the individual colors of the TMEM doorway stain, followed by the application of a median filter to improve signal clarity.
b. *Classification*: A linear Bayesian classifier is used to assign each pixel to one of five categories: macrophages, darkly and lightly stained tumor cells, blood vessels, and stroma. This classification is based on pathologist-marked examples from the stained slides. The training for this classifier was completed in our original report[32] and is thus applied consistently thereafter.
c. *Postprocessing*: Classified pixels are grouped as separate objects based on continuity, and then further manipulated according to their size, shape, and spatial relationships, to identify "TMEM doorway objects."
d. *Identification of "TMEM Doorway Objects"*: This step involves merging the generated classifications, adjusting object boundaries through dilation and erosion, and removing irrelevant structures based on size constraints. Through analysis of the spatial relationships between macrophages, vessels, and tumor cells, these objects are categorized into complexes based on their perimeter contact, leading to the identification of the "TMEM-Doorway Objects." These parameters are not listed here due to space restrictions but are detailed in our original report.[32]
e. *Calculation and Ranking*: The final steps involve quantifying the identified objects using metrics such as count and area, and comparing high-power field areas within the ROI for TMEM Doorway Object counts. These counts are used to rank the areas with the highest scoring regions presented for a pathologist's review. The cumulative TMEM doorway score for each sample is computed by summing counts from the top-ranked fields, essentially simulating the pathologist's manual procedure.

9.3.2 Digital Pathology Infrastructure Supporting Automation of TMEM Doorway Score

The reproducibility and performance of the automated TMEM Doorway Score have been rigorously evaluated using several methodologies to ensure reliability and accuracy in various settings. Initially, the algorithm's reproducibility was confirmed by scanning 12 different slides and analyzing them to determine TMEM doorway scores on two different days, achieving a near-perfect repeatability with a Pearson correlation coefficient of ~1.00.

Evaluation of the algorithm's quantification performance comprised three distinct approaches: first, the algorithm's ability to identify TMEM doorways was assessed by comparing the algorithm's quantification of TMEM doorway scores in 100 FOVs with the average of scores generated by two pathologists, showing strong correlation and concordance (Pearson r = 0.86), and high accuracy in classifying TMEM doorway scores into tertiles (AUCs of 0.90 and 0.96). Second, the algorithm's comprehensive capability to mimic complete pathological

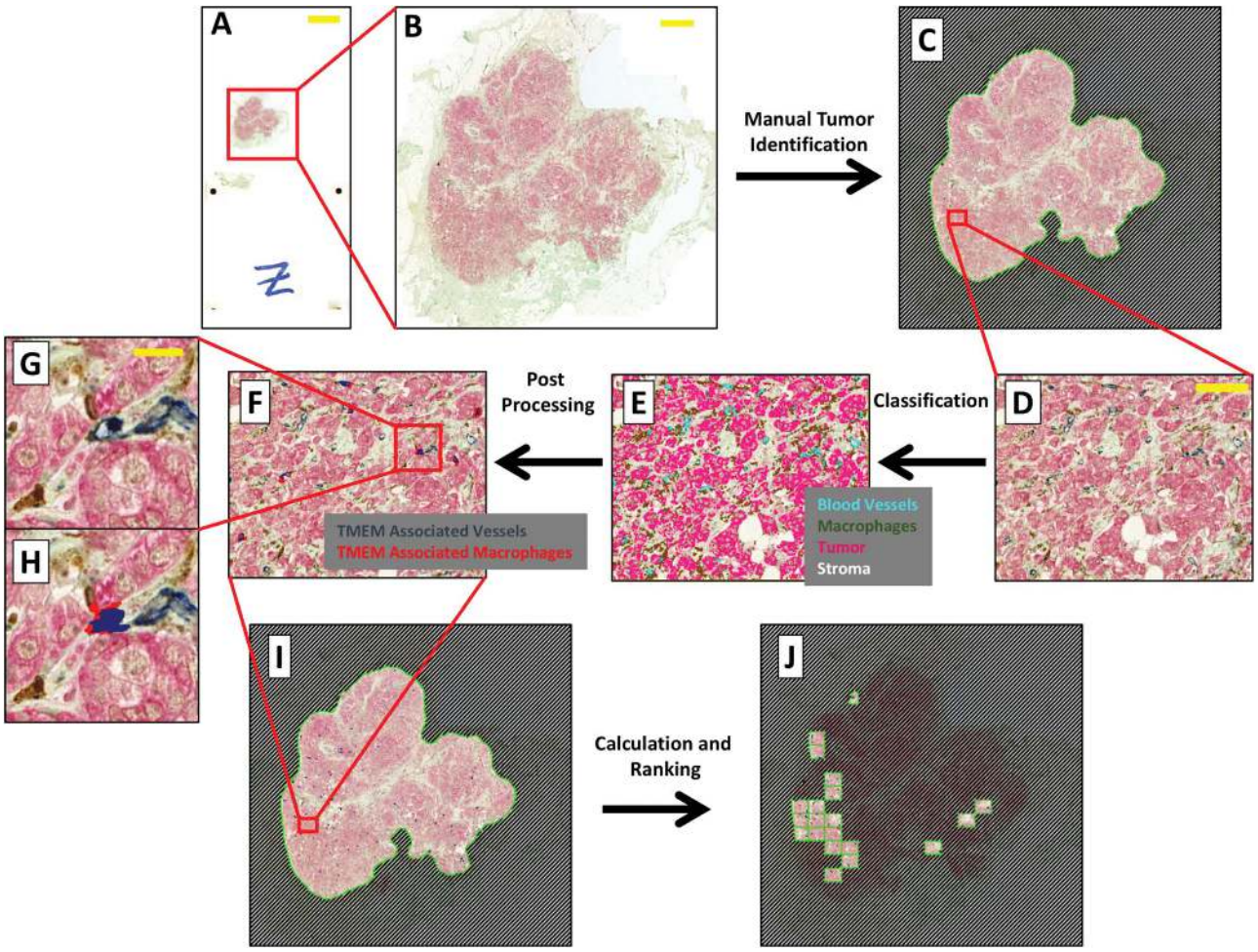


FIGURE 9.2 Algorithm for performing automated tumor microenvironment of metastasis (TMEM) doorway analysis and scoring on whole scanned slides: Automated identification of TMEM doorway in immunohistochemical (IHC)-stained slides requires several steps applied at different magnifications. (**A**) Digitized whole slide captured on the 3DHistech Pannoramic P250 slide scanner. Bar = 5000 µm. (**B**) Zoomed-in region of just the tissue. Bar = 1000 µm. (**C**) Pathologist identification of invasive tumor by drawing a region of interest (ROI) separating and masking off the tumor tissue from the surrounding stroma. (**D**) A 40× equivalent zoomed-in field-of-view within the tumor area. Bar = 100 µm. (**E**) Classification of each pixel as either a blood vessel (cyan), a macrophage (brown), a tumor cell (pink), or stroma (uncolored). (**F**) Identification of TMEM doorways highlights the TMEM-associated vessels in dark blue and the TMEM-associated macrophages in red. (**G**) Zoomed-in view of one identified TMEM doorway from Figure 1F. Bar = 20 µm. (**H**) Same as Figure 1G with the overlay indicating the TMEM-associated vessel (dark blue) and TMEM-associated macrophage (red). (**I**) Image of the whole tissue after application of both the tumor identification and TMEM identification algorithms. (**J**) The top ten scoring fields are highlighted by the TMEM ranking algorithm.

analysis was assessed using 60 breast cancer cases, resulting in very high correlation ($r = 0.94$) and discrimination accuracy (AUCs of 0.95 and 0.94). Finally, testing on a larger cohort of 468 breast cancer cases also demonstrated excellent concordance, further validating the algorithm's robustness across diverse testing scenarios (AUCs of 0.81 and 0.79 for low- and high-risk TMEM doorway scores, respectively). The algorithm for performing automated TMEM doorway analysis and scoring is shown in Figure 9.2.[32]

9.3.3 Methodology for the Analysis of TMEM Doorway Activity in Murine Breast Cancer[28]

9.3.3.1 Tumor and Sample Preparation

To determine TMEM doorway opening in response to various conditions we perform a tail-vein injection of 100 µL of 10 mg/mL 155 kDa tetramethylrhodamine isothiocyanate (TRITC)-dextran, which has been shown not to extravasate from tumor vessels except during TAVO,[18] one hour before sacrificing the mice and harvesting the breast tumors. The tumors are fixed in 10% formalin for 48–72 hours and then paraffin embedded. Two 5 µm thick sequential slides are cut from the formalin-fixed paraffin-embedded (FFPE) tissues. The first section is used for staining the TMR-dextran using immunofluorescence, and the second for TMEM doorways as described previously. After standard deparaffinization, we submerge the sections into citrate buffer (pH 6.0-adjusted) and perform antigen retrieval by heating (~95 °C). The sections are then cooled at room temperature for 15–20 minutes, washed in PBS three times (2 min each), and blocked for 60–90 min in blocking buffer (10% FBS, 1% BSA, 0.0025% fish skin gelatin in PBST (PBS + 0.05% Tween-20). Next, the slides are incubated with a mixture of rat polyclonal anti-endomucin and

rabbit polyclonal anti-tetramethylrhodamine (TRITC), and washed in PBST, three times (2 min each). The slides are then incubated with a mixture of secondary goat antibodies against rat IgG conjugated to Alexa-647 and rabbit IgG conjugated to Alexa-488 and washed in PBST, three times (2 min each). To visualize nuclei, slides are stained with DAPI diluted 1:1000 in PBS for 5–6 min. Finally, the slides are mounted using a glycerol-free "hard" mounting medium such as ProLong Gold Antifade mountant and stored in a dark place until scanning. For optimal results, slides are scanned on a digital whole slide scanner.

9.3.3.2 Image Analysis

Image analysis starts with capturing ten high-power fields (HPFs) per case, using any software suitable for digital pathology (e.g., Case Viewer). We save the "Endomucin" (Red) and "TMR" (Green) channels separately as TIFF images. Using ImageJ/Fiji, the TIFF files are opened and converted into 8-bit images. The 8-bit images are next thresholded to the level of the negative control, and two binarized images are generated to create the Endomucin and TMR "masks." From the "Binary" tools, the "Fill Holes" function is run on the Endomucin mask, and the following five regions of interest are generated and saved:

a. The thresholded dextran image; saved as "Dextran ROI" (ROI1).
b. The thresholded endomucin image; saved as "Vascular ROI" (ROI2).
c. The inverted endomucin image; saved as "Extravascular ROI" (ROI3).
d. The intersected "Extravascular ROI" and "Dextran ROI" image (ROI1 ∩ ROI3); saved as "Extravascular Dextran ROI" (ROI4).
e. The entire image; saved as "Tumor ROI" (ROI5).

The ROI4 area is then divided by the ROI5 area and multiplied by 100 to generate the percent area that the extravascular dextran covers in the entire tumor. This process is repeated for all 10–20 HPFs per case (depending on tissue availability) to generate average extravascular dextran (% area) for each case.

9.4 Results

9.4.1 The Use of TMEM Doorway Score and Activity in Preclinical Research

Using manual analysis of TMEM doorway score we showed that in the MMTV-PyMT mouse model of breast carcinoma, as well as in ER^+ and ER^- breast cancer patient-derived xenografts, treatment with neoadjuvant chemotherapy (i.e., paclitaxel, doxorubicin, and cyclophosphamide) leads to a significant increase in TMEM doorway score.[23] Mechanistic analysis of this increase in TMEM doorway score has collectively suggested that TMEM doorway score and potentially metastasis is increased in response to a wound-healing response, triggered by the effects of cytotoxic chemotherapy on the tumor microenvironment, followed by the accrual of bone marrow-derived $TIE2^+$ monocytes in response to the chemotherapy-driven alarmins.[25–27,33] Our subsequent studies focused on delineating TAVO events in MMTV-PyMT mice treated with neoadjuvant chemotherapy, either through the high-molecular dextran assay described earlier (Figure 9.3)[23] or via the MRI approach.[34] Both approaches showed a significantly increased TMEM doorway activity upon treatment with neoadjuvant chemotherapy, leading to enhanced metastatic dissemination and metastatic recurrence.

To demonstrate the strategic potential of targeting TMEM doorway-dependent cancer cell dissemination after treatment with neoadjuvant chemotherapy, we co-administered chemotherapy with the TIE2 inhibitor, rebastinib, in MMTV-PyMT mice and observed a decrease in TAVO and the number of CTCs.[23] Together, these observations establish the potential of TMEM doorway score and its complementary TAVO assessments, as key investigational tools for studying the cancer cell dissemination machinery, including during intratumoral remodeling after treatment with established anticancer therapies.

9.4.2 Potential Use of TMEM Doorway Score and Activity as Companion Diagnostic

The manual analysis of TMEM doorway score was initially used to demonstrate the association of high TMEM doorway score with metastatic outcome in patients.[21] Subsequently, we performed an automated analysis of TMEM doorway score to evaluate if the TMEM doorway score correlates with a common prognostic marker of cancer cell proliferation, Oncotype DX. In a cohort of 600 patients with localized, or locally advanced, breast cancer from Eastern Cooperative Oncology Group (ECOG) trial 2197, we demonstrated that there is no correlation between the two prognostic markers, further supporting the premise that they indeed measure different cancer hallmarks.[22] Interestingly however, we demonstrated that high TMEM doorway score identifies a subgroup of patients in the low-Oncotype DX score presenting with an up to nine-fold increased risk of developing distant metastasis (Figure 9.4). This observation suggests that the TMEM doorway score may be a complementary companion diagnostic, when co-assessed with established clinical panels, such as Oncotype DX. Therefore, measurement of both proliferation and dissemination potential of breast cancer may generate more optimal prognostication in breast cancer patients at putative risk for developing metastatic disease.

9.4.3 TMEM Doorway Score in Racial Disparities of Metastatic Potential

The presence of TMEM doorways may be mechanistically associated with racial disparity in outcome for breast cancer patients who have estrogen receptor positive HER2 negative (ER^+ $HER2^-$) residual disease after neoadjuvant chemotherapy. Our group and others demonstrated that Black compared to White patients have worse outcomes only in

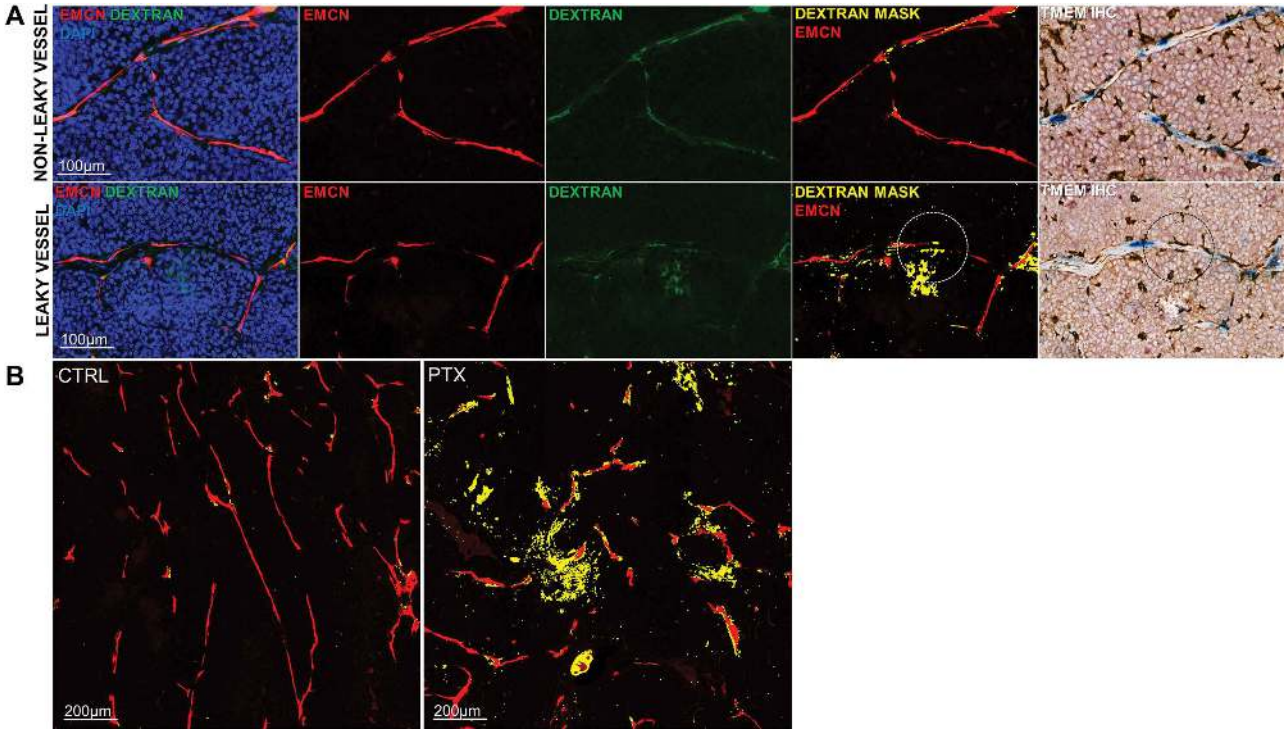


FIGURE 9.3 *TMEM doorway activity measured in formalin-fixed paraffin-embedded breast tumors using TMR-dextran.* **(A)** Multichannel immunofluorescence of endomucin with nuclear stain DAPI (first column), endomucin only (second column), and dextran-TMR only (third column). The fourth column shows demarcated extravascular dextran masks in non-leaky (top row) and leaky (bottom row) vessels. Corresponding sequential section of TMEM IHC is shown in the fifth column. **(B)** Breast cancer tissues from MMTV-PyMT mice treated with vehicle (CTRL) or paclitaxel (PTX). Note increase extravascular dextran in paclitaxel-treated tumors.

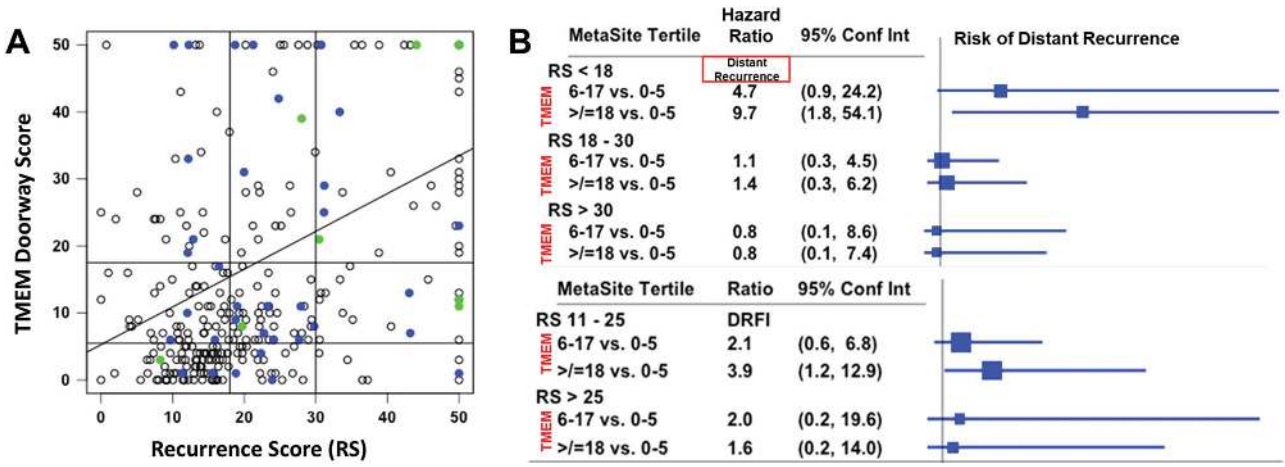


FIGURE 9.4 TMEM doorway score provides complementary information to Oncotype DX recurrence score (RS). **(A)** Correlation between TMEM doorway score and recurrence score, with both scores capped at a maximum value of 50. The raw Pearson correlation coefficient (r) is 0.29. Recurrences occurring within 5 years are shown in colors (local recurrence – green; distant recurrence – blue), including all recurrences (RFI) and distant recurrences (DRFI). **(B)** Complementary information regarding the risk of distant recurrence when combining TMEM doorway score with recurrence score (RS). Note the significant increase in the risk of distant recurrence for patients with low RS who have intermediate (6–17) or high (≥18) TMEM doorway score compared to patients with low (0–5) TMEM doorway score. Similar findings are noted with updated RS scoring classification showing increased risk of distant recurrence for patients with intermediate RS (11–25) who also have intermediate (6–17) or high (≥18) TMEM doorway score compared to patients with low (0–5) TMEM doorway score (excerpt from Sparano et al. *npj Breast Cancer* 2017).

ER+ HER2− disease,[35–37] but the mechanistic reason for this disparity was unknown. We found that Black compared to white patients have higher densities of TMEM doorways in their residual ER+ HER2−cancers and that high TMEM doorway score is an independent prognostic indicator of distant recurrence.[38]

The prior examples show the importance of evaluating pro-metastatic changes in tumor microenvironment in addition to proliferative potential of cancer. Moreover, they show that, in some breast cancer patients, treatment may adversely affect outcome and the likelihood of this can be assessed using the TMEM doorway score.

9.5 Discussion

We outlined here our approach to determining the dissemination potential of breast cancer using the TMEM doorway score and how this information can be used clinically to identify patients who are at risk of developing distant recurrence. This is an attractive approach because it uses FFPE tissue samples that are routinely collected as part of standard clinical practice. We also developed an automated staining approach which can likewise be applied in any clinical laboratory equipped with an autostainer.[32] However, TMEM doorway staining employs three antibodies applied sequentially, which is not typical for clinical laboratories and may present an obstacle to daily workflow. TMEM doorway scoring can be done by hand but requires training.[21] Once training is complete, manual analysis of a single slide may require about 30 minutes to an hour of a pathologist's time, which may limit the number of cases that can be signed out per time period. On the other hand, automated TMEM doorway scoring requires only a few minutes of a pathologist's time to determine the best area on a slide (called the region of interest or ROI) to which the scoring algorithm is then applied.[32] The automated scoring can only be done if the laboratory is equipped with a scanner, which is now readily available in most pathology laboratories. However, the Visiopharm software is custom designed and needs to be supplied to individual laboratories. These are all minor adjustments to the workflow that could be easily done to determine the cancer dissemination prognosis for each patient based on an enumeration of TMEM doorways in FFPE tissue sections. However, TMEM doorway score measured in FFPE tissue samples is inadequate for monitoring tissue response to various systemic treatments in real time.

Our group showed that, in mice, chemotherapy-mediated increases in TMEM doorway activity can be measured by quantifying the amount of extravascular high molecular weight dextran which has extravasated as a result of TMEM doorway openings.[23] However, injecting high molecular weight dextran to measure the effect of drugs on TMEM doorway activity in real time is not a feasible approach for clinical use. Since providing immediate information on cancer cell dissemination in response to systemic treatment would of be of great benefit to patients (because it would potentially lessen therapy-induced metastatic spread), we sought an alternative approach to high molecular weight dextran that can be applied in clinical setting. To this end, we developed a gadolinium-based MRI method which can measure TMEM doorway opening in patients' breast cancers in real time, without the need for tissue sampling. We demonstrate that TMEM-doorway activity MRI (TDAM) correlates with TMEM doorway score in patients as measured in FFPE tissue sections by IHC.[34] In addition, we showed that chemotherapy-induced increases in TMEM doorway activity, as well as reversal of this increase by TMEM-doorway activity blocking agents (small molecule inhibitor of TIE2), can be measured by TDAM in mice.[34] Given the foregoing, we envision the use of TDAM as companion diagnostic in patients undergoing preoperative systemic therapies to downstage the tumor. This approach would provide patients with the benefit of chemotherapy-mediated shrinkage of tumor along with usage of anti-dissemination treatment when needed. Given that Black patients with residual ER+ HER2− breast cancers after neoadjuvant chemotherapy have worse outcomes compared to White patients, this approach may help reduce racial disparities in breast cancer outcomes,[38] by recommending discontinuation of neoadjuvant treatment if it is associated with increased TAVO.

In addition to being prognostic biomarker in ER+ HER2− breast cancer, TMEM doorway score could have prognostic value in other cancers of glandular origin that show late recurrence in some instances such as prostate cancer. In other more aggressive cancers of glandular origin such pancreatic cancer, TMEM doorway score may not be prognostic due to overwhelming early recurrence. However, measuring TMEM doorway activity in response to neoadjuvant treatment in pancreatic cancer by TDAM would be extremely helpful to adjust treatment in real time and prevent cancer cell dissemination. Although TDAM can be used successfully in breast cancer and will most likely be useful in pancreatic cancer due to similar tissue quality, it is unclear that it could be used in other organs such as the lungs.

In summary, breast cancer progression to the currently incurable metastatic stage includes not only cancer cell proliferation but also cancer cell dissemination to and colonization of distant organs. Here we described a clinically validated marker of cancer cell dissemination to distant sites, the TMEM doorway score, as well as our approach to utilize TMEM doorway activity MRI (TDAM) as a companion diagnostic to predict the TMEM doorway activity (opening) in patients receiving neoadjuvant chemotherapy.[21,22,34] We envision that this approach combined with drugs that can stop TMEM doorway opening will help to decrease cancer cell dissemination, increase survival of patients with locally advanced breast cancer, and potentially aid patients with metastatic disease.

9.6 Future Trends

The field of cancer care has placed great emphasis on providing personalized care for cancer patients. Collectively, we are making important strides towards personalized care, but we are not yet there. Given that most cancer-related deaths are due

to metastases, future efforts should focus on developing individualized prognostic and predictive biomarkers for, as well as therapies that can stop (or at least slow), the metastatic process. If research can make progress along this path, we will start to witness the transformation of metastatic cancer from a fatal disease to a chronic ailment.

Currently, cancer treatments are chosen based on variety of parameters, including breast cancer subtype and the results of prognostic tests, such as Oncotype DX Recurrence Score, which measure a cancer's proliferative ability. In the future, we envision combining proliferation-based prognostics with other non-invasive prognostic approaches, such as TDAM, that measure a cancer's dissemination potential. In addition to its prognostic power, we envision TDAM will enable us to modify treatment in real time, based on its ability to assess not only tumor size, but also the effect of therapy on disease spread. Furthermore, TMEM doorway biology may represent a mechanism of cancer spread common not only to primary breast tumors and their local and distant metastases, but also to multiple other cancers of epithelial and mesenchymal origin. Therefore, TDAM could be optimized as a prognostic and predictive biomarker for use at metastatic sites, and well as in malignancies other than breast cancer. This advance would revolutionize the way we approach the evaluation and treatment of patients with many types of locally advanced and metastatic cancer.

Acknowledgments

This work was supported by NCI grants (R01 CA255153, R01 CA240646, F32 CA243350, and K99/R00 CA237851), the Breast Cancer Research Foundation (BCRF-23–140), SIG OD019961, the Gruss-Lipper Biophotonics Center, the Integrated Imaging Program, the Integrated Imaging Program for Cancer Research, the Evelyn Gruss-Lipper Charitable Foundation, and the Helen and Irving Spatz Family Foundation.

REFERENCES

1. Brown, M. *et al*. Lymph node blood vessels provide exit routes for metastatic tumor cell dissemination in mice. *Science* 359, 1408–1411, doi:10.1126/science.aal3662 (2018).
2. Pereira, E. R. *et al*. Lymph node metastases can invade local blood vessels, exit the node, and colonize distant organs in mice. *Science* 359, 1403–1407, doi:10.1126/science.aal3622 (2018).
3. Borriello, L., Condeelis, J., Entenberg, D. & Oktay, M. H. Breast cancer cell re-dissemination from lung metastases-a mechanism for enhancing metastatic burden. *J Clin Med* 10, doi:10.3390/jcm10112340 (2021).
4. Hoda, S. A., Brogi, E., Koerner, F. C. & Rosen, P. P. *Rosen's Breast Pathology*. Lippincott Williams & Wilkins (LWW) (2020).
5. Rakha, E. A. *et al*. Prognostic significance of Nottingham histologic grade in invasive breast carcinoma. *J Clin Oncol* 26, 3153–3158, doi:10.1200/JCO.2007.15.5986 (2008).
6. Sun, L., Zhu, Y., Qian, Q. & Tang, L. Body mass index and prognosis of breast cancer: An analysis by menstruation status when breast cancer diagnosis. *Medicine (Baltimore)* 97, e11220, doi:10.1097/MD.0000000000011220 (2018).
7. Coughlin, S. S. Social determinants of breast cancer risk, stage, and survival. *Breast Cancer Res Treat* 177, 537–548, doi:10.1007/s10549-019-05340-7 (2019).
8. Kim, G. *et al*. Racial disparity in distant recurrence-free survival in patients with localized breast cancer: a pooled analysis of national surgical adjuvant breast and bowel project trials. *Cancer* 128, 2728–2735, doi:10.1002/cncr.34241 (2022).
9. Sparano, J. A. *et al*. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *N Engl J Med* 380, 2395–2405, doi:10.1056/NEJMoa1904819 (2019).
10. Paik, S. *et al*. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351, 2817–2826, doi:10.1056/NEJMoa041588 (2004).
11. van 't Veer, L. J. *et al*. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530–536, doi:10.1038/415530a (2002).
12. Wallden, B. *et al*. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. *BMC Med Genom* 8, 54, doi:10.1186/s12920-015-0129-6 (2015).
13. Sparano, J. A. *et al*. Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N Engl J Med* 379, 111–121, doi:10.1056/NEJMoa1804710 (2018).
14. Sparano, J. A. *et al*. Prospective validation of a 21-gene expression assay in breast cancer. *N Engl J Med* 373, 2005–2014, doi:10.1056/NEJMoa1510764 (2015).
15. Kalinsky, K. *et al*. 21-Gene assay to inform chemotherapy benefit in node-positive breast cancer. *N Engl J Med* 385, 2336–2347, doi:10.1056/NEJMoa2108873 (2021).
16. Buus, R. *et al*. Molecular drivers of oncotype DX, prosigna, endopredict, and the breast cancer index: a TransATAC study. *J Clin Oncol* 39, 126–135, doi:10.1200/JCO.20.00853 (2021).
17. Karagiannis, G. S., Goswami, S., Jones, J. G., Oktay, M. H. & Condeelis, J. S. Signatures of breast cancer metastasis at a glance. *J Cell Sci* 129, 1751–1758, doi:10.1242/jcs.183129 (2016).
18. Harney, A. S. *et al*. Real-time imaging reveals local, transient vascular permeability, and tumor cell intravasation stimulated by TIE2hi macrophage-derived VEGFA. *Cancer Discov* 5, 932–943, doi:10.1158/2159–8290.CD-15–0012 (2015).
19. Robinson, B. D. *et al*. Tumor microenvironment of metastasis in human breast carcinoma: a potential prognostic marker linked to hematogenous dissemination. *Clin Cancer Res* 15, 2433–2441, doi:10.1158/1078–0432.CCR-08–2179 (2009).
20. Oktay, M. H. & Jones, J. G. TMEM: a novel breast cancer dissemination marker for the assessment of metastatic risk. *Biomark Med* 9, 81–84, doi:10.2217/bmm.14.104 (2015).
21. Rohan, T. E. *et al*. Tumor microenvironment of metastasis and risk of distant metastasis of breast cancer. *J Natl Cancer Inst* 106, doi:10.1093/jnci/dju136 (2014).
22. Sparano, J. A. *et al*. A metastasis biomarker (MetaSite Breast Score) is associated with distant recurrence in hormone receptor-positive, HER2-negative early-stage breast cancer. *Nature PJ Breast Cancer* 3, 42, doi:10.1038/s41523-017-0043-5 (2017).
23. Karagiannis, G. S. *et al*. Neoadjuvant chemotherapy induces breast cancer metastasis through a TMEM-mediated mechanism. *Sci Transl Med* 9, doi:10.1126/scitranslmed.aan0026 (2017).

24. Harney, A. S. *et al*. The selective Tie2 inhibitor rebastinib blocks recruitment and function of Tie2(Hi) macrophages in breast cancer and pancreatic neuroendocrine tumors. *Mol Cancer Therap* 16, 2486–2501, doi:10.1158/1535–7163. MCT-17–0241 (2017).
25. Karagiannis, G. S., Condeelis, J. S. & Oktay, M. H. Chemotherapy-induced metastasis: mechanisms and translational opportunities. *Clin Exp Metastasis*, doi:10.1007/s10585-017-9870-x (2018).
26. Karagiannis, G. S., Condeelis, J. S. & Oktay, M. H. Chemotherapy-induced metastasis in breast cancer. *Oncotarget* 8, 110733–110734, doi:10.18632/oncotarget.22717 (2017).
27. Ginter, P. S. *et al*. Tumor microenvironment of metastasis (TMEM) doorways are restricted to the blood vessel endothelium in both primary breast cancers and their lymph node metastases. *Cancers (Basel)* 11, doi:10.3390/cancers11101507 (2019).
28. Karagiannis, G. S. *et al*. Assessing TMEM doorway-mediated vascular permeability associated with cancer cell dissemination, using intravital imaging and fixed tissue analysis. *Jove* Jun 26;(148), doi:10.3791/59633. doi: 10.3791/59633. PMID: 31305525; PMCID: PMC6784529 (2019).
29. Canene-Adams, K. Preparation of formalin-fixed paraffin-embedded tissue for immunohistochemistry. *Methods Enzymol* 533, 225–233, doi:10.1016/B978-0-12-420067-8.00015-5 (2013).
30. Cardiff, R. D., Miller, C. H. & Munn, R. J. Manual hematoxylin and eosin staining of mouse tissue sections. *Cold Spring Harb Protoc* 2014, 655–658, doi:10.1101/pdb.prot073411 (2014).
31. Bankhead, P. *et al*. QuPath: open source software for digital pathology image analysis. *Sci Rep* 7, 16878, doi:10.1038/s41598-017-17204-5 (2017).
32. Entenberg, D. *et al*. Validation of an automated quantitative digital pathology approach for scoring TMEM, a prognostic biomarker for metastasis. *Cancers (Basel)* 12, doi:10.3390/cancers12040846 (2020).
33. Sanchez, L. R. *et al*. The emerging roles of macrophages in cancer metastasis and response to chemotherapy. *J Leukoc Biol*, doi:10.1002/JLB.MR0218–056RR (2019).
34. Karagiannis, G. S. *et al*. Assessment of MRI to estimate metastatic dissemination risk and prometastatic effects of chemotherapy. *NPJ Breast Cancer* 8, 101, doi:10.1038/s41523-022-00463-5 (2022).
35. Sparano, J. A. *et al*. Race and hormone receptor-positive breast cancer outcomes in a randomized chemotherapy trial. *J Natl Cancer Inst* 104, 406–414, doi:10.1093/jnci/djr543 (2012).
36. Schneider, B. P. *et al*. Impact of genetic ancestry on outcomes in ECOG-ACRIN-E5103. *JCO Precis Oncol* 2017, doi:10.1200/PO.17.00059 (2017).
37. Kabat, G. C., Ginsberg, M., Sparano, J. A. & Rohan, T. E. Risk of recurrence and mortality in a multi-ethnic breast cancer population. *J Racial Ethn Health Disparities* 4, 1181–1188, doi:10.1007/s40615-016-0324-y (2017).
38. Kim, G. *et al*. Racial disparity in tumor microenvironment and distant recurrence in residual breast cancer after neoadjuvant chemotherapy. *NPJ Breast Cancer* 9, 52, doi:10.1038/s41523-023-00547-w (2023).

10

Reading the Reader: Utilizing Eye Movements and Machine Learning to Enhance Accuracy during Diagnostic Visual Search

Gregory J. DiGirolamo, Federico Sorcini, Leo Joskowicz, Yakir Winograd, Assaf Hoog, and Max P. Rosen

10.1 Introduction

Lung cancer is one of the leading causes of cancer death in the US [1], and has more total deaths than breast and prostate cancer combined [2]. Lung cancer also has one of the lowest five-year survival rates of any cancer, at only 23% [2]. Most troubling, more than 80% of patients with lung cancer are diagnosed at an advanced stage [3]. Hence, finding small lung nodules in the early stages is critical for patient survival. The United States Preventive Services Task Force officially recommended yearly computed tomography scan (CT) screening for patients at risk for lung cancer [1]—more than 18 million people in the United States meet this criterion for screening [1].

A central and key limitation of these diagnostic screenings is the radiologist's ability to consciously recognize nodules present in these CT examinations. A major impediment for lung nodule detection is that difficult diagnostic conditions, such as small nodules, can result in many misses, even when experts perform the CT reading. The average miss rate for lung nodule detection can be as high as 50% and has remained at this rate for the last 60 years [4, 5]. The preponderance of radiological errors are misses [6, 7], and most medical malpractice suits stem from these failures to diagnose [8, 9].

A retrospective review by Xu and colleagues [10] of diagnosed lung cancers detected initially by a lung nodule found during annual CT screening reports that 54% of the patients diagnosed had visible nodules in the previous year's screening that had been missed. Hence, conscious misses of nodules in lung CTs during the visual assessment of medical images are a pervasive clinical phenomenon, and reducing these errors is a critical need. One method to improve detection rates is to have a better understanding of how medical providers assess a medical image for an abnormality (diagnostic visual search [DVS]) in radiology and develop alternative means to improve detection of otherwise missed lung nodules.

10.1.1 Diagnostic Visual Search (DVS)

When trying to locate a target within a complex visual scene (like one's keys on a very messy and crowded table, or a nodule in a chest CT), the human brain employs a visual search strategy to bring the eyes to specific portions of the scene that are more likely to contain the target that is being searched for (e.g., the area near where you were seated). Because attention in the human brain is limited [11], not every object within this specific region of the scene will be viewed and processed; rather, a smaller number of particular locations will be looked at and selected for further processing. The likelihood that a specific location in that area is looked at is determined, in part, by the similarity of features between the object in that location and the target you are looking for—your eyes are more likely to be guided to objects that are either metallic, have jagged edges, or are shiny on this part of the table in this lost key scenario.

There are multiple explanations for how these similarity computations are performed. One prominent theory by Wolfe and colleagues [12] suggests that when a scene is first presented, our brain performs an automatic, non-conscious global scan of the entire image, and, using a low-level similarity judgement, selects a set of locations that might contain the target. The eyes are then directed to each of these locations in turn, which *can* allow for further processing that *can* lead to a conscious awareness of the object and a decision as to whether that object at that location (which contains similar properties as the target) is the actual target [13]. With an unlimited amount of time, a static image, and a target being present, the process of visual search will dependably yield the location of the target [14]. However, in complex, dynamic scenes, with limited time, and without the certainty of a target being present, the task becomes much more difficult and the likelihood of a successful search decreases significantly [15].

One such example is DVS in radiology, in which the observer is attempting to locate a target (e.g., a nodule) in a medical image that can be very subtle and yet finding it can be mortally critical. DVS is a complex task for a variety of reasons. The type of abnormality might not be known; more than one abnormality may need to be identified; the features of abnormalities (and indeed, normal anatomy) are variable; and finally, there may be no abnormality present [16]. In practice, there is very rarely an unlimited amount of time for the radiologist to perform the task, preventing a complete and exhaustive search from being performed. In some modalities, such as searching for a lung nodule in a CT scan, the scene is dynamic and requires integration across space, and multiple slices of the lung must be scanned in a short time to determine if an abnormality is present, with many more CTs waiting to be read [17]. These constraints allow less time and efficiency for the brain processes required to analyze the scene and allocate attention (and eye movements) to the locations that share similarities to abnormalities. Given these difficulties present

DOI: 10.1201/9781003449942-12

for successful DVS, it is not surprising that miss rates for radiologists trying to find lung nodules in chest CT examinations are quite high at 30%–50% [18].

While these limitations of DVS are substantial, radiological training clearly enhances these visual search processes. Radiologists tend to accurately locate more abnormalities, make fewer false positive findings, spend less time searching, and are more confident in their decisions than novices [19, 20]. Expertise in DVS displays itself not only in diagnostic performance, but also in how the eyes of experts move across a medical image when trying to locate abnormalities. Experts tend to be more efficient in bringing their eyes to critical locations in the image, display more consistent patterns of eye movements when searching for an abnormality, and direct their eyes to the abnormality faster than novices [19–23]. Nevertheless, even experts miss abnormalities, and uncovering the reasons for these misses is necessary to minimize them [24].

10.1.2 Eye Tracking the Conscious Error

Eye tracking, measuring where and for how long radiologists are looking at a given location, has been a particularly useful tool for determining how errors are made [25]. The most critical of these errors are misses, where an abnormality is present in the scan but the radiologist does not consciously report it. Missed abnormalities have been grouped into three categories, with each being attributed to different underlying causes, but all are categorized as errors [25].

1. *Search errors*: The eyes have never landed on the location of the abnormality, so the observer has not seen the perceptual information at the location of the abnormality.
2. *Recognition errors*: The eyes have landed on the location of the abnormality but have stayed there below some arbitrarily defined duration threshold considered sufficient to process the perceptual information to consciously identify it as a potential abnormality.
3. *Decision errors*: The eyes have landed on the location of the abnormality, have remained at that location above this arbitrarily defined duration threshold considered sufficient to consciously process the perceptual information, but a conscious decision error is made by the radiologist classifying the abnormality as normal anatomy.

The threshold between recognition errors and decision errors is defined by the duration of time the eyes are looking at the nodule and has seen some variation in the amount of time the eyes are looking at the location (ranging from 0.5 to 1 second [26–29]) to determine if the error was a recognition error or a decision error. Using this duration threshold to determine whether a conscious decision was made may not accurately capture if the process was conscious or non-conscious, as the time it takes to become conscious of a visual stimulus varies greatly, but can be as short as ~150 ms [30]. Note that the non-conscious processes of selection of a potential target for an eye movement to that location is even faster, at ~120 ms [31].

Nevertheless, all three of these classifications are considered an error as none of them lead to a conscious recognition of a present abnormality. However, a more nuanced approach that takes the non-conscious processes involved in DVS as separate processes from the processes of conscious recognition would consider these recognition errors as successful non-conscious detection, rather than simply classifying them as a conscious failure. The ability to empirically parse decision errors from recognition errors may allow for isolation of non-conscious processes and determine if they are successful in detecting the abnormality even when conscious recognition fails to occur.

This framework of differentiation of the non-conscious processes from conscious processes as indexed by changes in eye movements during DVS might also allow for more parameters for machine learning (ML) to be used to help improve nodule detection. Rather than just relying on the ML to learn the image, detection accuracy might be improved by using the eye movements of the radiologist during DVS to teach ML to use these non-conscious processes and detect abnormalities that were missed by the conscious processes.

10.1.3 Machine Learning of DVS to Detect Abnormalities

ML has been used extensively in radiology but normally for analyzing images for classification segmentation and anomaly detection [32, 33]. Recently, eye-tracking data (ETD) has been incorporated into ML to analyze and improve the radiologist's performance [34–39]. ETD has been used [38, 39] to train a convolutional neural network (CNN) on the location of the eye of radiologists when viewing a brain scan and found that ML segmentation of the anatomy using ETD was equal to hand annotation. The importance of adding ETD to ML to improve radiologist performance is highlighted by a *Nature* article [35] that reports the creation of a database for ETD and chest X-rays to be used to train models. Mall and colleagues [37] used ETD to train a CNN to accurately predict the radiologist's diagnostic decision, confidence of decision, and location of attention in mammography DVS. However, this model was applied only to the conscious report of the radiologist.

Recently, using ETD to train ML has been applied to lung nodule detection [34, 36]. One study [36] combined a CNN ML of the CT image with ETD and achieved very high nodule detection accuracy. Aresta and colleagues [34] also used ETD and CT image analysis with a CNN and were able to train a model with a similar nodule sensitivity of 0.69 compared to 0.67 for the radiologists. However, for the ML to achieve this high nodule detection rate, it required a high false positive rate of 0.34 (matching that of the radiologists). In the studies [34–39] to date that have used ETD and ML, they have used CNN ML of the CT image as well as the ETD. No modeling to date has parsed the conscious and non-conscious ETD, and none have used ensemble learning ML methods exclusively on the ETD.

We hypothesize that separating the conscious from non-conscious ETD will lead to better training of the ML as these factors will not be confounded in model training and hence the ensemble model of the ETD will have good classification performance [40].

The purposes of the present study were (1) to determine if non-conscious processes are successfully detecting the abnormality even when conscious recognition fails and (2) to assess whether these non-conscious processes, as indexed by eye movement changes, can be used to train ensemble ML to detect abnormalities that were missed by conscious recognition processes.

10.2 Materials

10.2.1 Participants

A total of six radiologists (five males), who were currently licensed and practicing radiologists, from the University of Massachusetts Memorial Medical Centers' Department of Radiology were recruited. All participants gave informed written consent. The protocol was approved by the IRB of our institution.

10.2.2 Stimuli and Apparatus

CT examinations were obtained from a database of lung nodules that were missed during initial interpretation in the clinic. The nodules in these abnormal images were all solid and round, with well-defined borders, and an average diameter of 6.7 mm (SD = 6.8 mm, range = 27 mm). They were present in both lungs, in the upper, lower, and middle sections of lung. A four-radiologist consensus panel confirmed the presence and location of these nodules in the abnormal CT examinations and ensured that normal CT examinations did not contain lung nodules. There were nine abnormal CT examinations (with 16 total nodules) and nine normal CT examinations with, on average, 282 slices (SD = 50) with a slice thickness of 3 mm. For the purposes of accurate eye tracking, each CT slice was presented on a 21" monitor set at 1920 × 1080 resolution without the ability to change the resolution, maximize the intensity, or magnify the image. Clinical information for each CT was not provided to the reader.

Eye movements and pupil size were recorded with an SR Research EyeLink 1000 high-speed video eye tracker with 1000 Hz sampling and ~0.1° of resolution [41, 42]. Prior to the start of the experiment, after every CT examination, as well as after every break, eye position was calibrated to ensure accurate tracking of the eye throughout the experiment.

10.3 Methods

Participants were permitted unlimited time to read each CT examination. Detection of solid lung nodules was the sole task. Participants were specifically instructed to disregard any other abnormalities that might be present, for example emphysema. Each participant's head rested in a chin rest placed on the edge of the table. Participants had the opportunity to take short breaks in between each CT examination and those who decided to do so were recalibrated prior to the start of the next CT examination. Instructions for the task were given to the participant through a display screen. After reading the instructions, the experimenter ensured the participant understood the task before beginning. Participants scrolled through stacks of CT slice images using the right and left arrow keys of the keyboard. Participants indicated a lung nodule by left mouse clicking on it, with click locations compared to the known locations of the 16 lung nodules to compute accuracy.

In this experiment, participants were also instructed to right click on anything they even momentarily considered to be a lung nodule but then decided it was not. We instructed the participants that they should right click on any area that they even momentarily considered might include a nodule, no matter how briefly. These considerations were stressed to ensure that participants clicked on any location they consciously considered as a nodule but then decided it was not no matter how briefly (see Figure 10.1).

Hence, in this protocol, decision errors are when the radiologist looks at a nodule and right clicks on it (signaling they thought it may have been a nodule but changed their mind). Recognition errors would be when the radiologist neither right nor left clicked on a location that contained a nodule, even though they looked directly at it.

Radiologists received one practice trial prior to experimental data recording. When the radiologist had finished viewing each CT examination and had located the presence of any nodules in that examination, they pressed the space bar to end the trial and move on to the next CT examination.

10.3.1 Data Analysis for Eye Movements

To analyze the location of each radiologist's eye position relative to the lung nodules in the study, we superimposed, for data analyses only, a 12 × 12 grid entailing 144 locations on each CT slice of 85 × 85 pixels each. Each eye fixation corresponded to a box in the grid on the screen, which was not visible to the viewer. Each of the identified parenchymal lung nodules was located and marked prior to testing according to its corresponding grid location.

Using the grid in the experiment and the nodule location, we compared dwell times and pupil size (as pupil size is a direct measure of physiological arousal [41, 43]) of each radiologist's eye-tracking data to the nodule location for our analyses. The locations of mouse clicks in the analysis grid were also used to determine whether the radiologist had consciously located or considered the nodule. Dwell time and pupil diameter for locations where nodules were "missed" and locations on normal CTs (no nodules) were compared with paired *t*-tests. If non-conscious processes were capable of detecting nodules that are neither recognized nor considered, we would expect significantly longer dwell times and pupil size for the "missed" nodules compared to normal tissue.

10.3.2 Machine Learning

10.3.2.1 Ensemble Machine Learning

In machine learning, ensemble learning algorithms stand out, and this approach has not been used previously with ETD and non-conscious processes during DVS. The premise of the ensemble models is to improve model performance by

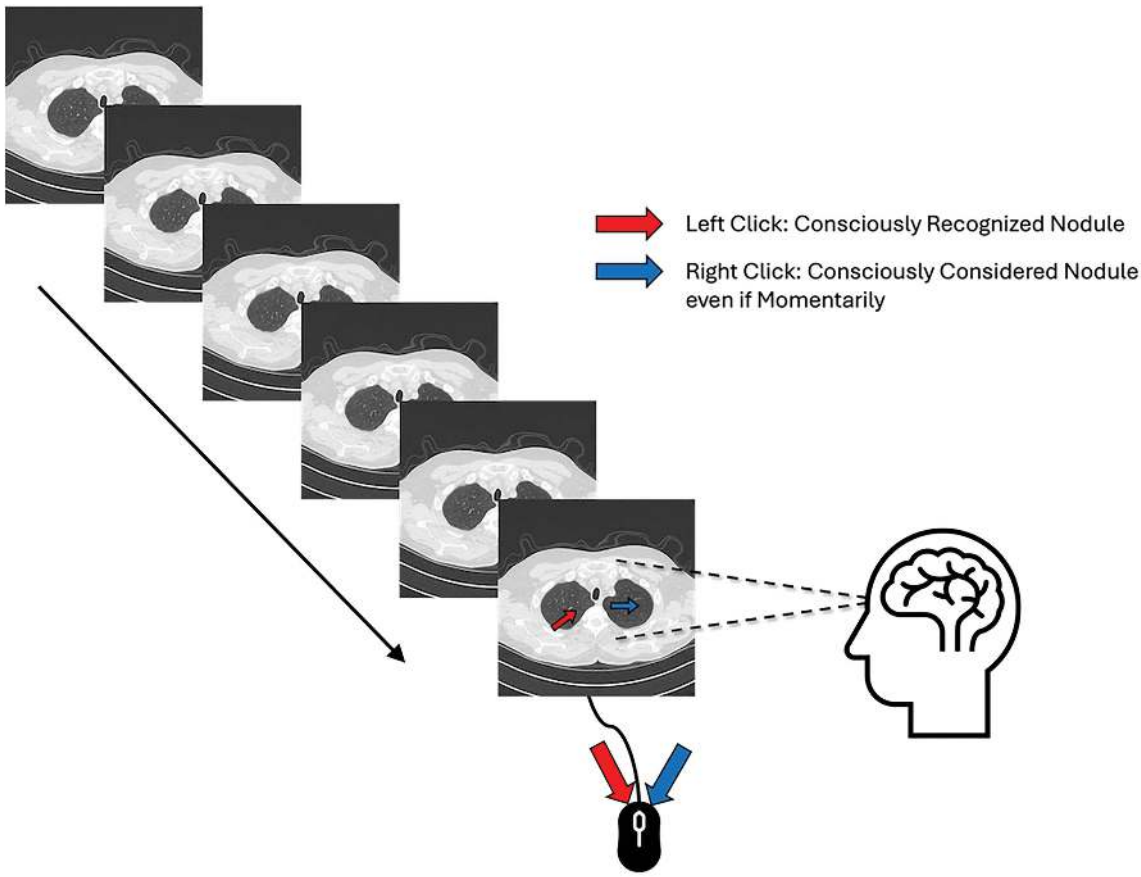


FIGURE 10.1 Radiologists indicated the location of a nodule with a left mouse click and the location of any region that they even momentarily considered as a nodule with a right mouse click.

combining the predictions from different models to reduce error.

Light gradient boosting machine classifier (LightGBM) is the ensemble learning algorithm used in the present study. LightGBM incorporates boosting that involves training individual models sequentially, where each model attempts to correct the errors made by the previous models. It uses decision tree algorithms for classification. LightGBM classifier uses a technique of gradient-based one-side sampling (GOSS) and exclusive feature bundling (EFB) to handle large-scale data with accuracy. LightGBM's GOSS performs random sampling on the instances with small gradients. GOSS introduces a constant multiplier for the data instances with small gradients to compensate for the information loss during sampling. EFB is a near-lossless algorithm that combines mutually exclusive features (to reduce the number of dimensions) to accelerate training. Since these features are "exclusive," the original feature space is retained without significant information loss.

Ensemble deep neural networks harness the collective power of multiple deep learning models to achieve superior performance in various tasks. One of their primary strengths lies in their ability to mitigate the weaknesses of individual models. Deep neural networks are prone to overfitting, especially when dealing with complex datasets or limited training data. Ensemble methods, such as bagging and boosting, help address this issue by aggregating the predictions of multiple models, thereby reducing variance and improving generalization. Additionally, ensemble deep nets often exhibit robustness and resilience to noisy data or outliers. By leveraging the diversity of individual models within the ensemble, they can effectively filter out irrelevant information and focus on learning the underlying patterns in the data. This characteristic makes them particularly valuable in real-world scenarios where data may be incomplete. Furthermore, ensemble models can capture complex relationships and nonlinearities in the data, enabling them to model intricate structures with high accuracy. As a result, they are well suited for tasks that demand high precision and reliability, such as medical diagnosis. Hence, we used this LightGBM ensemble ML in our current study. LightGBM was created using the Microsoft Python application programming interface.

10.3.2.2 Features of Model

Seven features were used for training the ML model: (1) dwell time: how long the radiologist looked at a specific location; (2) pupil dilation: the size of the radiologist's pupil when looking at that location; (3) fixation component duration: the duration of the current fixation split by the amount of time spent on each slice; (4) fixation index: the current fixation number

within the current trial; (5) fixation grid location: the grid location of the current fixation; (6) fixation component count: the number of CT slices that the current fixation lasts for; and (7) fixation component index: the current fixation component number within each fixation. While dwell time has been used in other ML with ETD works, we are uniquely using physiological arousal indexed by pupil diameter and isolating non-conscious processes in the ETD.

The model was evaluated with optimized hyperparameters, using the single parameter range search method, detailed later. The data set consists of 125,262 target data rows from six radiologists. The training consisted of 80% of the data randomly chosen, with the remaining 20% used for testing. This process was repeated ten times in a cross-validation framework. We evaluated the true positive, false positive, true negative, and false negative results.

10.3.2.3 Data Pre-Processing

Normalization is a crucial step in data pre-processing, particularly when dealing with features that exhibit varying scales. Fixation index, pupil size, and fixation duration were normalized to a range between 0 and 1. This normalization process ensures that each feature contributes proportionally to the model training, regardless of its original scale.

Choosing the right hyperparameters is crucial for optimal model performance. These parameters help the model effectively understand the task without merely memorizing the training data, thus preventing overfitting. In our study, we tuned the following parameters using a method called single parameter range searches. This approach involves optimizing one parameter at a time by testing a wide range of its values and selecting the value that yields the best results on test data. We repeated this process for multiple parameters and iteratively performed it several times until no further improvements in performance were observed. The optimized hyperparameters included:

- Number of leaves—the maximum number of terminal nodes (leaves) in each decision tree. A higher number of leaves can capture more detailed data patterns but may lead to overfitting. The final used value was 300 (range of 10–350 in increments of 1).
- Learning rate—determines the impact of each tree on the final outcome. A lower rate requires more trees but can result in better generalization. The final used value was 0.1 (range of 0.01–0.8 in increments of 0.01).
- Minimum "children" samples—the minimum number of samples required to split an internal node. Setting this parameter helps control the depth of the trees, influencing both performance and overfitting. The final used value was 44 (range of 1–50 in increments of 1).
- Number of estimators—the total number of trees built before taking the average of predictions. More trees can improve accuracy but also increase computational cost. The final used value was 100 (range of 50–200 in increments of 10).
- Class weights—adjustments to the model to correct for imbalances in the training data, ensuring that each class is treated appropriately during learning. The final "balanced" used value was 171.3 (range of 1–300 in increments of 1) for the true label and 1 for the false label, equalizing the weight of the true and the false labels.

The goal was to evaluate whether ensemble machine learning (LightGBM) could use the non-conscious processes as indexed by changes in eye movements during DVS to determine the location of consciously missed lung nodules. Importantly, we removed all eye-tracking data where the radiologist had registered a mouse click; hence, any ETD where the readers consciously reported a nodule or consciously considered a nodule were removed, since our goal was to predict the location of the missed lung nodule in the CT scans using only the non-conscious processes. Targets (marked as 1) for classification for the model were fixations on the nodule where the radiologists' eyes were looking directly at the nodule even though they had not consciously reported or consciously considered. All other fixations to non-nodule locations were non-targets (marked as 0).

10.4 Results

As is common in harder-to-detect lung nodules, radiologists missed many of these nodules (M = 59%). Radiologists successfully detected, on average, 6.5 of the 16 nodules and made more than double (M = 15.8) the number of false positive errors. Finally, radiologists consciously considered (i.e., right mouse clicked) more than 41 locations on average across the 18 CT scans where they, even if only momentarily, thought there might have been a nodule (on average, three of these locations were nodules on which the radiologists made decision errors).

Most important is the comparison between the dwell times when looking at nodules that were neither consciously recognized (left mouse clicked) nor consciously considered (right mouse clicked), and the dwell times when looking at normal lung tissue. In this direct comparison, the radiologists have not mouse clicked in either CT slice and are classifying both normal slices and slices with missed nodules as being normal. As Figure 10.2 shows, radiologists looked significantly longer ($t(5) = 5.38$, $p < 0.004$) at the locations of the nodules than on normal lung locations. Further, as Figure 10.3 shows, every radiologist showed this phenomenon, as it was highly reliable and robust (Cohen's D = 2.20).

In addition, we also measured changes in pupil size for missed lung nodules and normal tissue. Changes in pupil size are reflective of physiological arousal and correlate with heart rate and skin conductance [41, 44]. Radiologists also demonstrated (see Figure 10.4) significantly greater arousal to the locations of nodules that were neither consciously recognized nor consciously considered compared to locations with normal tissue, with $t(5) = 4.98$, $p < 0.005$. As Figure 10.5 demonstrates, every radiologist showed this phenomenon. Again, it was highly reliable and robust (Cohen's D = 2.03).

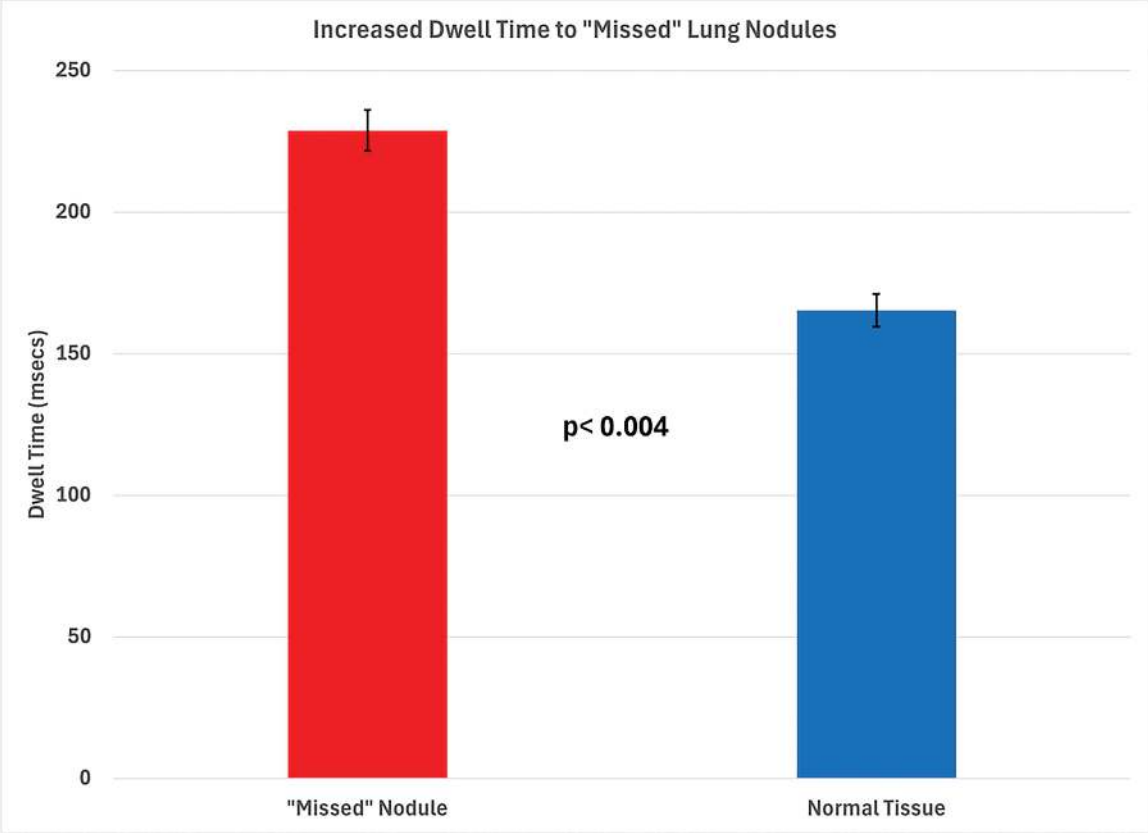


FIGURE 10.2 Increased looking time for the location of nodules that were neither consciously recognized nor considered compared to looking time at normal tissue.

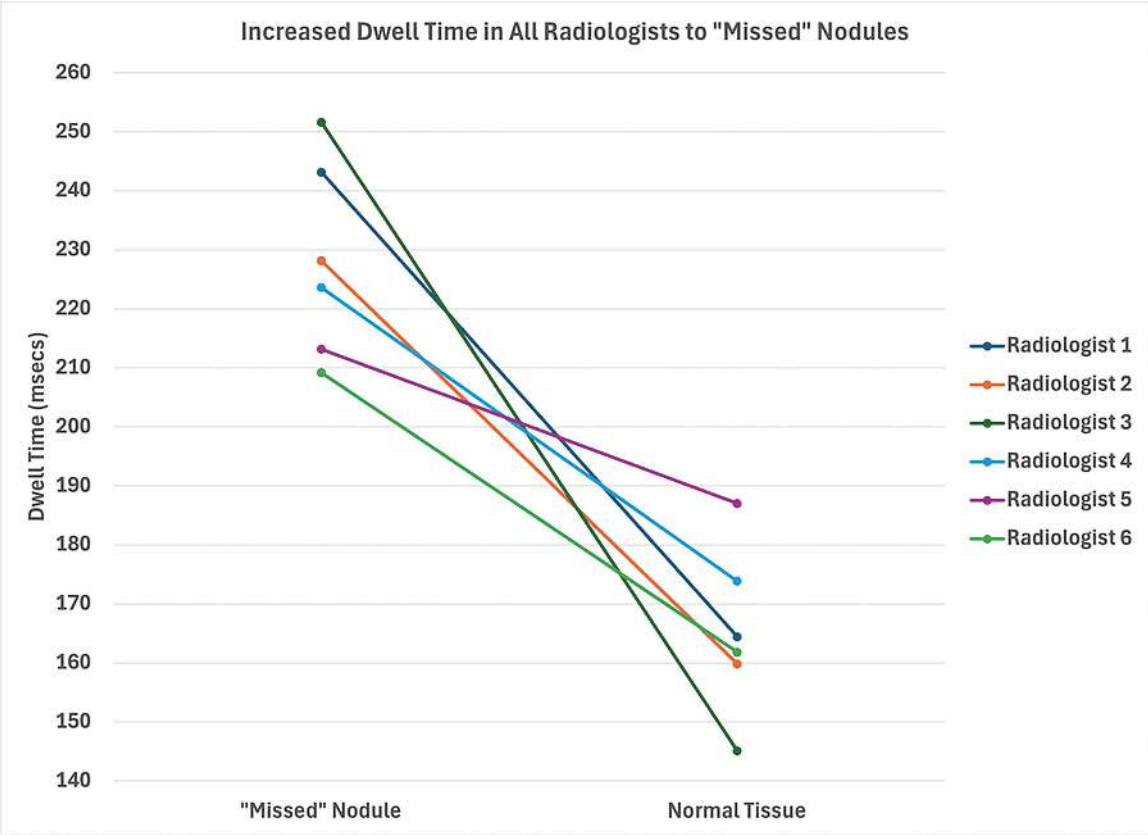


FIGURE 10.3 All radiologists showed longer looking times for nodules that were neither consciously recognized nor considered compared to normal tissue.

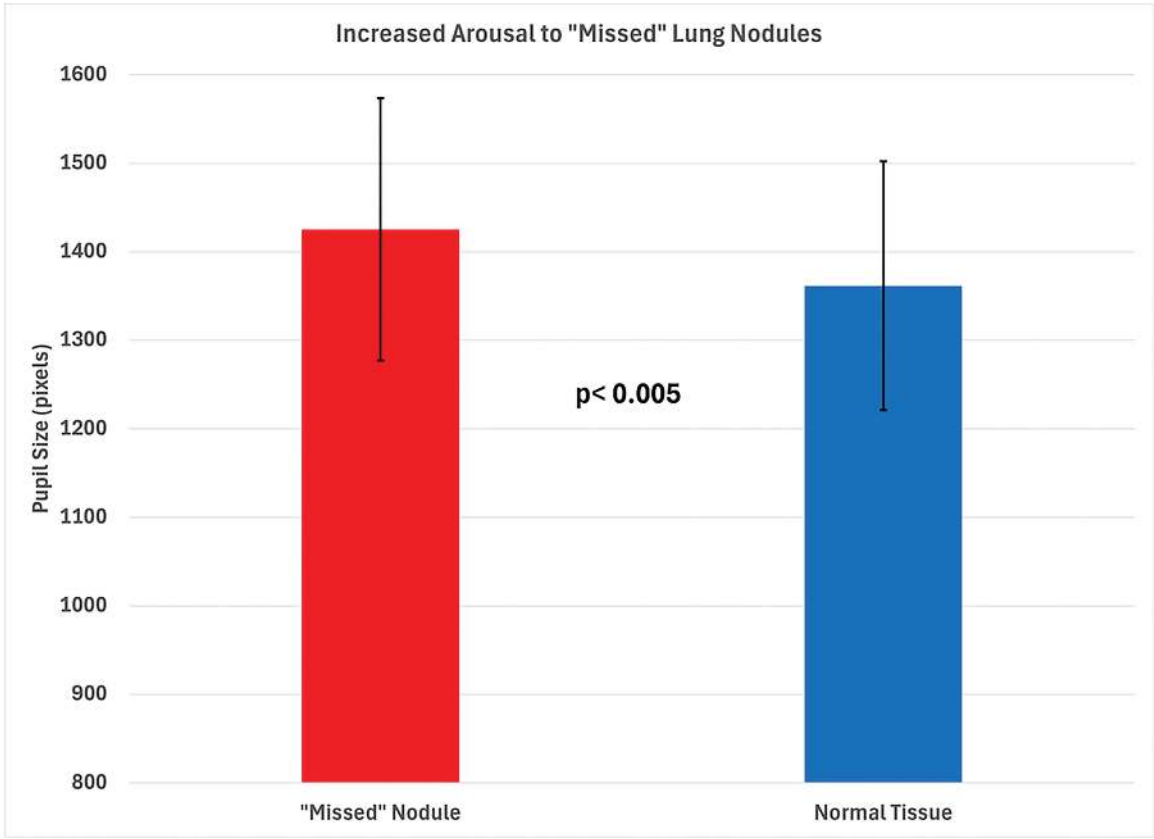


FIGURE 10.4 Increased arousal for the location of nodules that were neither consciously recognized nor considered compared to looking time at normal tissue.

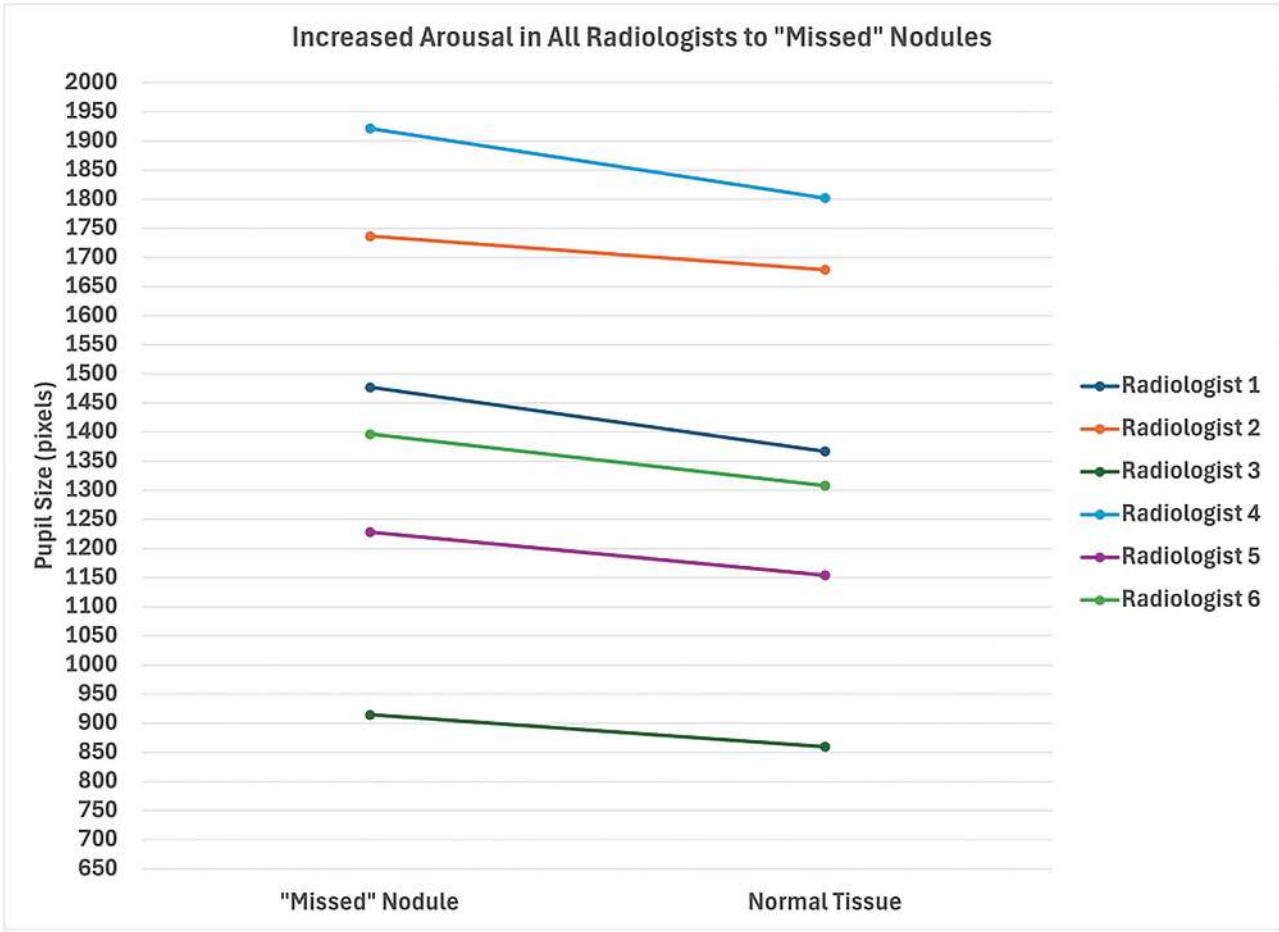


FIGURE 10.5 All radiologists showed increased arousal for nodules that were neither consciously recognized nor considered compared to normal tissue.

Non-conscious processes successfully detecting the nodule is startling; but unless these processes can be used also to successfully increase nodule detection, they have no practical application. Hence, we asked if these non-conscious processes, as indexed in the eye-tracking data (ETD), could be used with ensemble machine learning to determine the location of the missed nodules. Importantly, we removed all ETD where the radiologist had registered a mouse click; hence, any ETD where they consciously reported a nodule or consciously considered a nodule was excluded. Our goal was to create a computational model to predict the location of the missed lung nodule in the CT examinations using only the non-conscious processes as indexed in the ETD. Note that the model was trained on the ETD and neither the CT image information nor the radiologists' conscious reports was used. Hence only ETD from nodules that were missed and ETD from normal CT examinations were used.

The ensemble ML did extremely well (see Table 10.1) with a mean detection rate of missed nodule locations of 66%, with a mean of 0.1% false positives. Note that the radiologist had consciously determined that all these CT slices and locations were normal and did not contain a nodule or a location the radiologist even considered as a nodule. Yet, non-conscious processes as indexed in the ETD can be used to train ML to detect the missed nodules. These models showed good sensitivity (0.66) and with excellent specificity (0.99). Hence, our ensemble LightGBM could find a majority of the locations of nodules that the radiologist had consciously missed using only the eye-tracking data without use of the image or the radiologist's conscious report.

TABLE 10.1

Confusion Matrix of ML of Missed Nodules

	Predicted Nodule	**Predicted Normal**
True Nodule	0.66	0.34
True Normal	0.001	0.99

It is also critical to realize that this ML model is finding the exact location on a 12 × 12 grid on the CT image where the nodule is located (see Figure 10.6) and is doing so using solely the eye-tracking data of the non-conscious processes. It is not simply determining a nodule is present somewhere in the CT or somewhere on a slice of the CT image. Rather, the LightGBM model is determining the location of the nodule in the CT image from these non-conscious processes indexed in the eye movements that the radiologist had looked at and missed.

10.5 Discussion and Conclusion

Our experimental results demonstrate two important points. First, there are a significant number of locations during DVS where radiologists consider there might be a nodule but decide against it [45], and these processes are clearly conscious as radiologists are able to report them via mouse click. Second, when a nodule is present in the CT and the radiologist neither consciously detects it nor considers it, their brains are still detecting the nodule, as indexed by longer looking times and greater arousal at the location of the nodule than normal tissue. The disconnect between the conscious process as reported by the radiologist (via mouse clicks) and these non-conscious processes during DVS are telling. In both cases, the radiologist has consciously identified each of these CT slices as being normal even though one contains a missed nodule that was not consciously recognized. And yet, independent of these erroneous conscious processes are clear indications of successful non-conscious detection of the lung nodule as indicated in the eye movements by longer looking times and pupil diameter increases.

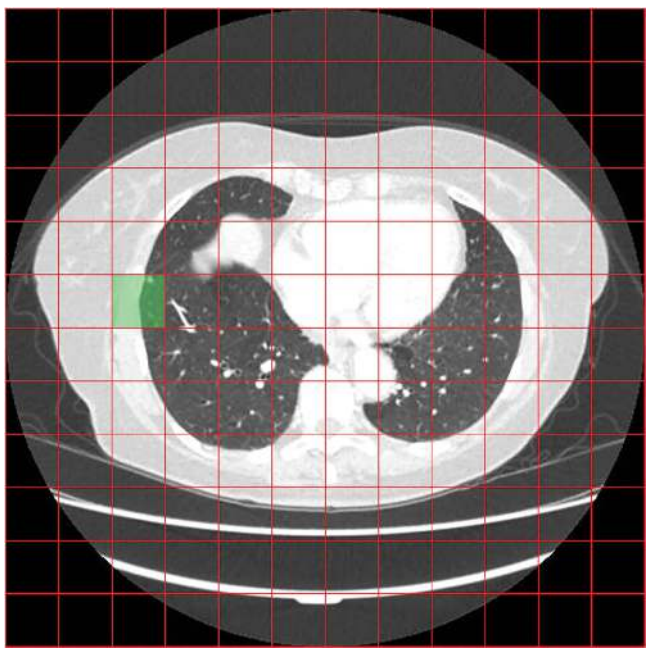


FIGURE 10.6 Example of a "missed" nodule located by the ensemble machine learning using the non-conscious eye-tracking data.

This phenomena of "looking without seeing" [46] has, until now, been seen as the non-conscious precursor to a conscious error: what the literature has termed a *recognition error* [20, 23, 25, 47–51]. We are suggesting a reinterpretation of this explanation and to consider this process as a successful finding of the nodule by the brain, which happens to be non-conscious. The eye movement data indicates clear and reliable biomarkers of these successful non-conscious processes, and these biomarkers were present in every radiologist because of their training.

When CT scans with known nodule locations are shown to radiologists during empirical studies, their miss rate has been consistent since the 1960s at approximately 30% [4, 5]. Advancements in both imaging technology and training have not translated into concomitantly improved detection accuracy. The lack of improved performance might occur because of the limitations in conscious processes during DVS. But the current data suggest that radiologists' brains are detecting far more nodules than they can consciously report.

These results are consistent with previous work showing non-conscious processes (as indexed by eye movement) succeeding in the absence of conscious success [52–55]. When shown a group of letters that spell out a word, subjects move their eyes to the correct letters in the correct order to spell the word even when they can't report the word [52]. In non-diagnostic visual search, subjects' eyes will move to a target even when the subjects don't report the presence of the target or have any awareness that their eyes moved to it [54]. These data demonstrate that non-conscious perceptual processes are frequently successful in the absence of conscious success on the task.

Moreover, these successful non-conscious processes have proven useful in training machine learning to detect nodules that radiologists would otherwise miss. Using ensemble ML of these non-conscious processes, our model was able to detect 66% of the times when the radiologists' eyes were on a nodule that they failed to consciously detect or consider. These results are consistent with previous CNN modeling of nodule detection using eye movements [34], though those data included both conscious and non-conscious eye-tracking data as well as ML analysis of the CT image itself.

It is important to point out that this model by Aresta and colleagues [34] actually obtained a higher nodule sensitivity (0.69) than our current model. But this previous model also had a much higher high false positive rate (0.34) than our current model which rarely (0.001) called a normal image abnormal. The current data are suggestive of slightly less sensitivity using solely the non-conscious processes indexed in the eye movements, but with very high specificity. These data indicate a means of improving detection of nodules using eye tracking and ML by tapping into the non-conscious processes that are successfully finding the nodule in the absence of conscious recognition.

We believe that this discovery of successfully non-conscious processes that can be used to train ML to detect otherwise missed nodules finally allows a breaking of the barrier of the limitations of conscious recognition in medical DVS that has impeded nodule detection and impaired health outcomes for the last 60 years.

During diagnostic visual search for a potentially cancerous lung nodule, the human brain employs both conscious and non-conscious processes to detect the abnormality. Current methods of detecting a lung nodule have relied solely on the conscious recognition of the nodule by the trained radiologist. Unfortunately, this conscious process frequently fails [4, 5, 10], and in lung cancer, the failure to detect a nodule has disastrous consequences. The majority of detection of lung nodules occurs at more advanced stages, when treatment is less successful. Earlier detection of smaller nodules would greatly improve outcomes, but this process frequently fails because of the limits of conscious recognition in DVS. Here we show that these limitations might be overcome by using eye-tracking and ensemble ML to use the successful non-conscious processes of the radiologist's highly trained brain to detect lung nodules that would have been otherwise missed by the radiologist. It is our hope and goal to improve lung nodule detection rates using this innovative approach to improve lung cancer survival rates.

10.6 Future Trends

Tapping into the full training of the radiologist's brain by harnessing the non-conscious process in diagnostic visual search using eye tracking conveys several paths for future improvements in nodule detection. With lung cancer being the number one cause of cancer deaths in the United States [2] and the initial presentation as a pulmonary nodule, it is not surprising that several very good ML methods for analyzing medical images and detecting nodules have been developed [56]. These tools rely on the ML being trained on specific medical images (e.g., radiograph) to find specific abnormalities (e.g., lung nodule) while other ML models are necessary if the imaging type (e.g., CT) or abnormality change. Hence separate ML models are required for trying to detect lung nodules in CT vs radiography. These ML models can be very useful for flagging potential nodules that then must be viewed by the radiologist to determine if the flagged location is actually a nodule. These models are less effective at detecting nodules that were missed; with one ML model detecting one out of ten missed nodules [56]. In contrast, the ML of the non-conscious processes in the eye movements of the radiologist allows for 66% detection of missed lung nodules.

The first path for improving detection of lung nodules is a disruption of standard methods of ML. Rather than training ML to read the images, we suggest training ML to "read the reader." Our method of using ensemble ML of the radiologist's eye-tracking data has led to the detection of more missed nodules than previous ML. This disruption of ML to "reading the reader" takes advantage of the extensive training of the radiologist and taps into the full expertise of their highly skilled and experienced brain to detect nodules that are missed by the conscious processes. Even using standard ML of the image to detect lung nodules relies on the radiologist's conscious detection to confirm the presence of a lung nodule. And, as we and others have shown, this conscious process fails frequently. Hence, using ML of the non-conscious processes in the eye-tracking data can overcome these limitations of conscious detection that are present in other ML methods of detecting nodules.

Further, since this "reading of the reader" ML relies on the eye-tracking data and not the image, the current ML presented in this chapter might be able to learn to detect abnormalities independent of the image or abnormality creating a single general abnormality detection ML model. Given the expense and time to develop ML for different modality types and different abnormalities, a single generalized ML model capable of detecting abnormalities in any image would be a sizable benefit. Using ML of the eye-tracking data of the radiologist suggests a path to create this single generalizable ML model for abnormality detection.

Finally, one more path to increase nodule detection might be the creation of *personalized AI* for each radiologist. Using solely an individual radiologist's own eye movement patterns when detecting potentially cancerous lung nodules to train the ML would create a personalized ML model for that radiologist that could excel at detecting lung nodules beyond the general model presented in this chapter. This unique ability to create individualized AI for each radiologist to detect lung nodules increases the likelihood of increased detection by taking advantage of the specific patterns of looking. A curated AI for abnormality detection using the non-conscious processes of the individual radiologist opens the door for further improvement in abnormality detection beyond lung nodules.

All of these potential paths of using the non-conscious processes involved in diagnostic visual search to detect abnormalities have one specific goal: increase early detection of cancer to increase survival rates. This chapter describes the first step down this path to reach this goal.

REFERENCES

1. Force, U.P.S.T., Screening for lung cancer: us preventive services task force recommendation statement. *JAMA*, 2021. 325(10): p. 962–970.
2. Siegel, R.L., et al., Cancer statistics 2023. *CA Cancer J Clin*, 2022. 73: p. 17–48.
3. Gould, M.K., Clinical practice. Lung-cancer screening with low-dose computed tomography. *N Engl J Med*, 2014. 368: p. 1980–1991.
4. Rubin, G.D., et al., Characterizing search, recognition, and decision in the detection of lung nodules on CT scans: elucidation with eye tracking. *Radiology*, 2015. 274(1): p. 276–286.
5. Manning, D., S. Ethell, T. Donovan, and T. Crawford, How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*, 2006. 12(2): p. 134–142.
6. Bruno, M.A., E.A. Walker, and H.H. Abujudeh, Understanding and confronting our mistakes: the epidemiology of error in radiology and strategies for error reduction. *RadioGraphics*, 2015. 35(6): p. 1668–1676.
7. Berlin, L., Malpractice and radiologists, update 1986: an 11.5-year perspective. *Am J Roentgenol*, 1986. 147: p. 1291–1298.
8. Berlin, L., Medicolegal-malpractice and ethical issues in radiologyfaster radiologic interpretation, errors, and malpractice: an unavoidable triad? *Am J Roentgenol*, 2018. 210(2): p. W92–W93.
9. Siegal, D., L.M. Stratchko, and C. DeRoo, The role of radiology in diagnostic error: a medical malpractice claims review. *Diagnosis (Berl)*, 2017. 4(3): p. 125–131.
10. Xu, D.M., et al., Retrospective review of lung cancers diagnosed in annual rounds of CT screening. *Am J Roentgenol*, 2014. 203: p. 965–972.
11. Posner, M.I. and G.J. DiGirolamo, Cognitive neuroscience: origins and promise. *Psychol Bull*, 2000. 126(6): p. 873.
12. Wolfe, J.M., et al., Visual search in scenes involves selective and nonselective pathways. *Trends Cogn Sci*, 2011. 15(2): p. 77–84.
13. Wolfe, J.M. and T.S. Horowitz, What attributes guide the deploymment of visual attention and how do they do it? *Nat Rev Neurosci*, 2004. 5: p. 1–7.
14. Treisman, A.M. and G. Gelade, A feature-integration theory of attention. *Cogn Psychol*, 1980. 12: p. 97–136.
15. Wolfe, J.M., et al., How do radiologists use the human search engine? *Radiat Prot Dosimetry*, 2015: p. ncv501.
16. Wolfe, J.M., et al., Low target prevalence is a stubborn source of errors in visual search tasks. *J Exp Psychol Gen*, 2007. 136(4): p. 623–638.
17. Bhargavan, M. and J.H. Sunshine, Workload of radiologists in the United States in 2002–2003 and trends since 1991–1992. *Radiology*, 2005. 236(3): p. 920–931.
18. Berlin, L., Accuracy of diagnostic procedures: has it improved over the past five decades? *Am J Roentgenol*, 2007. 188(5): p. 1173–1178.
19. Bertram, R., et al., Eye movements of radiologists reflect expertise in CT study interpretation: a potential tool to measure resident development. *Radiology*, 2016. 281(3): p. 805–815.
20. Manning, D., et al., How do radiologists do it? The influence of experience and training on searching for chest nodules. *Radiography*, 2006. 12(2): p. 134–142.
21. Kelly, B.S., et al., The development of expertise in radiology: in chest radiograph interpretation, "expert" search pattern may predate "expert" levels of diagnostic accuracy for pneumothorax identification. *Radiology*, 2016. 280(1): p. 252–260.
22. Nodine, C.F., et al., How experience and training influence mammography expertise. *Acad Radiol*, 1999. 6(10): p. 575–585.
23. Drew, T., et al., Scanners and drillers: characterizing expert visual search through volumetric images. *J Vision*, 2013. 13(10): p. 1–13.
24. Berlin, L., Errors of omission. *Am J Roentgenol*, 2005. 185: p. 1416–1421.
25. Kundel, H.L., C.F. Nodine, and D. Carmody, Visual scanning, pattern recognition and decision-making in pulmonary nodule detection. *Invest Radiol*, 1978. 13(3): p. 175–181.
26. Kundel, H.L., C.F. Nodine, and E.A. Krupinski, Searching for lung nodules. Visual dwell indicates locations of false-positive and false-negative decisions. *Invest Radiol*, 1989. 24(6): p. 472–478.
27. Donovan, T. and D. Litchfield, Looking for cancer: expertise related differences in searching and decision making. *Appl Cognit Psychol*, 2012. 27(1): p. 43–49.
28. Manning, D.J., S. C. Ethell, and T. Donovan, Detection or decision errors? Missed lung cancer from the posteroanterior chest radiograph. *Br J Radiol*, 2004. 77: p. 231–235.

29. Mello-Thoms, C., et al., Effects of lesion conspicuity on visual search in mammogram reading. *Acad Radiol*, 2005. 12(7): p. 830–840.
30. Thorpe, S., D. Fize, and C. Marlot, Speed of processing in the human visual system. *Nature*, 1996. 381: p. 520–522.
31. Kirchner, H. and S.J. Thorpe, Ultra-rapid object detection with saccadic eye movements: visual processing speed revisited. *Vision Res*, 2006. 46: p. 1762–1776.
32. Bi, W.L., et al., Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin*, 2019. 69(2): p. 127–157.
33. Xu, X., et al., DeepLN: a framework for automatic lung nodule detection using multi-resolution CT screening images. *Knowl Based Syst*, 2020. 189: p. 105128.
34. Aresta, G., et al., Automatic lung nodule detection combined with gaze information improves radiologists screening performance. *IEEE J Biomed Health Inf*, 2019. 24(10): p. 2894–2901.
35. Karargyris, A., et al., Creation and validation of a chest X-ray dataset with eye-tracking and report dictation for AI development. *Sci Data*, 2021. 8(1): p. 92.
36. Khosravan, N., et al., A collaborative computer aided diagnosis (C-CAD) system with eye-tracking, sparse attentional model, and deep learning. *Med Image Anal*, 2019. 51: p. 101–115.
37. Mall, S., P.C. Brennan, and C. Mello-Thoms, Modeling visual search behavior of breast radiologists using a deep convolution neural network. *J Med Imaging*, 2018. 5: p. 035502.
38. Stember, J.N., et al., Integrating eye tracking and speech recognition accurately annotates MR brain images for deep learning: proof of principle. *Radiol Artificial Intell*, 2020. 3(1): p. e200047.
39. Stember, J.N., et al., Eye tracking for deep learning segmentation using convolutional neural networks. *J Digital Imaging*, 2019. 32(4): p. 597–604.
40. Wang, H., Z. Wu, and E.P. Xing, Removing confounding factors associated weights in deep neural networks improves the prediction accuracy for healthcare applications. *bioRxiv*, 2018: p. 442442.
41. DiGirolamo, G.J., N. Patel, and C.L. Blaukopf, Arousal facilitates involuntary eye movements. *Exp Brain Res*, 2016. 234(7): p. 1967–1976. doi: 10.1007/s00221-016-4599-3. Epub 2016 Feb 29. PMID: 26928432; PMCID: PMC4893927.
42. DiGirolamo, G.J., et al., Breakdowns of eye movement control toward smoking cues in young adult light smokers. *Addict Behav*, 2016. 52: p. 98–102.
43. Bradley, M., et al., The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 2008. 45(4): p. 602–607.
44. Bradley, M.M., et al., The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, 2008. 45: p. 602–607.
45. Chakraborty, D.P. and K.S. Berbaum, Observer studies involving detection and localization: modeling, analysis, and validation. *Med Phys*, 2004. 31(8): p. 2313–2330.
46. Wolfe, J.M., A. Kosovicheva, and B. Wolfe, Normal blindness: when we look but fail to see. *Trends Cogn Sci*, 2022. 26(9): p. 809–819.
47. Krupinski, E.A., et al., Eye-movement study and human performance using telepathology virtual slides. Implications for medical education and differences with experience. *Hum Pathol*, 2006. 37(12): p. 1543–1556.
48. Kundel, H.L., Visual search and lung nodule detection on CT scans. *Radiology*, 2015. 274(1): p. 14–16.
49. Nodine, C.F. and H.L. Kundel, Using eye movements to study visual search and to improve tumor detection. *Radiographics*, 1987. 7(6): p. 1241–1250.
50. Rubin, G.D. and E.A. Krupinski, Tracking eye movements during CT interpretation: inferences of reader performance and clinical competency require clinically realistic procedures for unconstrained search. *Radiology*, 2017. 283(3): p. 920.
51. Llewellyn-Thomas, E. and E.L. Lansdown, Visual search patterns of radiologists. *Radiology*, 1963. 81: p. 288–291.
52. Ellis, J.J., M.G. Glaholt, and E.M. Reingold, Eye movements reveal solution knowledge prior to insight. *Conscious Cogn*, 2011. 20: p. 768–776.
53. Gibson, E.J., *Principles of Perceptual Learning and Development*. 1969, New York: Prentice-Hill.
54. van Zoest, W. and M. Donk, Awareness of the saccade goal in oculomotor selection: your eyes go before you know. *Conscious Cogn*, 2010. 19: p. 861–871.
55. Kramer, A.F., et al., Age differences in the control of looking behavior: do you know where your eyes have been? *Psychol Sci*, 2000. 11: p. 210–217.
56. Milam, M.E. and C.W. Koo, The current status and future of FDA-approved artificial intelligence tools in chest radiology in the United States. *Clin Radiol*, 2023. 78(2): p. 115–122.

11

Optical Imaging Technology for In Vivo *Tumor Detection*

Radka Bokorova, Duy Binh Tran, and Tohru Yamada

11.1 Introduction

The primary objective of surgery is to completely remove tumors, but it is also important to preserve healthy tissue. In several types of solid tumors, it is difficult to accurately identify tumor margins during surgery; for instance, 20%–70% positive margin in breast tumors, 22% in vulvar melanoma tumors, 7%–75% positive margin in prostate cancer, and 4%–17% positive margin in uterine corpus (1–3). Moreover, in the oral cavity, oropharyngeal squamous cell carcinoma samples occurred in ~23% of resections with positive margins (4). Preoperative imaging such as MRI, PET/CT, ultrasound, and so on detects tumor size and location. Although these traditional imaging modalities have been of immense help in advancing tumor detection and determining the extent of resection(s) (5), the need for accurate intraoperative imaging still remains one of the most important unmet needs in the management of such cancer types.

Fluorescent-guided surgery (FGS) uses two types of fluorescent probes: near-infrared (NIR; always on probes) that exhibit fluorescent signals and "activable or smart probes" that show fluorescence activation on reaction with targeted biomolecules such as tumors (6). FGS has been tested in various types of tumors such as head/neck cancer, breast cancer, esophagus cancer, hepatocellular carcinoma, colorectal cancer, and melanoma (7). One of the most popular fluorescent probes in FGS is indocyanine green (ICG). ICG absorbs NIR light at ~800 nm and emits fluorescence at ~830 nm (5). It has excellent safety records in humans, and its lethal dose for humans is 50-80 mg/kg (8). However, NIR imaging clinical studies have reported that ICG has a relatively short circulation half-life and is a nonspecific contrast agent that limits discerning malignant from normal tissue (9, 10). This has limited its use for tumor imaging. To achieve a significant contrast of tumor vs. normal tissue, ICG needs to be targeted, delivered to, and retained in cancer cells over an extended period.

We identified a cupredoxin protein, azurin, from cultured media when *P. aeruginosa* was co-incubated with human cancer cells (11–13). Secreted azurin preferentially enters a variety of human cancer cells and induces p53-mediated apoptosis (14–16). Our findings (11, 14–22) and subsequent publications by independent research groups form the basis for exploring the cupredoxin protein family as a new, unique, and significant platform for targeting cancers (23–31). We further identified a fragment of azurin, amino acids 50–77 (p28) that is responsible for azurin's preferential penetration and anti-proliferative activity (14, 16, 32, 33). Cellular entry of cell-penetrating peptide (CPP) p28 is faster into cancer cells than normal cells and is retained significantly longer in cancer cells (34–36). Upon entry, p28 induces p53-mediated anti-proliferative effects (35, 36). Clinically, p28 (NSC745104) as a single therapeutic agent was tested in two phase I clinical trials. In the trials, p28 showed preliminary efficacy without apparent adverse effects, toxicity, or immunogenicity in patients with advanced solid tumors and pediatric patients with recurrent and refractory CNS tumors (37–39). Since p28 is highly water soluble and stable, preferentially penetrates cancer cells, crosses the blood brain barrier, and clinically exhibits no significant adverse effects, toxicity, or immunogenicity (34, 35, 38–41), unlike other CPPs (42–44), p28 is an ideal CPP as a tumor-targeting carrier. In this chapter, we describe the methods of preclinical imaging of human melanoma *in vitro* and in vivo.

11.2 Materials

11.2.1 Mice and Cell Culture

Human melanoma cells (UISO-Mel-6, UISO-Mel-23, and UISO-Mel-29) were cultured in Minimum Essential Medium Hanks (Gibco) supplemented with 10% non-heat fetal bovine serum (R&D Systems), 1% Antibiotic-Antimycotic (100x) (Gibco), 7.5% sodium bicarbonate (Gibco),200 mM L-glutamine(100x) (Gibco), and MEM Non-Essential Amino Acids (100x) (Gibco) at 37 °C (45). Four- to five-week-old athymic mice (Jackson Laboratory) were housed in the animal facility. $1x10^6$ Mel–23 cells were injected subcutaneously into the flank in a volume of 100 µl per mouse. All experiments were approved by the University of Illinois at Chicago (UIC) Institutional Animal Care and Use Committee (IACUC) and conformed to the guidelines set by the United States Animal Welfare Act and the National Institutes of Health.

11.3 Methods

11.3.1 Confocal Microscopy

Melanoma cells were seeded at 1×10^4/well on glass coverslips in 24 wells plates overnight at 37 °C and rinsed with fresh medium. Cells were incubated at 37 °C for two hours in pre-warmed medium containing Alexa Fluor 568 conjugated p28 following washing. Coverslips were mounted in medium containing DAPI (VECTASHIELD, Vector Laboratories). Cellular entry of p28 was evaluated and imaged by confocal microscope (Meta 710, Carl Zeiss Inc.).

DOI: 10.1201/9781003449942-13

11.3.2 *In Vivo* Animal Imaging by Odyssey CLx

Athymic mice developing subcutaneous (s.c.) tumors were injected with ICG-p28 at 0.5 mg/kg intravenously through tail vein. After 24 hours, mice were anesthetized using a mixture of 4:1 ketamine:xylazine (100µl/mouse, i.p.), and they were placed against scanning surface of the Odyssey CLx Imager (LI-COR Bioscience).

11.4 Results

11.4.1 Cellular Entry of p28 *in Vitro*

Human melanoma cells (UISO-Mel-6, UISO-Mel-23, and UISO-Mel-29) were incubated with Alexa Fluor 568 conjugated p28. Confocal images showed that p28 penetrated human melanoma (Figure 11.1). As described previously (35, 36), amphipathic, alpha-helical peptide, p28 penetrates cancer cells via endocytotic, caveosome-directed, and caveosome-independent pathways.

By using fragments of p28 (e.g., p18b: a short α-helix and partial β-sheet with highly hydrophilic property; p12: lack of a secondary α-helical structure with hydrophilic property), overall hydrophobicity/hydrophilicity can contribute to preferential cell penetration (35, 36). As a single therapeutic agent, not as an imaging carrier/agent, p28 (NSC745104) was tested in two phase I clinical trials and granted the FDA Orphan Drug and Rare Pediatric Disease Designations since it showed preliminary efficacy without apparent adverse effects, toxicity, or immunogenicity in patients with advanced solid tumors and pediatric patients with recurrent and refractory CNS tumors (NCI, Pediatric Brain Tumor Consortium) (37–39). These results supported the idea that p28 is a potentially ideal vehicle for tumor imaging.

11.4.2 *In Vivo* Imaging

We tested ICG conjugated p28 (ICG-p28) in a xenograft mouse model. Based on our previous preclinical studies, a range of 0.2−0.5 mg/kg is an optimum dose to obtain clear contrast (46–48). Under these conditions, we tested our imaging approach on athymic mice carrying s.c. UISO-Mel-23 melanoma tumors. NIR fluorescence was detected using the Odyssey imaging system (Licor). NIR fluorescence image of mice after a single i.v. injection of 0.5 mg/kg ICG-p28 showed preferential uptake of the imaging agent in the tumor (Figure 11.2). We have reported that our imaging approach with ICG-p28 allows accurate identification of tumors in the margin to increase the complete resection rate, thereby improving tumor recurrence rates in multiple breast cancer animal models (46, 47). For melanoma patients, the overall frequency of positive margins is 4.2% (56/1345), ranging from 2.2% to 22.6%. This depends on the size of the surgical margins, patient characteristics, biopsy history, and the clinicopathology of the melanoma (3). For example, a recent report showed that T4 vulvar melanoma tumors were most commonly excised with positive margins (22.9%) (1). Based on the results, our imaging approach can also be applied to melanoma.

11.5 Discussion

Several NIR fluorescent probes are currently in preclinical development and clinical studies (Table 11.1). NIR fluorescent

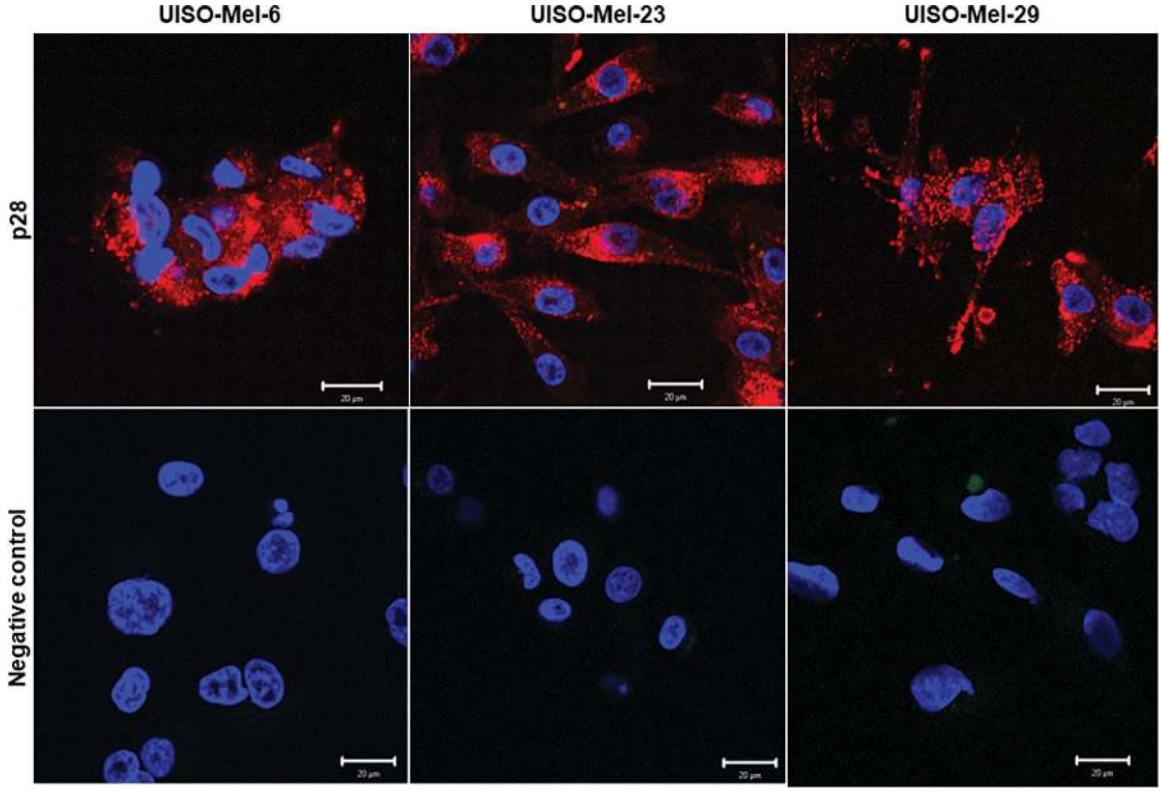


FIGURE 11.1 *p28 penetration into human melanoma cells.* UISO-Mel-6, UISO-Mel-23, and UISO-Mel-29 were cultured with Alexa Fluor 568–labeled p28, and images were recorded by confocal microscopy. Red: Alexa Fluor 568– labeled p28, blue: DAPI (nucleus).

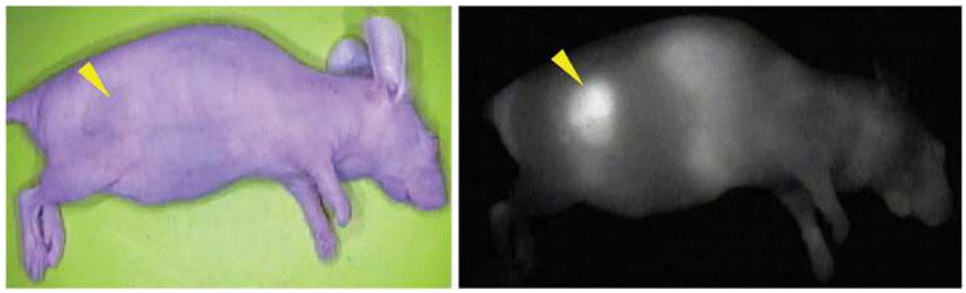


FIGURE 11.2 *In vivo imaging with ICG-p28*. UISO-Mel-23 melanoma cells were injected s.c. into mice. ICG-p28 at 0.5 mg/kg was injected i.v. tail-vein into the mice. Twenty-four hours post-ICG-p28 injection, the mice were imaged with the NIR scanner (Odyssey, Licor). White light photographs (left) and NIR images (right) are shown (grayscale at 800nm). Yellow-arrowheads: tumors.

TABLE 11.1

NIR Fluorescent Probes in Clinical Studies

Intervention/Drug	Disease	Phases	Completed/ Ongoing	Trial	References
Combination product: • IC2000 and SPY-PHI • Tc-99m radioactive colloid and Gamma probe	Breast cancer, lymph node mapping, sentinel lymph node biopsy	3	2023	NCT03200704: A Study Assessing the Safety and Effectiveness of IC2000 and SPY Fluorescence Imaging Systems (SPY-PHI) in the Visualization of Lymphatic Vessels and Identification of Lymph Nodes During Sentinel Lymph Node Biopsy in Subjects with Breast Cancer (FILM-B)	(58)
bevacizumab-IRDye800CW	Breast cancer, breast cancer invasive	2 3	Ongoing	NCT05939310:Fluorescence Guided Surgery in Breast Cancer—The MARGIN-2 Study	(59)
Fluorescein + ALA	Brain cancer	2	Ongoing	NCT02691923: Diagnostic Performance of Fluorescein as an Intraoperative Brain Tumor Biomarker	(60)
SGM-101	Colorectal cancer, low-grade dysplasia of colon	2	Ongoing	NCT06280690: SGM-101 Tumor Targeted Fluorescence Endoscopy in Rectal Polyps with Suspected T1 Adenocarcinoma or High-Grade Dysplasia	(61)
CRGD-ZW800–1	Pancreas adenocarcinoma, cholangiocarcinoma	2	Ongoing	NCT05518071: FLUOPANC-trial—Fluorescence-Guided Surgery of Pancreatic and Bile Duct Tumors Using cRGD-ZW800–1	(62)
5-Aminolevulinic acid hydrochloride	Brain tumor, pediatric	2	Ongoing	NCT04738162: Clinical Safety Study on 5-Aminolevulinic Acid (5-ALA) in Children and Adolescents with Supratentorial Brain Tumors	(63)
CEA-targeted fluorophore SGM-101	Pancreas adenocarcinoma	2	Ongoing	NCT05984810: NIR-Fluorescence Guided Surgical Resection of Neoadjuvant Treated Localized Pancreatic Cancer Using SGM-101 (FLUOPANC-II)	(64)
VGT-309	Lung cancer, lung metastases	2	Ongoing	NCT06145048: Phase 2B Safety and Efficacy Study of VGT-309 in Subjects with Cancer in the Lung (VISUALIZE)	(65)
IRDye800CW-nimotuzumab	Lung cancer	1 2	Ongoing	NCT04459065: Evaluation of IRDye800CW-nimotuzumab in Lung Cancer Surgery	(66)
pHLIP® ICG NIRF imaging	Breast cancer	1 2	Ongoing	NCT05130801: A Study of Multiparametric MRI and pHLIP® ICG in Breast Cancer Imaging during Surgery	(67)
RD-Cy7 fluorophore	Hepatocellular carcinoma	1	Ongoing	NCT06204835: ITGA6 Targeting NIR-II Fluorescence Image Guided Surgery	(68)
Cetuximab-IRDye800	Glioblastoma, high-grade glioma	1	Ongoing	NCT05929456: Multispectral Bimodal Fluorescence Guided Surgery of High-Grade Glioma with Cetuximab-800CW and 5-ALA (5-aminolevulinic Acid)	(69)
HS-196	Solid tumor	1	2023	NCT03333031: A Study of HS-196, an HSP90 Inhibitor-Linked NIR Probe for Solid Malignancies	(70)
EMI-137	Papillary thyroid cancer, lymph node metastases	1	2020	NCT03470259: Precision Thyroid Cancer Surgery with Molecular Fluorescent Guided Imaging	(71)

dyes need to be bright and have high signal-to-background ratios with limited photobleaching. There is a limited number of NIR probes in FGS because it is difficult to develop dyes that are bright and with long-wavelength emission (49). Moreover, ICG is one of these dyes that are commonly tested in FGS. ICG can also be used in the dual-tracker method in combination with methylene blue (MB) and creates a strong tracker in the sentinel lymph node. This technique helps avoid ICG spillage and compensates for the low recognition rate and high false negative results of MB alone (50). ICG has a short half-life, and it does not have a tumor-specific interaction as a passive fluorescent probe. Creating new ICG-derived probes such as ICG-p28 or ICG-cathepsin (VGT-309) leads to specific tumor targeting and will be more useful in surgery and tumor imaging (46). A VGT-309 probe showed signal penetration of 1.7 cm below the pleural surface in lung cancer resection (51).BLZ-100 (tozuleristide), ICG conjugated with peptide isolated from the venom of the scorpion is also in clinical studies (52, 53). Several other fluorescent dyes have been developed for tumor imaging. For example, fluorescence lifetime imaging (FLIM) is autofluorescence molecular imaging that is employed in cancer detection and efficacy studies (54, 55). The potential for FLIM to enhance intraoperative decision-making during robotic-assisted surgery of oropharyngeal cancer was also evaluated (56). Additionally, FLIM can be used for multimodal systems with optical coherence tomography and provide biochemical and morphological maps of the tissue at imaging rates compatible with *in vivo* imaging (57). Surgical microscopes are enriched with various intraoperative imaging modules, not only with fluorescence imaging or optical coherence tomography, but can also be adapted to other imaging modalities such as hyperspectral imaging, photoacoustic microscopy, and laser speckle contrast imaging.

11.6 Limitation of Fluorescence Imaging and Future Trends

Currently, biological imaging techniques use three wavelength ranges—"traditional" (400-700 nm), NIR-I (700-900 nm), and NIR-II (1,000-1,700 nm). The "traditional" probes generally cause a relatively high background that interferes with imaging. It leads to a decrease in the signal-to-background ratio. Another limitation of traditional probes is that these fluorescent molecules need higher photon energy. When biological tissues are exposed to strong excitation light, it can result in tissue damage (72). The agents utilize 700-900 nm wavelength (NIR-I) that can generally penetrate several centimeters depth through blood and tissue (73). Also, imaging in this region provides a high signal-to-noise ratio ("low-noise") because intrinsic autofluorescence from tissues is quite low in this wavelength (74, 75), although autofluorescence imaging with shorter wavelength excitation (e.g., UV) provides a sensitive and noninvasive procedure as an aid to detection/resection in some types of cancer (e.g., oral cancer) (76–78). The NIR-II region can achieve even higher penetration depth (~2 cm) and spatial-temporal resolution because of negligible tissue absorption, weakened scattering, and minimum autofluorescence (79,

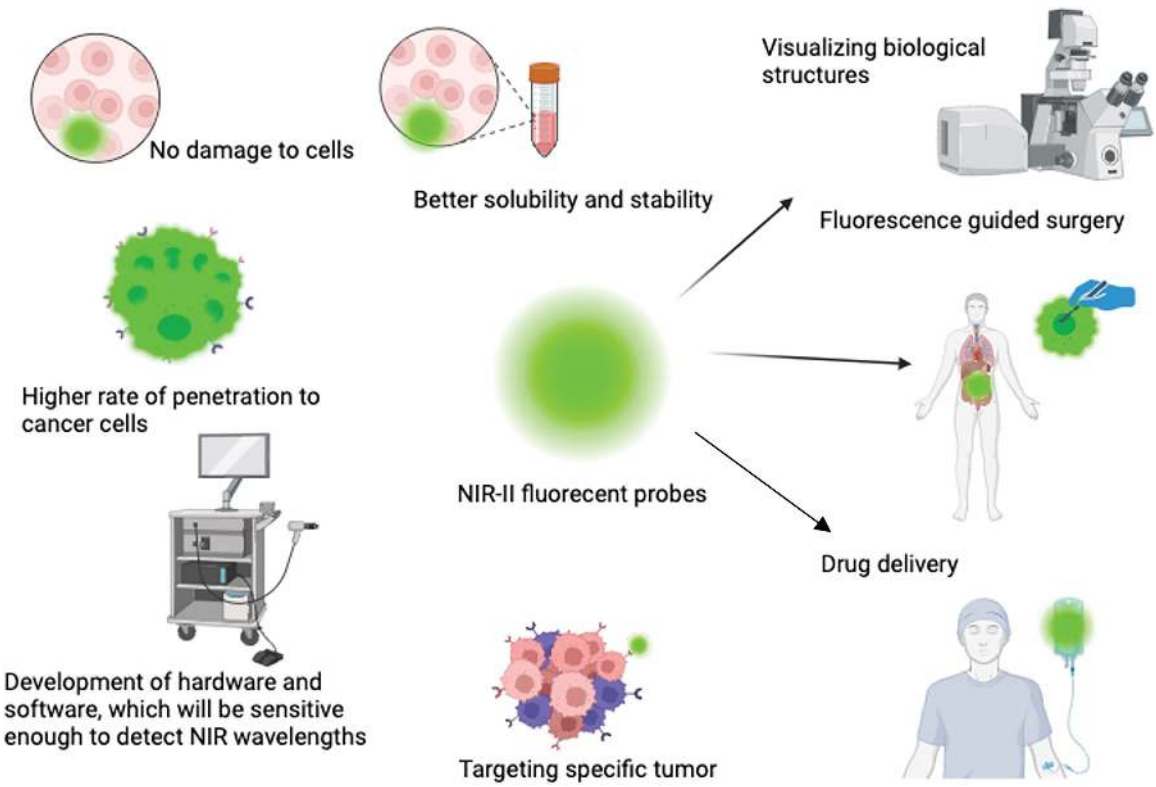


FIGURE 11.3 Development of NIR fluorescent probes leads to clear visualization of tumors in fluorescence-guided surgery or visualizing specific biological structures. Also, a therapeutic agent conjugated with an NIR fluorescent probe can help to track the biodistribution of the agent.

80). However, they have poor biocompatibility, low water solubility, slow metabolism, and relatively high toxicity (80).

NIR fluorescent probes have shown a promising future in surgery and biomedical imaging. Also, the further development of hardware and software for NIR detection and analysis will help to evaluate results in the clinic more easily and quickly. Moreover, when fluorescent probes are conjugated with therapeutic agents, it can be possible to follow drug delivery (Figure 11.3) (81).

Acknowledgments

This work made use of facilities in the Biologic Resource Library (BRL) and Research Resource Center (Fluorescence Imaging) at the University of Illinois at Chicago. This research was supported by NIH/National Cancer Institute (R01CA272564 and R21CA280814) to TY.

11.8 Competing Interests

The authors declare no competing interests.

REFERENCES

1. Cheraghlou S, Doudican NA, Criscito MC, Stevenson ML, Carucci JA. Assessing rates of positive surgical margins after standard excision of vulvar melanomas. *Dermatologic Surgery: Official Publication for American Society for Dermatologic Surgery [et al]*. 2023;49(5):437–44.
2. Lauwerends LJ, Abbasi H, Bakker Schut TC, Van Driel P, Hardillo JAU, Santos IP, et al. The complementary value of intraoperative fluorescence imaging and Raman spectroscopy for cancer surgery: combining the incompatibles. *European Journal of Nuclear Medicine and Molecular Imaging*. 2022;49(7):2364–76.
3. Miller CJ, Shin TM, Sobanko JF, Sharkey JM, Grunyk JW, Elenitsas R, et al. Risk factors for positive or equivocal margins after wide local excision of 1345 cutaneous melanomas. *Journal of the American Academy of Dermatology*. 2017;77(2):333–40.e1.
4. Li MM, Puram SV, Silverman DA, Old MO, Rocco JW, Kang SY. Margin analysis in head and neck cancer: state of the art and future directions. *Annals of Surgical Oncology*. 2019;26(12):4070–80.
5. Sutton PA, van Dam MA, Cahill RA, Mieog S, Polom K, Vahrmeijer AL, et al. Fluorescence-guided surgery: comprehensive review. *BJS Open*. 2023;7(3).
6. Ito R, Kamiya M, Urano Y. Molecular probes for fluorescence image-guided cancer surgery. *Current Opinion in Chemical Biology*. 2022;67:102112.
7. Zheng Y, Yang H, Wang H, Kang K, Zhang W, Ma G, et al. Fluorescence-guided surgery in cancer treatment: current status and future perspectives. *Annals of Translational Medicine*. 2019:S6.
8. Reinhart MB, Huntington CR, Blair LJ, Heniford BT, Augenstein VA. Indocyanine green: historical context, current applications, and future considerations. *Surgical Innovation*. 2016;23(2):166–75.
9. Withrow KP, Gleysteen JP, Safavy A, Skipper J, Desmond RA, Zinn K, et al. Assessment of indocyanine green-labeled cetuximab to detect xenografted head and neck cancer cell lines. *Otolaryngology Head Neck Surgery*. 2007;137(5):729–34.
10. Marshall MV, Rasmussen JC, Tan IC, Aldrich MB, Adams KE, Wang X, et al. Near-infrared fluorescence imaging in humans with indocyanine green: a review and update. *Open Surgical Oncology Journal*. 2010;2(2):12–25.
11. Yamada T, Goto M, Punj V, Zaborina O, Kimbara K, Das Gupta TK, et al. The bacterial redox protein azurin induces apoptosis in J774 macrophages through complex formation and stabilization of the tumor suppressor protein p53. *Infection and Immunity*. 2002;70(12):7054–62.
12. Mahfouz M, Hashimoto W, Das Gupta TK, Chakrabarty AM. Bacterial proteins and CpG-rich extrachromosomal DNA in potential cancer therapy. *Plasmid*. 2007;57(1):4–17.
13. Choi JK, Naffouje SA, Goto M, Wang J, Christov K, Rademacher DJ, et al. Cross-talk between cancer and Pseudomonas aeruginosa mediates tumor suppression. *Communications Biology*. 2023;6(1):16.
14. Yamada T, Goto M, Punj V, Zaborina O, Chen ML, Kimbara K, et al. Bacterial redox protein azurin, tumor suppressor protein p53, and regression of cancer. *Proceedings of the National Academy of Sciences of the United States of America*. 2002;99(22):14098–103.
15. Yamada T, Hiraoka Y, Ikehata M, Kimbara K, Avner BS, Das Gupta TK, et al. Apoptosis or growth arrest: modulation of tumor suppressor p53's specificity by bacterial redox protein azurin. *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(14):4770–5.
16. Punj V, Bhattacharyya S, Saint-Dic D, Vasu C, Cunningham EA, Graves J, et al. Bacterial cupredoxin azurin as an inducer of apoptosis and regression in human breast cancer. *Oncogene*. 2004;23(13):2367–78.
17. Yamada T, Fialho AM, Punj V, Bratescu L, Gupta TK, Chakrabarty AM. Internalization of bacterial redox protein azurin in mammalian cells: entry domain and specificity. *Cellular Microbiology*. 2005;7(10):1418–31.
18. Goto M, Yamada T, Kimbara K, Horner J, Newcomb M, Gupta TK, et al. Induction of apoptosis in macrophages by Pseudomonas aeruginosa azurin: tumour-suppressor protein p53 and reactive oxygen species, but not redox activity, as critical elements in cytotoxicity. *Molecular Microbiology*. 2003;47(2):549–59.
19. Yamada T, Hiraoka Y, Das Gupta TK, Chakrabarty AM. Rusticyanin, a bacterial electron transfer protein, causes G1 arrest in J774 and apoptosis in human cancer cells. *Cell Cycle*. 2004;3(9):1182–7.
20. Hong CS, Yamada T, Hashimoto W, Fialho AM, Das Gupta TK, Chakrabarty AM. Disrupting the entry barrier and attacking brain tumors: the role of the Neisseria H.8 epitope and the Laz protein. *Cell Cycle*. 2006;5(15):1633–41.
21. Chaudhari A, Mahfouz M, Fialho AM, Yamada T, Granja AT, Zhu Y, et al. Cupredoxin-cancer interrelationship: azurin binding with EphB2, interference in EphB2 tyrosine phosphorylation, and inhibition of cancer growth. *Biochemistry*. 2007;46(7):1799–810.
22. Kwan JM, Fialho AM, Kundu M, Thomas J, Hong CS, Das Gupta TK, et al. Bacterial proteins as potential drugs in the treatment of leukemia. *Leukemia Research*. 2009;33(10):1392–9.

23. Yan X, Feng J, Ye Z, Miao X, Li W, Yang D. p53 siRNA inhibits apoptosis of U2OS cells treated with azurin. *Molecular Medicine Reports*. 2011;4(6):1089–94.
24. Zhang Y, Zhang Y, Xia L, Zhang X, Ding X, Yan F, et al. Escherichia coli Nissle 1917 targets and restrains mouse B16 melanoma and 4T1 breast tumors through expression of azurin protein. *Applied and Environmental Microbiology*. 2012;78(21):7603–10.
25. Shaikh F, Abhinand P, Ragunath P. Identification & characterization of lactobacillus salavarius bacteriocins and its relevance in cancer therapeutics. *Bioinformation*. 2012; 8(13):589–94.
26. Ramachandran S, Mandal M. Induction of apoptosis of azurin synthesized from P. aeruginosa MTCC 2453 against Dalton's lymphoma ascites model. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie*. 2011;65(7):461–6.
27. Cho JH, Lee MH, Cho YJ, Park BS, Kim S, Kim GC. The bacterial protein azurin enhances sensitivity of oral squamous carcinoma cells to anticancer drugs. *Yonsei Medical Journal*. 2011;52(5):773–8.
28. Yang DS, Miao XD, Ye ZM, Feng J, Xu RZ, Huang X, et al. Bacterial redox protein azurin induce apoptosis in human osteosarcoma U2OS cells. *Pharmacological Research*. 2005;52(5):413–21.
29. Bernardes N, Abreu S, Carvalho FA, Fernandes F, Santos NC, Fialho AM. Modulation of membrane properties of lung cancer cells by azurin enhances the sensitivity to EGFR-targeted therapy and decreased beta1 integrin-mediated adhesion. *Cell Cycle (Georgetown, Tex)*. 2016;15(11):1415–24.
30. Van Dessel N, Swofford CA, Forbes NS. Potent and tumor specific: arming bacteria with therapeutic proteins. *Therapeutic Delivery*. 2015;6(3):385–99.
31. Gao M, Zhou J, Su Z, Huang Y. Bacterial cupredoxin azurin hijacks cellular signaling networks: protein-protein interactions and cancer therapy. *Protein Science*. 2017;26(12):2334–41.
32. Yamada T, Das Gupta TK, Beattie CW. p28, an anionic cell-penetrating peptide, increases the activity of wild type and mutated p53 without altering its conformation. *Molecular Pharmaceutics*. 2013;10(9):3375–83.
33. Yamada T, Christov K, Shilkaitis A, Bratescu L, Green A, Santini S, et al. p28, a first in class peptide inhibitor of cop1 binding to p53. *British Journal of Cancer*. 2013;108(12):2495–504.
34. Yamada T, Mehta RR, Lekmine F, Christov K, King ML, Majumdar D, et al. A peptide fragment of azurin induces a p53-mediated cell cycle arrest in human breast cancer cells. *Molecular Cancer Therapeutics*. 2009;8(10):2947–58.
35. Taylor BN, Mehta RR, Yamada T, Lekmine F, Christov K, Chakrabarty AM, et al. Noncationic peptides obtained from azurin preferentially enter cancer cells. *Cancer Research*. 2009;69(2):537–46.
36. Yamada T, Signorelli S, Cannistraro S, Beattie CW, Bizzarri AR. Chirality switching within an anionic cell-penetrating peptide inhibits translocation without affecting preferential entry. *Molecular Pharmaceutics*. 2015;12(1):140–9.
37. Razzak M. Targeted therapies: one step closer to drugging p53. *Nature Reviews Clinical Oncology*. 2013; 10(5):246.
38. Warso MA, Richards JM, Mehta D, Christov K, Schaeffer C, Rae Bressler L, et al. A first-in-class, first-in-human, phase I trial of p28, a non-HDM2-mediated peptide inhibitor of p53 ubiquitination in patients with advanced solid tumours. *Br J Cancer*. 2013;108(5):1061–70.
39. Lulla R, Goldman S, Beattie CW, Yamada T, Pollack I, Fisher P, et al. Phase 1 trial of p28 (NSC745104), a non-HDM2 mediated peptide inhibitor of p53 ubiquitination in children with recurrent or progressive CNS tumors: a final report from the pediatric brain tumor consortium. *Neuro Oncology*. 2016. doi: 10.1093/neuonc/now047.
40. Mander S, Naffouje SA, Gao J, Li W, Christov K, Green A, et al. Tumor-targeting cell-penetrating peptide, p28, for glioblastoma imaging and therapy. *Frontiers in Oncology*. 2022;12:940001.
41. Mander S, Gorman GS, Coward LU, Christov K, Green A, Das Gupta TK, et al. The brain-penetrant cell-cycle inhibitor p28 sensitizes brain metastases to DNA-damaging agents. *Neuro-oncology Advances*. 2023;5(1):vdad042.
42. Sample CJ, Hudak KE, Barefoot BE, Koci MD, Wanyonyi MS, Abraham S, et al. A mastoparan-derived peptide has broad-spectrum antiviral activity against enveloped viruses. *Peptides*. 2013;48:96–105.
43. Moreno M, Giralt E. Three valuable peptides from bee and wasp venoms for therapeutic and biotechnological use: melittin, apamin and mastoparan. *Toxins*. 2015;7(4):1126–50.
44. Richard JP, Melikov K, Vives E, Ramos C, Verbeure B, Gait MJ, et al. Cell-penetrating peptides: a reevaluation of the mechanism of cellular uptake. *The Journal of Biological Chemistry*. 2003;278(1):585–90.
45. Kichina JV, Rauth S, Das Gupta TK, Gudkov AV. Melanoma cells can tolerate high levels of transcriptionally active endogenous p53 but are sensitive to retrovirus-transduced p53. *Oncogene*. 2003;22(31):4911–7.
46. Goto M, Ryoo I, Naffouje S, Mander S, Christov K, Wang J, et al. Image-guided surgery with a new tumour-targeting probe improves identification of positive margins. *eBioMedicine—The Lancet Discovery Science*. 2022;76.
47. Naffouje S, Goto M, Gorman G, Konstantin C, Wang J, Green A, et al. Non-toxic tumor-targeting probe, ICG-p28, for near-infrared image-guided surgery. *Journal of Medicinal Chemistry*. 2022;65(10):7371–9.
48. Naffouje S, Goto M, Ryoo I, Green A, Das Gupta TK, Yamada T. A method of tumor in vivo imaging with a new peptide-based fluorescent probe. In *Methods in Molecular Biology*, Vol. 2394. New York, NY: Springer US; 2022. p. 857–65.
49. Seah D, Cheng Z, Vendrell M. Fluorescent probes for imaging in humans: where are we now? *ACS Nano*. 2023;17(20):19478–90.
50. Yang R, Dong C, Jiang T, Zhang X, Zhang F, Fan Z. Indocyanine green and methylene blue dye guided sentinel lymph node biopsy in early breast cancer: a single-center retrospective survival study in 1574 patients. *Clinical Breast Cancer*. 2023;23(4):408–14.
51. Kennedy GT, Holt DE, Azari FS, Bernstein E, Nadeem B, Chang A, et al. A cathepsin-targeted quenched activity–based probe facilitates enhanced detection of human tumors during resection. *Clinical Cancer Research*. 2022;28(17):3729–41.
52. Leary S, Blatt JE, Cohen AR, Cohen KJ, Cole B, Governale L, et al. A phase II/III randomized, blinded study of tozuleristide for fluorescence imaging detection during

neurosurgical resection of pediatric primary central nervous system (CNS) tumors: PNOC012 (Pacific Pediatric Neuro-oncology Consortium). *Journal of Clinical Oncology*. 2020;38(15_suppl):TPS2575.
53. Parrish-Novak J, Holland EC, Olson JM. Image-guided tumor resection. *Cancer Journal (Sudbury, Mass)*. 2015;21(3):206–12.
54. Datta R, Heaster TM, Sharick JT, Gillette AA, Skala MC. Fluorescence lifetime imaging microscopy: fundamentals and advances in instrumentation, analysis, and applications. *Journal of Biomedical Optics*. 2020;25(7):1–43.
55. Zickus V, Wu M-L, Morimoto K, Kapitany V, Fatima A, Turpin A, et al. Fluorescence lifetime imaging with a megapixel SPAD camera and neural network lifetime estimation. *Scientific Reports*. 2020;10(1):20986.
56. Weyers BW, Marsden M, Sun T, Bec J, Bewley AF, Gandour-Edwards RF, et al. Fluorescence lifetime imaging for intraoperative cancer delineation in transoral robotic surgery. *Translational Biophotonics*. 2019;1(1–2).
57. Shrestha S, Serafino MJ, Rico-Jimenez J, Park J, Chen X, Zhaorigetu S, et al. Multimodal optical coherence tomography and fluorescence lifetime imaging with interleaved excitation sources for simultaneous endogenous and exogenous fluorescence. *Biomedical Optics Express*. 2016;7(9):3184–97.
58. Leong SP. Detection of melanoma, breast cancer and head and neck squamous cell cancer sentinel lymph nodes by Tc-99m Tilmanocept (Lymphoseek®). *Clinical & Experimental Metastasis*. 2022;39(1):39–50.
59. Lamberts LE, Koch M, de Jong JS, Adams ALL, Glatz J, Kranendonk MEG, et al. Tumor-specific uptake of fluorescent bevacizumab-IRDye800CW microdosing in patients with primary breast cancer: a phase I feasibility study. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*. 2017;23(11):2730–41.
60. Ahrens LC, Krabbenhøft MG, Hansen RW, Mikic N, Pedersen CB, Poulsen FR, et al. Effect of 5-aminolevulinic acid and sodium fluorescein on the extent of resection in high-grade gliomas and brain metastasis. *Cancers*. 2022;14(3).
61. de Valk KS, Deken MM, Schaap DP, Meijer RP, Boogerd LS, Hoogstins CE, et al. Dose-finding study of a CEA-targeting agent, SGM-101, for intraoperative fluorescence imaging of colorectal cancer. *Annals of Surgical Oncology*. 2021;28(3):1832–44.
62. Handgraaf HJM, Boonstra MC, Prevoo H, Kuil J, Bordo MW, Boogerd LSF, et al. Real-time near-infrared fluorescence imaging using cRGD-ZW800–1 for intraoperative visualization of multiple cancer types. *Oncotarget*. 2017;8(13):21054–66.
63. McCracken DJ, Schupper AJ, Lakomkin N, Malcolm J, Painton Bray D, Hadjipanayis CG. Turning on the light for brain tumor surgery: a 5-aminolevulinic acid story. *Neuro Oncology*. 2022;24(Suppl 6):S52–S61.
64. van Manen L, de Muynck L, Baart VM, Bhairosingh S, Debie P, Vahrmeijer AL, et al. Near-infrared fluorescence imaging of pancreatic cancer using a fluorescently labelled anti-CEA nanobody probe: a preclinical study. *Biomolecules*. 2023;13(4).
65. Bou-Samra P, Chang A, Guo E, Azari F, Kennedy G, Santini JT, Jr, et al. Cathepsin detection to identify malignant cells during robotic pulmonary resection. *Translational Lung Cancer Research*. 2023;12(12):2370–80.
66. Bernhard W, Barreto K, El-Sayed A, Gonzalez C, Viswas RS, Toledo D, et al. Pre-clinical study of IRDye800CW-nimotuzumab formulation, stability, pharmacokinetics, and safety. *BMC Cancer*. 2021;21(1):270.
67. Bargon CA, Huibers A, Young-Afat DA, Jansen BAM, Borel-Rinkes IHM, Lavalaye J, et al. Sentinel lymph node mapping in breast cancer patients through fluorescent imaging using indocyanine green: the INFLUENCE trial. *Annals of Surgery*. 2022;276(5):913–20.
68. Lin YZ, Wu Y, Cao DH, Peng YJ, Deng J, Lin WJ, et al. Integrin α6 targeted near infrared fluorescent imaging and photoacoustic imaging of hepatocellular carcinoma in mice. *Journal of Clinical and Translational Hepatology*. 2023;11(1):110–7.
69. Zhou Q, van den Berg NS, Rosenthal EL, Iv M, Zhang M, Vega Leonel JCM, et al. EGFR-targeted intraoperative fluorescence imaging detects high-grade glioma with panitumumab-IRDye800 in a phase 1 clinical trial. *Theranostics*. 2021;11(15):7130–43.
70. Osada T, Crosby EJ, Kaneko K, Snyder JC, Ginzel JD, Acharya CR, et al. HSP90-Specific nIR probe identifies aggressive prostate cancers: translation from preclinical models to a human phase I study. *Molecular Cancer Therapeutics*. 2022;21(1):217–26.
71. Armstrong GR, Khot MI, Portal C, West NP, Perry SL, Maisey TI, et al. A novel fluorescent c-met targeted imaging agent for intra-operative colonic tumour mapping: translation from the laboratory into a clinical trial. *Surgical Oncology*. 2022;40:101679.
72. Wang T, Chen Y, Wang B, Wu M. Recent progress of second near-infrared (NIR-II) fluorescence microscopy in bioimaging. *Frontiers in Physiology*. 2023;14:1126805.
73. Dang X, Bardhan M, Qi J, et al. Deep-tissue optical imaging of near cellular-sized features. *Scientific Report*. 2019; 9:3873.
74. Gibbs SL. Near infrared fluorescence for image-guided surgery. *Quantitative Imaging in Medicine and Surgery*. 2012;2(3):177–87.
75. Quek C, Leong K. Near-infrared fluorescent nanoprobes for in vivo optical imaging. *Nanomaterials* 2012;2(2):92–112.
76. Pavlova I, Williams M, El-Naggar A, Richards-Kortum R, Gillenwater A. Understanding the biological basis of autofluorescence imaging for oral cancer detection: high-resolution fluorescence microscopy in viable tissue. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*. 2008;14(8):2396–404.
77. Schwarz RA, Gao W, Redden Weber C, Kurachi C, Lee JJ, El-Naggar AK, et al. Noninvasive evaluation of oral lesions using depth-sensitive optical spectroscopy. *Cancer*. 2009;115(8):1669–79.
78. Roblyer D, Kurachi C, Stepanek V, Williams MD, El-Naggar AK, Lee JJ, et al. Objective detection and delineation of oral neoplasia using autofluorescence imaging. *Cancer Prevention Research (Philadelphia, PA)*. 2009;2(5):423–31.
79. Li S, Cheng D, He L, Yuan L. Recent progresses in NIR-I/II fluorescence imaging for surgical navigation. *Frontiers in Bioengineering and Biotechnology*. 2021;9:768698.
80. Chen Y, Xue L, Zhu Q, Feng Y, Wu M. Recent advances in second near-infrared region (NIR-II) fluorophores and biomedical applications. *Frontiers in Chemistry*. 2021;9:750404.
81. Georgiev NI, Bakov VV, Anichina KK, Bojinov VB. Fluorescent probes as a tool in diagnostic and drug delivery systems. *Pharmaceuticals (Basel, Switzerland)*. 2023;16(3).

12

Adapting Image Foundational Model to Identify Tumor Budding from H&E Images in Colorectal Cancer Diagnostics

Ziyu Su, Usama Sajjad, Wendy L. Frankel, Metin N. Gurcan, Wei Chen, and Muhammad Khalid Khan Niazi

12.1 Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related mortality in the United States, affecting approximately 52,550 patients in 2023 [1]. Unfortunately, a significant portion of CRC cases are diagnosed at advanced stages, with global distribution findings indicating diagnosis rates of 14%, 28%, 37%, and 21% for Stages I, II, III, and IV, respectively [2]. Routine screenings, such as colonoscopies, play a vital role in CRC detection and surveillance. Between 2016 and 2020 in the US, 38.3% of CRC cases were locally advanced (including invasion into surrounding tissue), while 21.8% had distant metastasis (spreading to distant body parts). These locally advanced and distant metastasis cases revealed relative five-year survival rates of 71% and 16%, respectively [3], emphasizing the crucial need for diagnostic tools to aid in prognosis and potentially offer predictive insights for additional treatments. Tumor budding (TB) has recently emerged as a significant factor across various cancer types, such as CRC, head and neck squamous cell carcinoma [4, 5], breast cancer [6], esophageal cancer [7], gastric cancer [8, 9], and cervical cancer [10, 11]. A tumor bud is defined as the presence of a single tumor cell or small tumor clusters (up to four cells) that "bud" from the invasive front of the main tumor [12]. Multiple studies focusing on CRC have linked TB [13–20] with advanced tumor stage, lymph node metastasis, and reduced disease-free survival. TB has been incorporated as a notable feature in the CRC Protocol established by the College of American Pathologists [21].

TB represents a type of epithelial-to-mesenchymal transition, in that the tumor cells gain migratory capability by loss of cell-to-cell adhesions and polarity, the first biologic steps towards metastasis. While TB is acknowledged as a valuable prognostic indicator in CRC, its routine clinical application faces challenges due to a lack of consensus on the optimal assessment method. The College of American Pathologists Cancer Protocol and the International Tumor Budding Consensus Conference (ITBCC) guidelines recommend evaluating TB on hematoxylin and eosin (H&E) slides. However, several studies have demonstrated the difficulty and near irreproducibility of identifying tumor buds (TBs) on H&E slides [17, 22]. Manual identification of TB requires meticulous examination of the tumor's invasive front at intermediate to high magnification, along with the exclusion of tumor bud mimics, which is time-intensive and subjective. The evaluation of TBs can be hindered by inflammation, poorly differentiated clusters (PDC—a cluster consisting of five tumor cells or more), and fibrosis at the tumor invasive front [23]. These factors present a significant obstacle to effectively utilizing this crucial prognostic factor in cancer assessment.

Advancements in network infrastructure and the availability of more cost-effective storage solutions have streamlined the process for pathologists in handling digital slide images and facilitating their sharing for clinical use. Simultaneously, remarkable advancements in machine learning have enabled the use of artificial intelligence (AI) with digital pathology, offering new opportunities for image-based diagnoses. Specifically, whole-slide imaging serves as a foundational platform for the application of AI in digital pathology [24]. Significant progress in AI offers considerable benefits to computational histopathology, effectively addressing a range of diagnostic and prognostic challenges, including predicting gene expression, detecting lymph node metastasis, and performing nuclei detection and segmentation from whole slide images (WSI) [25–27].

Several studies have aimed to identify TBs using deep learning techniques automatically. Pai and colleagues [28] utilized the commercially available Aiforia platform to segment various anatomical regions from H&E-stained tissue microarrays (TMAs) of CRC. They utilized a total of 11,383 regions within TMAs annotated using the Aiforia platform. The segmentation of anatomical regions was divided into four convolutional neural network (CNN) tasks: first, CNN segmented the tissue into carcinoma, TBs/PDCs, stroma, mucin, necrosis, fat, and muscle tissue. Second, CNN segmented the stroma into immature, mature, and inflammatory components. Third, CNN segmented carcinoma into low-grade, high-grade, and signet ring cell carcinoma. Fourth, CNN was responsible for identifying tumor-infiltrating lymphocytes within the carcinoma layer. They achieved an average classification accuracy of 71% for detecting TBs/PDCs. Despite potential differences in prognostic value between TBs and PDCs for certain cancers [29], this study treated them as a single segmentation feature. The study's scope is constrained by the use of tissue microarray images, which involve the extraction of tissue from various regions that could introduce sampling bias.

In our prior work (Tavolara and colleagues [30]), we proposed an automated technique for generating ground truth for H&E images, leveraging pan-cytokeratin staining. Pathologists manually selected high-power fields from 120 pairs of tissue slides (comprising H&E and pan-cytokeratin staining) in this

DOI: 10.1201/9781003449942-14

algorithm. They then aligned all pairs to extract 648 adjacent high-power fields. Subsequently, they employed Swin transformers [31] for semantic segmentation of tumor bulk. Unlike conventional CNNs, transformer modules utilize attention mechanisms instead of convolution operations [32]. Swin transformers extend transformers by learning at multiple scales of an input image through a shifted-window strategy, incorporating a decoding module for semantic segmentation. Annotating the tumor bulk using a Swin transformer facilitated the identification of the tumor invasive front, which is essential for locating TBs. Due to the potential overestimation of TBs by pan-cytokeratin staining [33], we empirically determined to retain positive objects with 1000 to 10,000 pixels as potential TBs in cytokeratin stain. Subsequently, they overlaid the potential TBs onto H&E slides and conducted nuclei detection within potentially positive regions to determine the presence of zero to four tumor cells. When compared to pathologists' H&E annotations, their method demonstrated a precision of 0.3856 ± 0.2678 and a recall of 0.3254 ± 0.2312. This algorithm is subject to several limitations: while pan-cytokeratin staining is utilized to correct the ground truths only, it is typically not employed for tumor bud identification in clinical settings [33]. Furthermore, determining a threshold empirically is very challenging considering intra- and inter-reader variability among pathologists, which may hinder generalization across large cohorts.

Bokhorst and colleagues [34] proposed a semi-supervised learning algorithm for detecting TBs through pan-cytokeratin-stained WSI. In this approach, they used a student-teacher framework [33], where a set of 1765 manually annotated TBs was utilized to train a deep learning teacher network. This teacher network aided in identifying potential tumor bud candidates to train the student network iteratively. Following the training of the student-teacher network, the authors applied the student network to the training sets to generate a combined dataset. The authors reported a teacher sensitivity of 0.47 and a student sensitivity of 0.91. However, this network does not provide information regarding the location or count of TBs. To address this, Faster R-CNN was employed to segment TBs on 240 evaluation dataset slides. Furthermore, they conducted an analysis to identify TB hotspots and found no correlation between survival and TB using survival analysis. Nevertheless, this algorithm possesses a few limitations: (i) pan-cytokeratin annotations may be noisy, as staining might not exclusively target the antibodies present in TBs but also other proteins in the tissue; (ii) the study is constrained by the limited number of annotated labeled dataset, and the specificity of the algorithms, which accounts for the total number of predicted false positives, was not reported.

Another method from Bokhorst and colleagues [35] performed the automated assessment of TBs through nuclei segmentation in a dataset comprising 981 patients. Initially, the authors employed a U-Net deep learning model to segment 14 distinct morphological regions within WSIs. Later, they utilized the U-Net architecture [36], initialized with EfficientNetB4 weights [37], to outline epithelial regions at 40x magnification. Following this, HoVerNet [27] was applied to identify neoplastic cells. A circular disk of 0.785mm^2 was moved over each location at the tumor border to identify TB hotspots. Survival analysis revealed a low TB count in patients exhibiting improved overall survival. The study involved five observers, and the comparison of results indicated that the algorithm has low agreement with the human observers in identifying the TBs.

Tavolara and colleagues [33] presented a weakly-supervised deep learning algorithm for identifying TB regions from routine H&E slides, marking a pioneering effort in approaching TB identification as a weakly supervised problem. The authors annotated 15,000 regions of interest (ROIs) corresponding to TBs and non-tumor regions. They then extracted a 512×512 pixel region around each pathologist annotation as a TB or non-TB. Following the multiple instance learning (MIL) terminology, these regions are termed "bags," from which they cropped 96 × 96-pixel regions with a stride of 32, referred to as "instances." To ensure at least one TB patch was present within each bag, they generated overlapping patches. Features of instances were extracted utilizing ResNet [38] pretrained on ImageNet [39]. They computed the attention weights of instances using a neural network signifying the importance score for each instance. Weighted aggregation was performed to produce bag representations and perform the classification. The reported precision was 0.7698 ± 0.1104, with a recall of 0.9080±0.0535 in six-fold cross-validation study. This algorithm is limited by accurately identifying TB locations and it treats the TB detection as a binary class classification. Furthermore, the algorithm's validation was limited to annotated regions, highlighting the necessity for automated validation across the entire tumor invasive front to ensure comprehensive results.

Observing the high standard deviation observed in the results of Tavolara's work [33], Sajjad and colleagues [40] developed Bayesian-MIL—an enhanced weakly supervised algorithm. Since the features extractor [38] used was trained on natural images [39], it differs significantly in extracting the features from histopathological images. They replaced the feature extractor module with the one that was trained on histopathological images [41]. Additionally, they proposed a Bayesian MIL approach, wherein bags are generated using a Bayesian framework. Specifically, they randomly combined 512×512 regions around each pathologist annotation from the same class and slide to reinforce the concept that this process eliminates variability in ground truths from the training dataset. This algorithm aims to ensure that each bag contains at least one correct annotation, thus enhancing the robustness of the training. The reported precision was 0.83 ± 0.06, with a recall of 0.91 ± 0.04. The results indicate an improvement in performance and a reduction in standard deviation in a six-fold cross-validation study. This algorithm is also as limited as Tavolara's study [33].

12.2 Materials

The Institutional Review Board (IRB), under protocol number 2018C0098, approved this study at The Ohio State University. In consideration of data anonymization, the IRB waived the requirement for informed consent. Our dataset comprises 86 H&E-stained WSI obtained from CRC patients at The Ohio

State University Wexner Medical Center (OSUWMC). These WSIs were scanned at a magnification of 40x, with a resolution of 0.061 microns/pixel^2. A pathologist annotated the primary tumor bulk and subsets of TBs within the tumor front regions because of the challenge of fully annotating all TBs across the WSIs. Subsequently, we extracted regions containing TBs and resized them into small 512×512 images.

12.3 Methods

12.3.1 Segment Anything Model

Segmentation anything model (SAM) is the first foundational model for image segmentation tasks. It is trained based on 11 million natural images with 1.1 billion masks of various objects. SAM architecture comprises three primary components: image encoder, prompt encoder, and mask decoder. In the image encoder segment, the model generates image embeddings utilizing a pretrained vision transformer (ViT) model [42, 43]. The prompt encoder, on the other hand, translates human prompts—such as points, bounding boxes, and texts—into an embedding vector. This encoded vector then serves as a trigger for the mask decoder, which is responsible for pointing out the target objects within the image. Notably, the mask decoder is designed to be lightweight, predicting segmentation masks based on the embeddings derived from both the image and prompt encoders.

12.3.2 Adapting Foundation AI model for TB Segmentation

To adapt SAM for our specific task, we employed SAM-Adapter [44] to fine-tune the SAM on our TB segmentation dataset. Our decision to employ SAM-Adapter is driven by two primary considerations. First, SAM is a large-scale model that poses challenges for retraining on downstream tasks, particularly given our university-level computation resources. To solve this issue, SAM-Adapter offers a lightweight neural network module that seamlessly integrates with SAM, facilitating efficient fine-tuning [44]. Second, the original SAM necessitates manual prompts to accurately segment target objects, which proves impractical for our task. Given the complexity of histopathology images and the impracticality of expecting human annotators to identify all TBs, achieving fully automatic segmentation becomes necessary. SAM-Adapter addresses this challenge by autonomously generating task-specific prompts, eliminating the need for human involvement in the process.

SAM-Adapter contains a two-layer multi-layer perceptron (MLP). It takes the element-wise addition of two kinds of task-specific prompts, the images' high-frequency components and patch embeddings tune, based on the method proposed in [45] (see Figure 12.1a). We used multiple adapter modules and added corresponding embedded prompts between different image encoder blocks (see Figure 12.1b). The MLP Layer 2 is shared across all the adapter modules, while the MLP Layer 1 is independent for different adapter modules (see Figure 12.1a). We froze the SAM image encoder to preserve

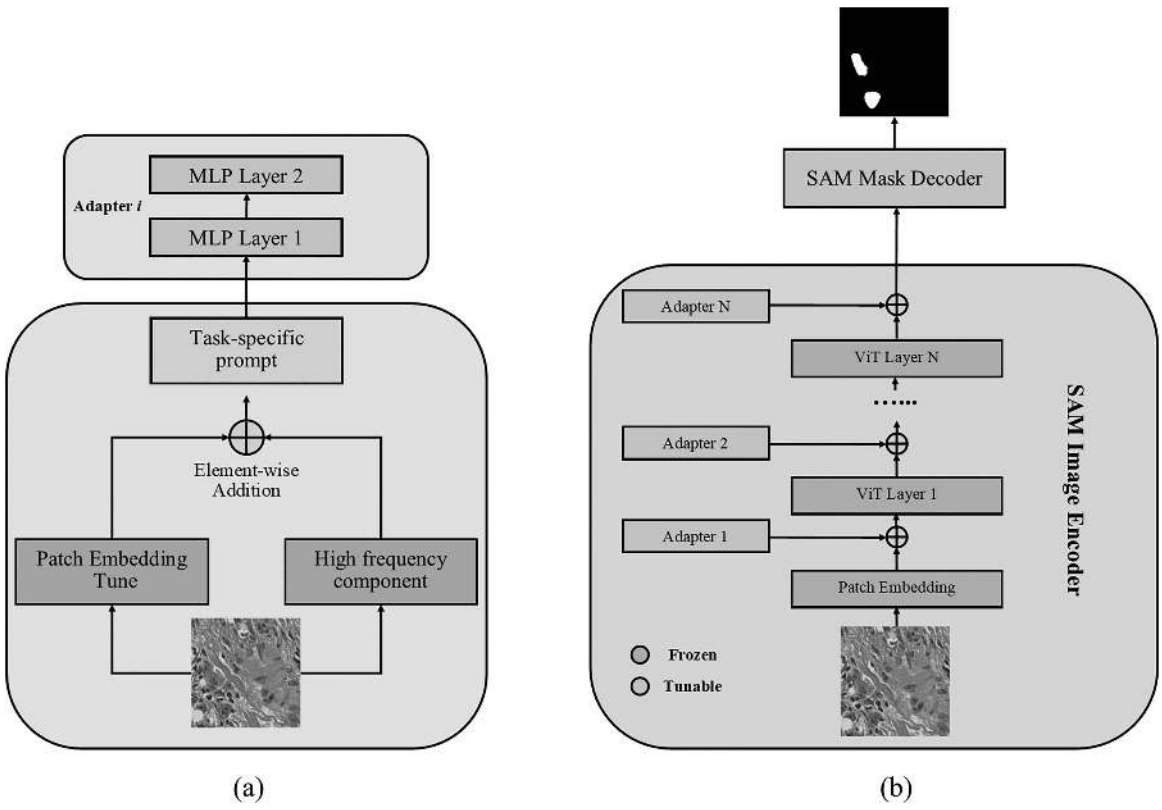


FIGURE 12.1 Overview of SAM-Adapter. (a) Architecture of a single adapter module. (b) Schematics of adapter's employment in SAM.

the semantic understanding acquired through large-scale pretraining. However, we kept the adapters and the mask decoder tunable, allowing us to incorporate explicit visual characteristics from CRC images into the image encoder and generate mask for TBs.

12.3.3 Task-Specific Prompt

The task-specific prompt is the input to the adapters containing information on CRC images and TBs. This prompt is generated by combining tuned image embeddings and the high-frequency components of the original images. The tuned image embeddings refer to the image embeddings of the vision transformer (ViT) [42], fine-tuned through a linear layer. Meanwhile, the high-frequency component is derived by applying a high-pass filter to the image spectrum obtained via fast Fourier transformation (FFT), as illustrated in Figure 12.2. High-frequency components retain the texture details crucial for segmentation tasks. By providing the combined prompts to the adapter, we encouraged the SAM model to transition from the distribution of its pretrained dataset to that of the task-specific dataset [45].

12.3.4 Mask Refinement

Even though the generated masks covered most of the TBs, there were still some errors, such as those tiny false positive objects and masks with small holes in the middle. To resolve those errors, we applied a processing pipeline onto the outputs from the mask decoder. Given one output binary mask, we first applied morphological closing to fill small holes, and then filtered out the small area regions below 500 pixels.

12.3.5 Implementation Details

In our experiment, we randomly split all of the cropped 512×512 images into training and testing sets, resulting in 8,606 training images and 2,036 testing images. We split the images in the case-level to make sure images from the same slide were not assigned for training and testing separately. We trained our model for 35 epochs in a batch size of two. We optimized our model using a combined loss function of binary cross-entropy loss and intersection over union (IoU) loss with an Adam optimizer [46].

For the evaluation, we randomly selected 176 images from our testing set and asked our experienced pathologist to finely reannotate all TBs from those images. These 176 fully annotated images served as our final testing set to evaluate our model. We did this because only partial TBs were annotated in our dataset due to the difficulty of exhaustively annotating all TBs on the WSIs, which was a common issue [47]. We used IoU and instance-level Dice score ($\text{Dice}_{\text{inst}}$) as our metrics. The IoU is defined as:

$$IoU = \frac{TP}{TP + FP + FN} \tag{12.1}$$

where TP, FP, and FN indicate the number of true positive, false positive, and false negative predicted pixels. The $\text{Dice}_{\text{inst}}$ is defined as:

$$Dice_{inst} = \frac{2 \times TP}{2TP + FP + FN} \tag{12.2}$$

where we considered a 0.5 IoU between a predicted and a ground-truth object as a TP, classifying the rest as FNs or FPs. This approach accommodates the challenge of determining TBs' boundaries for pathologists, where coarse annotations are common, making pixel-level metrics like IoU inadequate in reflecting true performance. In addition, the segmentation results will be used for TB quantification at the instance level, which makes $\text{Dice}_{\text{inst}}$ a proper metric. We compared our model with UNet++ [48], which was trained from scratch, and another SAM-based adaption model [49], which utilized random click as a prompting method [49].

12.4 Results

The quantitative findings are presented in Table 12.1, where we conduct a comparative analysis between our model and two counterparts: a model trained from scratch, UNet++ [48], and another SAM-based adaptation model, Medical SAM-Adapter (MSA) [49]. Across all metrics, our SAM-Adapter model demonstrates superior performance compared to both UNet++ and MSA, underscoring the effectiveness of our approach. These

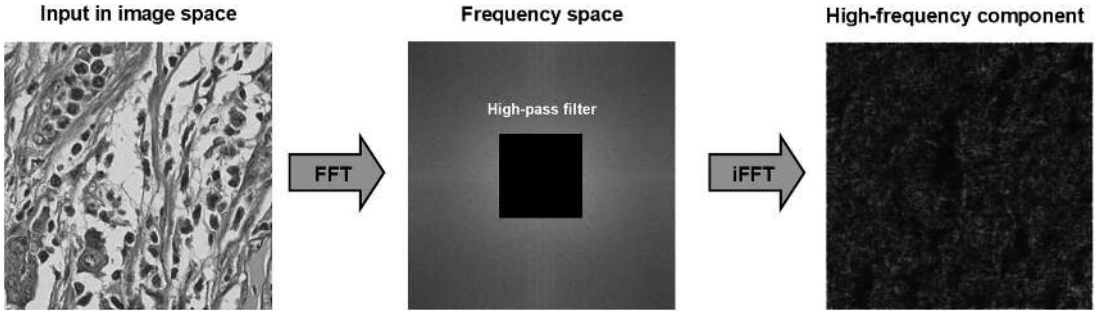


FIGURE 12.2 The procedure of extracting high-frequency components from the image. We employ FFT to convert the image into frequency space and remove the low-frequency parts using a high-pass filter. Then, we employ inverse FFT (iFFT) to revert the image back to spatial space, thereby emphasizing the high-frequency component.

TABLE 12.1

TB Segmentation Results Based on the Testing Images

Architecture	IoU	$\text{Dice}_{\text{inst}}$
UNet++ [48]	0.60	0.71
MSA [49]	0.59	0.68
SAM-Adapter (ours)	**0.62**	**0.76**

results highlight the potential of adapting foundational models to specific medical imaging tasks, showcasing their capability to outperform models trained from scratch. This underscores the significance of prioritizing model adaptation as a pivotal avenue for future advancements in the domain.

We further showcase our segmentation results in Figure 12.3A–J, which illustrates the efficacy of our model in accurately segmenting the majority of TBs. For example, the segmentation matches with the ground truth for all TBs in Figure 12.3A, B, and D. However, we note occasional instances where our model divides a larger TB containing multiple tumor cells into smaller, distinct TBs, as depicted in Figure 12.3C. Additionally, there are cases where the model detects only part of a TB, particularly when tumor cells within a TB are less cohesive rather than closely clustered, as evidenced by examples in Figure 12.3H and J.

12.5 Discussion and Conclusions

Annotating CRC TB can be a time-consuming and challenging task in pathology. In this investigation, we employed the SAM model alongside adapters [44] by passing task-specific models to train our TB segmentation model, leveraging task-specific prompts. Consequently, our model achieved notable performance metrics, with an IoU of 0.62 and a $\text{Dice}_{\text{inst}}$ of 0.76, surpassing the performance of the trained-from-scratch model (UNet++[48]). To our knowledge, this study is the first to showcase such robust performance utilizing SAM for cellular-level object segmentation tasks. Our findings underscore the potential of adapting foundational segmentation models to specific tasks, demonstrating the superiority of this approach over training task-specific models. This suggests a promising avenue for future developments in segmentation models. Furthermore, our visualizations (see Figure 12.3) affirm the proficiency of our model in detecting the majority of TBs within the 512×512 -sized histopathology images.

Our study's exploration into the automated segmentation of TB in CRC using the SAM-Adapter model has yielded promising results, signifying a substantial step forward in integrating AI within the realm of digital/computational pathology. The SAM-Adapter's advanced capabilities in discerning the nuanced morphology of TBs have potential ramifications not only for patient risk stratification and treatment planning but also for our understanding of cancer biology. Particularly, the ability of our model to accurately quantify TB may provide insights into the mechanisms of tumor progression and metastasis.

In light of existing literature, the strength of our model lies in its robustness and adaptability. Previous attempts to automate TB detection have grappled with the inherent complexity of histopathological images and the subtleties of TB presentation. Our study circumvents the limitations of models trained on more restrictive datasets by employing a foundational model originally pretrained on a vast array of non-medical images. This pretraining allows SAM-Adapter to leverage a broad spectrum of visual features, increasing its discriminative power. Such pretraining is particularly pertinent given the diverse and heterogeneous nature of pathological images, which often contain artifacts and variations that can confound less sophisticated models. Furthermore, our findings contribute to a critical discourse on the operational challenges faced by pathologists. Manual TB detection is notoriously labor

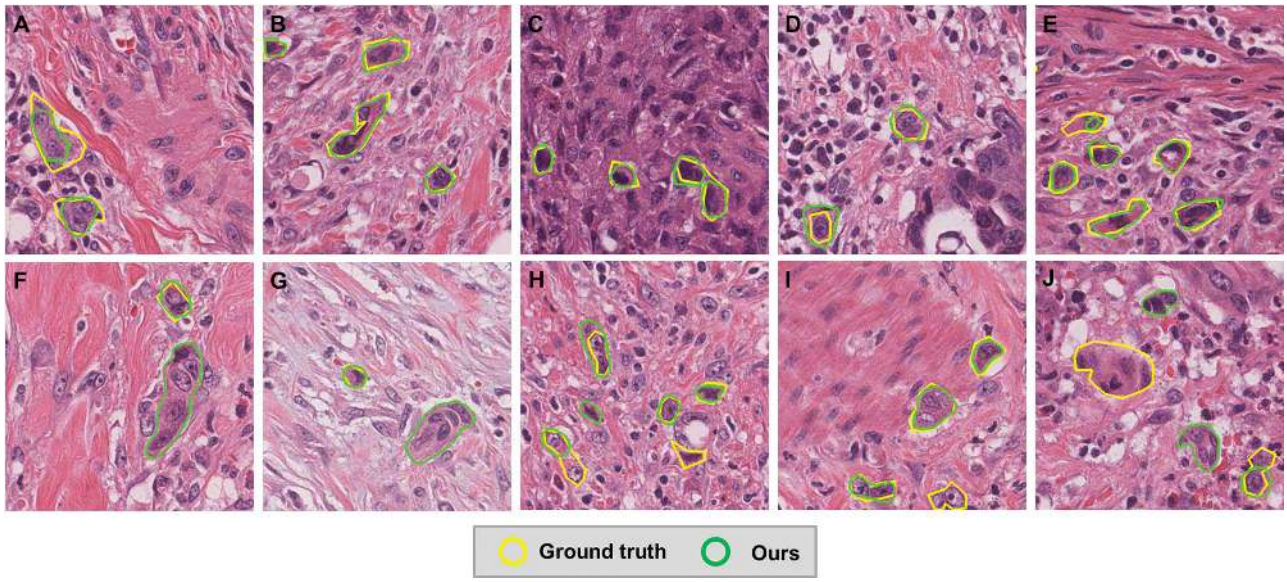


FIGURE 12.3 Visualization results of our method compared with the ground truth. Yellow annotations are ground-truth annotated by our pathologist, and green annotations are segmentations from our model.

intensive and subject to interobserver variability, leading to potential discrepancies in patient prognoses. By streamlining this process through automation, our model not only may enhance diagnostic quantification of TB but may also alleviate the cognitive burden on pathologists. This augmentation of the diagnostic process has the potential to harmonize TB assessments across institutions, promoting consistency in clinical practice and potentially influencing treatment management.

Nonetheless, there are shortcomings observed. Specifically, in cases where tumor cells are dispersed rather than tightly clustered, our model tends to segment one large TB into multiple smaller TBs or detects only partial TBs, as evidenced in Figure 12.3c, h, and j. Additionally, our model occasionally detects PDCs, defined as clusters of five or more tumor cells lacking glandular structure (see Figure 12.3f and g). We attribute this to the variability in annotations within our dataset. Due to the inter-/intra-reader variability in manual TB annotations, our model encounters inconsistencies during training. Consequently, during testing, these inconsistencies may lead to confusion regarding the definition of TBs, resulting in some randomness in our model's segmentation outputs. To solve this issue, we are planning to integrate our model into a semi-automatic image annotation pipeline with pathologists in the loop. This strategy has been shown to be efficient in generating massive amount of accurate image labels in the computer vision field [50].

12.6 Future Trends

Our future endeavors will focus on enhancing the model's discriminative capabilities between TB and PDCs, given their potential distinct prognostic implications. We aim to refine our algorithm's accuracy by incorporating additional morphological parameters and utilizing more extensive, diverse training datasets. The goal is to create an AI-assisted tool that reliably identifies TB across a broad spectrum of CRC cases, reducing the margin of error to a minimum. In parallel, we advocate for the integration of AI models like SAM-Adapter into semi-automatic annotation pipelines, which will allow for iterative model improvement through pathologist validation. Such collaborative efforts could streamline the annotation process, generating a more comprehensive and precise dataset for model training. This integration could also facilitate the exploration of AI's role in other complex pathological assessments, expanding the utility of digital/computational pathology beyond TB quantification. In conclusion, our study not only showcases the capability of AI to enhance the prognostic assessment of CRC but also illuminates the path forward for digital/computational pathology. As we continue refining our model, we envision a future where AI-powered tools are seamlessly integrated into pathology workflows, offering high-throughput, precise assessments that could transform cancer diagnostics and patient care globally.

Acknowledgments

The project described was supported in part by National Institutes of Health R01CA276301 (PIs: Niazi, Chen) and Alliance Clinical Trials in Oncology GR125886 (PIs: Frankel and Niazi). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. We appreciate Ms. Deborah Knight's assistance with imaging handling and annotation.

REFERENCES

1. Siegel, R.L.; Miller, K.D.; Wagle, N.S.; Jemal, A. Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians* 2023, *73*, 17–48, doi:10.3322/caac.21763.
2. Hultcrantz, R. Aspects of colorectal cancer screening, methods, age and gender. *Journal of Internal Medicine* 2021, *289*, 493–507, doi:10.1111/joim.13171.
3. Control, C.f.D.; Prevention. *US Cancer Statistics Colorectal Cancer Stat Bite*. US Department of Health and Human Services. 2022.
4. Boxberg, M.; Jesinghaus, M.; Dorfner, C.; Mogler, C.; Drecoll, E.; Warth, A.; Steiger, K.; Bollwein, C.; Meyer, P.; Wolff, K.D. Tumour budding activity and cell nest size determine patient outcome in oral squamous cell carcinoma: proposal for an adjusted grading system. *Histopathology* 2017, *70*, 1125–1137.
5. Angadi, P.V.; Patil, P.V.; Hallikeri, K.; Mallapur, M.; Hallikerimath, S.; Kale, A.D. Tumor budding is an independent prognostic factor for prediction of lymph node metastasis in oral squamous cell carcinoma. *International Journal of Surgical Pathology* 2015, *23*, 102–110.
6. Li, X.; Wei, B.; Sonmez, C.; Li, Z.; Peng, L. High tumor budding count is associated with adverse clinicopathologic features and poor prognosis in breast carcinoma. *Human Pathology* 2017, *66*, 222–229.
7. Jesinghaus, M.; Boxberg, M.; Konukiewitz, B.; Slotta-Huspenina, J.; Schlitter, A.M.; Steiger, K.; Specht, K.; Wieczorek, K.; Warth, A.; Schmidt, T. A novel grading system based on tumor budding and cell nest size is a strong predictor of patient outcome in esophageal squamous cell carcinoma. *The American Journal of Surgical Pathology* 2017, *41*, 1112–1120.
8. Che, K.; Zhao, Y.; Qu, X.; Pang, Z.; Ni, Y.; Zhang, T.; Du, J.; Shen, H. Prognostic significance of tumor budding and single cell invasion in gastric adenocarcinoma. *OncoTargets and Therapy* 2017, 1039–1047.
9. Olsen, S.; Jin, L.; Fields, R.C.; Yan, Y.; Nalbantoglu, I. Tumor budding in intestinal-type gastric adenocarcinoma is associated with nodal metastasis and recurrence. *Human Pathology* 2017, *68*, 26–33.
10. Huang, B.; Cai, J.; Xu, X.; Guo, S.; Wang, Z. High-grade tumor budding stratifies early-stage cervical cancer with recurrence risk. *PLoS One* 2016, *11*, e0166311.
11. Satabongkoch, N.; Khunamornpong, S.; Pongsuvareeyakul, T.; Settakorn, J.; Sukpan, K.; Soongkhaw, A.; Suprasert, P.; Siriaunkgul, S. Prognostic value of tumor budding in early-stage cervical adenocarcinomas. *Asian Pacific Journal of Cancer Prevention: APJCP* 2017, *18*, 1717.
12. Lugli, A.; Kirsch, R.; Ajioka, Y.; Bosman, F.; Cathomas, G.; Dawson, H.; El Zimaity, H.; Fléjou, J.-F.; Hansen, T.P.; Hartmann, A. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. *Modern Pathology* 2017, *30*, 1299–1311.

13. Mezheyeuski, A.; Hrynchyk, I.; Karlberg, M.; Portyanko, A.; Egevad, L.; Ragnhammar, P.; Edler, D.; Glimelius, B.; Östman, A. Image analysis-derived metrics of histomorphological complexity predicts prognosis and treatment response in stage II-III colon cancer. *Scientific Reports* 2016, *6*, 36149.
14. Lai, Y.H.; Wu, L.C.; Li, P.S.; Wu, W.H.; Yang, S.B.; Xia, P.; He, X.X.; Xiao, L.B. Tumour budding is a reproducible index for risk stratification of patients with Stage II colon cancer. *Colorectal Disease* 2014, *16*, 259–264.
15. Cappellesso, R.; Luchini, C.; Veronese, N.; Mele, M.L.; Rosa-Rizzotto, E.; Guido, E.; De Lazzari, F.; Pilati, P.; Farinati, F.; Realdon, S. Tumor budding as a risk factor for nodal metastasis in pT1 colorectal cancers: a meta-analysis. *Human Pathology* 2017, *65*, 62–70.
16. Zlobec, I.; Hädrich, M.; Dawson, H.; Koelzer, V.H.; Borner, M.; Mallaev, M.; Schnüriger, B.; Inderbitzin, D.; Lugli, A. Intratumoural budding (ITB) in preoperative biopsies predicts the presence of lymph node and distant metastases in colon and rectal cancer patients. *British Journal of Cancer* 2014, *110*, 1008–1013.
17. Koelzer, V.H.; Zlobec, I.; Berger, M.D.; Cathomas, G.; Dawson, H.; Dirschmid, K.; Hädrich, M.; Inderbitzin, D.; Offner, F.; Puppa, G. Tumor budding in colorectal cancer revisited: results of a multicenter interobserver study. *Virchows Archiv* 2015, *466*, 485–493.
18. Graham, R.P.; Vierkant, R.A.; Tillmans, L.S.; Wang, A.H.; Laird, P.W.; Weisenberger, D.J.; Lynch, C.F.; French, A.J.; Slager, S.L.; Raissian, Y. Tumor budding in colorectal carcinoma: confirmation of prognostic significance and histologic cutoff in a population-based cohort. *The American Journal of Surgical Pathology* 2015, *39*, 1340–1346.
19. Koelzer, V.H.; Zlobec, I.; Lugli, A. Tumor budding in colorectal cancer—ready for diagnostic practice? *Human Pathology* 2016, *47*, 4–19.
20. Petrelli, F.; Pezzica, E.; Cabiddu, M.; Coinu, A.; Borgonovo, K.; Ghilardi, M.; Lonati, V.; Corti, D.; Barni, S. Tumour budding and survival in stage II colorectal cancer: a systematic review and pooled analysis. *Journal of Gastrointestinal Cancer* 2015, *46*, 212–218.
21. Grizzi, F.; Celesti, G.; Basso, G.; Laghi, L. Tumor budding as a potential histopathological biomarker in colorectal cancer: hype or hope? *World Journal of Gastroenterology: WJG* 2012, *18*, 6532.
22. Kai, K.; Aishima, S.; Aoki, S.; Takase, Y.; Uchihashi, K.; Masuda, M.; Nishijima-Matsunobu, A.; Yamamoto, M.; Ide, K.; Nakayama, A. Cytokeratin immunohistochemistry improves interobserver variability between unskilled pathologists in the evaluation of tumor budding in T1 colorectal cancer. *Pathology International* 2016, *66*, 75–82.
23. Zlobec, I.; Lugli, A. Epithelial mesenchymal transition and tumor budding in aggressive colorectal cancer: tumor budding as oncotarget. *Oncotarget* 2010, *1*, 651.
24. Niazi, M.K.K.; Parwani, A.V.; Gurcan, M.N. Digital pathology and artificial intelligence. *The Lancet Oncology* 2019, *20*, e253-e261, doi:10.1016/s1470–2045(19)30154–8.
25. Tavolara, T.E.; Niazi, M.K.K.; Gower, A.C.; Ginese, M.; Beamer, G.; Gurcan, M.N. Deep learning predicts gene expression as an intermediate data modality to identify susceptibility patterns in Mycobacterium tuberculosis infected diversity outbred mice. *EBioMedicine* 2021, *67*, 103388, doi:10.1016/j.ebiom.2021.103388.
26. Lu, M.Y.; Williamson, D.F.K.; Chen, T.Y.; Chen, R.J.; Barbieri, M.; Mahmood, F. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature Biomedical Engineering* 2021, *5*, 555–570, doi:10.1038/s41551-020-00682-w.
27. Graham, S.; Vu, Q.D.; Raza, S.E.A.; Azam, A.; Tsang, Y.W.; Kwak, J.T.; Rajpoot, N. Hover-Net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis* 2019, *58*, 101563, doi:10.1016/j.media.2019.101563.
28. Pai, R.K.; Hartman, D.; Schaeffer, D.F.; Rosty, C.; Shivji, S.; Kirsch, R.; Pai, R.K. Development and initial validation of a deep learning algorithm to quantify histological features in colorectal carcinoma including tumour budding/poorly differentiated clusters. *Histopathology* 2021, *79*, 391–405.
29. Szalai, L.; Jakab, Á.; Kocsmár, I.; Szirtes, I.; Kenessey, I.; Szijártó, A.; Schaff, Z.; Kiss, A.; Lotz, G.; Kocsmár, É. Prognostic ability of tumor budding outperforms poorly differentiated clusters in gastric cancer. *Cancers* 2022, *14*, 4731.
30. Tavolara, T.E.; Dutta, A.; Burks, M.V.; Chen, W.; Frankel, W.; Gurcan, M.N.; Niazi, M.K.K. Automatic generation of the ground truth for tumor budding using H&E stained slides. In *Proceedings of the Medical Imaging 2022: Digital and Computational Pathology*, 2022; pp. 40–46.
31. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: hierarchical vision transformer using shifted windows. In *Proceedings of the Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021; pp. 10012–10022.
32. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Advances in Neural Information Processing Systems* 2017, *30*.
33. Tavolara, T.E.; Chen, W.; Frankel, W.L.; Gurcan, M.N.; Niazi, M.K.K. Minimizing the intra-pathologist disagreement for tumor bud detection on H and E images using weakly supervised learning. In *Proceedings of the Medical Imaging 2023: Digital and Computational Pathology*, 2023; pp. 277–283.
34. Bokhorst, J.-M.; Nagtegaal, I.D.; Zlobec, I.; Dawson, H.; Sheahan, K.; Simmer, F.; Kirsch, R.; Vieth, M.; Lugli, A.; van der Laak, J.; et al. Semi-supervised learning to automate tumor bud detection in cytokeratin-stained whole-slide images of colorectal cancer. *Cancers* 2023, *15*, 2079.
35. Bokhorst, J.-M.; Ciompi, F.; Öztürk, S.K.; Oguz Erdogan, A.S.; Vieth, M.; Dawson, H.; Kirsch, R.; Simmer, F.; Sheahan, K.; Lugli, A.; et al. Fully automated tumor bud assessment in hematoxylin and eosin-stained whole slide images of colorectal cancer. *Modern Pathology* 2023, *36*, 100233, doi:10.1016/j.modpat.2023.100233.
36. Ronneberger, O.; Fischer, P.; Brox, T. U-net: convolutional networks for biomedical image segmentation. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015, Proceedings*, Part III 18, 2015; pp. 234–241.
37. Tan, M.; Le, Q. EfficientNet: rethinking model scaling for convolutional neural networks. In *Proceedings of the International Conference on Machine Learning*, 2019; pp. 6105–6114.

38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 27–30 June 2016*, 2016; pp. 770–778.
39. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Kai, L.; Li, F.-F. ImageNet: a large-scale hierarchical image database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, 20–25 June 2009*, 2009; pp. 248–255.
40. Usama Sajjad, W.C., Mostafa Rezapour, Ziyu Su, Thomas Tavolara, Wendy L. Frankel, Metin N. Gurcan, M. Khalid Khan Niazi. Enhancing colorectal cancer tumor bud detection using deep learning from routine H&E-stained slides. *Proceedings of SPIE—International Society for Optical Engineering A* 2024, 12933.
41. Wang, X.; Yang, S.; Zhang, J.; Wang, M.; Zhang, J.; Yang, W.; Huang, J.; Han, X. Transformer-based unsupervised contrastive learning for histopathological image classification. *Medical Image Analysis* 2022, *81*, 102559, doi:10.1016/j.media.2022.102559.
42. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S. An image is worth 16x16 words: transformers for image recognition at scale. In *Proceedings of the International Conference on Learning Representations*, 2020.
43. He, K.; Chen, X.; Xie, S.; Li, Y.; Dollár, P.; Girshick, R. Masked autoencoders are scalable vision learners. In *Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022; pp. 16000–16009.
44. Chen, T.; Zhu, L.; Ding, C.; Cao, R.; Zhang, S.; Wang, Y.; Li, Z.; Sun, L.; Mao, P.; Zang, Y. SAM fails to segment anything?--SAM-adapter: adapting SAM in underperformed scenes: camouflage, shadow, and more. *arXiv Preprint* arXiv:2304.09148, 2023.
45. Liu, W.; Shen, X.; Pun, C.-M.; Cun, X. Explicit visual prompting for low-level structure segmentations. In *Proceedings of the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023; pp. 19434–19445.
46. Kingma, D.P.; Ba, J. Adam: a method for stochastic optimization. *arXiv Preprint* arXiv:1412.6980, 2014.
47. Bokhorst, J.-M.; Ciompi, F.; Öztürk, S.K.; Oguz Erdogan, A.S.; Vieth, M.; Dawson, H.; Kirsch, R.; Simmer, F.; Sheahan, K.; Lugli, A.; et al. Fully automated tumor bud assessment in hematoxylin and eosin stained whole slide images of colorectal cancer. *Modern Pathology* 2023, doi:10.1016/j.modpat.2023.100233.
48. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: redesigning skip connections to exploit multiscale features in image segmentation. *IEEE Transactions on Medical Imaging* 2019, *39*, 1856–1867.
49. Wu, J.; Fu, R.; Fang, H.; Liu, Y.; Wang, Z.; Xu, Y.; Jin, Y.; Arbel, T. Medical SAM adapter: adapting segment anything model for medical image segmentation. *arXiv Preprint* arXiv:2304.12620, 2023.
50. Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A.C.; Lo, W.-Y. Segment anything. *arXiv Preprint* arXiv:2304.02643, 2023.

13

Quantitative Phase Imaging for Assessing Tumor Cell Adaptability

Rebecca G. Zitnay, Shukran Alizada, Anastasia Prokofyeva, Thomas A. Zangle, and Robert L. Judson-Torres

13.1 Introduction

13.1.1 Tumor Cell Adaptability

Phenotypic plasticity underlies tumor cell escape from "normal" behavior and has gained recognition as an emerging hallmark of cancer [1]. Phenotypic plasticity manifests as an increased landscape of potential transcriptomes, morphologies, and behaviors among clonal populations of cells [2]. The traits associated with phenotypic states can be used to classify individual cells within a population (Figure 13.1a and b). States associated with metastasis or targeted therapy resistance provide a survival advantage in the presence of stressors associated with tumorigenesis. In the absence of a uniform stimulus or selective pressure, plasticity may present as heterogeneity in a population. However, in the presence of a stressor, such as exposure to a toxic compound or change in the environment, cells with the ability to adapt more rapidly and completely have an advantage over slower to respond cell populations.

Currently, the gold standard tool for precision oncology is genomic sequencing, a static diagnostic that identifies tumors with oncogenic variants likely to respond to available therapeutics. However, this approach is limited; less than 40% of patients have a mutation that matches a targeted therapy [3]. Additionally, the presence of an actionable mutation is insufficient to predict the depth and duration of response since epigenetic reprogramming, another emerging hallmark of cancer, underlies many phenotypic changes that drive tumor plasticity and is not fully captured by genomic sequencing [1]. Instead, a genetic diagnosis describes the probability of cells existing in specific phenotypic states associated with distinct disease phenotypes and resistance mechanisms [2].

Functional assays that measure adaptation to perturbation can resolve the depth and duration of response better than static measurements. By plotting the phenotype over time, a general trend emerges and reveals three discrete stages in the adaptive response: (1) the initial latent period in which no change in phenotype can be detected, (2) the transition period, and (3) the completion of response to the new state (Figure 13.1c). A single cell can transition faster or slower than its neighbors and follow several different trajectories to the final state (Figure 13.1d). Trajectories A, B, and C all reach the same final phenotype but at different rates. Alternatively, cells may not achieve a full transition; here, trajectory D demonstrates a partial phenotype switch. Cells may have no response, following trajectory E, such as those that are drug resistant. While generalized as a positive change in phenotype here, the same features can be present in phenotypes that change negatively in response to a perturbation, such as trajectory F. This is common for a phenotype like growth rate, which we expect to decrease when cells are exposed to a compound that inhibits growth. In this general form, the phenotype is anything that can be measured, including examples like morphology, motility, mechanics, and growth dynamics [4–8]. When there is a more favorable accessible state, cells with increased adaptability can switch to a more advantageous phenotype faster than less adaptable cells (Figure 13.1e). Importantly, cancer cells exhibit substantial inter- and intra-tumoral heterogeneity of both phenotype and adaptability. In the context of cancer cells responding to a library of potential therapies, identifying and tracking such heterogeneity may be critical for predicting response to therapy.

Kinetic functional precision oncology assays take advantage of fast, label-free, live imaging to capture the dynamic changes in cell mass, morphology, and cellular organization [9, 10]. QPI is a technology with the temporal and spatial resolution necessary for defining the adaptive response both as a population and with single-cell resolution.

13.1.2 Quantitative Phase Imaging (QPI)

QPI measures the phase shift of light as it passes through a sample. When applied to living cells, this phase shift is linearly proportional to cell mass [11, 12]. QPI provides data on the mass distribution within individual cells or cell clusters. Many QPI methods have been applied to biological samples [13, 14]. Applications in this chapter focus on two distinct methods of QPI: Fourier ptychography and phase retrieval from differential phase contrast (DPC). Fourier ptychography captures multiple images to cover a broader range of imaging Fourier space than possible with a single image capture and uses iterative phase retrieval to compute phase shifts through the sample. This allows Fourier ptychography to capture quantitative phase information over large fields of view at high resolution. DPC is a phase retrieval method that combines just four intensity images to retrieve the phase shift of light through a sample [15, 16].

QPI offers advantages over alternative approaches. Compared to traditional proliferation assays, QPI measurements are much faster at measuring changes in cell growth rates and provide single-cell data, therefore enabling measurements of population heterogeneity. As an approach based on brightfield microscopy, the illumination intensity in QPI is typically low with minimal phototoxicity. As a label-free technique, QPI can be applied for long-term monitoring of freshly isolated primary cell populations or lineages with minimal impact on cell behavior. Finally,

DOI: 10.1201/9781003449942-15

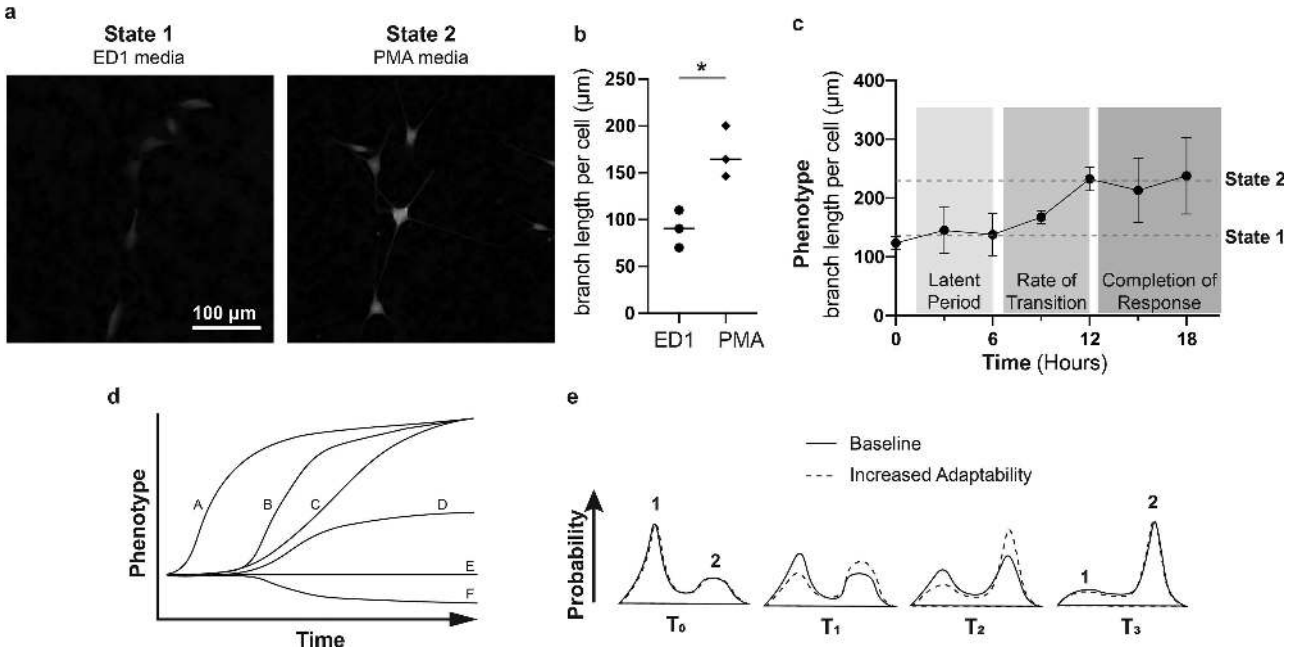


FIGURE 13.1 Adaptability in phenotypic state transitions. (a) Representative QPI images of the phenotypic cell states in primary human melanocytes. (b) Quantification of a, branch length normalized to number of cells, each point represents average measurement from one field of view. (c) Population kinetics of the rate of change of branch length over time after switching the media from ED1 to PMA. Points are the mean and standard deviation (SD) from three fields of view. Colored blocks indicate three distinct stages of an adaptive transition. (d) Example trajectories of an individual cell may take getting from an initial phenotype to a final phenotype. A–C all start and end at the same phenotype but have a different rate of change. D is an incomplete transition. E represents no response. F represents a phenotype where the final state phenotype is decreased compared to the initial phenotype. (e) Example probability curves for a two-state phenotype landscape.

the data from QPI is quantitative, enabling new information, such as cell mass and growth rates, to be derived from the data that is not possible with other approaches.

Many of the applications of QPI in biology take advantage of the specific features of this technology. For example, the fact that QPI directly measures changes in cell mass has been used to monitor cell growth or changes in growth due to drug treatments [17–21], with one application demonstrating a change in growth rate measurement within six hours, versus seven days for proliferation-based measurements [22]. Single-cell data from QPI have been used to measure cell-to-cell heterogeneity in growth rates [23]. The quantitative data provided by QPI have been used to characterize cell states [4, 5, 24, 25], measure cell mechanical properties [6, 26, 27], and characterize multiple measures of cell response to drugs [7].

Here, we highlight new and emerging strategies for QPI kinetic assays to evaluate the transition of cells in the presence of a stimulus that pushes cells to adapt to new conditions. As a tool for functional precision medicine, measuring adaptability across a population with single-cell resolution will improve our mechanistic understanding of cancer adaptation and is a predictive method that could be used to match patients *a priori* to the most effective available therapy.

13.2 Materials

13.2.1 Device Materials and Instrumentation

Ptychography was performed for using a Livecyte imager equipped with the cell analysis toolbox (Phasefocus). DPC was performed using a custom LED array microscope [28] in a cell culture incubator (Thermo Fisher Heracell 150i) and image processing was performed in MatLab (Mathworks).

13.2.2 Reagents

For the state switching assay, we use primary human melanocytes that exhibit a growth factor-dependent change in cell shape and phenotypic state [29, 30]. Cells are cultured in a base media of Cascade Biologics Medium 254 (Gibco) supplemented with either 10 nM Endothelin-1 (ET1, Thermo Fisher Scientific, M254500 and S0165) or 10 ng/ml phorbol 12-myristate 13-acetate (PMA, Thermo Fisher Scientific, M254500 and S0025). Cells were seeded on six well plates (Costar®, 3516).

The spheroid outgrowth assay requires adherent, sphere-forming cells. Here, spheroids of two clonal populations of melanoma cell lines from parental 624MELs and WM793s were prepared in a 96-well Nunc Sphera round bottom non-adherent microplate (Thermo Scientific 174925). A 24-well plate (Corning Costar 3526 or Greiner bio-one cellstar 662 160) was coated with 10 mg/ml fibronectin (Corning 354008), in PBS for 1 h at 37 °C in a cell culture incubator, washed two times with PBS, and stored at 4 °C for one week until ready to plate cells. Cells were cultured in F10 media (Gibco), with 10% FBS and NEAA. 1% penicillin-streptomycin was added prior to imaging.

MDA-MB-231 cells were cultured in RPMI with 10% FBS and 1% penicillin-streptomycin. Cells were plated on 96-well plates (Eppendorf 0030 730.119).

13.3 Methods

13.3.1 General Methods and Considerations for QPI Kinetic Assays

Confluence can greatly impact gene expression, growth dynamics, and differentiation status. Cells were passaged twice a week to maintain a confluence between 30% and 80%. All QPI imaging was performed at 37 °C with 5% CO_2. For phase imaging to be quantitative, it is essential to use high-quality flat-bottom plates and validate the microplates on the system prior to use with cells.

13.3.2 Rate of Phenotypic State Transition

Phenotypic cell state transitions can be monitored for any cell population whose phenotypic states can be distinguished with morphological or behavioral features. Primary human melanocytes naturalized to either state 1 (ED1 media) or state 2 (PMA media) for at least two passages were plated at 40,000 cells per well and allowed to adhere for 24 h. Immediately before imaging, media was either refreshed or switched from ED1 media to PMA-containing media. Cells were imaged by ptychography every 30 min for 24–48 h. Raw images were processed in Fiji to calculate the branch length per cell. Images were thresholded manually to isolate cells from the background. Cells were then skeletonized and branches >15 mm were measured and summed across each field of view. Finally, the image was thresholded to detect cell bodies, and objects greater than 50 μm^2 were counted.

13.3.3 Spheroid Outgrowth

Spheroids at 50 cells/well were cultured for four days before preparation for imaging. First, a 10 µl drop of media was placed in the center of a well of a coated 24-well dish. The media around each sphere was then aspirated. The sphere was then resuspended in ~10 µl of media and added to the droplet on the dish. Spheres were allowed to settle and adhere for 2–4 h. Prior to imaging, media was added dropwise to the side of the well. Spheres were imaged by ptychography every 15–20 min by QPI for 24 h. The number of individual objects was measured using Fiji. First, images were converted to 8-bit, smoothed, and binarized using the Huang2 threshold method. Binary images were then dilated, and holes were filled. The images were eroded to initial size and then residual spots were removed by despeckle. Particles larger than 200 nm were then counted. Sphere spreading was evaluated at 6, 12, 18, and 24 h after the spheres were plated on the dish. A Gaussian blur ($\sigma = 20$) was applied in imageJ. The intensity profile was plotted horizontally across the image across the diameter of the sphere and the intensity values were converted to phase shift (where 2π radians = 65536). Optical mass was determined using an assumed specific refractive increment of 1.8×10^{-4} m^3/kg [12].

13.3.4 Multiparametric Analysis of Cell Growth Dynamics

MDA-MB-231 breast cancer cells were plated at 1500 cells/well 24 h prior to imaging. DPC images were computed using Tikhonov regularization [16]. The resulting phase images were processed using custom Matlab scripts described previously [7] and available on GitHub (github.com/Zangle-Lab/MultiparametricQPI). Dry mass was determined using an assumed specific refractive increment of 1.8×10^{-4} m^3/kg [12]. Cell mass over time was calculated by threshold-based image segmentation followed by tracking using the Crocker and Grier algorithm [31] for x position, y position, and cell mass. Specific growth rate (SGR) was calculated as the slope of a least squares fit line to mass over time data normalized by the mass of each cell. SGR was tracked over time in overlapping sliding windows of 24 h for plotting dynamic changes in populations.

13.4 Results

13.4.1 QPI Captures the Rate of Phenotypic Cell State Change during Melanoma Adaptation

The ability of cells to exist in different phenotypic states most obviously manifests as a change in cellular morphology. Human melanocytes have a notable switch in response to the media supplement PMA that regulates a variety of cell programs including differentiation state, cell cycle, cell adhesion, and susceptibility to oncogenic transformation [29]. This programmatic state change is accompanied by a change in cellular morphology. Cells cultured with ED1-containing media that is free of PMA are characterized by larger and more polygonal cell bodies with small dendrites. A second phenotypic state, observed in PMA-containing media, is distinguished by a smaller cell body and longer dendrites (Figure 13.1a). Importantly, these thin dendrites can be sufficiently resolved by QPI for accurate segmentation and downstream imaging processing. We measured dendrite length by reducing the dendrites to branches using a skeletonize method and tabulated the total branch length per cell. Melanocytes cultured in ED1 have an average branch length of ~100 μm while melanocytes in PMA-containing media have a branch length >150 µm (Figure 13.1b). Notably, this phenotype state switching is reversible, and QPI imaging enables us to evaluate the morphological change over time and capture the latent period, rate of transition, and completion of response (Figure 13.1c).

13.4.2 Spheroid Outgrowth Measures Rate and Depth of Population Response

Migration and subsequent metastasis require cells to leave the tumor and adapt to an array of new environmental conditions. This process can be modeled *in vitro* as outgrowth from a spheroid [32–34]. QPI can capture the rate at which a tumor cell spheroid disassembles into single cells (Figure 13.2a). Segmentation of objects >200 nm successfully captures cell departure from the main spheroid (Figure 13.2b and c). This assay captures the behavior of the most migratory subpopulation of cells as they shift from collective migration as part of the main sphere to single-cell migration.

We extended this assay to capture the rate of sphere disassembly to measure the depth of adaptation in the population

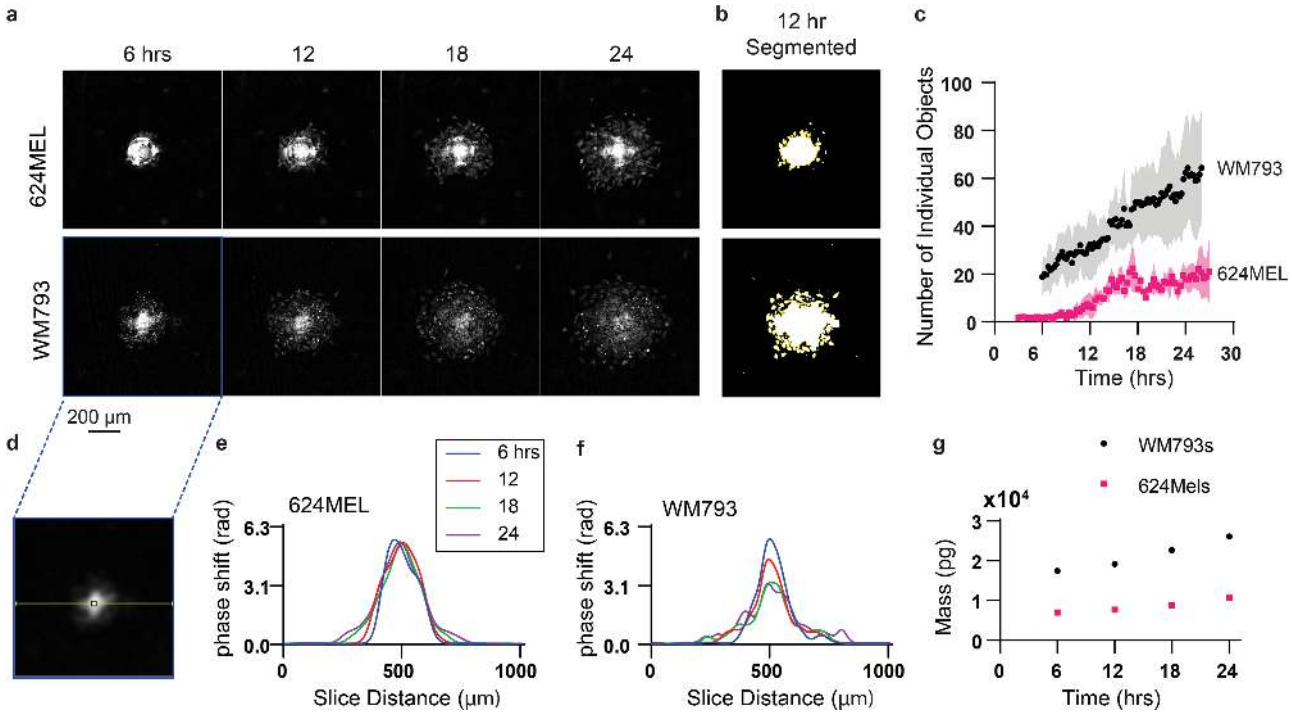


FIGURE 13.2 Spheroid outgrowth measures environmental adaptation. (a) Representative images over 24 h of spheroid outgrowth ($n = 4$). Scale bar is 200 mm. (b) Corresponding segmentation mask for 12 h timepoint. (c) Overall number of objects over time plots for WM793 (black) and 624MEL (pink). Mean values are shown as solid points, error bands are SD. (d) Processed example of WM793 after Gaussian blur. Yellow line shows the section used for e–f. (e–f) Phase shift measured across widest part of sphere. (g) Total dry mass of spheroids over time.

(Figure 13.2d, e, and f). Since QPI quantitatively evaluates cell dry mass by measuring the phase shift of the light as it travels through the object, it provides a topography of the object of interest. In our case, the phase shift across a cross section of the sphere enables visualization of the sphere disassembly over time. By evaluating the profile, we found that 624MELs maintained a spherical organization as cells left (Figure 13.2e), however, in the case of WM793 cells, the sphere fully disassembled over the 24 h imaging period (Figure 13.2f). One limitation of QPI is that if objects cause a phase shift greater than the source wavelength, phase wrapping artifacts introduce quantitative inaccuracies. We empirically determined that spheres prepared at 50 cells approach the upper limit of QPI feasibility for our instrument setup, resulting in a phase shift of $>2\pi$ radians in the sphere center when the sphere was plated three to four days later. However, QPI could still be used effectively for segmentation and sphere flattening as well as mass quantification of departed cells.

13.4.3 Multiparametric Approach to Drug Screening Reveals Cell Dynamics and Heterogeneity

QPI data on cell mass provides the opportunity to monitor the growth response of individual cells in response to therapy [17–21]. Of special interest for clinical applications, QPI data on single-cell drug responses can be used to develop multiple parameters that describe cell response to therapy in terms of more than just EC50 [7]. In this assay, QPI images of individual cells are collected over a period of observation ranging from 24 to 72 h (Figure 13.3a and b). Single cells are then tracked individually, yielding single-cell data showing total mass as a function of time. These data reveal distinct patterns in growing and dying cells at each therapy in the assay (Figure 13.3c). In terms of population response, we can group QPI response data into overlapping 24 h windows and track the time required for the treated population to separate from control. Here, the treated population mean growth rate decreases steadily over the assay relative to control (Figure 13.3d). From the same data, we can track individual cells based on specific features and track their growth rates as a function of time (Figure 13.3d, green/red individual data points). Population response data (Figure 13.3d) can also be characterized at each timepoint in terms of a population mean and standard deviation of specific growth rate. When the mean and standard deviation of SGR are plotted against one another, and so "parameterized" by time as a hidden variable, they reveal distinct trends in dose response behavior (Figure 13.3e). At the lowest drug concentration, the heterogeneity and growth rates of the treated cells are comparable to that of the control group. However, with increased concentrations, the time-dependent response can then be characterized in terms of population mean response and heterogeneity, potentially revealing novel patterns of drug response. Overall, these multiparametric data that can be extracted from QPI dose response experiments yield a more complete picture of tumor cell response to therapy than is provided by standard endpoint, population-level dose response measurements alone.

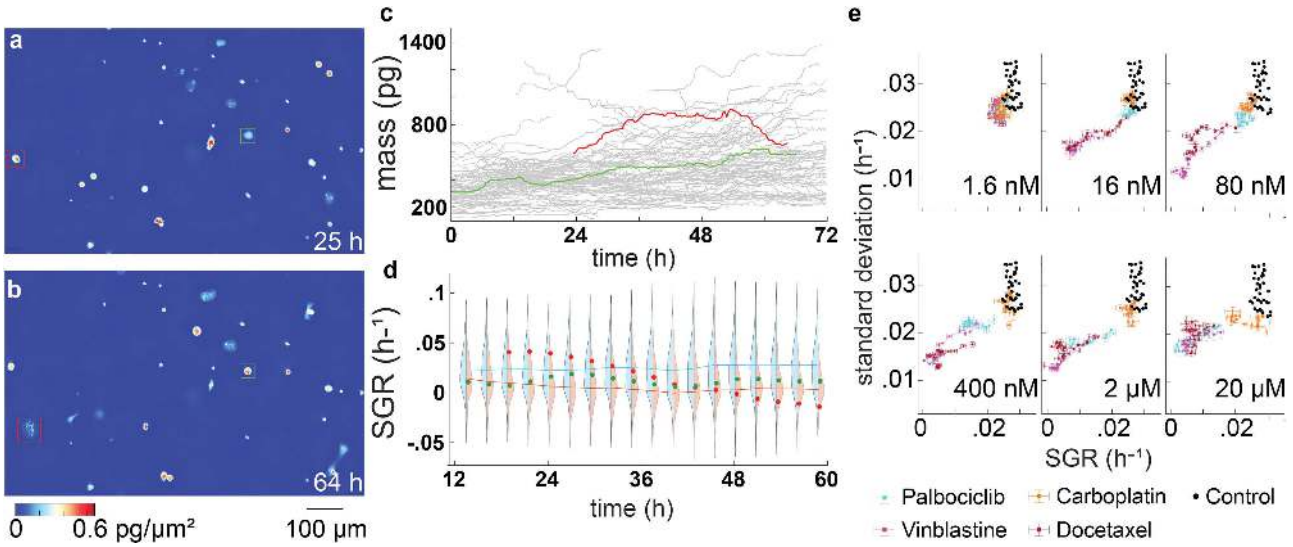


FIGURE 13.3 Measuring response to therapy using QPI. (a) QPI image of MDA-MB-231 cells treated with 20 μM of docetaxel 25 h and (b) 64 h after the start of imaging. Representative growing and dying cells are outlined as green and red, respectively. (c) Overall mass over time plots of the growing (green), dying (red), and other (gray) cells in the corresponding treatment well. (d) Distribution of specific growth rate, growth rate normalized by mass, of 20 μM docetaxel (orange) versus DMSO control (blue) over time. Mean values of SGR are shown as solid lines. The green and red dots show the specific growth rates of the outlined growing and dying cells from panels a–c, respectively. (e) Time parameterized plots of standard deviation of SGR plotted against SGR for MDA-MB-231 cells at different drug concentrations.

13.5 Discussion

13.5.1 Summary of Results

QPI is a versatile label-free platform well suited to image processing workflows for measuring the adaptive response of cancer cells. Here we describe example applications, ranging from both measurement of cell morphology to more complex analysis using custom-built segmentation and tracking tools developed specifically for QPI datasets.

The assays demonstrated here illustrate the basic concepts of kinetic phenotyping (as illustrated in Figure 13.1). By evaluating adaptability as the rate of phenotypic state transition, it is possible to quantify distinctions in the adaptability of two cell populations that are otherwise indistinguishable at the initial and final timepoints of the experiment. The general population change is analogous to trajectory A, however, individual cells may follow different trajectories within that bulk measurement (Figure 13.1c). In the sphere outgrowth assay (Figure 13.2), we show that different melanoma cell lines have different depth of response, similar to trajectories A and D (Figure 13.2c). This is highlighted by the rapid collapse of the sphere in WM793 cells, indicative of a rapid and complete transition to single cells. In contrast, 624MEL cells maintain the sphere with only a portion of the cells returning to adherent culture. In the cancer cell drug response assay (Figure 13.3), where we evaluate single-cell trajectories, there is rich information on both the magnitude and heterogeneity of the response. Single-cell tracks reveal distinct subpopulations of cells with different growth rate phenotypes and can identify both surviving and dying cells, similar to trajectories E and F, respectively.

13.5.2 Clinical Translation

QPI has potential as a method for functional precision oncology but will require rigorous preclinical testing and standardization. With high temporal resolution and picogram sensitivity, QPI can longitudinally track cells in a heterogeneous population to provide meaningful results in just 24–72 h. Critically, QPI assays can be run using relatively small samples of primary cells, reducing the time and cost required for cell expansion. With this quick turnaround, QPI has the potential to provide meaningful data to a clinician within days or weeks of sample acquisition. Current proof-of-concept experiments have been performed on cell lines with low cell-to-cell variability [7, 17]. For translation of this assay, segmentation and image processing workflows will need to accommodate the variation expected in primary clinical samples. Alongside the image processing pipeline, image quality control and visualization tools will be essential for high-quality data to enable successful subsequent analysis [35, 36].

QPI methods vary and will require standardization of instrumentation as well as standardized methods for calibration [37]. Most importantly, the *in vitro* conditions used for QPI testing are distinct from the environment *in vivo* in which tumors grow and adapt. Rigorous comparison to clinical outcomes and combination with other available assays will, therefore, be required to determine which QPI features provide clinically useful predictions of response *in vivo*. Since QPI is a flexible platform that can be used to acquire multidimensional data of a wide variety of phenotypes, these features may take many forms including observable traits like morphology [4, 8], dynamic behaviors like cell mechanical properties and growth rates [6, 7], or patterns of behavior like motility or cell death [5, 30]. New techniques in data mining, modeling,

image processing, and deep learning could all contribute to the utility of QPI to identify and evaluate features that may be useful for classification. In turn, these developments will help develop QPI as a powerful tool for understanding the basic science behind cell adaptability and as a workhorse for functional precision medicine.

13.6 Future Trends in QPI—Illuminating Mechanisms Underlying Adaptability

Beyond identification of adaptable subpopulations and predicting drug response, QPI will likely be useful for revealing mechanisms underlying adaptability. In the most basic form, QPI already provides a detailed spatial map of cell dry mass that can be used to accurately segment cells. QPI can also be readily combined with other imaging modalities such as fluorescence, darkfield, and quadrant darkfield [38], which provide additional information that can be tracked longitudinally and attributed to individual cells. Thus, in more complex workflows, features extracted from QPI data, such as growth dynamics, mechanical properties, cell morphology, and migration, could be correlated with signals from these additional imaging modalities as a powerful tool for both probing basic biological questions and quantifying more nuanced cell responses and behaviors in response to therapies.

Acknowledgments

This work was supported by NCI R01CA276653. We thank Dr. Rachel Belote for helpful conversations and the human melanocyte cell line.

REFERENCES

1. Hanahan D. Hallmarks of cancer: New dimensions. *Cancer Discov.* 2022;12(1):31–46. doi:10.1158/2159-8290.CD-21-1059
2. Feinberg AP, Levchenko A. Epigenetics as a mediator of plasticity in cancer. *Science*. 2023;379(6632). doi:10.1126/science.aaw3835
3. Flaherty KT, Gray RJ, Chen AP, et al. Molecular landscape and actionable alterations in a genomically guided cancer clinical trial: National cancer institute molecular analysis for therapy choice (NCI-MATCH). *J Clin Oncol Off J Am Soc Clin Oncol*. 2020;38(33):3883–3894. doi:10.1200/JCO.19.03010
4. Kamlund S, Janicke B, Alm K, Judson-Torres RL, Oredsson S. Quantifying the rate, degree, and heterogeneity of morphological change during an epithelial to mesenchymal transition using digital holographic cytometry. *Appl Sci Switz*. 2020;10(14). doi:10.3390/app10144726
5. Barker KL, Boucher KM, Judson-Torres RL. Label-free classification of apoptosis, ferroptosis and necroptosis using digital holographic cytometry. *Appl Sci*. 2020;10(13):4439–4439. doi:10.3390/app10134439
6. Nguyen TL, Polanco ER, Patananan AN, Zangle TA, Teitell MA. Cell viscoelasticity is linked to fluctuations in cell biomass distributions. *Sci Rep*. 2020;10(1):1–11. doi:10.1038/s41598-020-64259-y
7. Polanco ER, Moustafa TE, Butterfield A, et al. Multiparametric quantitative phase imaging for real-time, single cell, drug screening in breast cancer. *Commun Biol*. 2022;5(1):794–794. doi:10.1038/s42003-022-03759-1
8. Pradeep S, Tasnim T, Zhang H, Zangle TA. Simultaneous measurement of neurite and neural body mass accumulation: Via quantitative phase imaging. *Analyst*. 2021;146(4):1361–1368. doi:10.1039/d0an01961e
9. van Renterghem AWJ, van de Haar J, Voest EE. Functional precision oncology using patient-derived assays: bridging genotype and phenotype. *Nat Rev Clin Oncol*. 2023;20(5):305–317. doi:10.1038/s41571-023-00745-2
10. Letai A, Bhola P, Welm AL. Functional precision oncology: Testing tumors with drugs to identify vulnerabilities and novel combinations. *Cancer Cell*. 2022;40(1):26–35. doi:10.1016/j.ccell.2021.12.004
11. Davies HG, Wilkins MHF. Interference microscopy and mass determination. *Nature*. 1952;169(4300):541–541. doi:10.1038/169541a0
12. Zangle TA, Teitell MA. Live-cell mass profiling: An emerging approach in quantitative biophysics. *Nat Methods*. 2014;11(12):1221–1228. doi:10.1038/nmeth.3175
13. Nguyen TL, Pradeep S, Judson-Torres RL, Reed J, Teitell MA, Zangle TA. Quantitative phase imaging: recent advances and expanding potential in biomedicine. *ACS Nano*. 2022;16(8):11516–11544. doi:10.1021/acsnano.1c11507
14. Park YK, Depeursinge C, Popescu G. Quantitative phase imaging in biomedicine. *Nat Photonics*. 2018;12(10):578–589. doi:10.1038/s41566-018-0253-x
15. Mehta SB, Sheppard CJR. Quantitative phase-gradient imaging at high resolution with asymmetric illumination-based differential phase contrast. *Opt Lett*. 2009;34(13):1924–1924. doi:10.1364/ol.34.001924
16. Tian L, Waller L. Quantitative differential phase contrast imaging in an LED array microscope. *Opt Express*. 2015;23(9):11394–11394. doi:10.1364/oe.23.011394
17. Reed J, Chun J, Zangle TA, et al. Rapid, massively parallel single-cell drug response measurements via live cell interferometry. *Biophys J*. 2011;101(5):1025–1031. doi:10.1016/j.bpj.2011.07.022
18. Cooper KL, Oh S, Sung Y, Dasari RR, Kirschner MW, Tabin CJ. Multiple phases of chondrocyte enlargement underlie differences in skeletal proportions. *Nature*. 2013;495(7441):375–378. doi:10.1038/nature11940
19. Mir M, Wang Z, Shen Z, et al. Optical measurement of cycle-dependent cell growth. *Proc Natl Acad Sci U S A*. 2011;108(32):13124–13129. doi:10.1073/pnas.1100506108
20. Kemper B, Bauwens A, Vollmer A, et al. Label-free quantitative cell division monitoring of endothelial cells by digital holographic microscopy. *J Biomed Opt*. 2010;15(03):1–6. doi:10.1117/1.3431712
21. Popescu G, Park YK, Lue N, et al. Optical imaging of cell mass and growth dynamics. *Am J Physiol - Cell Physiol*. 2008;295(2):538–544. doi:10.1152/ajpcell.00121.2008
22. Chun J, Zangle TA, Kolarova T, Finn RS, Teitell MA, Reed J. Rapidly quantifying drug sensitivity of dispersed and clumped breast cancer cells by mass profiling. *Analyst*. 2012;137(23):5495–5498. doi:10.1039/c2an36058f

23. Huang D, Leslie KA, Guest D, et al. High-speed live-cell interferometry: A new method for quantifying tumor drug resistance and heterogeneity. *Anal Chem*. 2018;90(5):3299–3306. doi:10.1021/acs.analchem.7b04828
24. Hejna M, Jorapur A, Song JS, Judson RL. High accuracy label-free classification of single-cell kinetic states from holographic cytometry of human melanoma cells. *Sci Rep*. 2017;7(1):1–12. doi:10.1038/s41598-017-12165-1
25. Huang D, Roy IJ, Murray GF, Reed J, Zangle TA, Teitell MA. Identifying fates of cancer cells exposed to mitotic inhibitors by quantitative phase imaging. *Analyst*. 2020; 145(1):97–106. doi:10.1039/c9an01346f
26. Park YK, Best CA, Badizadegan K, et al. Measurement of red blood cell mechanics during morphological changes. *Proc Natl Acad Sci U S A*. 2010;107(15):6731–6736. doi:10.1073/pnas.0909533107
27. Eldridge WJ, Ceballos S, Shah T, et al. Shear modulus measurement by quantitative phase imaging and correlation with atomic force microscopy. *Biophys J*. 2019;117(4):696–705. doi:10.1016/j.bpj.2019.07.008
28. Moustafa TE, Polanco ER, Belote RL, Judson-Torres RL, Zangle TA. Fabrication and validation of an LED array microscope for multimodal, quantitative imaging. *HardwareX*. 2023;13:e00399. doi:10.1016/j.ohx.2023.e00399
29. McNeal AS, Belote RL, Zeng H, et al. BRAFV600E induces reversible mitotic arrest in human melanocytes via microRNA-mediated suppression of AURKB. *eLife*. 2021;10:1–26. doi:10.7554/eLife.70385
30. Zeng H, Jorapur A, Shain AH, et al. Bi-allelic loss of CDKN2A initiates melanoma invasion via BRN2 activation. *Cancer Cell*. 2018;34(1):56–68.e9. doi:10.1016/j.ccell.2018.05.014
31. Crocker JC, Grier DG. Methods of digital video microscopy for colloidal studies. *J Colloid Interface Sci*. 1996;179: 298–310.
32. Weiswald LB, Bellet D, Dangles-Marie V. Spherical cancer models in tumor biology. *Neoplasia U S*. 2015;17(1):1–15. doi:10.1016/j.neo.2014.12.004
33. Vinci M, Gowan S, Boxall F, et al. Advances in establishment and analysis of three-dimensional tumor spheroid-based functional assays for target validation and drug evaluation. *BMC Biol*. 2012;10. doi:10.1186/1741-7007-10-29
34. Hou Y, Konen J, Brat DJ, Marcus AI, Cooper LAD. TASI: A software tool for spatial-temporal quantification of tumor spheroid dynamics. *Sci Rep*. 2018;8(1):7248–7248. doi:10.1038/s41598-018-25337-4
35. Lange D, Polanco E, Judson-Torres R, Zangle T, Lex A. Loon: Using exemplars to visualize large-scale microscopy data. *IEEE Trans Vis Comput Graph*. 2022;28(1):248–258. doi:10.1109/TVCG.2021.3114766
36. Lange D, Judson-Torres R, Zangle TA, Lex A. Aardvark: Composite Visualizations of Trees, Time-Series, and Images. *IEEE Trans Vis Comput Graph*. 2025;31(1):1290–1300. doi:10.1109/TVCG.2024.3456193
37. Kwee E, Peterson A, Halter M, Elliott J. Practical application of microsphere samples for benchmarking a quantitative phase imaging system. *Cytometry A*. 2021;99(10):1022–1032. doi:10.1002/cyto.a.24291
38. Moustafa TE, Belote RL, Polanco ER, Judson-Torres RL, Zangle TA. Quadrant darkfield (QDF) for label-free imaging of intracellular puncta. *bioRxiv*. Published online January 1, 2024:2024.08.05.606686. doi:10.1117/1.JBO.29.11.116501

14

An Exo-PROS Biosensor Simultaneously Detects Tumor-Derived Exosomal Protein-MicroRNA Pairs for Lung and Breast Cancer Diagnosis

Chang-Chieh Hsu and Yun Wu

14.1 Introduction

Cancer remains one of the leading causes of death worldwide. Effective screening and early detection of cancer have proven to improve patient outcome and reduce cancer mortality [1–4]. Current cancer screening and diagnostic methods, such as low-dose computed tomography (CT) for lung cancer and mammography for breast cancer, face significant challenges including high false positive/negative rates, limited tumor information, and potential radiation exposure risks. Although tissue biopsy stands as the gold standard for cancer diagnosis, it is limited by the invasive procedure and difficulties in repeated sampling. Liquid biopsy represents a promising alternative. By detecting circulating cancer biomarkers, it supplements medical imaging and risk factor data and permits sequential monitoring of cancer progression. Therefore, liquid biopsy holds great potential as an *in vitro* diagnostic tool to improve cancer care.

Exosomes are small extracellular vesicles secreted by the cells. They transport various biomolecules including DNAs, RNAs, proteins, and lipids between parental cells and recipient cells to facilitate intercellular communication. Exosomes are actively involved in cancer growth [5, 6], angiogenesis [7], and metastasis [8], making them promising biomarkers in cancer liquid biopsy [9, 10]. However, the development of exosome-based cancer liquid biopsy is limited by low sensitivity [11], labor-intensive and time consuming processes [12], large sample consumption, and high cost associated with conventional characterization methods, including enzyme-linked immunosorbent assay (ELISA), western blot, flow cytometry, liquid chromatography-mass spectrometry (LC-MS), quantitative reverse transcription polymerase chain reaction (qRT-PCR), next-generation sequencing, and micro-array. Moreover, a significant challenge encountered by these approaches is their inability to differentiate tumor-derived exosomes (TEXs) from non-tumor exosomes. Given that TEXs represent only a small fraction of total exosomes, interference from non-TEXs can significantly compromise the sensitivity and specificity of these tests in cancer diagnosis.

To address these challenges, emerging technologies have been developed to detect exosomal biomarkers including SPR sensors [13–16], SERS sensors [17–20], nPLEX sensor [21], Exo-Screen [22], EV array [23, 24], iMEX sensor [25] uNMR sensor [26], hMFEX sensor [27], FET sensors [28, 29], tethered cationic nanoparticles (tCLN)-based sensors [30–32], electrochemical biosensors [33–35], and microfluidic chips [36–40]. However, these new sensing technologies mainly detect a single type of exosomal biomarkers, either proteins or RNAs at each time. Recent studies have suggested that using different types of exosomal biomarkers can further increase the sensitivity, specificity, and accuracy in cancer diagnosis [41, 42]. Therefore, we developed an exosome protein microRNA one-stop (Exo-PROS) biosensor to enable simultaneous and *in situ* detection of TEX protein-microRNA pairs in a simple, fast, reliable, low sample consumption and user-friendly way [43].

The Exo-PROS biosensor utilizes surface plasmon resonance (SPR) and surface plasmon enhanced fluorescence spectroscopy (SPEFS) as sensing mechanisms for the detection of TEX proteins and microRNAs, respectively (Figure 14.1). To perform the Exo-PROS assay, a biochip (Figure 14.1b insert) was fabricated by depositing a 2 nm Ti layer followed by a 49 nm Au film onto a glass slide. Then the surface of the biochip was coated with a mixture of PEG200-SH and biotin-PEG1000-SH followed by neutravidin. Lastly, biotinylated antibodies targeting overexpressed tumor proteins were added to bind the antibodies onto the biochip surface via biotin-avidin interactions.

To detect TEX proteins, background signals from water and PBS (I_{water} and I_{PBS}) were first collected. Next, exosomes were applied to facilitate the capture of TEXs by the antibodies. This capture altered the local refractive index, changed the optical characteristics of the surface plasmon modes, and enabled the optical detection of TEX proteins. Following the removal of unbound exosomes using PBS washes, the intensity of reflected laser light (I_{TEX}) was recorded. The expression of TEX proteins was calculated using the following equation:

$$Expression\ of\ TEX_{protein} = \frac{I_{TEX} - I_{PBS}}{I_{PBS} - I_{water}} \tag{14.1}$$

where I_{TEX}, I_{PBS}, and I_{water} were the reflected laser intensities of TEXs, PBS, and water. The difference between I_{PBS}

DOI: 10.1201/9781003449942-16

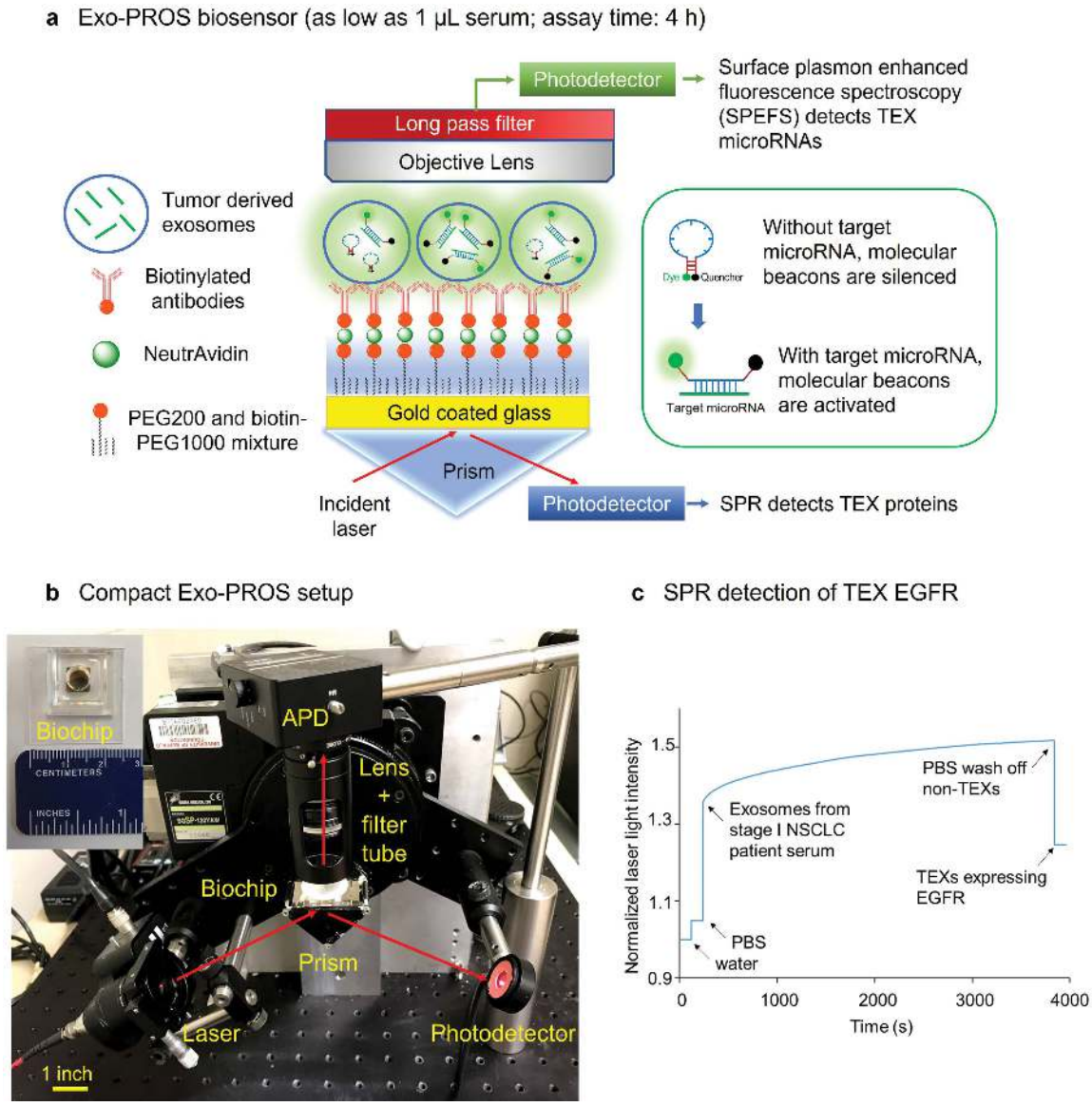


FIGURE 14.1 *Overview of the Exo-PROS assay.* (a) Schematics and sensing mechanism of Exo-PROS biosensor. Insert: sensing mechanism of molecular beacons. (b) A photo of the Exo-PROS biosensor. APD: avalanche photodetector. Insert: a photo of the biochip. (c) A representative SPR curve for measuring TEX EGFR expression in the serum sample from a stage I non-small cell lung cancer (NSCLC) patient. (Reproduced from Hsu et al., 2023 [43] with permission from ACS Publications.)

and I_{water} was used as the normalization factor to remove the biochip-to-biochip variation. A representative SPR curve for the detection of TEX EGFR in exosomes isolated from a stage I non-small cell lung cancer (NSCLC) patient serum is shown in Figure 14.1c.

Next, to detect TEX microRNAs, molecular beacons (MBs), the oligonucleotide hybridization probes for microRNA sensing, were applied. After MBs diffused into captured TEXs, they hybridized to target microRNAs and generated fluorescence signals (Figure 14.1a). The fluorescence signals were significantly amplified by the SPR system, enhancing them by approximately 100 times. Consequently, they could be readily detected using an avalanche photodetector (APD) positioned above the biochip (Figure 14.1b). The levels of TEX microRNAs were calculated based on the fluorescence intensities before and after the addition of MBs via the following equation:

$$Expression\ of\ TEX_{microRNA} = \frac{I_{after\ MB} - I_{before\ MB}}{I_{before\ MB}} \quad (14.2)$$

where $I_{before\ MB}$ and $I_{after\ MB}$ were fluorescence intensities collected before and after the addition of MBs.

Lung cancer and breast cancer were selected as the disease models to demonstrate the clinical potential of Exo-PROS assay. In lung cancer, when TEX epidermal growth factor receptor (EGFR)-miR-21 pair [44–47] and galectin-3-binding

protein (LG3BP)-miR-210 pair [48–50] were used as the biomarkers, we successfully distinguished lung cancer patients (n = 60, stage I–IV) from normal controls (n = 35, 15 low-risk and 20 high-risk controls) with accuracy of 99%. In breast cancer, TEX Annexin A8 (ANXA8)-miR-342 pair was selected as the biomarkers [51–54]. We successfully differentiated breast cancer patients (n = 60, stage 0–IV) from normal controls (n = 20) with accuracy of 96%.

Our results demonstrated that evaluating two types of TEX biomarkers simultaneously with Exo-PROS assay is an effective strategy to achieve highly sensitive and specific cancer diagnosis. Exo-PROS assay is a highly sensitive, fast (assay time: 4 h), reliable (coefficient of variation < 10%), low sample consumption (10 µL serum needed), user-friendly (only a pipette is required), and cost-effective test, which holds great clinical translational potential.

14.2 Materials

14.2.1 Optical Components of the Exo-PROS Biosensor

1. A solid-state 647 nm laser (Coherent, Cat #: OBIS 647 LX, 100 mW).
2. A trapezoid prism, regular glass (25.4 mm width × 35 mm length × 17 mm height. Trapezoid angle: top 73°, bottom 107°. Custom made by Jianbolaite Company, China).
3. A photodetector 400-1100 nm, 500 mV (Thorlabs, Cat #: S121C).
4. Analog handheld laser power meter console (Thorlabs, Cat #: PM100A).
5. A Si Variable-Gain Avalanche Detector, temperature compensated, wavelength range of 400 to 1000 nm, DC—10 MHz, M4 Taps. Detection area diameter of 1 mm (Thorlabs, Cat #: APD410A/M).
6. Two bi-convex lenses, uncoated, f = 25.3 mm, Ø 25.4 mm, thickness 1.9 mm (min) ~ 9.0 mm (max) (Thorlabs, Cat #: LB1761-N-BK7).
7. A 647 nm EdgeBasic™ long-pass edge filter (Semrock, Cat #: BLP01–647R-25).

14.2.2 Instrumentation

1. A two-axis rotational stage (Sigma Koki, Cat #: SGSP-120YAW).
2. A two-axis stage controller (Sigma Koki, Cat #: SHOT-302GS).
3. A digital multimeter (OWON, Cat #: XDM2041).
4. A shutter (Thorlabs, Cat #: SH-05).
5. A K-Cube™ Solenoid Controller (Thorlabs, Cat #: KSC10).
6. Ultrasonic cleaner (Branson Ultrasonics, Cat #: Branson 2800).
7. NanoSight, nanoparticle tracking analysis (NTA) system (Malvern Instruments Ltd., Cat #: LM10).
8. AXXIS electron-beam evaporator (Kurt J. Lesker Company).
9. Eppendorf 5804R and 5424R centrifuges.
10. Beckman coulter L-90K optima ultracentrifuge.
11. Benchtop low-pressure plasma system (Plasma Etch, Inc., Cat #: PE-50 XL).

14.2.3 Materials and Supplies

1. Glass slides (Fisher Scientific, Cat #: 12-550-15).
2. Centrifuge tubes and microcentrifuge tubes (Fisher Scientific).
3. Deionized water.
4. Acetone (VWR, Cat #: BDH1101–4LP).
5. Alcohol (Decon Laboratories, Inc., Cat #: 2716).
6. SYLGARD® 184 silicone elastomer kit (Dow Corning Corporation, Cat #: 4019862).
7. P150 Petri dishes (Fisher Scientific).
8. Biopunch tissue punches (Ted Pella Inc.).
9. Immerse oil, n=1.518, Olympus Type F, 30 mL (Thorlabs, Cat #: MOIL-30).
10. 10X PBS buffer (Fisher Scientific, Cat #: BP3994).
11. Methyl-polyethylene glycol-thiol PEG200 (Thermo Fisher Scientific, Cat #: 26132).
12. Biotinylated-polyethylene glycol-thiol PEG1000 (Nanocs, Cat #: PG2-BNTH-1K).
13. NeutrAvidin (Thermo Fisher Scientific, Cat #: 31000).
14. Biotinylated anti-EGFR antibodies (Abcam, Cat #: ab24293).
15. Biotinylated anti-LG3BP antibodies (LSBio, Cat #: LS-C686310–1).
16. Anti-ANXA8 antibodies (LSBio, Cat #: LS-C403822–200).
17. miR-21-sensing molecular beacon (MB) 5'-Cy5-CGCGATCTCAACATCAGTCTGATAAGCTAG-ATCGCG-BHQ3–3' (custom synthesized by Alpha DNA).
18. miR-210-sensing MB 5'-Cy5-CGCGATCTCAGC CGCTGTCACACGCACAGGATCGCG-BHQ3–3' (custom synthesized by Alpha DNA).
19. miR-342-sensing MB 5'-Cy5-CGCGAT-CACGGGT GCGATTTCTGTGTGAGAGATCGCG-BHQ3–3' (custom synthesized by Alpha DNA).
20. RPMI 1640 medium (Thermo Fisher Scientific, Cat #: 11875093).
21. DMEM medium (Thermo Fisher Scientific, Cat #: 11995073).
22. DMEM/F-12 medium (Thermo Fisher Scientific, Cat #: 11330032).
23. Fetal bovine serum (Thermo Fisher Scientific, Cat #: 26140–079).
24. Penicillin-streptomycin (Thermo Fisher Scientific, Cat #: 15140–122).

25. EZ-Link Micro Sulfo-NHS-LC-Biotinylation kit (Thermo Fisher Scientific, Cat #: 21935).
26. Deidentified human serum samples are obtained from Roswell Park Comprehensive Cancer Center (Buffalo, NY).
27. Total exosome isolation reagent (from serum, Thermo Fisher Scientific, Cat #: 4478360).
28. Exosome isolation kit (from cell culture medium, Thermo Fisher Scientific, Cat#: 4478359).

14.2.4 Software for Statistical Data Analysis

1. R software to perform binary logistic regression and receiver operating characteristic (ROC) analysis.
2. PASS software to perform power analysis.

14.3 Methods

14.3.1 Preparation of Exo-PROS Biochip

1. The glass slides are immersed in acetone in a coplin jar and cleaned with an ultrasonic cleaner for 10 min. Then the glass slides are blow dried with the filtered air.
2. Then the glass slides are immersed in ethanol in a coplin jar and cleaned with an ultrasonic cleaner for another 10 min. Then the glass slides are blow dried with the filtered air.
3. Finally, the glass slides are immersed in DI water in a coplin jar and cleaned with an ultrasonic cleaner for 10 min. Then the glass slides are blow dried with the filtered air.
4. The cleaned glass slides are secured on the nickel plates and placed inside the vacuum chamber of the electron-beam evaporator. When the pressure of the vacuum chamber reaches 7.5×10^{-6} mmHg, Ti is deposited on the glass slide surface at the deposition rate of 0.8 Å/sec. After a 2 nm Ti adhesion layer is deposited onto the glass slides, a 49 nm Au thin film is deposited at the same rate of 0.8 Å/sec.
5. A square polydimethylsiloxane (PDMS) cube with a hole of 6 mm diameter in the center is prepared to serve as the sample well.
6. The PDMS cube with the 6 mm hole and the gold-coated glass slide are treated with oxygen plasma for 75 seconds. The PDMS cube is bound onto the glass slide by pressing it firmly onto the slide. The biochip is ready for further use.

14.3.2 Surface Modification of Exo-PROS Biochip

1. The biochip is first cleaned with 100 µL 1X PBS three times.
2. A mixture of PEG200-SH and biotin-PEG1000-SH (100 µL, molar ratio of 3:1, 10 mM in 1X PBS) is added into the sample well and incubated at room temperature for 1 h.
3. After 1 h incubation, unbound PEG is removed by washing with 100 µL 1X PBS each time, total three times.
4. Then, neutravidin (50 µg/mL, 100µL) is added to the sample well and incubated at room temperature for 1 h.
5. After 1 h incubation, unbound neutravidin is removed by washing with 100 µL 1X PBS each time, total three times.
6. Finally, biotinylated antibodies, such as biotinylated anti-EGFR antibodies, biotinylated anti-LG3BP antibodies, or biotinylated anti-ANXA8 antibodies (50 µg/mL, 50 µL), are added to the sample well and incubated at room temperature for 1 h to allow the antibodies to attach on the surface of the biochip through biotin-avidin interaction.
7. Unbound antibodies are removed by washing with 100 µL 1X PBS each time, total three times. The biochip is ready for the detection of TEX biomarkers.

14.3.3 Isolation of Exosomes from Cell Culture Media

1. A549 and BEAS-2B cells are cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1× penicillin-streptomycin (PS).
2. MDA-MB-231 cells are maintained in DMEM medium supplemented with 10% FBS and 1× PS. MCF-10A cells are cultured in DMEM/F-12 medium supplemented with 10% FBS and 1× PS.
3. All the cells are seeded in P100 dishes at a density of 1.0×10^6 cells for A549 cells, 1.5×10^6 cells for BEAS-2B cells, 1.0×10^6 cells for MDA-MB-231 cells, and 2.0×10^6 cells for MCF-10A cells.
4. When the cell density reaches 80% confluence, the cells are starved for 48 hours by replacing the regular cell culture media with plain cell culture media containing no FBS and PS.
5. After 48 hours, the cell culture conditioned media are collected and centrifuged at 4,000 g for 30 min at 4 °C to remove cells.
6. The supernatant is transferred to a new tube and centrifuged at 10,000 g for 1 h at 4 °C to remove cell debris.
7. The supernatant is collected and mixed thoroughly with the total exosome isolation kit (from cell culture medium) at a volume ratio of 2:1 and incubated at 4 °C overnight.
8. The exosomes are pelleted by 10,000 g centrifugation at 4 °C for 1 h.
9. The exosomes are resuspended in 100 µL 1X PBS for further analysis.

14.3.4 Isolation of Exosomes from Human Serum Samples

1. Human serum samples are centrifuged at 10,000 g for 30 min at 4 °C to remove debris. The supernatant is transferred to a new tube.

2. Following the manufacturer's instruction, the total exosome isolation reagent (from serum, 10 μL) is mixed thoroughly with serum (50 μL) at a volume ratio of 1:5.
3. After the mixture is incubated at 4 °C for 30 min, it is centrifuged at 10,000 g for 10 min at room temperature to pellet exosomes.
4. The supernatant is removed, and the exosome pellet is resuspended in 50 μL 1X PBS for downstream applications.
5. To demonstrate the Exo-PROS assay can directly capture exosomes from serum without exosome isolation, 10 μL serum is diluted with 40 μL PBS and immediately applied on the biochip for TEX biomarkers detection.

14.3.5 Detection of TEX Protein-MicroRNA Pairs via the Exo-PROS Biosensor

1. A 647 nm laser module and a heat sink block are installed on the optical table. The laser beam is collimated and passed through a polarizer. The output of the laser power is set at 30 mW.
2. On the two-axis rotational stage, the optical fiber of the laser is installed on one axis and the photodetector is installed on another axis.
3. One drop of immerse oil is applied on the prism, then the biochip is placed on the prism. The biochip can be secured on the prism with a piece of invisible tape if necessary.
4. The sample well of the biochip is rinsed with 100 μL 1X PBS three times. Then 100 μL fresh DI water is added to the sample well.
5. The optical fiber of the laser and the photodetector are rotated using the two-axis rotational stage to allow the incident laser beam to shed at the center of the sample well and the reflected laser beam to shed at the center of the photodetector.
6. The incident angle of the laser beam is then fine-tuned to the surface plasmon resonance angle. Briefly, the incident angle of the laser beam is adjusted to achieve the lowest intensity readout of the reflected laser beam. When SPR is established, most of the incident laser's energy is deposited on the gold surface; therefore, the intensity of reflected laser beam is the lowest.
7. The intensity of the reflected light (I_{water}) is recorded for 2 min.
8. The DI water is removed and 100 μL 1X PBS is added. The intensity of the reflected light (I_{PBS}) is recorded for another 2 min.
9. After 1X PBS is removed, exosomes (50 μL in 1X PBS) or serum sample (10 μL serum diluted in 40 μL in 1X PBS, totaling 50 μL) are added to the sample well using a micropipette. The intensity of the reflected light is recorded for 60 min.
10. After 60 min incubation, nonspecific bound exosomes are removed by three times of 100 μL 1X PBS washing.
11. 100 μL 1X PBS is added to the sample well. The intensity of the reflected light (I_{TEX}) is recorded for 2 min.
12. The expression of TEX proteins ($TEX_{protein}$) is calculated using Equation 14.1.

$$Expression\ of\ TEX_{protein} = \frac{I_{TEX} - I_{PBS}}{I_{PBS} - I_{Water}} \quad (14.1)$$

A representative real-time SPR curve for the detection of TEX EGFR in exosomes isolated from the serum of a NSCLC patient is shown in Figure 14.1c.

13. To detect TEX microRNAs, an avalanche photodiode detector is placed on the top of the biochip, and the baseline fluorescence intensity ($I_{Before\ MB}$) is recorded.
14. The molecular beacons (100 μL, 3μM) that detect specific microRNA target, such as miR-21, are added and incubated with TEX for 1.5 h.
15. Excess molecular beacons are removed by three times of 100 μL 1X PBS washing. The fluorescence intensity of MBs ($I_{after\ MB}$) is recorded.
16. The expression of TEX microRNAs ($TEX_{microRNA}$) is calculated using Equation 14.2.

$$Expression\ of\ TEX_{microRNA} = \frac{I_{after\ MB} - I_{before\ MB}}{I_{before\ MB}} \quad (14.2)$$

14.4 Results

14.4.1 Demonstration of Clinical Potential of Exo-PROS Assay in Cancer Diagnosis Using Cell-Derived Exosomes

The clinical potential of Exo-PROS assay in detecting lung cancer and breast cancer was demonstrated using cell-derived exosomes. For lung cancer detection, exosomes obtained from A549 NSCLC cells and BEAS-2B human normal bronchial epithelial cells were applied onto the biochips at the same concentration of 1.25×10^{12} exosomes/mL. The expression of exosomal EGFR-miR-21 and LG3BP-miR-210 pairs was quantified. As illustrated in Figure 14.2a and b, the levels of exosomal EGFR-miR-21 pair from A549 cells were 4.12-fold and 3.56-fold higher compared to those from BEAS-2B cells. Similarly, the expression of exosomal LG3BP-miR-210 pair from A549 cells was 2.49-fold and 2.82-fold higher than those from BEAS-2B cells.

For breast cancer diagnosis, exosomes isolated from MDA-MB-231 breast cancer cells and MCF-10A normal cells were applied onto the biochips at the same concentration of 10^{11} exosomes/mL. The expression of exosomal ANXA8 from MDA-MB-231 cells was 2.07-fold higher than those from

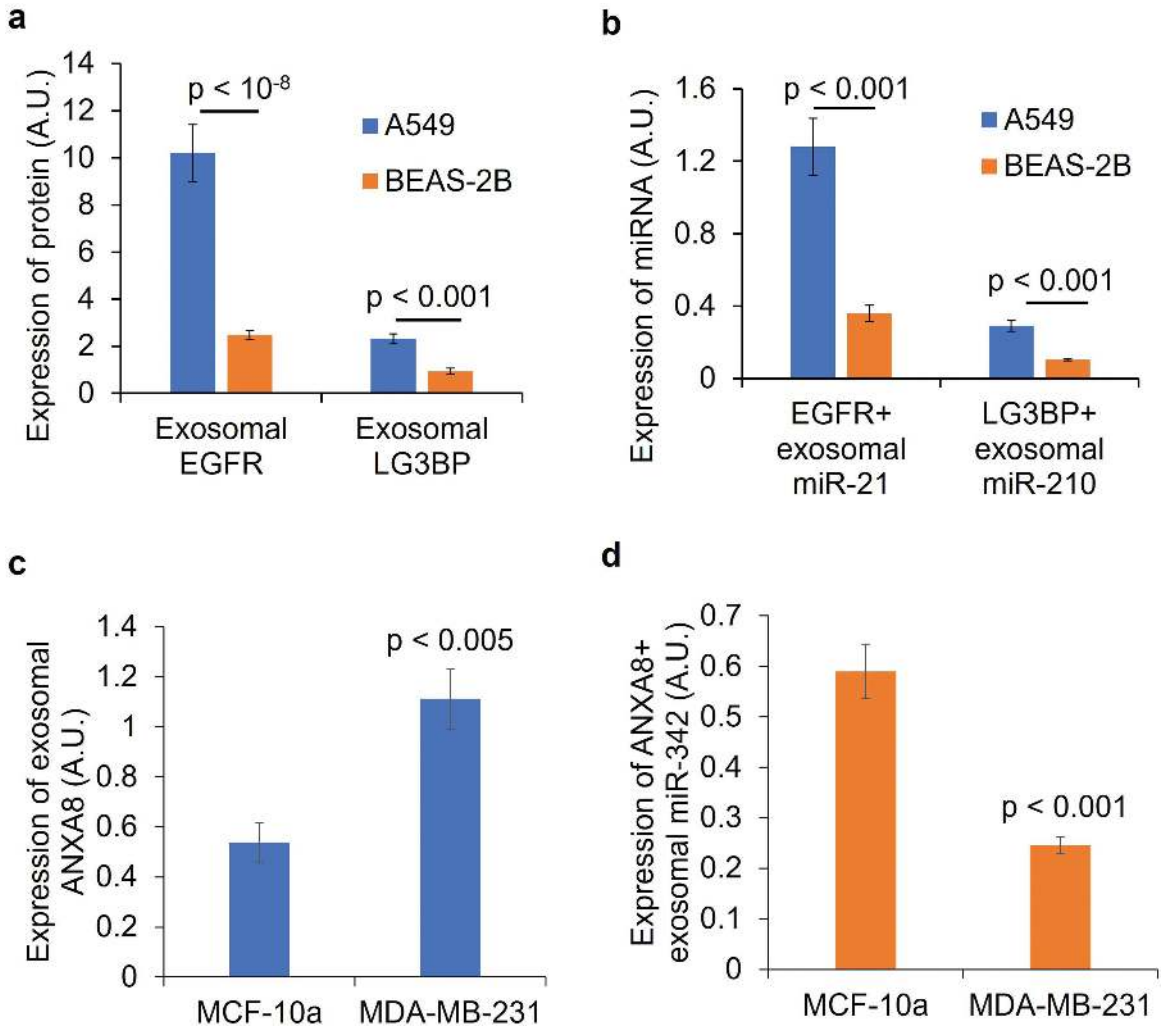
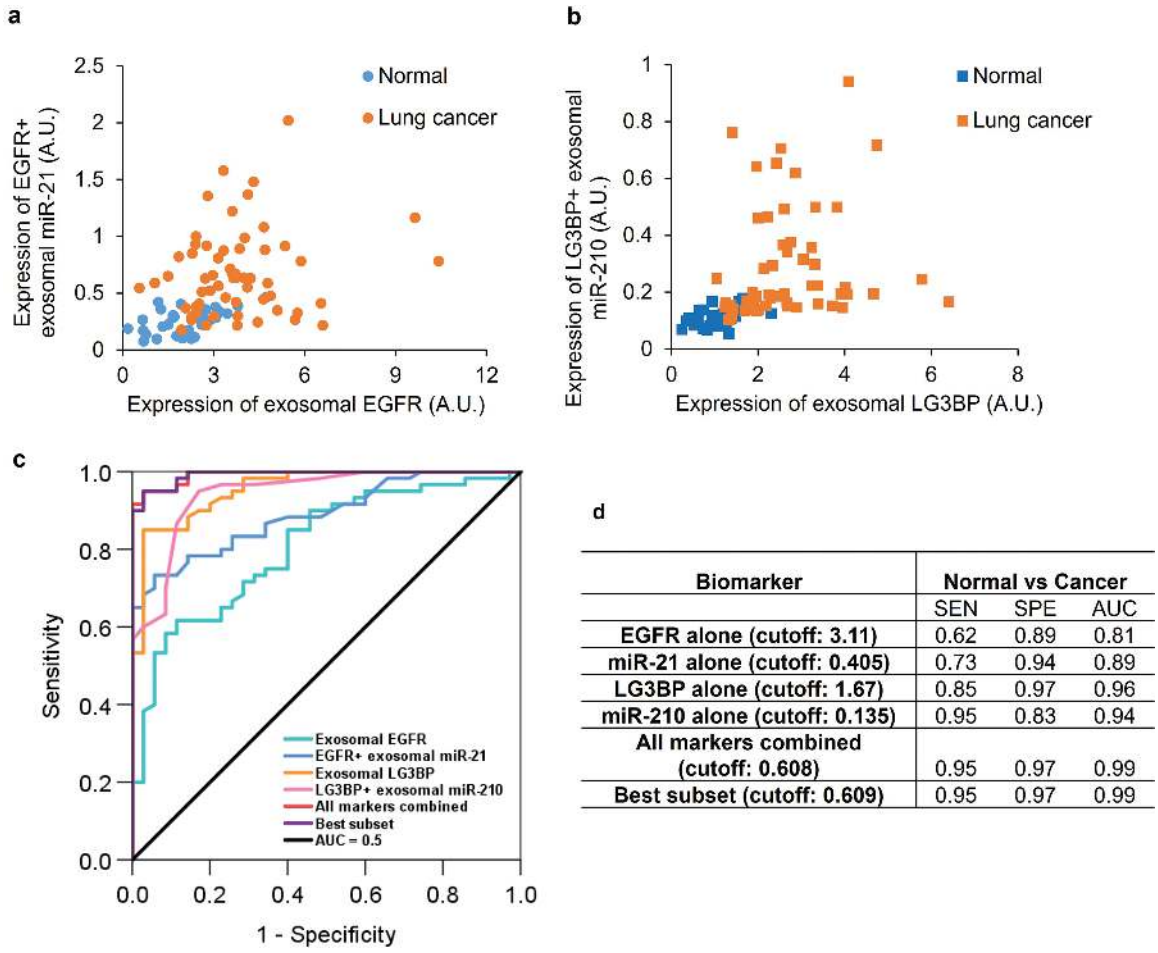


FIGURE 14.2 *Demonstration of clinical potential of Exo-PROS assay in cancer diagnosis using cell-derived exosomes.* (a, b) Exo-PROS assay detected significantly higher levels of EGFR-miR-21 pair and LG3BP-miR-210 pair in A549 cell-derived exosomes than BEAS-2B cell-derived exosomes at the concentration of $1.25\acute{}10^{12}$ exosomes/mL (n = 3). (c, d) Exo-PROS assay detected significantly higher expression of ANXA8 and lower expression of miR-342 in MDA-MB-232 cell-derived exosomes than those from MCF-10a cells at concentration of 10^{11} exosomes/mL (n = 3). (Reproduced from Hsu et al., 2023 [43] with permission from ACS Publications.)

MCF-10A cells (Figure 14.2c). MiR-342 is one of the microRNAs that directly down-regulates the expression of ANXA8. As expected, the expression of ANXA8+ exosomal miR-342 from MDA-MB-231 cells was 2.40-fold lower than those from MCF-10A cells (Figure 14.2d). These findings demonstrated the efficacy of Exo-PROS assay in discriminating cancer groups from normal controls.

14.4.2 Demonstration of Clinical Potential of Exo-PROS Assay in Cancer Diagnosis Using Human Serum Samples

Lung cancer and breast cancer were selected as the model diseases to demonstrate the clinical potential of using Exo-PROS assay for cancer diagnosis. For lung cancer detection, 35 normal controls and 60 treatment naïve patients with stage I–IV lung cancer were included. The normal controls included 15 controls at low risk of lung cancer and 20 controls at high risk of lung cancer. The lung cancer patients included 52 NSCLC patients and eight small cell lung cancer (SCLC) patients. Exosomes were isolated from 50 µL serum samples and resuspended in the same volume of PBS. The levels of TEX EGFR-miR-21 and LG3BP-miR-210 pairs were measured using the Exo-PROS assay. The expression of all four biomarkers was significantly higher in exosomes from lung cancer patients than normal controls. Specifically, the expression of EGFR, miR-21, LG3BP, and miR-210 was 1.76-fold, 2.64-fold, 2.74-fold, and 2.64-fold higher for lung cancer patients than normal controls (Figure 14.3a and b). These results demonstrated that TEX EGFR-miR-21 and LG3BP-miR-210 pairs measured by Exo-PROS assay were able to distinguish lung cancer patients from normal controls. Receiver operating characteristic (ROC) analysis was performed to first determine the cutoff value for each and combined biomarkers, then the sensitivity, specificity,



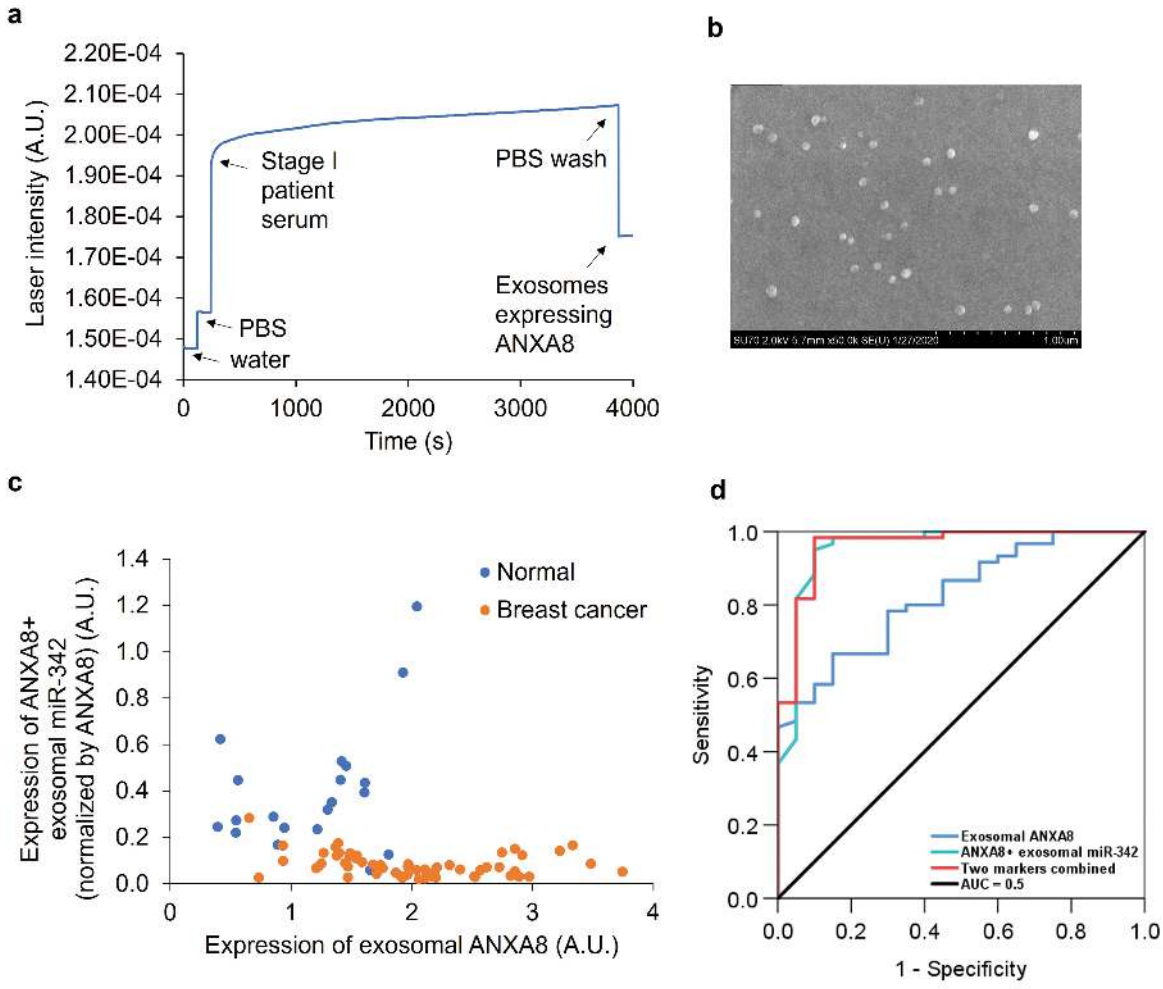
d

Biomarker	Normal vs Cancer		
	SEN	SPE	AUC
EGFR alone (cutoff: 3.11)	0.62	0.89	0.81
miR-21 alone (cutoff: 0.405)	0.73	0.94	0.89
LG3BP alone (cutoff: 1.67)	0.85	0.97	0.96
miR-210 alone (cutoff: 0.135)	0.95	0.83	0.94
All markers combined (cutoff: 0.608)	0.95	0.97	0.99
Best subset (cutoff: 0.609)	0.95	0.97	0.99

FIGURE 14.3 *Demonstration of clinical potential of Exo-PROS assay in lung cancer diagnosis using human serum samples.* Exo-PROS assay detected higher levels of TEX EGFR-miR-21 pair (a) and LG3BP-miR-210 pair (b) in sera from lung cancer patients (stage I–IV, $n = 60$, 52 NSCLC, and 8 SCLC) than normal controls ($n = 35$ including 15 normal controls at high risk of lung cancer). (c, d) ROC analysis was performed to determine the cutoff value, sensitivity (SEN), specificity (SPE), and AUC of each biomarker alone, all markers combined and the best subset biomarkers (i.e., TEX EGFR-miR-21 pair and LG3BP+ exosomal miR-210). (Reproduced from Hsu et al., 2023 [43] with permission from ACS Publications.)

and area under curve (AUC) of each biomarker and combined biomarkers in distinguishing normal controls from lung cancer patients were determined. As shown in Figure 14.3c and d, TEX EGFR (cutoff: 3.11), miR-21 (cutoff: 0.405), LG3BP (cutoff: 1.67), and miR-210 (cutoff: 0.135) distinguished lung cancer patients from normal controls with AUC of 0.81, 0.89, 0.96, and 0.94, respectively. When all four biomarkers were combined, the Exo-PROS assay successfully discriminated lung cancer patients from normal controls with sensitivity of 0.95, specificity of 0.97, and AUC of 0.99 at cutoff value of 0.608. Additionally, the best subset selection analysis showed that TEX EGFR-miR-21 pair and LG3BP+ exosomal miR-210 were the best biomarker combination that provided same diagnostic performance as all four biomarkers combined; that is, sensitivity of 0.95, specificity of 0.97, and AUC of 0.99 at the cutoff value of 0.609.

For breast cancer diagnosis, 20 normal controls and 60 treatment naïve patients with stages 0–IV breast cancer were evaluated in the study. For breast cancer patients, HR+/HER2−, HR−/HR2+, and triple negative breast cancer were included. To demonstrate that Exo-PROS assay was able to directly capture TEX from serum without the exosome isolation step, 10 μL serum were diluted in 40 μL PBS and directly applied on the biochip to measure the levels of exosomal ANXA8-miR-342 pair (Figure 14.4a). SEM image (Figure 14.4b) showed the exosomes captured on the surface of the biochip, confirming the direct capture of exosomes from serum without requiring the exosome isolation step prior to the measurement. As shown in Figure 14.4c, the levels of TEX ANXA8 and miR-342 were 1.69-fold higher and 5.00-fold lower in sera from breast cancer patients than those from normal controls. Results from ROC analysis (Figure 14.4d and e) showed that TEX ANXA8 (cutoff: 0.6) and miR-342 (cutoff: 0.165) distinguished breast cancer patients from normal controls with AUC of 0.83 and 0.96, respectively. When these two biomarkers combined, the sensitivity, specificity, and AUC were 0.98, 0.90, and 0.96 respectively at the cutoff value of 0.378. All results from both lung cancer and breast cancer demonstrated that Exo-PROS assay is an effective assay for cancer diagnosis.



e

Biomarker	ANXA8 alone (cutoff: 0.6)			miR-342 alone (cutoff: 0.165)			Two markers combined (cutoff: 0.378)		
	SEN	SPE	AUC	SEN	SPE	AUC	SEN	SPE	AUC
Normal vs Cancer	0.67	0.85	0.83	0.95	0.90	0.96	0.98	0.90	0.96

FIGURE 14.4 *Demonstration of clinical potential of Exo-PROS assay in breast cancer diagnosis using human serum samples.* (a) A representative SPR curve for the detection of TEX ANXA8 in the serum of a stage I breast cancer patient. 10 μL serum were diluted in 40 μL PBS and directly applied on the biochip for the measurement. (b) SEM confirmed the capture of ANXA8+ exosomes on the biochip. (c) The Exo-PROS assay detected significantly higher levels of exosomal ANXA8 and lower levels of ANXA8+ exosomal miR-342 in sera from breast cancer patients (stage 0–IV, $n = 60$) than normal controls ($n = 20$) ($p < 10^{-5}$). (d, e) ROC analysis was performed to determine the cutoff value, sensitivity (SEN), specificity (SPE), and AUC of each biomarker and two markers combined. (Reproduced from Hsu et al., 2023 [43] with permission from ACS Publications.)

14.5 Discussion

We developed a fast (~4 h) and user-friendly (only requires a micropipette) Exo-PROS assay to enable sensitive detection of TEX protein-microRNA pairs to further improve cancer diagnosis accuracy. Exo-PROS assay first captured TEXs from all exosomes using cancer overexpressed markers. With SPR sensing mechanism, it allowed one-stop, *in situ* characterization of TEX protein and microRNA biomarkers simultaneously with high sensitivity and specificity. The Exo-PROS assay has high analytical sensitivity with the LOD as low as 3.5×10^9 exosomes/mL, suggesting that this assay is highly sensitive and may require as low as 1 μL serum for each measurement [43]. Compared with conventional ELISA and qRT-PCR methods, the Exo-PROS assay was approximately 14-fold more sensitive (3.5×10^9 vs 5×10^{10} exosomes/mL) than ELISA, 10-fold faster than qRT-PCR (4 h vs 2 days), and required 10-fold less sample consumption for each modality(10 μL vs 100 μL) [43]. The superior sensitivity of the Exo-PROS assay compared to ELISA and qRT-PCR assays can be attributed to the physical characteristics of sensing target (i.e., exosomes) and detection mechanisms involved. Exosomes, with diameters ranging from 50 nm to 150 nm, may be more easily detected by SPR mechanism used in Exo-PROS assay than the colorimetric detection employed in ELISA. The qRT-PCR workflow

involves multiple steps: initially, immunomagnetic separation is used to isolate TEXs. RNA is then extracted from these isolated TEXs, and their microRNAs are detected using qRT-PCR. This multi-step process introduces significant material loss, which affects the sensitivity of qRT-PCR assay. In our study, even with 100 µL of serum (ten times the volume required for the Exo-PROS assay), the qRT-PCR assay yielded high cycle quantification (Cq) values (~35), approaching the detection limit, and thus showed inferior sensitivity to Exo-PROS assay. Although the LOD of Exo-PROS assay is not the lowest among all emerging biosensors, the major advantage of Exo-PROS assay is detecting two different types of TEX biomarkers (i.e., proteins and microRNAs) simultaneously to increase diagnostic accuracy. And the use of 10 µL serum per test is quite feasible in clinical settings. Exo-PROS assay also showed great analytical specificity and repeatability with < 10% intra- and inter-day coefficient of variation [43]. With a total of 175 patients and controls, we demonstrated that the Exo-PROS assay detected both lung cancer and breast cancer with high accuracy (AUC ≥ 0.96). The Exo-PROS assay also showed superior sensitivity, specificity, and AUC to ELISA and qRT-PCR methods in distinguishing normal controls from cancer patients [43].

14.6 Future Trends

The development of exosome-based diagnostic devices has the potential to make liquid biopsies a routine procedure in clinical settings, offering minimally or noninvasive, accurate, and cost-effective cancer screening, early detection, and treatment monitoring. To achieve this, we will investigate different combinations of TEX protein-microRNA pairs to enhance diagnostic accuracy. We will optimize the design of Exo-PROS biochip, for example, incorporating gold nanostructures to improve its analytical sensitivity and specificity. We will develop Exo-PROS biosensor into a high throughput and multiplex assay to meet the screening and diagnostic needs of large patient populations. Since the antibodies and molecular beacons can be easily changed to detect different TEX protein-microRNA pairs, we aim to transform Exo-PROS assay into a universal sensing platform for broader clinical applications, such as personalized treatment monitoring and prognosis assessment. We will validate the Exo-PROS assay in a larger patient cohort to establish its clinical utility in cancer care. Ultimately, obtaining FDA approval is crucial to ensure the successful transition of Exo-PROS assay into routine clinical practice. In short, the future of exosome-based cancer liquid biopsies lies in the synergy of technological innovation and improved diagnostic accuracy and their seamless integration into personalized cancer care.

REFERENCES

1. Carlsson, S.V., et al., Young age on starting prostate-specific antigen testing is associated with a greater reduction in prostate cancer mortality: 24-year follow-up of the Göteborg randomized population-based prostate cancer screening trial. *European Urology*, 2023. 83(2): p. 103–109.
2. Concepcion, J., et al., Trends of cancer screenings, diagnoses, and mortalities during the COVID-19 pandemic: implications and future recommendations. *The American Surgeon™*, 2023. 89(6): p. 2276–2283.
3. Moon, S., et al., Lower thyroid cancer mortality in patients detected by screening: a meta-analysis. *Endocrinology and Metabolism*, 2023. 38(1): p. 93–103.
4. De Jesus, C., et al., The benefits of screening mammography. *Current Breast Cancer Reports*, 2023. 15(2): p. 103–107.
5. Kok, V.C. and C.-C. Yu, Cancer-derived exosomes: their role in cancer biology and biomarker development. *International Journal of Nanomedicine*, 2020: p. 8019–8036.
6. Paskeh, M.D.A., et al., Emerging role of exosomes in cancer progression and tumor microenvironment remodeling. *Journal of Hematology & Oncology*, 2022. 15(1): p. 1–39.
7. Olejarz, W., et al., Exosomes in angiogenesis and anti-angiogenic therapy in cancers. *International Journal of Molecular Sciences*, 2020. 21(16): p. 5840.
8. Akoto, T. and S. Saini, Role of exosomes in prostate cancer metastasis. *International Journal of Molecular Sciences*, 2021. 22(7): p. 3528.
9. Armakolas, A., M. Kotsari, and J. Koskinas, Liquid biopsies, novel approaches and future directions. *Cancers*, 2023. 15(5): p. 1579.
10. Li, L., et al., Advanced technologies for molecular diagnosis of cancer: state of pre-clinical tumor-derived exosome liquid biopsies. *Materials Today Bio*, 2023. 18: p. 100538.
11. Logozzi, M., et al., High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients. *PLoS One*, 2009. 4(4): p. e5219.
12. Zhu, H., et al., PCR past, present and future. *Biotechniques*, 2020. 69(4): p. 317–325.
13. Liu, C., et al., Sensitive detection of exosomal proteins via a compact surface plasmon resonance biosensor for cancer diagnosis. *ACS Sensors*, 2018. 3(8): p. 1471–1479.
14. Sina, A.A.I., et al., Label-free detection of exosomes using a surface plasmon resonance biosensor. *Analytical and Bioanalytical Chemistry*, 2019. 411: p. 1311–1318.
15. Joshi, G.K., et al., Highly specific plasmonic biosensors for ultrasensitive microRNA detection in plasma from pancreatic cancer patients. *Nano Letters*, 2014. 14(12): p. 6955–6963.
16. Wu, W., et al., Surface plasmon resonance imaging-based biosensor for multiplex and ultrasensitive detection of NSCLC-associated exosomal miRNAs using DNA programmed heterostructure of Au-on-Ag. *Biosensors and Bioelectronics*, 2021. 175: p. 112835.
17. Zong, S., et al., Facile detection of tumor-derived exosomes using magnetic nanobeads and SERS nanoprobes. *Analytical Methods*, 2016. 8(25): p. 5001–5008.
18. Fan, C., et al., Ultrasensitive exosome detection by modularized SERS labeling for postoperative recurrence surveillance. *ACS Sensors*, 2021. 6(9): p. 3234–3241.
19. Lee, J.U., et al., Quantitative and specific detection of exosomal miRNAs for accurate diagnosis of breast cancer using a surface-enhanced Raman scattering sensor based on plasmonic head-flocked gold nanopillars. *Small*, 2019. 15(17): p. 1804968.
20. Pang, Y., et al., Dual-SERS biosensor for one-step detection of microRNAs in exosome and residual plasma of blood samples for diagnosing pancreatic cancer. *Biosensors and Bioelectronics*, 2019. 130: p. 204–213.

21. Im, H., et al., Label-free detection and molecular profiling of exosomes with a nano-plasmonic sensor. *Nature Biotechnology*, 2014. 32(5): p. 490–495.
22. Yoshioka, Y., et al., Ultra-sensitive liquid biopsy of circulating extracellular vesicles using ExoScreen. *Nature Communications*, 2014. 5(1): p. 3591.
23. Jorgensen, M., et al., Extracellular Vesicle (EV) Array: microarray capturing of exosomes and other 39 extracellular vesicles for multiplexed phenotyping. *J Extracell Vesicles*, (2013). 2: p. 40.
24. Jørgensen, M.M., R. Bæk, and K. Varming, Potentials and capabilities of the extracellular vesicle (EV) array. *Journal of Extracellular Vesicles*, 2015. 4(1): p. 26048.
25. Jeong, S., et al., Integrated magneto–electrochemical sensor for exosome analysis. *ACS Nano*, 2016. 10(2): p. 1802–1809.
26. Shao, H., et al., Protein typing of circulating microvesicles allows real-time monitoring of glioblastoma therapy. *Nature Medicine*, 2012. 18(12): p. 1835–1840.
27. Li, B., et al., Homogenous magneto-fluorescent nanosensor for tumor-derived exosome isolation and analysis. *ACS Sensors*, 2020. 5(7): p. 2052–2060.
28. Yu, Y., et al., Electrical and label-free quantification of exosomes with a reduced graphene oxide field effect transistor biosensor. *Analytical Chemistry*, 2019. 91(16): p. 10679–10686.
29. Kwong Hong Tsang, D., et al., Chemically functionalised graphene FET biosensor for the label-free sensing of exosomes. *Scientific Reports*, 2019. 9(1): p. 1–10.
30. Liu, C., et al., Non-invasive detection of exosomal MicroRNAs via tethered cationic lipoplex nanoparticles (tCLN) biochip for lung cancer early detection. *Frontiers in Genetics*, 2020. 11: p. 482368.
31. Wu, Y., et al., Detection of extracellular RNAs in cancer and viral infection via tethered cationic lipoplex nanoparticles containing molecular beacons. *Analytical Chemistry*, 2013. 85(23): p. 11265–11274.
32. Hu, J., et al., A signal-amplifiable biochip quantifies extracellular vesicle-associated RNAs for early cancer detection. *Nature Communications*, 2017. 8(1): p. 1683.
33. Zhang, J., et al., A ratiometric electrochemical biosensor for the exosomal microRNAs detection based on bipedal DNA walkers propelled by locked nucleic acid modified toehold mediate strand displacement reaction. *Biosensors and Bioelectronics*, 2018. 102: p. 33–40.
34. Boriachek, K., et al., An amplification-free electrochemical detection of exosomal miRNA-21 in serum samples. *Analyst*, 2018. 143(7): p. 1662–1669.
35. Luo, L., et al., A ratiometric electrochemical DNA biosensor for detection of exosomal microRNA. *Talanta*, 2020. 207: p. 120298.
36. Zhang, P., et al., Ultrasensitive detection of circulating exosomes with a 3D-nanopatterned microfluidic chip. *Nature Biomedical Engineering*, 2019. 3(6): p. 438–451.
37. Zhang, J., et al., Sensitive signal amplifying a diagnostic biochip based on a biomimetic periodic nanostructure for detecting cancer exosomes. *ACS Applied Materials & Interfaces*, 2020. 12(30): p. 33473–33482.
38. Ramshani, Z., et al., Extracellular vesicle microRNA quantification from plasma using an integrated microfluidic device. *Communications Biology*, 2019. 2(1): p. 189.
39. Cao, H., X. Zhou, and Y. Zeng, Microfluidic exponential rolling circle amplification for sensitive microRNA detection directly from biological samples. *Sensors and Actuators B: Chemical*, 2019. 279: p. 447–457.
40. Yang, Y., et al., Ultrafast detection of exosomal rnas via cationic lipoplex nanoparticles in a micromixer biochip for cancer diagnosis. *ACS Applied Nano Materials*, 2021. 4(3): p. 2806–2819.
41. Cho, S., H.C. Yang, and W.J. Rhee, *S*imultaneous multiplexed detection of exosomal microRNAs and surface proteins for prostate cancer diagnosis. *Biosensors and Bioelectronics*, 2019. 146: p. 111749.
42. Du, S., et al., Applying serum proteins and microRNA as novel biomarkers for early-stage cervical cancer detection. *Scientific Reports*, 2020. 10(1): p. 9033.
43. Hsu, C.-C., et al., Simultaneous detection of tumor derived exosomal protein–microRNA pairs with an exo-PROS biosensor for cancer diagnosis. *ACS Nano*, 2023. 17(9): p. 8108–8122.
44. Bethune, G., et al., Epidermal growth factor receptor (EGFR) in lung cancer: an overview and update. *Journal of Thoracic Disease*, 2010. 2(1): p. 48.
45. Capone, E., S. Iacobelli, and G. Sala, Role of galectin 3 binding protein in cancer progression: a potential novel therapeutic target. *Journal of Translational Medicine*, 2021. 19(1): p. 1–18.
46. Bica-Pop, C., et al., Overview upon miR-21 in lung cancer: focus on NSCLC. *Cellular and Molecular Life Sciences*, 2018. 75: p. 3539–3551.
47. Zhang, Y., Y. Hu, and F. Guo, miR-21 inhibits non-small cell lung cancer cell proliferation and promotes apoptosis. *Journal of Biomaterials and Tissue Engineering*, 2023. 13(1): p. 18–23.
48. Cao, G., et al., MiR-210 regulates lung adenocarcinoma by targeting HIF-1α. *Heliyon*, 2023. 9(5).
49. Bhalla, S., S. Yi, and D.E. Gerber, Emerging strategies in lung cancer screening: blood and beyond. *Clinical Chemistry*, 2024. 70(1): p. 60–67.
50. van den Broek, D. and H.J. Groen, Screening approaches for lung cancer by blood-based biomarkers: challenges and opportunities. *Tumor Biology*, 2023(Preprint): p. 1–16.
51. Rossetti, S., et al., Harnessing 3D models of mammary epithelial morphogenesis: an off the beaten path approach to identify candidate biomarkers of early stage breast cancer. *Cancer Letters*, 2016. 380(2): p. 375–383.
52. Stein, T., et al., Annexin A8 is up-regulated during mouse mammary gland involution and predicts poor survival in breast cancer. *Clinical Cancer Research*, 2005. 11(19): p. 6872–6879.
53. Young, J., et al., Tamoxifen sensitivity-related microRNA-342 is a useful biomarker for breast cancer survival. *Oncotarget*, 2017. 8(59): p. 99978.
54. Bitaraf, A., S. Babashah, and M. Garshasbi, Aberrant expression of a five-microRNA signature in breast carcinoma as a promising biomarker for diagnosis. *Journal of Clinical Laboratory Analysis*, 2020. 34(2): p. e23063.

Section 2

Imaging Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

15

Epigenetic Tools for Guiding Low Dose Computerized Tomography (LDCT) Screening Decision-Making

Winter Philibert and Robert Philibert

15.1 Introduction

15.1.1 Lung Cancer Screening Strategies

Lung cancer is the leading type of cancer death in the United States [1]. The overall five-year survival for patients diagnosed with lung cancer is only 15%, largely because nearly 60% of cancers are diagnosed at a late stage [2]. In contrast, the five-year survival for patients diagnosed with an early-stage lung cancer is around 56%, which strongly supports the need for early diagnosis and treatment.

Because smoking accounts for 90% of lung cancer mortality, screening efforts have traditionally targeted heavy smokers using imaging-based methods. Clinical trials have demonstrated that screening heavy smokers with low-dose CT (LDCT), though not chest radiographs or sputum cytology, can reduce lung cancer mortality [3, 4]. In 2011, the National Lung Screening Trial (NLST) consortium found that tree rounds of screening with LDCT scans reduced lung cancer mortality by 16% after nearly seven years of follow-up compared to screening with chest radiographs [3, 4]. Similarly, in 2018, the European Nederlands–Leuvens Longkanker Screenings Onderzoek (NELSON) trial investigators presented data at the World Conference on Lung Cancer showing that LDCT screening reduced lung cancer mortality by 26% for men and 61% for women over a ten-year follow up [5].

Based on the NLST, NELSON, and other study results, numerous professional organizations have endorsed LDCT screening [6–8]. Most notably, in 2013, the United States Preventive Services Task Force (USPSTF) recommended screening all of those who met criteria for inclusion into the NLST study. Because in the managed care environment screening decisions are often cost-sensitive, LDCT based methods took a quantum leap forward when, in 2015, the Centers for Medicare and Medicaid Services (CMS) issued a decision memo indicating lung cancer screening was a covered preventive service benefit for adults ages 55 to 77 meeting the NLST smoking criteria [9]. *Critically, all of the guidelines in this 2015 memo highlighted the importance of supporting shared decision-making (i.e., patient* and *doctor) for screening decisions.* As a consequence, CMS took the unprecedented step of mandating that a screening LDCT cannot be performed until a patient undergoes a counseling and shared decision-making visit that comprehensively addresses the potential benefits and harms of lung cancer screening.

15.1.2 Drawbacks of LDCT Screening

Why is there a need for counseling prior to engaging in lung cancer screening? Taken at face value, the decision for a smoker to choose potentially lifesaving LDCT screening should be an easy one. However, operationally, LDCT screening pathway is considerably more complex than one might assume. As a partial result, in the United States, screening rates are low with only 4%–7% of those eligible ever having received LDCT screening [10, 11]. The first step of the pathway is patient engagement. Potential patients need to first establish a therapeutic relationship with a clinical provider. Second, they have a dialogue about the possibility of LDCT examination. Third and finally, after LDCT examination, they have to get the LDCT exam conducted.

15.1.2.1 Exposure to Radiation

Fortunately, once the decision is made to engage in screening, there are more than adequate LDCT resources, and the coverage for indicated LDCT exams is universal. So, payment for screening itself is rarely an issue. Instead, a principal barrier to completion of the LDCT screening process is fear. Fear of potential medical and financial consequences of having lung cancer, and surprisingly, fear of suffering the consequences of LDCT screening itself.

Both sets of fears are not unreasonable. First, the fear of getting cancer is real. Between 10% and 20% of all smokers will eventually get lung cancer [12]. Second, LDCT screening itself is not without risks. Indeed, the reason that CMS mandates the counseling visit is that the screening decision is considered very complex, requiring patients to balance an absolute survival benefit of about three in 1,000 against a number of potential harms. For example, in the pivotal NLST study, the overall false positive rate for LDCT was 23%, and in three rounds of screening nearly 40% of the subjects had nodules detected on a scan [4]. However, 96% of the nodules detected in these abnormal scans were benign [13]. This high rate of suspicious nodules can have consequences because most abnormal scans are followed by further imaging studies. Given the current recommendations to screen eligible smokers between ages 50 and 80, the cumulative radiation exposure from the screening and diagnostic studies can be significant. Based on data from the Italian Continuous Observation of Smoking Subjects (COSMOS) cohort, one modeling study estimated one radiation-induced lung cancer death for every 2500 people who were screened over a period of 10 to 20 years [14].

DOI: 10.1201/9781003449942-18

15.1.2.2 Medical Complications from Biopsy

Excessive radiation exposure may not be the only problem associated with false-positive results. In the NLST, 2.7% of the subjects with a false-positive test result underwent invasive diagnostic testing. Trial sites, which were predominantly academic medical centers, performed these diagnostic procedures with a relative high degree of safety. Still nearly 10% of procedures had major complications [15]. However, this relatively low rate of complications has not translated to standard clinical practice. In 2019, Shih and colleagues examined a nationally representative sample of 344,510 patients and found that the risk of complications from invasive diagnostic procedures was twice as high in community practice, which has important implications for those considering the LDCT procedure in standard clinical settings [16].

15.1.2.3 Psychological Toll of Screening

A third problem for patients considering screening is anxiety, with the apprehension of those contemplating screening and potential for lung cancer being considerable. The suicide rate of lung cancer patients is four times that of the general population risk and is significantly higher than those with breast, colon, or prostate cancer [17]. Therefore, from the potential LDCT recipient's viewpoint, there is considerable reason to fear LDCT screening.

15.1.2.4 Update on the Costs of Screening

The concerns of patients are not the only barrier. The costs of screening are also a drawback to both policymakers and payers seeking to optimize their preventative service offerings. A cost effectiveness analysis conducted in 2014 by Black and colleagues concluded that lung cancer screening has an acceptable cost-effectiveness of $81,000 per quality-adjusted life year (QALY). [18]. But achieving these benefits as outlined by Black would entail Medicare expenditures of nearly $7 billion over five years. However, Medicare is not the only payer in the United States who would incur expense. In 2017, the Veterans Health Administration estimated that its first-year costs of a national lung cancer screening program could approach $900 million [19, 20].

However, these earlier estimates of costs were based on the 2013 screening guidelines. But in 2021, the USPSTF expanded their prior recommendations to include all of those between ages of 50 and 80 years old who have smoked within the past 15 years and now have at least a 20-pack-year history of smoking [21]. As a result, the potential pool of those eligible has nearly doubled. These changes, combined with the recent acceleration in healthcare inflation, strongly suggest that the current costs for conducting LDCT screening may greatly exceed prior estimates.

15.1.3 Using Algorithms to Select the Right Patients

Given the fear of potential patients and the potential high costs of LDCT screening, the "secret sauce" of any recipe for implementation of effective screening procedures in the future must include a more accurate identification of those most at risk for cancer. Prior to 2021, many clinicians used the Tammemagi $PLCOM_{2012}$ lung cancer prediction model. This algorithm was developed using data from 14,144 high-risk smokers (i.e., met NLST screening guidelines) from the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial [22]. Although a substantial improvement over prior methods, the Tammemagi model incorporated racial and other demographics features, which when incorporated into predictive algorithms can promote inequality and, from a purely analytic viewpoint, tend be unstable [23, 24]. As a result, the use of this algorithm is no longer recommended. Instead, in 2021, the USPSTF issued a new call for new biomarker-guided methods for improving prediction [21].

In the next several sections, we will detail how the use of epigenetics can potentially answer that summons for better methods.

15.2 Materials

In order to develop better methods of predicting lung cancer, scientists must have the requisite clinical materials. For those seeking to use DNA methylation to predict outcomes, it is necessary to have adequate supplies of DNA from the clinical intake and rigorous follow-up of the cohorts to ensure complete capture of clinical outcomes. Although almost all the studies of lung cancer collected imaging information, not all collected the DNA specimens needed for large-scale epigenetic studies. Perhaps two of the studies that collected the best combination of LDCT and DNA were the NLST and the Copenhagen City Heart Study.

15.2.1 The National Lung Screening Trial

The NLST enrolled 53,454 US smokers, adults ages 55 to 74 with a 30-pack-year smoking history who were either current smokers or had quit within the previous 15 years [3]. For the purpose of understanding the value of imaging, the subjects were equally divided into chest X-ray and LDCT arms at the 33 sites that participated in the study. In addition, the NLST constructed a DNA and serum repository for nearly 5000 subjects from both the X-ray only and LDCT arms of the study. These biological materials have been made available through the National Institutes of Health for studies by accredited investigators.

15.2.2 The Copenhagen City Heart Study

Using a population registry approach, the Copenhagen City Heart Study similarly has both the clinical and biological materials necessary for epigenetics. This Danish study includes two cohorts, 10,135 subjects collected between 1991 and 1994, and 6238 subjects recruited between 2001 and 2003 [25]. Because the study was not designed to test the ability of LDCT to detect lung cancer, well-powered analyses of the effectiveness of imaging cannot be conducted. However, because it is a population study, more in-depth analyses of lower smoking exposures and comparisons to non-smoking subjects can be more readily conducted.

15.3 Methods

15.3.1 Estimating Smoking Intensity Is Essential for Accurate Prediction

Why have prior attempts to improve prediction failed? Traditionally, most biomarker strategies attempting to improve prediction of lung cancer have focused on methods to objectively assess current and past cigarette consumption. This is exceedingly logical because most current risk models use self-reported tobacco use, even though these measures are known to be unreliable [26–31]. For example, many patients deny being current smokers despite biochemical testing results indicating otherwise, and even acknowledged smokers have difficulty accurately estimating pack-year exposures [27, 32–34]. At the same time, and in contrast to those situations in which social pressure or medical disapproval diminishes self-report [35, 36], there may be a tendency by some individuals to overreport in order to qualify for examinations such as LDCT [37].

However, even in the absence of denial or embellishment, a more important concern about using smoking history in cancer risk prediction models is that even an accurate history without recall bias might not be the best measure of actual tobacco smoke exposure. Pack-year consumption calculations depend on cigarettes per day (CPD) estimates by patients. However, CPD estimates by patients poorly correlate with nicotine metabolites in blood, serum, or saliva [38–40]. The amount inhaled (total puff volume) is a substantially better predictor of nicotine metabolite levels and total tobacco smoke exposure than CPD estimates [41]. Unfortunately, neither patients nor research subjects can readily ascertain their individual cigarette puff volumes. This limits the ability of self-report to accurately infer consumption histories.

15.3.2 Using DNA Methylation to Assess Smoking Intensity

One approach that has gathered general acceptance in the research community for assessing smoking history is the use of DNA methylation [42, 43]. Still, the pathway to this general acceptance was slow until 2012 when our group published the first genome-wide study of the epigenetics of smoking using the newly released Illumina 450K array [44]. This landmark study showed that the DNA methylation status in the aryl hydrocarbon receptor repressor (AHRR) gene, and in particular, at probe tagging CpG residue cg05575921, was a predictor of smoking status. Since that time, over 100 groups have replicated that finding (for review please see [45, 46]). Critical to the use of this technology to assess smoking behaviors are the findings that cg05575921 methylation status is a markedly better predictor of cigarette consumption than self-report measures [47, 48].

How does cg05575921 measure smoking consumption? Assessment of cg05575921 methylation differs from most prior approaches for assessing smoking status in that it does not rely on measuring nicotine consumption. Instead, methylation status at cg05575921 reflects the cumulative inhalation of polyaromatic hydrocarbons (PAH) and dioxins, both of which are well-established carcinogens.

PAH are the main offenders in tobacco smoke. These molecules consist of a large class of organic compounds that consist of two or more fused aromatic rings of carbon that can occur naturally in soil, air, and outer space, but are more frequently encountered by humans as a consequence of incomplete combustion of organic products such as gasoline and wood [49]. If ingested orally, PAH can be absorbed, metabolized by gut bacteria, or excreted in the stool with the effects on human health being dependent on the amount and type of PAH. Although it is difficult to be certain on the exact effects of low levels of oral PAH ingestion, association with altered gut biota, diarrhea, and colon cancer have been reported [50–52].

However, the primary threat to humans is when PAH are inhaled. Fortunately, in the United States and other well-regulated areas, the concentration of PAH in the air is generally less than 10 pg/m^3 [53]. However, in polluted areas of the world, PAH concentrations can rise to well over 1000 pg/m^3. Still, these levels of PAH are dwarfed by the concentration of PAH in mainstream tobacco smoke, which can exceed 1 ng/ml (i.e., 10^8 > normal air) with each cigarette delivering up to 1000ng of PAH to the lungs [54]. Therefore, to those who smoke, cigarette smoke is the largest source of not only PAH but dioxins as well.

15.3.3 The Biology of the Aryl Hydrocarbon Receptor Pathway

Fortunately for humans, most of the ingested PAH is detoxified through the aryl hydrocarbon receptor (AHR), or as it is often known, the xenobiotic pathway [55]. An overview of this pathway is illustrated in Figure 15.1. In brief, after being absorbed through the lung, dioxins and PAH molecules bind to AHR, which then binds to its cognate nuclear receptor partner, the aryl hydrocarbon nuclear translocator (ARNT). This AHR/ARNT heterodimer pair then translocates to the nucleus where it turns on cytochromes such as CYP1A1, CYP1A2, and CYP1B1 that are necessary for the catabolism of dioxin, PAH, and similarly structured molecules. Unfortunately, unrestrained AHR activity is also hazardous. Therefore, AHR/ARNT also activates the transcription the aryl hydrocarbon receptor repressor. The translated AHRR protein then serves as a decoy receptor and competes with AHR for binding ARNT while also sterically blocking xenobiotic response elements in AHR regulated genes [56]. This feedback modulation of AHR activation by AHRR fulfills a pivotal role in keeping the level of the AHR-regulated p450 enzymes at a non-harmful level and plays a large role in preventing increases in AHR activity that are associated with oncogenesis [57].

Interestingly, the cg05575921 site is not found in the promoter of AHRR. Instead, the cg05575921 CpG site is situated in a canonical enhancer element in intron 5 of AHRR. Demethylation of this site is associated with the recruitment of transcriptional activators and increased transcription of AHRR [44, 47]. Critically, for those interested in separating the effects of smoking from those of drinking on healthcare outcomes, alcohol has no discernible effect on the methylation

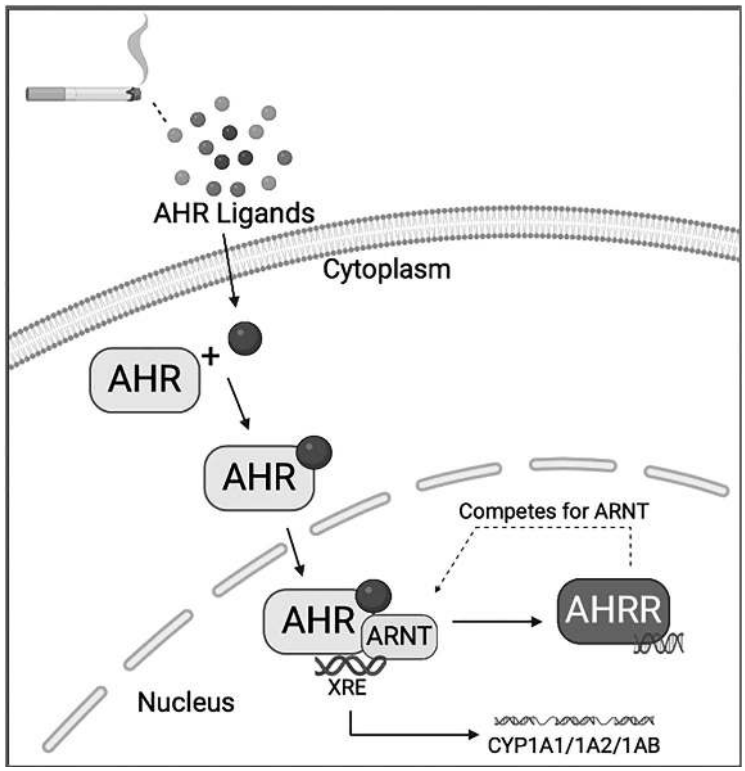


FIGURE 15.1 The aryl hydrocarbon receptor (AHR) Pathway. The AHR pathway in white blood cells and the liver is activated by the polyaromatic hydrocarbons (PAH) in tobacco smoke. These PAH molecules bind to AHR, which then dimerizes with the aryl hydrocarbon nuclear translocator (ARNT). The nuclear receptor (NR) heterodimer then translocates to the nucleus where it binds to NR motifs in the promoters of CYP1A1, CYP1A2, and CYP1B1 to activate transcription of these key catabolic cytochromes. The heterodimer pair also binds to an intron enhancer in AHRR containing the cg05575921 CpG site resulting in increased transcription of AHRR. The translated protein then serves as "decoy receptor" preventing unrestrained activation of the xenobiotic pathway. This activation of AHRR transcription is associated with demethylation of the cg05575921 site.

of AHRR [44, 58]. Furthermore, in part because nicotine itself is metabolized by a separate pathway (CYP2A6), the use of nicotine patches, vaping products, and non-combustible forms of tobacco also have no effect on cg05575921 methylation [59–62].

15.4 Clinical Examinations of Cg05575921 Methylation

First discovered in 2012, the relationship of cg05575921 methylation to smoking status has now been examined in hundreds of basic science and clinical studies [42, 44, 46]. The findings are sufficiently robust that it is regarded as generally accepted science [42, 43]. Still, the pathway to clinical translation for basic science findings is not always straightforward. With respect to the use of cg05575921 methylation for guiding LDCT decision-making, the following set of findings are particularly critical.

15.4.1 Cg05575921 Methylation Predicts Smoking Intensity

Perhaps one of the most critical properties of the cg05575921 locus is that the demethylation response to PAH exposure is dose dependent. Just as a hemoglobin A1c value indicates the steady state serum glucose level, the amount of demethylation at cg05575921 reflects the current intensity of smoking. Figure 15.2 illustrates the relationship between self-reported cigarette consumption in 421 biochemical confirmed current daily smokers and 423 non-smokers reported previously [63]. The methylation value for cg05575921 methylation in white blood cell DNA in adults who do not smoke is 86.6% ± 2.9 [63]. However, in those who do smoke, there is a dose-dependent, nonlinear decrease in methylation. Furthermore, just as a high HbA1c value responds to successful treatment by regressing toward the normal range, this PAH-induced demethylation reverts in response to sustained smoking cessation, making this a potentially excellent biomarker for monitoring the success of smoking cessation therapy [42].

15.4.2 Cg05575921 Methylation in the Copenhagen City Heart Study

The wealth of clinical and basic science clearly establishes the ability of cg05575921 to assess smoking intensity. However, the pivotal question for lung cancer prevention specialists is whether these objective assessments of cg05575921 methylation can also improve the prediction of risk for lung cancer. Fortunately, the answer appears to be "yes!"

The first firm demonstration of this possibility came from the work of the Copenhagen City Heart Study in 2017. The Bojesen-led group used a qPCR-based method to assess DNA

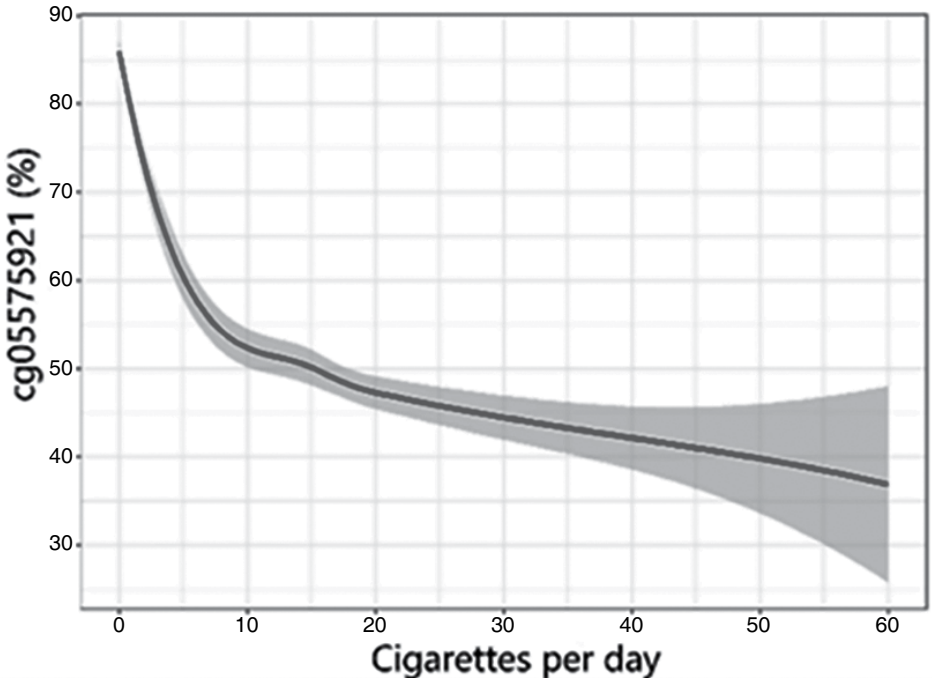


FIGURE 15.2 The relationship of digital methylation sensitive PCR assessments of cg05575921 to daily cigarette consumption ($n = 421$ smokers and 423 non-smoking controls). Shaded area represents the 95% confidence intervals for the estimate.

methylation at cg05575921 in DNA from 2576 high-risk smokers [64]. In this cohort, they found that cg05575921 methylation but not $PLCOM_{2012}$ formula strongly predicted six-year risk for lung cancer.

The next step forward then took place when in 2022, the Copenhagen City group next conducted an extended series of analyses of the ability of a base model that incorporated age, sex, and cumulative consumption together with cg05575921 methylation to predict five-year likelihood of lung cancer in groups of high-risk smokers as defined by the inclusion criteria for screening in six well-known longitudinal studies of the phenomenology of lung cancer [25]. For each and every definition of high-risk smoker used in these six studies, the addition of cg05575921 methylation to the base model delineated per the aforementioned significantly increased the prediction of lung cancer. In particular, for those 1979 subjects (including 59 lung cases) eligible by NLST criteria, the cg05575921 enhanced model led to a 22% decreased screening burden with a specificity of 84%.

15.4.3 Cg05575921 Methylation in the National Lung Cancer Screening Trial

Several months later in 2022, using an improved methylation sensitive digital PCR assay, our group reported the results of our examination the relationship of cg05575921 methylation to lung cancer in a cohort of subjects from the NLST study itself. Table 15.1 describes the 3081 NLST subjects whose data were included in the study. All subjects were between the ages of 55 and 74 years old, with a mean age of approximately 61 years for both males and females. Approximately half of the subjects undergoing screening were still smoking with the average smoker consuming nearly 1-1/2 packs of cigarettes per day. A total of 135 subjects had lung cancer during either screening or follow-up with 920 of the subjects (30%) having nodules noted during the initial three-year screening window. In this cohort, a model that included

TABLE 15.1

Clinical and Demographic Characteristics of the NLST Subjects Included in the 2022 Study.

Variable	Male	Female
N	1820	1261
Age	61.8 ± 5.0	61.3 ± 4.8
Race		
White	1691	1190
African American	73	46
Asian	22	9
Other	34	16
Current Smokers		
N	840	654
Cg05575921 (%)	46.1 ± 14.1	49.3 ± 13.8
Former Smokers		
N	980	607
Years quit	7.5 ± 5.1	6.9 ± 4.7
Cg05575921 (%)	65.6 ± 12.6	69.2 ± 11.1
Current Cigs per day	29.1 ± 11.3	26.5 ± 9.8
Pack year consumption	58.5 ± 24.5	51.6 ± 19.8
Screening status		
T0-T2 cancer	57	29
Post screen cancer	26	23
Nodules, no cancer	533	397
No nodules, no cancer	1204	812

just age, pack-year history, and cg05575921 methylation had an area under the curve of 0.66 for predicting the eight-year likelihood of lung cancer.

Although the results of such a screening could be useful for all, our results indicate that these assessments will be particularly used for those at the extremes of the pack-year consumption and cg05575921 distribution. Figure 15.3 illustrates the eight-year risk for smokers at the 10th and 90th percentile of the distribution in the NLST cohort. As the figure demonstrates, the greatest impact on differential risk will be among those at the extremes of the distribution. For example, a smoker with 88 pack-years of smoking history with 35% methylation at cg05575921 stands a 9% chance of developing lung cancer over the next eight years. In contrast, a smoker with 34 pack-years of smoking history with 78% methylation only stands a 2.8% chance of developing cancer. Notably, because these higher methylation values represent those individuals in the cohort who have successfully quit smoking, these data further illustrate the benefits of smoking cessation even for those patients with 30 or more years of smoking history.

15.5 Discussion

In summary, the decision to engage in LDCT screening is a complex decision with decided risks and benefits. The USPSTF recommends the consideration of screening for those over the age of 50 who have smoked in the past 15 years who have at least 20 pack-years of cigarette consumption. At the same time, the USPSTF recognizes that only a fraction of those meeting the criteria will benefit from screening and has called for better methods for assessing need for LDCT. In response to that call, the current data presented herein shows that two groups have shown that the assessment of DNA methylation at cg05575921 can help identify those most likely to benefit from LDCT

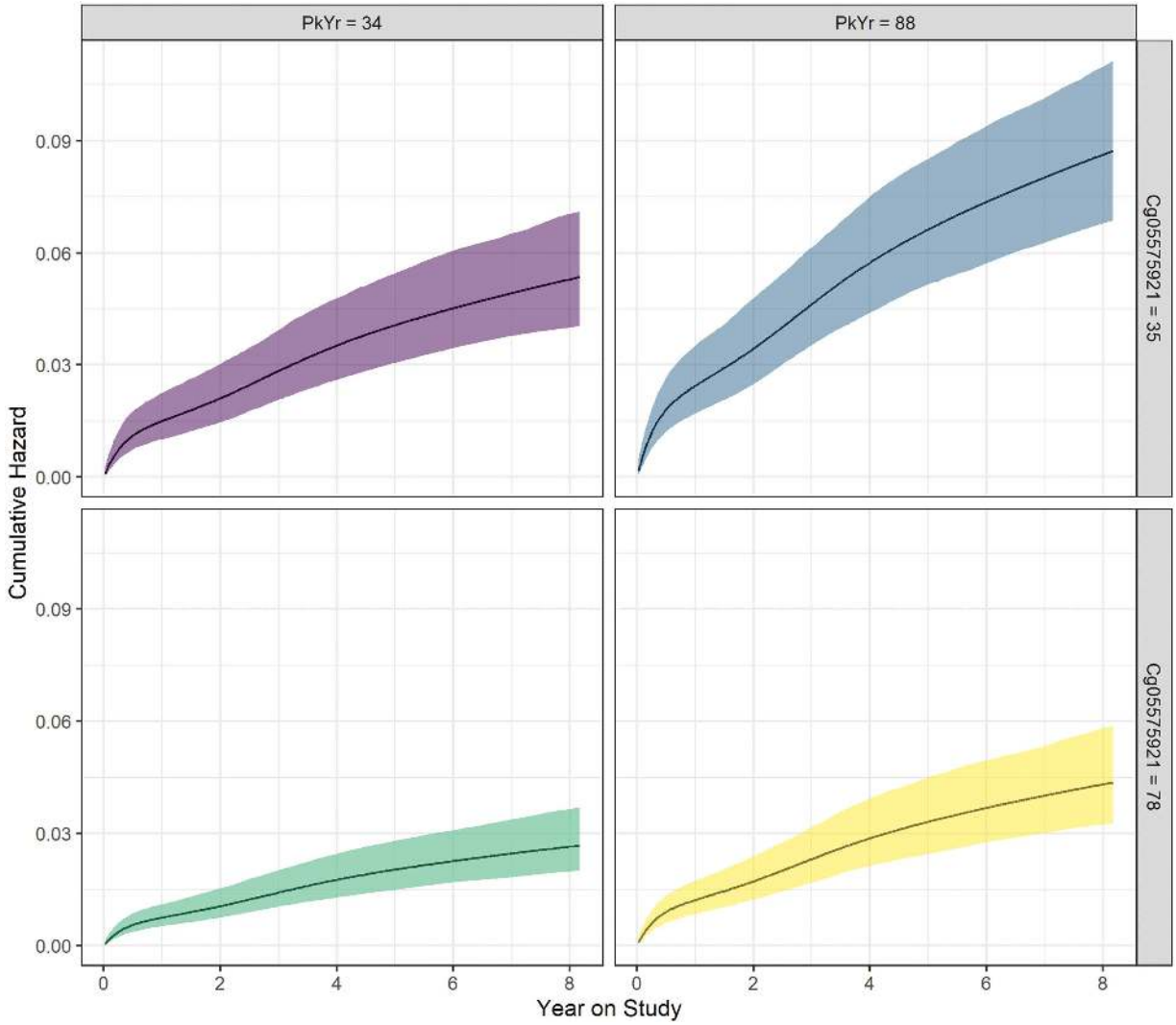


FIGURE 15.3 Cumulative hazard (95% confidence ribbons) as pack years consumption (PY) and cg05575921 methylation (%). Values for the latter two are set at the 10th and 90th percentiles of the population distribution. Age is set to the mean of the sample (62 years old). Curves are estimated based on the best fitting survival model.

screening. Future trials to test the real-world implementation of these screening methods and explore additional uses of DNA methylation in the LDCT pathway are indicated.

Still, there are two limitations of these NLST findings for clinical practice. First, the 2021 guidelines expanded both the age and pack-year range for LDCT screening beyond the inclusion criteria for enrollment in the Copenhagen City and NLST cohort studies. Fortunately, in our analyses, age was not associated with risk of lung cancer independently of cg05575921 methylation and pack-year history. Furthermore, the strongly defined curves for the relationship of pack-year and cg05575921 to lung cancer risk likely will allow imputation of risk to the lower boundary of cigarette consumption (20 pack-year) delineated in the 2021 guidelines. However, a second additional limitation is the lack of diversity in the NLST subject population. Fortunately, thanks to generous funding from the National Cancer Institute (R44CA285136), a project to augment this formula with data from an additional 6000 subjects (4500 controls, 1500 cases) including nearly 700 non-White ancestry and an expanded range of smoking history, is underway and is expected to be completed by the middle of 2026.

15.6 Future Trends

The aforementioned findings from the Copenhagen City and NLST cohorts strongly support the supposition that DNA methylation can be used to identify those most likely to benefit from screening. When these data are combined with the National Cancer Institute's funding of the characterization of additional subjects, it is likely that a robust method for assisting patient and clinicians in their decision-making will emerge. Even so, *are there additional uses for cg05575921 assessments in the LDCT screening process*?

15.6.1 Biomarker Guided Smoking Cessation

We believe that the answer is "yes!" For example, for many patients, the LDCT screening appointment is a "teachable moment" that can serve as an opportunity for clinicians and patients to gain traction in smoking cessation efforts [65]. We believe that precision epigenetic assessments of cg05575921 could also be used as part of a contingency management approach to help patients quit smoking [42]. However, there may be additional pathways through which epigenetic information can improve the care and treatment of those who smoke.

15.6.2 Improving Nodule Selection for Biopsy

Improving the selection of nodules for biopsy may be a second avenue for increasing the value of epigenetic assessments. Even under the best of circumstances, the preoperative assessment of nodules detected through LDCT is plagued with false positives. Approximately 30% of all LDCT images have at least one pulmonary nodule, and in the NLST study, 96% of these nodules selected for biopsy were found to be benign [66, 67]. Although in an ideal world clinicians tend to err on the side of caution by choosing to biopsy, and in a perfect world would elect to biopsy all nodules, the decision to biopsy lung nodules is not a decision without financial and human cost. The average cost for bronchoscopy with transbronchial lung biopsy in 2023 in the United States was $12,139 [68]. But that is only a fraction of the true cost of biopsies. As noted earlier, in Shih and colleagues' 2019 examination of a nationally representative sample of 344,510 patients who underwent LDCT screening, the median costs of minor, intermediate, and major complications was $5573, $19,470, and $57,893, respectively. Many patients had long intensive care unit stays with occasional deaths being noted [69]. Because inflation has preferentially affected healthcare costs, the current costs of these complications could actually be double those reported by Shih. Taken together with other evidence, this suggests that the development of techniques to decrease the rate of false positive nodule biopsies could cut costs and improve patient outcomes.

Currently, clinicians use nodule characteristics such as size and opacity to evaluate nodules greater than 8 mm in size and solid opacity as predictive of being cancerous [66]. Clinicians can also choose to use other imaging modalities such as positron emission tomography (PET) and magnetic resonance imaging (MRI) to further evaluate the tumor and search for evidence of metastasis.

15.6.3 Potential for Integration with Genetic Information

Could epigenetic features, together or integrated with genetic assessments, additionally predict the likelihood of malignancy? If so, adding epigenetic information into more complex algorithms for these purposes is not new ground. In 2014, Exact Science introduced the use of DNA methylation, together with assessment of genotype and stool hemoglobin levels, in an algorithm for predicting likelihood of colon cancer [70]. Since that time, dozens of methylation-based metrics to predict cancer have been introduced [71]. Could methylation be used in combination with other variables to predict the likelihood of nodule malignancy?

Conceivably, incorporation of methylation information into algorithm for selecting nodules for biopsy could be accomplished through two ways. The first method is simple. Since DNA methylation predicts the likelihood of lung cancer, DNA methylation at cg05575921 could conceivably be incorporated in a Bayesian approach that incorporates existing morphologic methods [72] to predict the likelihood of cancer. The second method that relies on machine learning (ML) methods is even potentially more powerful. Already deep learning techniques to interpret LDCT images have been developed [73]. Typically, these methods only use features based on imaging data to predict malignancy. Whereas these methods can perform well, it is conceivable that they could be further improved by the use of DNA methylation or genetic variation information. For example, it may well be that DNA methylation status may be subtly related to the position, shape, or size of lung nodules. If so, the incorporation of methylation data into ML algorithms could have additional impact on the LDCT screening pathway.

15.7 Conflict of Interest

Dr. Philibert is the Chief Executive Officer of Behavioral Diagnostics (Coralville, IA, 52242) and inventor on the use of cg05575921 status to determine smoking status, which is protected by US Patents 8,637,652 and 9,273,358.

Acknowledgments

The authors would like to thank Dr. Jeffrey Long for preparation of Figure 15.2. This effort was supported by a grant from the National Institutes of Health (R44CA285136).

REFERENCES

1. Center for Disease Control, Annual smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 1997–2001. *Morbidity and Mortality Weekly*, 2005. 54(25): p. 625–628.
2. Husmann, L. and P. Stolzmann, Staging, restaging and response evaluation of non-small-cell lung cancer, in *Diseases of the Chest and Heart 2015–2018: Diagnostic Imaging and Interventional Techniques*, J. Hodler, et al., Editors. 2015, Springer Milan: Milano. p. 183–188.
3. The National Lung Screening Trial Research Team, Reduced lung-cancer mortality with low-dose computed tomographic screening. *New England Journal of Medicine*, 2011. 365(5): p. 395–409.
4. The National Lung Screening Trial Research Team, Results of initial low-dose computed tomographic screening for lung cancer. *New England Journal of Medicine*, 2013. 368(21): p. 1980–1991.
5. De Koning, H.J., NELSON study: CT screening for early lung cancer reduces lung cancer mortality, in *World Congress on Lung Cancer*.2018, The ASCO Post: Toronto.
6. Moyer, V.A., Screening for lung cancer: US Preventive Services Task Force recommendation statement. *Annals of Internal Medicine*, 2014. 160(5): p. 330–338.
7. Wender, R., et al., American Cancer Society lung cancer screening guidelines. *CA: A Cancer Journal for Clinicians*, 2013. 63(2): p. 106–117.
8. Jaklitsch, M.T., et al., The American Association for Thoracic Surgery guidelines for lung cancer screening using low-dose computed tomography scans for lung cancer survivors and other high-risk groups. *The Journal of Thoracic and Cardiovascular Surgery*, 2012. 144(1): p. 33–38.
9. Centers for Medicare and Medicaid Services, *Decision Memo for Screening for Lung Cancer with Low Dose Computed Tomography (LDCT) (CAG-00439N).* 2015: Baltimore, M.D.
10. Li, J., et al., New recommendation and coverage of low-dose computed tomography for lung cancer screening: uptake has increased but is still low. *BMC Health Services Research*, 2018. 18: p. 1–10.
11. Jemal, A. and S.A. Fedewa, Lung cancer screening with low-dose computed tomography in the United States—2010 to 2015. *JAMA Oncology*, 2017. 3(9): p. 1278–1281.
12. Villeneuve, P.J. and Y. Mao, Lifetime probability of developing lung cancer, by smoking status, Canada. *Canadian Journal of Public Health*, 1994. 85(6): p. 385–8.
13. National Lung Screening Trial Research Team, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, Fagerstrom RM, Gareen IF, Gatsonis C, Marcus PM, Sicks JD. Reduced lung-cancer mortality with low-dose computed tomographic screening. *The New England Journal of Medicine*, 2011. 365(5): p. 395–409.
14. Rampinelli, C., et al., Exposure to low dose computed tomography for lung cancer screening and risk of cancer: secondary analysis of trial data and risk-benefit analysis. *BMJ*, 2017. 356.
15. Pinsky, P.F., et al., National Lung Screening Trial findings by age: medicare-eligible versus under-65 population. *Annals of Internal Medicine*, 2014. 161(9): p. 627–633.
16. Huo, J., et al., Complication rates and downstream medical costs associated with invasive diagnostic procedures for lung abnormalities in the community setting. *JAMA Interanal Medicine*, 2019. 179(3): p. 324–332. doi: 10.1001/jamainternmed.2018.6277. PMID: 30640382; PMCID: PMC6440230.
17. Rahouma, M., et al., Lung cancer patients have the highest malignancy-associated suicide rate in USA: a population-based analysis. *Ecancermedicalscience*, 2018. 12.
18. Black, W.C., et al., Cost-effectiveness of CT screening in the national lung screening trial. *New England Journal of Medicine*, 2014. 371(19): p. 1793–1802.
19. Kinsinger, L.S., et al., Implementation of lung cancer screening in the Veterans Health Administration. *JAMA Internal Medicine*, 2017. 177(3): p. 399–406.
20. Roth, J.A., et al., Projected clinical, resource use, and fiscal impacts of implementing low-dose computed tomography lung cancer screening in medicare. *Journal of Oncology Practice*, 2015. 11(4): p. 267–272.
21. Krist, A.H., et al., Screening for lung cancer: US preventive services task force recommendation statement. *JAMA*, 2021. 325(10): p. 962–970.
22. Gohagan, J.K., et al., The prostate, lung, colorectal and ovarian (PLCO) cancer screening trial of the National Cancer Institute: history, organization, and status. *Controlled Clinical Trials*, 2000. 21(6): p. 251S–272S.
23. Tammemägi, M.C., et al., Selection criteria for lung-cancer screening. *New England Journal of Medicine*, 2013. 368(8): p. 728–736.
24. Paulus, J.K. and D.M. Kent, Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digital Medicine*, 2020. 3(1): p. 99.
25. Jacobsen, K.K., et al., AHRR (cg5575921) methylation safely improves specificity of lung cancer screening eligibility criteria: a Cohort study. *Cancer Epidemiology and Prevention Biomarkers*, 2022.
26. Kandel, D.B., et al., Salivary cotinine concentration versus self-reported cigarette smoking: three patterns of inconsistency in adolescence. *Nicotine & Tobacco Research*, 2006. 8(4): p. 525–537.
27. Gorber, S.C., et al., The accuracy of self-reported smoking: a systematic review of the relationship between self-reported and cotinine-assessed smoking status. *Nicotine & Tobacco Research*, 2009. 11(1): p. 12–24.

28. Vasankari, T., et al., Serum cotinine predicts bronchial obstruction regardless of self-reported smoking history. *Scandinavian Journal of Public Health*, 2011. 39(5): p. 547–552.
29. Gerritsen, M., et al., Self-reporting of smoking cessation in cardiac patients: how reliable is it and is reliability associated with patient characteristics? *Journal of Addiction Medicine*, 2015. 9(4): p. 308–316.
30. Andersen, A.M., et al., Accuracy and utility of an epigenetic biomarker for smoking in populations with varying rates of false self-report. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 2017. 174(6): p. 641–650.
31. Goniewicz, M.L. and D.M. Smith, Are some E-cigarette users "blowing smoke"?: assessing the accuracy of self-reported smoking abstinence in exclusive E-cigarette users. *Nicotine & Tobacco Research*, 2018. 1: p. 2.
32. Caraballo, R.S., et al., Factors associated with discrepancies between self-reports on cigarette smoking and measured serum cotinine levels among persons aged 17 years or older: Third National Health and Nutrition Examination Survey, 1988–1994. *American Journal of Epidemiology*, 2001. 153(8): p. 807–814.
33. Caraballo, R.S., G.A. Giovino, and T.F. Pechacek, Self-reported cigarette smoking vs. serum cotinine among U.S. adolescents. *Nicotine & Tobacco Research*, 2004. 6(1): p. 19–25.
34. Caraballo, R.S., S. Sharapova, and K.J. Asman, Differences in current cigarette smoking between non-hispanic whites and non-hispanic blacks by gender and age group, United States, 2001–2013. *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco*, 2016. 18(Suppl 1): p. S41.
35. Murray, R.P., et al., Error in smoking measures: effects of intervention on relations of cotinine and carbon monoxide to self-reported smoking. The Lung Health Study Research Group. *American Journal of Public Health*, 1993. 83(9): p. 1251–1257.
36. Boyd, N.R., et al., Quality of measurement of smoking status by self-report and saliva cotinine among pregnant women. *Maternal and Child Health Journal*, 1998. 2(2): p. 77–83.
37. Studts, J.L., et al., Validity of self-reported smoking status among participants in a lung cancer screening trial. *Cancer Epidemiology Biomarkers & Prevention*, 2006. 15(10): p. 1825–1828.
38. Wagenknecht, L.E., et al., Racial differences in serum cotinine levels among smokers in the coronary artery risk development in (young) adults study. *American Journal of Public Health*, 1990. 80(9): p. 1053–1056.
39. Pickett, K.E., et al., Self-reported smoking, cotinine levels, and patterns of smoking in pregnancy. *Paediatric and Perinatal Epidemiology*, 2005. 19(5): p. 368–376.
40. Mustonen, T.K., et al., The influence of gender, race, and menthol content on tobacco exposure measures. *Nicotine & Tobacco Research*, 2005. 7(4): p. 581–590.
41. Krebs, N.M., et al., Comparison of puff volume with cigarettes per day in predicting nicotine uptake among daily smokers. *American Journal of Epidemiology*, 2016. 184(1): p. 48–57.
42. Fang, F., et al., Epigenetic biomarkers for smoking cessation. *Addiction Neuroscience*, 2023. 6: p. 100079.
43. Gelernter, J. and R. Polimanti, Genetics of substance use disorders in the era of big data. *Nature Reviews Genetics*, 2021. 22(11): p. 712–729. doi: 10.1038/s41576-021-00377-1. Epub 2021 Jul 1. PMID: 34211176; PMCID: PMC9210391.
44. Monick, M.M., et al., Coordinated changes in AHRR methylation in lymphoblasts and pulmonary macrophages from smokers. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, 2012. 159B(2): p. 141–151.
45. Andersen, A.M., et al., Current and future prospects for epigenetic biomarkers of substance use disorders. *Genes*, 2015. 6(4): p. 991–1022.
46. Gao, X., et al., DNA methylation changes of whole blood cells in response to active smoking exposure in adults: a systematic review of DNA methylation studies. *Clinical Epigenetics*, 2015. 7(1): p. 113.
47. Zeilinger, S., et al., Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*, 2013. 8(5): p. e63812.
48. Zhang, Y., et al., Smoking-associated DNA methylation markers predict lung cancer incidence. *Clinical Epigenetics*, 2016. 8: p. 127.
49. Patel, A.B., et al., Polycyclic aromatic hydrocarbons: sources, toxicity, and remediation approaches. *Frontiers in Microbiology*, 2020. 11.
50. Zhang, W., et al., Preliminary evidence for an influence of exposure to polycyclic aromatic hydrocarbons on the composition of the gut microbiota and neurodevelopment in three-year-old healthy children. *BMC Pediatrics*, 2021. 21(1): p. 86.
51. Abdel-Shafy, H.I. and M.S. Mansour, A review on polycyclic aromatic hydrocarbons: source, environmental impact, effect on human health and remediation. *Egyptian Journal of Petroleum*, 2016. 25(1): p. 107–123.
52. Cheng, T., et al., Polycyclic aromatic hydrocarbons detected in processed meats cause genetic changes in colorectal cancers. *International Journal of Molecular Sciences*, 2021. 22(20): p. 10959.
53. Zhang, X., et al., Characteristics of polycyclic aromatic hydrocarbons (PAHs) and common air pollutants at Wajima, a remote background site in Japan. *International Journal of Environmental Research and Public Health*, 2020. 17(3).
54. Adesina, O.A., T.I. Olowolafe, and A. Igbafe, Levels of polycyclic aromatic hydrocarbon from mainstream smoke of tobacco products and its risks assessment. *Journal of Hazardous Materials Advances*, 2022. 5: p. 100053.
55. Larigot, L., et al., AhR signaling pathways and regulatory functions. *Biochimie Open*, 2018. 7: p. 1–9.
56. Haarmann-Stemmann, T., et al., Analysis of the transcriptional regulation and molecular function of the aryl hydrocarbon receptor repressor in human cell lines. *Drug Metabolism and Disposition*, 2007. p. 2262–2269.
57. Opitz, C.A., et al., An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature*, 2011. 478(7368): p. 197–203.
58. Philibert, R., et al., The impact of recent alcohol use on genome wide methylation signatures. *Frontiers in Epigenetics*, 2012. 3: p. 54. doi: 10.3389/fgene.2012.00054. PMID: 22514556; PMCID: PMC3322340.

59. Besingi, W. and Å. Johansson, Smoke related DNA methylation changes in the etiology of human disease. *Human Molecular Genetics,* 2014. 23(9): p. 2290–2297. doi: 10.1093/hmg/ddt621. Epub 2013 Dec 11. PMID: 24334605.
60. Dawes, K., et al., Epigenetic analyses of alcohol consumption in combustible and non-combustible nicotine product users. *Epigenomes*, 2021. 5(3): p. 18.
61. Andersen, A., et al., DNA methylation differentiates smoking from vaping and non-combustible tobacco use. *Epigenetics*, 2021: p. 1–13.
62. Perez-Stable, E.J., et al., Nicotine metabolism and intake in black and white smokers. *JAMA*, 1998. 280(2): p. 152–156.
63. Dawes, K., et al., The relationship of smoking to cg05575921 methylation in blood and saliva DNA samples from several studies. *Scientific Reports*, 2021. 11(1): p. 21627.
64. Bojesen, S.E., et al., AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax*, 2017. 72(7): p. 646–653.
65. Cadham, C.J., et al., Smoking cessation interventions for potential use in the lung cancer screening setting: a systematic review and meta-analysis. *Lung Cancer*, 2019. 135: p. 205–216.
66. Mazzone, P.J. and L. Lam, Evaluating the patient with a pulmonary nodule: a review. *JAMA*, 2022. 327(3): p. 264–273.
67. Tammemägi, M.C., et al., Selection criteria for lung-cancer screening. *New England Journal of Medicine*, 2013. 368(8): p. 728–736.
68. MDsave.com, *Bronchoscopy with Transbronchial Lung Biopsy.* 2024. https://www.mdsave.com/procedures/bronchoscopy-with-transbronchial-lung-biopsy/d284f9cb
69. Huo, J., et al., Complication rates and downstream medical costs associated with invasive diagnostic procedures for lung abnormalities in the community setting complications and medical costs associated with diagnostic procedures for lung abnormalities complications and medical costs associated with diagnostic procedures for lung abnormalities. *JAMA Internal Medicine*, 2019. 179(3): p. 324–332.
70. Kisiel, J.B. and P.J. Limburg, Colorectal cancer screening with the multitarget stool DNA test. *Official Journal of the American College of Gastroenterology | ACG*, 2020. 115(11): p. 1737–1740.
71. Taryma-Leśniak, O., K.E. Sokolowska, and T.K. Wojdacz, Current status of development of methylation biomarkers for in vitro diagnostic IVD applications. *Clinical Epigenetics*, 2020. 12(1): p. 1–16.
72. Gould, M.K., et al., Evaluation of individuals with pulmonary nodules: when is it lung cancer? Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest*, 2013. 143(5 Suppl): p. e93S–e120S.
73. Thanoon, M.A., et al., A review of deep learning techniques for lung cancer screening and diagnosis based on CT images. *Diagnostics*, 2023. 13(16): p. 2617.

16

Multi-View Models for Colorectal Polyp Detection in CT Colonography

Mohamed Yousuf, Samir Harb, Asem Ali, Salwa Elshazly, and Aly Farag

16.1 Introduction

This chapter introduces a framework for automatic polyp detection in CT colonography (CTC). The framework consists of two stages: the first stage consists of two CNN-based polyps detectors. The CNN-based model is a programmable algorithm based on convolutional neural networks (CNN) that is trained to detect polyps within an image. The first detector is trained and validated to detect polyps in the axial CT slices. To narrow down the effective receptive field of the detector neurons, the colon regions are segmented and then fed into the network instead of the original CT slice. This drastically improves the detection and localization results, as reflected in increasing the mean average precision (mAP) by 36%. The second detector is a CNN-based model which is trained and validated to detect polyps in images, which are generated using the Fly-In visualization approach. To further enhance the polyp detection rate, we project 3D geometric feature maps, for example depth and curvature maps, on these images. The trained CNN-based model has provided polyp detection rate with accuracy reaching mAP ~97.1%. To reduce the false positives generated by the detectors, in the second stage, a multi-view network (MVN) is deployed and is trained using the sagittal and coronal views corresponding to the detected axial views, to reject false positives. Experimental results, using 50 CTC-annotated cases, showed that sessile-type polyps can be detected with a specificity measured by an area under the curve (AUC) of ~95.27%.

Early diagnosis of colorectal polyps, before they turn into cancer, is one of the main keys for treatments. Therefore, many research studies were conducted to investigate screening procedures to help in identifying polyp candidates. Radiologists distinguish true polyps in the colon from other structures on CT scans by analyzing their shape, size, surface characteristics, location, and density. True polyps typically have a round or oval shape, a smooth surface, and a homogeneous density. They can be pedunculated or sessile, as shown in Figure 16.1. The consistent appearance of polyps across multiple views also helps in identification. Advanced imaging techniques and computer-aided detection (CAD) systems further assist by performing volumetric measurements, texture analysis, and shape descriptors, helping to differentiate true polyps from artifacts or folds, with experienced radiologists making the final assessment based on a combination of these features and their clinical expertise. One method used to distinguish a polyp candidate from normal colon mucus tissue is the Hounsfield unit (HU), since each tissue type has an HU value range. The CT device generates images that take values of HU between −1000 and +3000 depending on the scanned tissue type; for example, air is −1000, soft tissue is between −200 and 300, and bone is 1000. Therefore, to increase the contrast between different soft tissues, a windowing is applied on CT scans.

Over the past two decades, many researchers exploited computer vision algorithms to enhance different stages of CTC pipeline, shown in Figure 16.2. Table 16.1 shows a summary of the literature review for algorithms that perform polyp analysis from CT scans. The summary categorizes these algorithms and shows the number of used CT scans as well as the reported sensitivity S% for each algorithm. Although the reported sensitivities listed in Table 16.1 are promising, there are many limitations and drawbacks in these approaches. The approaches in [1, 2] are shape-based algorithms but polyps come in a variety of shapes, sizes, and types. Wijk and colleagues [3] focused on a single type with lots of false positives. Furthermore, [4–8] classified polyps vs non polyps, and [9–11] classified polyp categories (e.g., adenomatous, serrated, etc.) However, these approaches do not detect polyps.

Currently, there exists no baseline dataset for polyps, and the authors, in Table 16.1, used their private datasets for evaluation. In this work, experiments are conducted on the dataset that was used by Zhang and colleagues [11]. However, they focused on distinguishing malignant from benign lesions, but we focus on the detection and localization of polyps. The only public dataset that has CT scans with colonic polyp annotation is presented by The Cancer Imaging Archive (TCIA), ACRIN-6664 [13]. However, only the number of slices that have polyps is provided, not the location or size of each polyp within each slice. Finally, none of the approaches in Table 16.1 detects the exact location and size of a polyp within a CT scan except Grosu and colleagues' approach [10], which detects polyps before classification, but with a detection rate of 73%.

The proposed framework aims to identify the accurate location and size of potential polyps by locating the slice and the bounding box for the detected polyp. Colonic polyp detection is a challenging problem due to the uncertainties in acquiring CT scans, for example preparation artifacts, polyp size, and location [14]. To develop an efficient approach that overcomes polyp detection and localization challenges, the proposed algorithm focuses on colon regions. So, segmented colon regions are fed

DOI: 10.1201/9781003449942-19

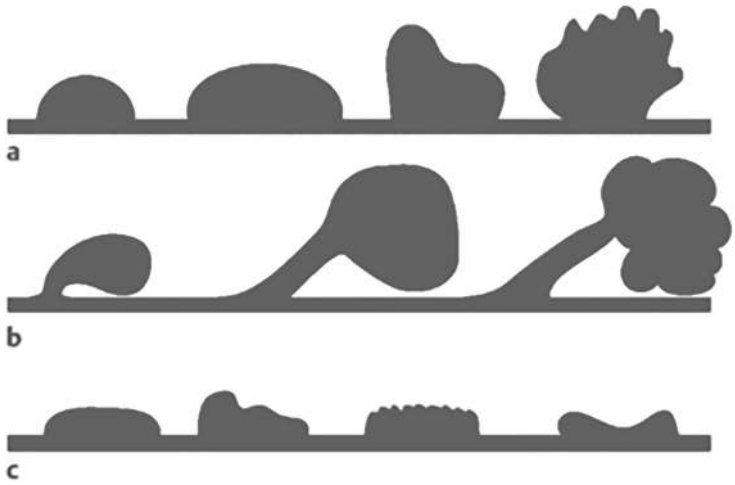
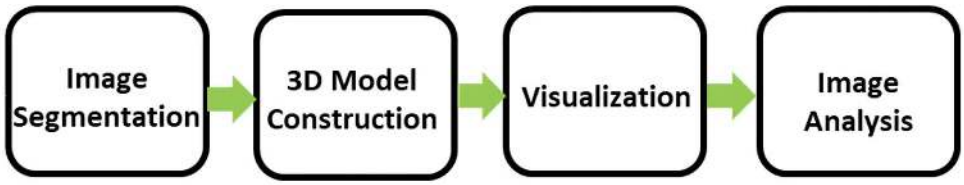
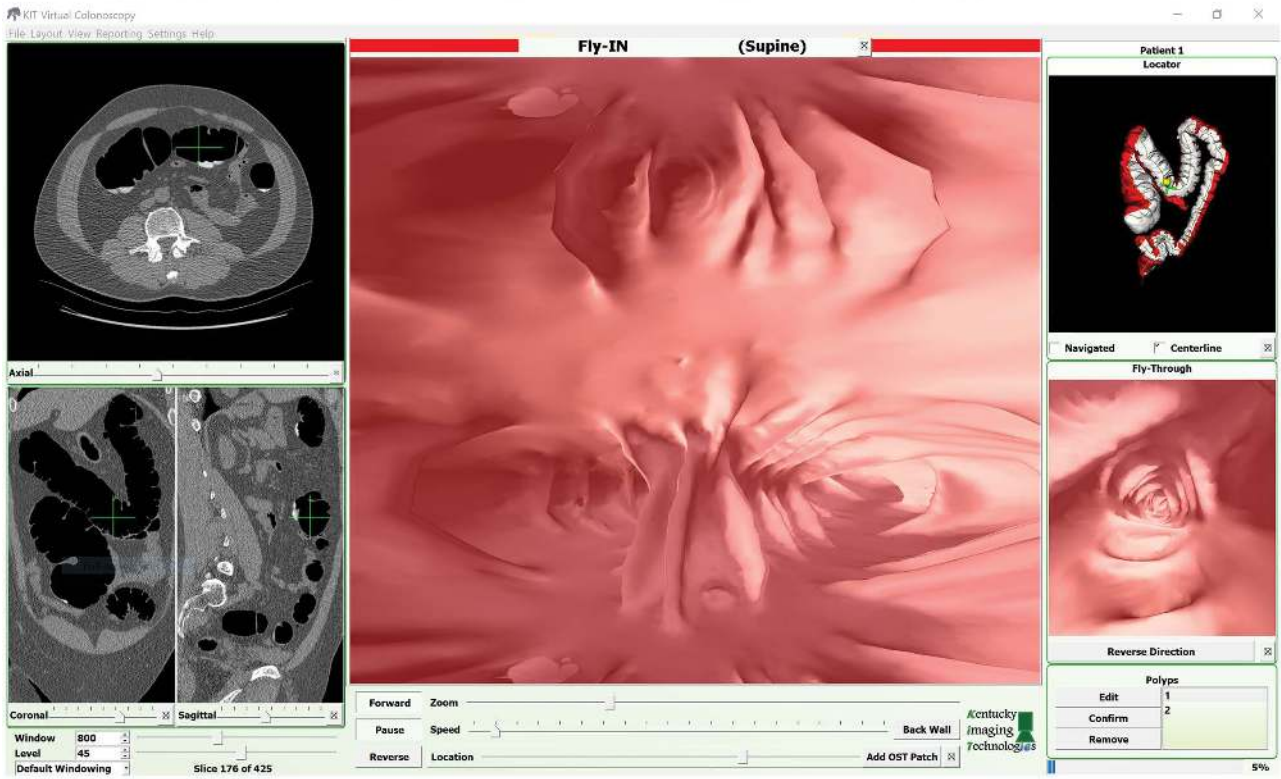


FIGURE 16.1 Polyps shape categories: (a) sessile polyps (round, ovoid, lobulated); (b) pedunculated polyps; (c) flat polyps (smooth, nodular, central depression).



(a) Typical four-step CTC system, starting from CT to polyp detection



(b) GUI for Research Platform for CTC

FIGURE 16.2 CTC Platform: (a) typical pipeline; (b) CTC components: 2D CT slices, 3D model reconstruction and visualizations: Fly-In and Fly-Through.

TABLE 16.1

Summary of Polyp Analysis Using CT Scans

Approach	# Scans	S%	Limitations
Shape-based model [1]	10	71	shape variant, no location, and size estimate
Texture and shape analysis [2]	56	100	
Shape and size invariant [3]	249	95	a single type, high false positives
DETALE transfer learning [4]	154	94	polyp or non-polyp only, no location or size estimate
— Deep ensemble CNN [7]	403	90.5	
— DCNN-CADe [8]	144 polyps	93	
— 2D CNN [7]	N/A	97	
— 3D-GLCM CNN [8]	63 polyps	90.1	
3D-Dense CNN [9]	403	95.1	categorization only, no location or size estimate
MM-GLCM CNN [12]	100	94.5	
Seg + 3D CNN [11]	107	80	
MM-GLCM CNN [10]	106	82	categorization with a low detection rate

into a CNN-based detector. We also exploit the 3D surface information in virtual colonoscopy (VC) to generate different geometric features that discriminate between polyp and non-polyp regions. These features are used to train another CNN-based detector. Both detectors feed their candidates to the proposed MVN module to reject false positives. The main contributions of this chapter are the following:

1. Developing a CT-based polyp detector that is trained and validated using 2D axial CT scans.
2. Developing a VC-based model that uses generated features encoding 3D surface geometry.
3. Developing an MVFN module that rejects false positives generated by the two detectors.

16.2 Materials

16.2.1 Dataset

CTC experts from the University of Wisconsin provided the dataset used to train and validate the proposed modules. The data consist of scans, in the supine and prone positions, for 50 patients. The scans have 59 annotated polyps larger than 6 mm. Lesion annotation was done by one of three experienced radiologists. This research study was conducted retrospectively using human subject data made available in open access by ACRIN, Walter Reed Medical Center, and the University of Wisconsin [11]. The research protocol is governed by the University of Louisville IRB No. 07.0252.

16.2.2 Hardware

The server hardware used for the training and validation has the following specs: processor: AMD EPYC 7702P 64-core processor, RAM size: 264GB, used GPU: two Nvidia TITAN X with 24 GB of VRAM each.

For the CT device to generate the dataset, the manufacturer was GE Medical Systems, with kilovoltage peak = 120, slice thickness 1.25, and X-ray tube current 180mA.

16.2.3 Software and Packages

The Ubuntu 20.04.6 LTS is the main operating system that had been used, with Python 3.9 as the main programming language. For the CNN-based detector, YOLO detector is used [15], and Pytorch 2 and Keras 3 libraries have been used for creating the model. In addition, Numpy, pandas, PyDicom, and OpenCV libraries have been used for preparing the CT images before processing.

16.3 Methods

In this chapter, we focus on the last stage of the CTC pipeline, shown in Figure 16.2, i.e., the polyp detection. First, we begin with the DICOM image slices generated by the CT scanner; an example of a slice is displayed in Figure 16.3b. Next, only the colon regions from each slice are extracted using a multi-step colon segmentation approach [16]. This approach deploys three classic algorithms: 1) Expectation-Maximization to estimate thresholds between dominant modes of HU in a DICOM volume from which an initial segmentation is generated. 2) Region Growing to extract an eroded colon starting from the rectum, and 3) Graph Cut to refine the initial segmentation and generate the optimal masks. Two distinct methods employ these output masks. The first method, depicted in Figure 16.3b, eliminates all non-colon regions from every CT slice which is fed into the CT-based detector that identifies the location and size of the polyp candidate. In the meantime, the second method reconstructs a 3D model of the colon using the masks for every slice. This three-dimensional model is fed into the VC-based detector, which, as seen in Figure 16.3a, makes use of the extracted three-dimensional features projected on Fly-In images [17] to determine the location and size of the polyp candidates. While these two detectors in Figure 16.3a and Figure 16.3b may generate many false positives, the CNN-based MVFN classifier, shown in Figure 16.3c, filters the polyp candidates from both prior polyp detectors by taking advantage of sagittal and coronal views of the specified axial slice of each polyp candidate.

Each model will be described in the following sections.

16.3.1 CT-Based Polyp Detection Approach

Given a CT scan, a segmentation step [16] is performed to identify the colon wall. Then each CT slice is multiplied by its dilated binary mask, as shown in Figure 16.3. Initially, the CNN-based detector is guided by segmentation, which modifies the effective receptive field of the detector's neurons to focus on colon regions. As shown in Figure 16.3b, the segmentation exports binary masks corresponding to colon and non-colon regions. However, the generated masks are tangent to the colon wall, to which polyps are often attached. To overcome this problem a dilation process is employed.

These segmented images have been used to train the CNN-based model [18] described in the material section. The goal of the trained model is to extract all the true polyps from the axial view, even with a high number of false positives. Therefore, the threshold of the confidence scores is chosen to have a low value to produce all possible true positive regions.

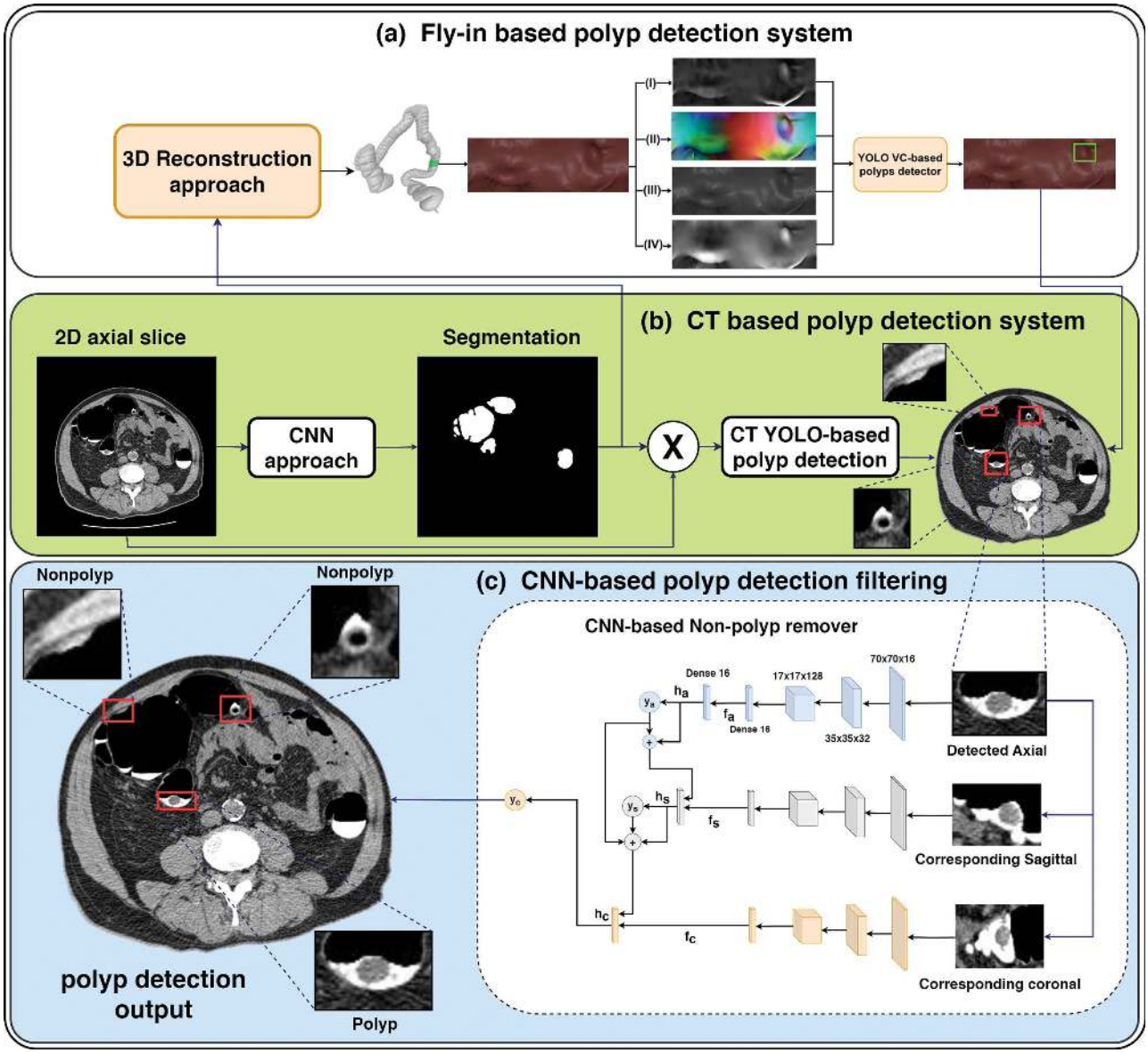


FIGURE 16.3 The proposed colonic polyp detection framework. (a) VC-based detection depends on (I) the curvature map, (II) the normal map, (III) the classical Fly-In visualization albedo and lightning image, and (IV) the depth map. (b) In the CT-based approach an input axial CT slice is segmented and then fed into a CNN-based detector to localize polyp candidates. (c) MVN classifier depends on a Markov chain network that uses detected axial candidate and its corresponding sagittal and coronal views to classify a candidate and reject false positives.

16.3.2 VC-Based Polyp Detection Approach

After a colon is segmented, 3D reconstruction to generate a surface representation of the colon is performed using the Delaunay triangulation method [19]. This technique reconstructs the 3D mesh surface of the colon as shown in Figure 16.3a using the extracted 2D segmentation mask, as illustrated in Figure 16.3. Next, images for luminal surfaces are created using the virtual colonoscopy Fly-In technique [17], which uses a ring of virtual cameras with optical centers on the colon centerline and the optical axes perpendicular to it, as shown in Figure 16.3a. A filet representation results by patching the images in the frustum of the cameras. In the proposed VC-based polyp detection framework, Fly-In is used to generate 2D images of the internal surface of the colon for better visualization of the colon wall.

In addition to the conventional images formed by Fly-In (i.e., albedo and lightning, Figure 16.3a(III)), we generate sets of images coding the 3D geometric surface features:

1. The surface curvature map, as shown in Figure 16.3a(I), in which the curvature is calculated using the algebraic point set surface [20]. The curvature is based on moving least squares (MLS) fitting algebraic spheres to the surface.
2. Normal map that could be computed at each vertex, on the surface, by the cross product of two

vectors from two neighbor vertices. As shown in Figure 16.3a(II), the normal map is represented as a three-channel (RGB) image to represent the normal vector (N_x, N_y, N_z) .

3. The depth map that reflects the smallest distance between each surface point and the centerline, as shown in Figure 16.3a(IV).

We hypothesize that fusing the 2D projections and the 3D colon representation in virtual colonoscopy can enhance polyp detection accuracy. The reason of this expected enhancement is that images formed by Fly-In can't accurately encode convex (e.g., polyps) and concave (e.g., folds) surface regions if improper lightning is used. Therefore, the generated feature maps are combined as multi-channel images and fed into a CNN-based model, which can be trained to extract polyp candidates in these images.

16.3.3 Multi-View Fusion Network (MVFN)

Since the detectors produce many false positives, an algorithm is needed to filter these false positives. We tried to mimic what radiologists do when they find a potential polyp candidate. They usually confirm or discard this finding by examining the different views of the polyp candidate. This is why we extract the corresponding sagittal and coronal views corresponding to the current axial view to be fed into our proposed MVFN model. Our hypothesis is based on that most polyps have the shape of a small protruding mound, and they should appear in the three views as a small protruding mound, unlike the non-polyp areas, for example colon folds. Figure 16.4 shows examples of different false positive samples. The top row illustrates the axial views, and each has a geometric feature like polyps, so they can deceive the detector. However, the sagittal and coronal views, in the bottom row, show that the geometric and appearance features of the polyp are not presented. Therefore, using the other views will filter the detector results. However, in some hard cases (e.g., Figure 16.4e) the three views show the presence of the geometric feature of a polyp while it is not a true polyp. To integrate information from the three different views, the joint probability over all the views factorizes into the conditional probabilities over the separate views, using the following chain rule.

$$P(Y \mid X) = P(Y_c \mid X, Y_s, Y_a) P(Y_s \mid X, Y_a) P(Y_a \mid X), \tag{16.1}$$

where $X = \{X_c, X_s, X_a\}$ is the input sequence of the three views and $Y = \{Y_c, Y_s, Y_a\}$ is the output sequence, which should be predicted. As shown in Figure 16.3c, the learned features $\{f_c, f_s, f_a\}$ are extracted from the three views $\{X_c, X_s, X_a\}$, respectively. So, the conditional probabilities can be calculated as follows: $P(Y_a \mid X) = Sigmoid(h_a)$, $P(Y_s \mid X, Y_a) = Sigmoid(h_s)$, and $P(Y_c \mid X, Y_a, Y_s) = Sigmoid(h_c)$. Where $h_a = Net(f_a)$, $h_s = Net([f_s, h_a, y_a])$, $h_c = Net([f_c, h_s, y_s, y_a])$, and $Net(.)$ represents a fully connected network.

16.4 Results

All the training was performed on Nvidia TITAN RTX 24 GB. Moreover, all models were trained to a high number of epochs (2000 epochs) to get the best validation score epoch.

16.4.1 Training and Testing Procedure for CT-Based Detector Stage

To evaluate the CT-based detection model, first, the proposed colon segmentation approach has been applied to the input axial view of CT slices to be fed into the next stage for detecting polyps. Next, data augmentation was used to make different variations to the dataset by flipping some of the images horizontally and applying different exposure, saturation, and brightness. These augmented and segmented images have been used to train the model. To investigate which SOTA CNN-based detector can be used, masked slices have been used to

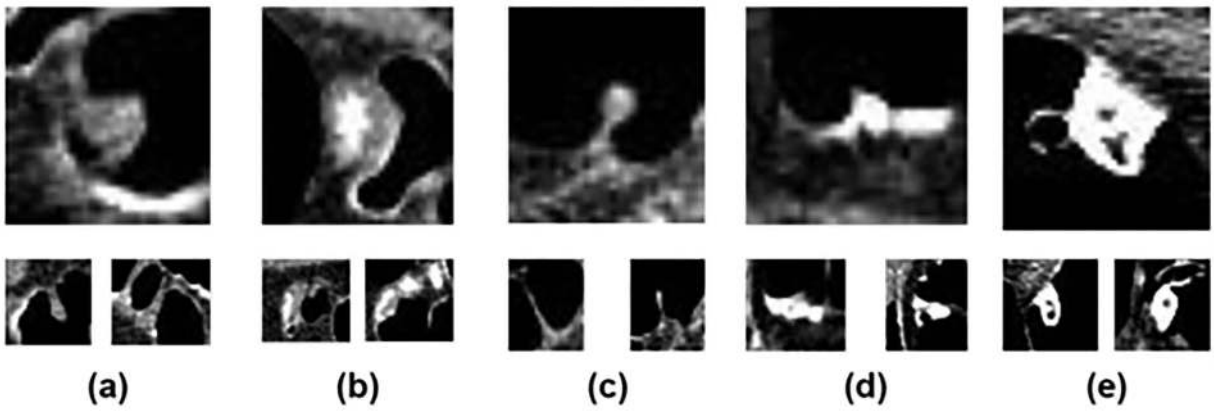


FIGURE 16.4 Different false positive samples. The top is the axial view showing the geometric appearance and features of polyps. (a–d) False positives that do not have consistent features in the three views. In hard cases (e.g., e) the three views confirm the presence of the geometric feature of a polyp, so the proposed approach may fail.

TABLE 16.2

The Results of the CT-Based Detection Stage

Model	Sensitivity	mAP
YOLO-V5 [21]	86.67%	79.7%
YOLO-V7 [15]	**88.05%**	**85.7%**
Faster RCNN R-50 [22]	69.1%	71%
Faster RCNN R-101 [22]	70.6%	71.7%

train different models; for example, YOLO-V5 [21], YOLO-V7 [15], Faster-RCNN with Resnet-50 [22], and Faster RCNN with Resnet-101 [22]. As shown in Table 16.2, the YOLO-V7 detector has the highest mean average precision, so it has been chosen to be the detector of the first stage in the proposed framework. To present the importance of the segmentation step on axial view images, a YOLO model has been trained with the original 2D CT slices without the segmentation step. This model gives only 44% mAP.

To explore other architectures, Dynamic RCNN with ResNet 50 [23], Retina Net with Efficient Net backbone [24], Sparse RCNN [25], and Swin Transformer [26] were implemented and trained for the polyp detection task. However, their results were not included in Table 16.2 as these models performed poorly, with a very low mAP. These models were unable to detect very small objects such as a polyp, which only occupy on average 4%–5% of the total image size. All the detectors training batch size is 16, but all other hyperparameters were the original hyperparameters from each model's original papers.

16.4.2 Training and Testing Procedure for VC-based Detector Stage

For the 3D VC-based polyp detection approach, we test each geometric feature separately, to know which one should be used. Table 16.3 shows the comparison for each proposed geometric feature as an input for different SOTA detection approaches like YOLO V7 [15], dynamic RCNN with ResNet 50 [23], faster RCNN with ResNet-50 or ResNet-101 as a backbone [22], Retina Net with Efficient Net backbone [24], Sparse RCNN [25], and Swin Transformer [26]. For the training procedure, all the detector's training batch size = 16 except the Swin transformer which has batch size = 8 due to the size of the model, but all other hyperparameters were the original hyperparameters from each model's original papers.

As shown in Table 16.3, the curvature feature shows the highest mAP score with the use of YOLOv7, which is the appearance feature that can best distinguish between convex curvature and concave curvature. The colon contains a lot of concave curvatures that can deceive the detector, so a feature that can distinguish concave and convex surfaces is needed to make the detector work best since polyps have mostly a convex curvature shape. That is why the curvature feature is the most important. The second-best feature is the depth feature with Dynamic RCNN network. Since depth can distinguish between near and far tissue surfaces, it can give the model the needed feature for distinguishing between polyp and nonpolyp parts. We believe that the Swin transformer has great potential to achieve better results, but it needs a larger polyp dataset than we used to train the model.

TABLE 16.3

The mAP Validation Results for Different Approaches for the 3D-Based Detection Model

Model	Curvature (I)	Normal (II)	Original (III)	Depth (IV)
YOLO-V7 [15]	**97.1%**	79.4%	84.5%	82.9%
Dynamic-RCNN R-50 [23]	77.1%	**94.7%**	91.5%	**95.4%**
Faster-RCNN R-50 [22]	76.2%	85.8%	91.0%	88.6%
Faster-RCNN R-101 [22]	81.6%	74.9%	92.0%	83.8%
RetinaNet Efficient-Net [24]	83.6%	93.0%	**93.7%**	92.5%
Sparse-RCNN [25]	45.1%	39.2%	64.3%	62.6%
Swin Transformer [26]	94.0%	90.3%	88.1%	89.2%

TABLE 16.4

The Results of the MVN Classification Stage

Model	Parameters	Sensitivity	AUC
Proposed MVN	**520,067**	**96.92%**	**95.27%**
3D CNN [9]	1,296,577	87.80%	93.52%
Depthwise conv [27]	399,314	84.73%	89.97%
ResNet-18 [28]	11,186,274	79.06%	84.90%
VGG-19 [29]	20,036,418	82.59%	86.31%
GoogleNet [30]	6,158,050	80.58%	85.12%
DenseNet [31]	6,956,298	83.85%	90.53%
SENet [32]	11,256,250	81.46%	88.94%
PNASNet [33]	450,594	79.57%	86.82%
MobileNetV2 [34]	2,286,674	79.82%	85.71%
ResNext [35]	9,120,578	78.56%	82.53%
ShuffleNetV2 [36]	1,263,854	79.31%	84.22%

16.4.3 Training and Testing Procedure for Classifier Stage

To reduce the false positives generated by both detectors, the proposed multi-view fusion network, Figure 16.3, is trained using sagittal and coronal views corresponding to the detected axial views to classify the candidates. Additional experiments have been conducted to investigate more classification architectures compared to the proposed MVFN model. In the first experiment, instead of using the detected 2D axial view and its corresponding sagittal and coronal views, that is, three 70×70 images, the corresponding volume ($70 \times 70 \times 15$) is extracted from the DICOM. Then, the 3D-CNN [9] model has been trained using the extracted volumes. The 3D-CNN model, as expected, needs high computational power to be trained and a larger dataset for the training. As shown in Table 16.4, it is almost seven times the size of the 2D-CNN model as shown in the number of parameters.

In the second experiment, instead of using each of the three corresponding views as a separate image, the three views were combined as three channels, then the three-channel images were used to train the depthwise convolution-based model [27]. The depthwise convolution layer had been used in this experiment because the spatial relations are different in the three views, so the standard CNN architecture cannot be used.

In the third experiment, we use some of the most successful classifiers to compare them with the previous results as

shown in Table 16.4. The implemented classifiers are ResNet 18 [28], VGG-19 [29], GoogleNet [30], DenseNet [31], SENet [32], PNASNet [33], MobileNetV2 [34], ResNext [35], and ShuffleNetV2 [36]. The results show that the proposed MVN classifier has the highest number in the sensitivity 94.80% and AUC 95.27%; that is, it rejected most of the false positive candidates that had been extracted in the previous detector stage, and most of the true positive results are correctly classified. We believe that our proposed MVN model outperforms other models because the model has been customized to fit perfectly on the polyp classification task with the lowest number of parameters needed, which leads to being the fastest model while performing the testing procedure. Moreover, it consumes less time and memory compared to the 3D network (e.g., 3D-CNN [9]). ADAM optimizer with early stopping was used with batch size =256.

To illustrate the effect of each stage in the proposed approach, an experiment has been conducted using 19 2D axial slices, each has a polyp. The standard YOLO-based detector missed 11 polyps— $FN = 11$—in addition to three false positive samples. So, we reduced its confidence score threshold, as in the previous experiments, to detect all the polyps. Although we make $FN = 0$, this configuration extracts three false positives. After applying MVFN classifier, these false positive samples have been decreased to only one sample.

We note that the proposed MVFN model may fail in hard conditions, which have geometric and appearance features like polyp in all the three views. As shown in Figure 16.4, cases (a) to (d) have been correctly classified because at least one of the views does not have polyp-like features, but as shown in case (e), the classifier identifies the sample as a polyp because its three views have polyp-like features, although it is not a polyp.

16.5 Discussion

In this chapter, we presented a deep learning-based framework to identify the locations and sizes of the polyps for CT colonography. The proposed approach consisted of two cascade stages: a detector, to detect polyp candidates, followed by a classifier, to reject false positives. The proposed CT-based detector was guided by changing the neuron's effective receptive field, using the segmented colon wall, which drastically enhanced the detection performance. Moreover, the modified Fly-In method was developed to generate images encoding 3D surface geometry features. Then a VC-based detector was trained using these images to extract potential polyp regions with high performance—mAP ~97.1%. Since both detectors produced false positives, each detected candidate was fed into the proposed multi-view fusion network (MVN) classifier to exploit the different views of the CT scans for each candidate from the two detectors. This step refined the detection by rejecting false positives. The high sensitivity ~94.8% and AUC ~95.3% of the proposed approach illustrate that it can be adopted by radiologists to examine CT scans in a short time. We should indicate what polyps we worked on, and we should provide a vision of how all types of polyps may be evaluated, even using simulation.

16.6 Future Trends

The methodologies described in this chapter can be further developed for various polyp types, in terms of size, location, and topology. Two specific fronts are being pursued:

1. Independent evaluation of supine and prone detection results. A potential workflow involves obtaining results for both positions (supine and prone) after applying registration to make connection between both positions, thereby allowing detection validation in one position (supine or prone) by verifying with the result from the other position.
2. The current framework does not distinguish between polyp types. A trained classification model can be added to classify the polyp type (adenomatous, serrated, villous adenoma polyps) to enable automatic diagnosis and quantify treatment.

REFERENCES

1. R. M. Summers, C. F. Beaulieu, L. M. Pusanik, J. D. Malley, R. B. Jeffrey Jr, D. I. Glazer and S. Napel, "Automated polyp detector for CT colonography: feasibility study," *Radiology,* 2000, vol. 216, no. 1: p. 284–290.
2. W. Hong, F. Qiu and others, "A pipeline for computer aided polyp detection," in *IEEE Trans Vis Comp Graph,* 2006.
3. C. v. Wijk, V. F. v. Ravesteijn, F. M. Vos, R. Truyen, A. H. d. Vries, J. Stoker and L. J. v. Vliet, "Detection of protrusions in curved folded surfaces applied to automated polyp detection in CT colonography," in *MICCAI*, 2006.
4. J. Näppi, T. Hironaka, D. Regge and H. Yoshida, "Deep transfer learning of virtual endoluminal views for the detection of polyps in CT colonography," in *Med Imag Comp Aided Diag*, 2016.
5. K. Umehara, J. J. Näppi, T. Hironaka, D. Regge, T. Ishida and H. Yoshida, "Deep ensemble learning of virtual endoluminal views for polyp detection in CT colonography," in *Med Imag 2017 Comp Aided Diag*, 2017.
6. J. Näppi, P. Pickhardt, D. Kim, T. Hironaka and H. Yoshida, "Deep learning of contrast-coated serrated polyps for computer-aided detection in CT colonography," in *Comp Aided Diag*, 2017.
7. Y. Chen, Y. Ren, L. Fu, J. Xiong, R. Larsson, X. Xu, J. Sun and J. Zhao, "A 3D convolutional neural network framework for polyp candidates detection on the limited dataset of CT colonography," in *IEEE Int Conf EMBC*, 2018.
8. J. Tan, Y. Gao, Z. Liang, W. Cao, M. J. Pomeroy, Y. Huo, L. Li, M. A. Barish, A. F. Abbasi and P. J. Pickhardt, "3D-GLCM CNN: a 3-dimensional gray-level Co-occurrence matrix-based CNN model for polyp classification via CT colonography," *IEEE Tran Med Imag,* 2019.
9. T. Uemura, J. J. Näppi, T. Hironaka, H. Kim and H. Yoshida, "Comparative performance of 3D-DenseNet, 3D-ResNet, and 3D-VGG models in polyp detection for CT colonography," in *Med Imag 2020 Comp Aided Diag*, 2020.
10. S. Grosu, P. Wesp, A. Graser, S. Maurus, C. Schulz, T. Knösel, C. Cyran, J. Ricke, M. Ingrisch and P. Kazmierczak, "Machine learning–based differentiation of benign and premalignant

colorectal polyps detected with CT Colonography in an asymptomatic screening population: a proof-of-concept study," *Radiology,* 2021, vol. 299, no. 2: p. 326–335.
11. P. Wesp, S. Grosu, A. Graser, S. Maurus, C. Schulz, T. Knösel, M. P. Fabritius, B. Schachtner, B. M. Yeh, C. C. Cyran and others, "Deep learning in CT colonography: differentiating premalignant from benign colorectal polyps," *Eur Radiol,* vol. 32, p. 4749–4759, 2022.
12. S. Zhang, J. Wu, E. Shi, S. Yu, Y. Gao, L. C. Li, L. R. Kuo, M. J. Pomeroy and Z. J. Liang, "MM-GLCM-CNN: a multi-scale and multi-level based GLCM-CNN for polyp classification," *Comp Med Imag and Graphics,* 2023, vol. 108: p. 102257.
13. K. Smith, K. Clark, W. Bennett, T. Nolan, J. Kirby, M. Wolfsberger, J. Moulton, B. Vendt and J. Freymann, Data from CT colonography. *The Cancer Imaging Archive* (2015). https://www.cancerimagingarchive.net/collection/ct-colonography/#citations
14. H. I. R. O. M. I. Shinya and W. I. Wolff, "Morphology, anatomic distribution and cancer potential of colonic polyps.," *Ann Surg,* 1979, vol. 190, p. 679.
15. C.-Y. Wang, A. Bochkovskiy and H.-Y. M. Liao, "YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *arXiv Preprint* arXiv:2207.02696, 2022.
16. I. Alkabbany, A. M. Ali, M. Mohamed, S. M. Elshazly and A. Farag, "An AI-based colonic polyp classifier for colorectal cancer screening using low-dose abdominal CT," *Sensors,* 2022, vol. 22, no. 24: p. 9761.
17. M. Mohamed, A. Farag, A. M. Ali, S. Elshazly, A. Farag and M. Ghanoum, "Fly-In visualization for virtual colonoscopy," in *IEEE ICIP*, 2018.
18. A. Bochkovskiy, C.-Y. Wang and H.-Y. M. Liao, "Yolov4: optimal speed and accuracy of object detection," *arXiv Preprint* arXiv:2004.10934, 2020.
19. M. S. Hassouna and A. A. Farag, "Robust centerline extraction framework using level sets," in *IEEE Comput Soc Conf Comp Vision Pattern Recog*, 2005.
20. G. Guennebaud and M. Gross, "Algebraic point set surfaces," in *ACM Siggraph*, 2007.
21. H. Liu, F. Sun, J. Gu and L. Deng, "Sf-yolov5: a lightweight small object detection algorithm based on improved feature fusion mode," *Sensors,* vol. 22, p. 5817, 2022.
22. S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *IEEE Trans Pattern Anal Mach Intell,* 2016, vol. 39, no. 6: p. 1137–1149.
23. H. Zhang, H. Chang, B. Ma, N. Wang and X. Chen, "Dynamic R-CNN: towards high quality object detection via dynamic training," in *ECCV*, 2020.
24. T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal loss for dense object detection," in *ICCV*, 2017.
25. P. Sun, R. Zhang, Y. Jiang, T. Kong, C. Xu, W. Zhan, M. Tomizuka, L. Li, Z. Yuan, C. Wang and others, "Sparse R-CNN: end-to-end object detection with learnable proposals," in *CVPR*, 2021.
26. Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin and B. Guo, "Swin transformer: hierarchical vision transformer using shifted windows," in *Proc IEEE ICCV*, 2021.
27. F. Chollet, "Xception: deep learning with depthwise separable convolutions," in *CVPR*, 2017.
28. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *CVPR*, 2016.
29. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv 1409.1556,* 2014.
30. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," in *CVPR*, 2015.
31. G. Huang, Z. Liu, L. Van Der Maaten and K. Weinberger, "Densely connected convolutional networks," in *CVPR*, 2017.
32. J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation networks," in *CVPR*, 2018.
33. C. Liu, B. Zoph, M. Neumann, J. Shlens, W. Hua, L. Li, L. Fei-Fei, A. Yuille, J. Huang and K. Murphy, "Progressive neural architecture search," in *ECCV*, 2018.
34. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "Mobilenetv2: inverted residuals and linear bottlenecks," in *CVPR*, 2018.
35. S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated residual transformations for deep neural networks," in *CVPR*, 2017.
36. N. Ma, X. Zhang, H. Zheng and J. Sun, "Shufflenet v2: practical guidelines for efficient cnn architecture design," in *ECCV*, 2018.

17

Hybrid Multi-Dimensional MRI of Prostate Cancer

Aritrick Chatterjee and Aytekin Oto

17.1 Introduction

Prostate cancer (PCa) is one of the leading cancers affecting men worldwide (1, 2), and magnetic resonance imaging (MRI) is increasingly playing a key role in the diagnosis of prostate cancer. A multi-parametric approach consisting of different MRI sequences such as T2-weighted imaging (T2W), diffusion weighted imaging (DWI), and dynamic contrast enhanced (DCE) MRI makes up multi-parametric MRI (mpMRI), which is widely used. Cancers are seen as hypointense foci with respect to surrounding benign tissue on T2W and apparent diffusion coefficient (ADC) maps derived from DWI, and hyperintense on high b-value DWI and early phase DCE-MRI. The consensus guidelines to interpret and report prostate mpMRI has been recommended by the Prostate Imaging Reporting and Data System (PI-RADS) (3) and involves the subjective interpretation of prostate MRI images from these multiple MRI sequences by radiologists using the guidelines. ADC/DWI and T2W are the primary sequence for interpretation in the peripheral and transition zones respectively, and DCE-MRI is a secondary sequence that works in conjunction with these primary sequences.

Recent studies have demonstrated that prostate MRI-guided biopsies significantly improve PCa diagnosis (4, 5). Nonetheless, even expert radiologists have been shown to overlook up to 30% of clinically significant cancers on mpMRI (6). Moreover, significant interobserver variability in the qualitative interpretation of prostate mpMRI among radiologists is a notable concern (7). Hence, there is a keen interest in novel quantitative MRI approaches for PCa diagnosis that could alleviate these challenges and enhance PCa diagnosis, ultimately resulting in improved patient outcomes.

One of these techniques is hybrid multi-dimensional MRI (HM-MRI), which was originally used in studies of neurological model systems (8, 9). Our group modified MRI data acquisition and analysis for prostate MRI. HM-MRI combines two fundamental contrast mechanisms utilized in prostate mpMRI: T2-relaxometry and diffusion. Conventional diffusion imaging samples only a series of b-values (degree of diffusion weighting applied) at a single echo time (TE)—a single row in Figure 17.1. Conventional T2 mapping samples only the left-hand column in the diagram (multiple TEs at b = 0). In contrast, HM-MRI samples all matrix elements (generally 3 × 3 or 4 × 4) in Figure 17.1. Unlike mpMRI, HM-MRI does not assume diffusion and T2 to be independent. Instead, it recognizes that each tissue component has distinct and coupled properties (10, 11)—for example, lumen has high diffusivity and T2 relaxation time, while epithelium has low diffusivity and T2 relaxation time (12). Prostate tissue composition of these gland components—stroma, epithelium, and lumen—change with the presence (13) and Gleason grade of PCa (14). HM-MRI leverages the unique MRI properties of prostate tissue components—stroma, epithelium, and lumen (12)—to noninvasively measure the tissue composition using MRI. These measurements serve as biomarkers for noninvasive diagnosis of PCa.

A feasibility study has demonstrated the noninvasive measurement of prostate tissue composition using HM-MRI, indicating its potential to enhance PCa diagnosis and assess its aggressiveness (15). Validation of prostate tissue composition measured noninvasively using HM-MRI has been achieved through comparison with established reference standards such as quantitative histology results (16) and pathologists' measures (17) from whole mount prostatectomy, suggesting this approach has the potential to improve PCa diagnosis. The clinical efficacy of HM-MRI has been validated through two retrospective observer studies. One study revealed superior diagnostic performance of HM-MRI as a standalone tool compared to mpMRI/radiologists, as well as showing improved interobserver agreement and reduced interpretation time using HM-MRI compared to mpMRI (18). In another study, the combination of HM-MRI with conventional mpMRI improved radiologists' performance (19). Recently, validation in a prospective clinical trial for guiding prostate biopsies has further solidified its clinical utility (20). This chapter will elucidate the imaging and analysis techniques and present findings from studies detailing these pivotal steps in integrating HM-MRI into routine clinical practice.

17.2 Materials

17.2.1 MRI

A 3T MR scanner (Philips Achieva and Ingenia scanners) was used for MR imaging. A 16-channel phased array anterior coil placed on the pelvis and a built-in multi-channel posterior coil, combined with an endorectal coil (Medrad; Bayer Healthcare), was used.

17.2.2 Histology Material and Equipment

Formalin and hematoxylin-eosin (H&E) were used to process and stain histologic slides. A digital microscope (Olympus VS120; Olympus) was used to digitize these histology slides.

DOI: 10.1201/9781003449942-20

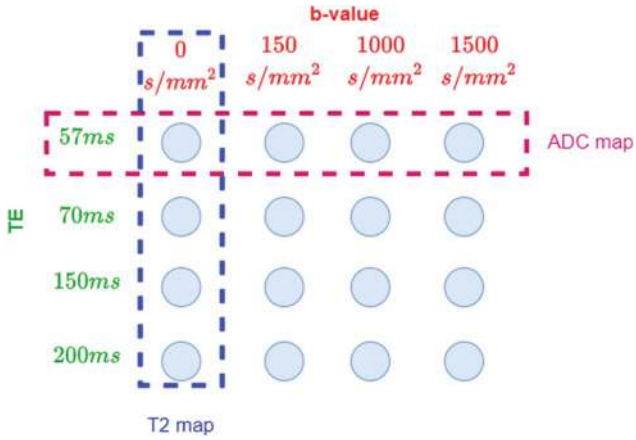


FIGURE 17.1 Sampling of the hybrid matrix of b and TE values is shown schematically. The first row and column, indicated by dashed lines, show the data that would be acquired for conventional ADC and T2 measurements. The additional matrix elements are only acquired in HM-MRI.

17.2.3 Biopsy Equipment

A Uronav MR-US fusion biopsy device was used to perform prostate biopsies.

17.2.4 Software

MR image analysis was performed on MATLAB (MathWorks, Natick, MA). Image Pro software (Media Cybernetics) was used to segment digitized histology images for validating HM-MRI with quantitative histology. Statistical analysis was done using SPSS (IBM Corporation, Armonk, NY) and R.

17.3 Methods

17.3.1 Patient Recruitment

These prospective studies were approved by the institutional review board and are compliant with the Health Insurance Portability and Accountability Act. All participants provided informed written consent. Patients with known or suspected PCa scheduled for a diagnostic MRI exam of the prostate, followed by either radical prostatectomy or MRI-guided fusion biopsy of the prostate, were recruited for these studies at our hospital and research center. Subjects with prior radiation or hormonal replacement therapy or subjects incapable of giving informed written consent or adhering to the experimental protocols for any reason were excluded.

17.3.2 MRI Imaging Protocol

The setup included subjects undergoing MR imaging with a 3T MR scanner. A 16-channel phased array anterior coil placed on the pelvis, and a built-in multi-channel posterior coil, combined with an endorectal coil, was used. The HM-MRI sequence consisted of a spin-echo module with diffusion sensitizing gradients placed symmetrically about the 180° pulse followed by single shot echo-planar imaging readout. HM-MRI was acquired at all combinations of TE (typically 57, 70, 150, 200 ms) and b-values (typically 0, 150, 1000, 1500 s/mm^2) similar to that shown in Figure 17.1. Fat saturation was done using spectral adiabatic inversion-recovery (SPAIR). Typical HM-MRI parameters used in these studies were field-of-view, 180 × 180 mm^2; scan matrix, 120 × 120; in-plane resolution, 1.5 × 1.5 mm^2; repetition time, 5 sec; typical number of slices, 25; slice thickness, 3 mm; reconstructed resolution, 1.41 × 1.41 mm^2; and reconstruction matrix, 128 × 128. The acquisition time was 10–12 minutes.

Subjects were scanned with a standard clinical mpMRI protocol: T2-weighted, DWI along with ADC maps, T1-weighted images, and DCE-MRI as per standard of care.

17.3.3 HM-MRI Analysis

Prostate tissue component volumes (V_n) of stroma, epithelium, and lumen can be calculated by fitting the HM-MRI data to a three-compartment signal model by exploiting the coupled and distinct T2 and ADC values associated with each tissue component. The analysis was performed using a custom software written in MATLAB. The HM-MRI signals were modeled as unmixed pools of water in three tissue components: stroma, epithelium, and lumen, each with distinct and coupled T2 and ADC properties. The fractional volumes of tissue component were calculated on a voxel-by-voxel basis by fitting the following equation using nonlinear least squares method:

$$\frac{S}{S_o} = \sum_{n=1}^{n=3} V_n \times exp\left(-ADC_n \times b - \frac{TE}{T2_n}\right)$$

where V_n, $T2_n$, and ADC_n are the volume fractions, T2 and diffusivity values for each tissue compartments (stroma, epithelium, and lumen), and S is the signal intensity at each combination of echo time and b-value and S_0 is the signal intensity at the lowest echo time and b-value. The sum of volumes of the three tissue components = 100%.

Cancer predicted map was created based off these results, where comjoint voxels with > 40% epithelium and < 20% lumen meeting a size criteria < 25 mm^2 were marked as predicted cancer lesions.

17.3.4 Histological Analysis and Validation

Some participants underwent radical prostatectomy after MRI. The excised prostate was fixed in formalin and serially sectioned in approximately the same plane as the axial MRI scans. Whole-mount tissue sections were stained with H&E to create histologic slides. The slides were evaluated for PCa by expert pathologists and PCa were marked. These whole mount slides were digitized using a digital microscope at 20× magnification. Rad-path correlation was done by co-registering MRI and histology slides visually by consensus of expert radiologist, pathologist, and medical physicist. Regions of interest (ROIs) were drawn for further analysis.

For histologic validation against quantitative histology, quantitative histologic evaluation was performed to calculate volume fractions for each tissue component for ROIs

corresponding to MRI using Image Pro software. Tissue composition measured by using HM-MRI and quantitative histologic evaluation were compared. For validating agreement with pathologists, three experienced genitourinary pathologists independently reviewed each histology image and provided an estimate of percentage of tissue composition for ROIs corresponding to MRI. After a month's gap they worked in consensus to come up with reference standard results. Agreement statistics using total deviation index ($TDI_{0.9}$) was performed for tissue composition measured using HM-MRI and reference standard results from pathologists' consensus.

17.3.5 Observer Studies

Four readers with 1–20 years of experience interpreted MRI data for 61 subjects. In the first study they interpreted mpMRI (triplanar T2-weighted, diffusion-weighted images along with ADC maps and dynamic contrast-enhanced images) and HM-MRI (predicted cancer maps) examinations independently, with a four-week washout period between interpretations. PI-RADS score, lesion location, and interpretation time were recorded. The radiologists' performance using mpMRI was compared with their performance using HM-MRI, along with interpretation time, interobserver agreement. In the second study, they interpreted mpMRI alone followed by the combination of mpMRI and HM-MRI.

17.3.6 Prospective Clinical Trial in Guiding Prostate Biopsies

We evaluated the use of an automated HM-MRI-based tool to prospectively identify targets before MR-US fusion biopsy in comparison with PI-RADS-based evaluation by expert radiologists. In this prospective clinical trial, 91 patients with known or suspected PCa underwent 3T MRI with a conventional mpMRI protocol and HM-MRI and subsequent biopsy. Using the HM-MRI tool, tissue composition was calculated using a three-compartment model, and suspected PCa regions with elevated epithelium (> 40%) and reduced lumen (< 20%) meeting the minimum size requirement of 25 mm^2 were identified in similar manner to what was described in Section 17.3.3. Up to two additional biopsy targets per patient were automatically selected by the HM-MRI tool, in addition to the biopsy targets selected based on an expert radiologist's mpMRI interpretation (≥ PI-RADS 3) using the Uronav MR-US fusion biopsy device. Additional 12-core TRUS-guided sextant random biopsy cores were also obtained.

17.4 Results

17.4.1 Feasibility Study

In our initial investigation to assess the viability of HM-MRI (15), we recruited 22 subjects diagnosed with PCa and scanned with HM-MRI followed by prostatectomy. Volumes of the tissue components: stroma and epithelium alongside measurements of the apparent diffusion coefficient and T2 values were taken for ROIs encompassing PCa ($n = 28$) and benign tissue ($n = 71$) that were confirmed by prostatectomy specimens. PCa exhibited a significantly higher volume of rapidly proliferating epithelial cells (49 ± 9 vs 23 ± 7%) compared to benign tissue. Moreover, the epithelial cells were observed to displace supporting stroma (37 ± 9 vs 51 ± 16%) and luminal space (14 ± 5 vs 26 ± 14%), as depicted in Figure 17.2. The results showed that it was feasible to measure prostate tissue composition accurately and noninvasively with HM-MRI.

These parameters demonstrate excellent diagnostic accuracy (area under the ROC curve) in distinguishing prostate cancer from a normal prostate: epithelium (0.991), lumen (0.800), and stroma volume (0.789). Furthermore, these HM-MRI parameters exhibit stronger correlation with PCa aggressiveness/Gleason score (epithelium: $\rho = 0.65$, lumen: $\rho = -0.39$, stroma: $\rho = -0.44$) compared to conventional MRI parameters: ADC ($\rho = -0.32$) and T2 ($\rho = -0.29$). Hence, HM-MRI holds promise for enhancing prostate cancer diagnosis and evaluating its aggressiveness.

17.4.2 Histologic Validation

17.4.2.1 Validation against Quantitative Histology

These noninvasive tissue measures were validated in an additional cohort of 25 subjects (30 cancers and 45 benign regions) utilizing quantitative histology of prostatectomy samples as the gold standard reference (16). Excellent agreement/concordance was observed between the results obtained from HM-MRI and reference standard results from histological evaluation of whole mount prostatectomy specimens. Our results showed an excellent correlation ($r = 0.91$, epithelium $r = 0.93$, lumen $r = 0.90$) in tissue measures from HM-MRI and pathology as shown in Figure 17.3. There were no significant differences in prostate tissue composition measurements by HM-MRI and quantitative histologic evaluation: stroma (45% vs 44%), epithelium (31% vs 34%), and lumen (24% vs 22%) (≥ 0.05). Notably, these tissue measures exhibited excellent differentiation (AUC for epithelium using HM-MRI = 0.87, pathologists = 0.97, and AUC for lumen volume using HM-MRI = 0.85, pathologists = 0.77) between cancerous and benign tissue.

17.4.2.2 Agreement with Expert Pathologists

HM-MRI underwent further evaluation in another 165 regions of interest (64 PCa and 101 benign tissue) in 36 subjects (17). The findings demonstrated that tissue composition measures from HM-MRI were found to show excellent agreement with the consensus of three expert pathologists. When comparing the results of tissue composition measurements obtained from HM-MRI to the ground truth determined by the consensus of three pathologists, the paired differences (assessed using the total deviation index or $TDI_{0.9}$) were 18.8% and 22.4% for measuring epithelium and lumen, respectively. These differences were lower than the agreement observed among the pathologists themselves (22.3% for epithelium and 24.2% for lumen), indicating that HM-MRI performed comparably to expert pathologists.

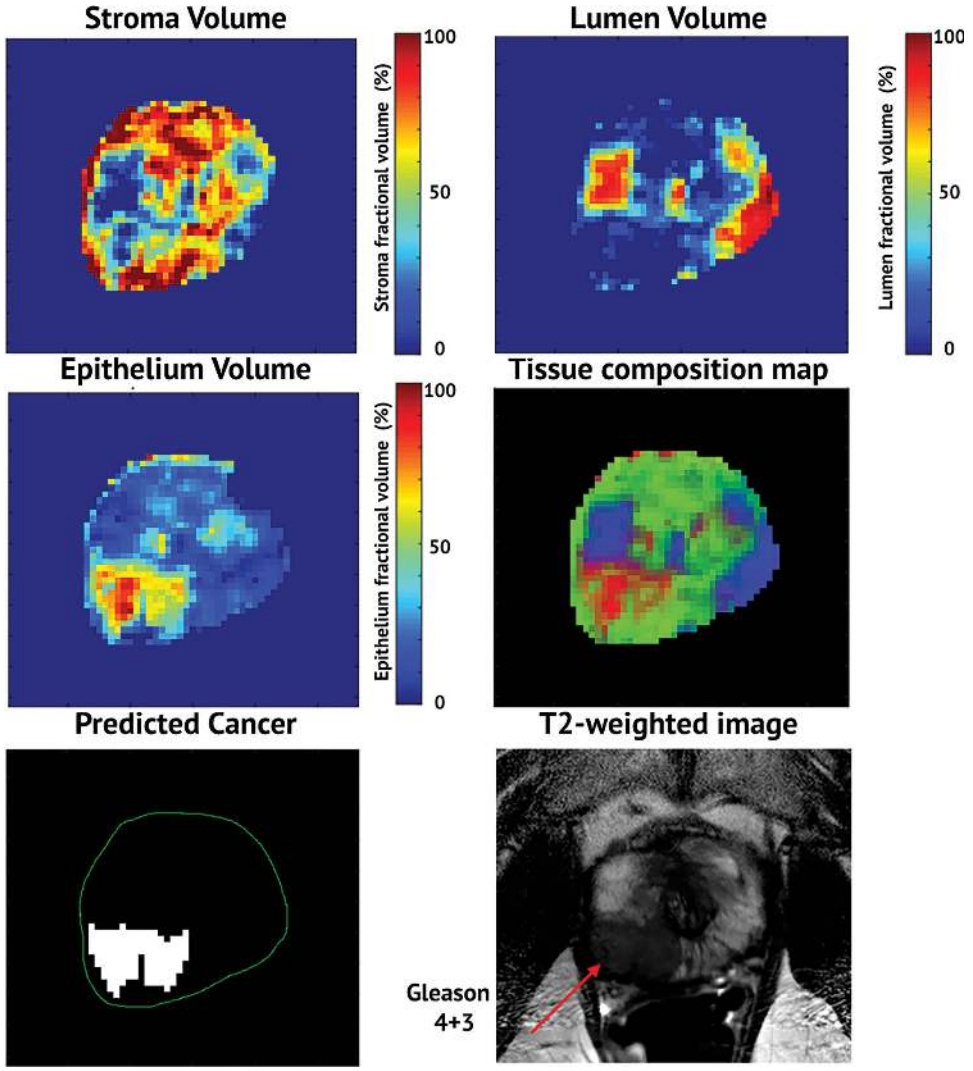


FIGURE 17.2 56-year-old patient with Gleason 4+3 cancer in right peripheral zone. The cancer lesion (arrows) shows reduced stroma (36.4%) and lumen (7.8%) volume and elevated epithelium volume (55.8%) compared to surrounding benign tissue. HM-MRI correctly predicts it to be a clinically significant cancer on the predicted cancer map, and seen as hypointense region in the corresponding T2-weighted image. (Reproduced from Chatterjee et. al., (21) with permission from Springer Nature.)

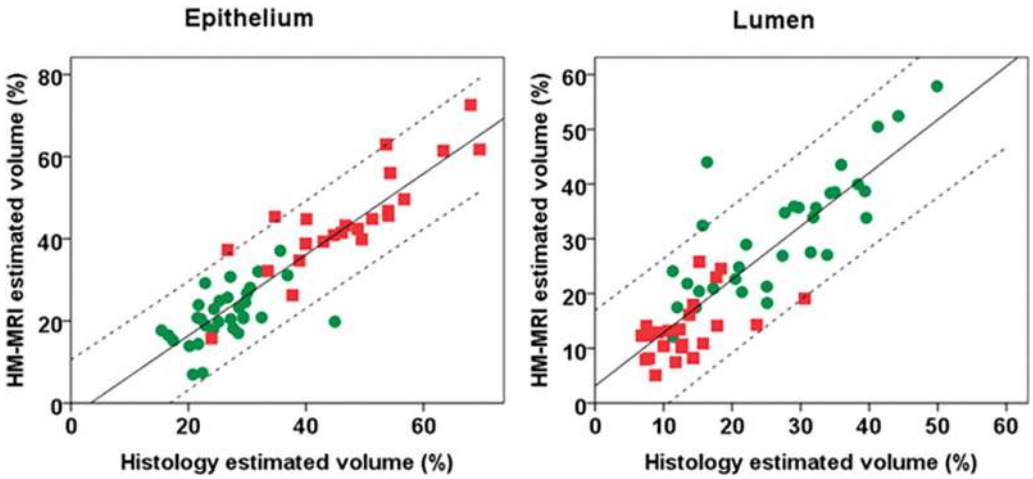


FIGURE 17.3 There is excellent correlation (epithelium: $r = 0.93$, lumen: $r = 0.90$) in tissue measures from HM-MRI and quantitate histology from whole mount pathology. Cancers (red dots) have higher epithelium and reduced lumen compared to benign ROIs (green dots).

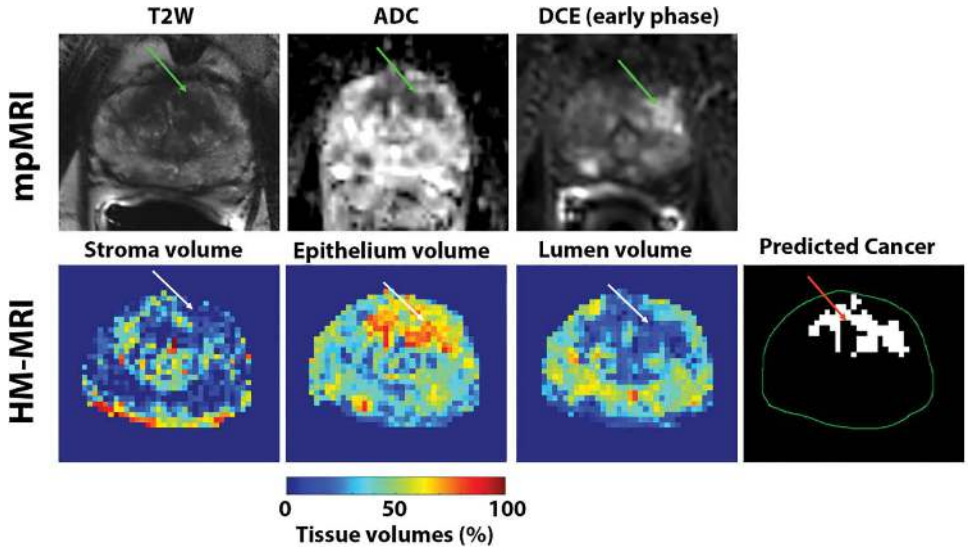


FIGURE 17.4 75-year-old subject that was part of the prospective clinical trial. Hybrid MRI found a lesion in the anterior apex that was missed by radiologists on mpMRI. This was biopsied as part of the clinical trial and found to be a Gleason 3+4 cancer.

17.4.3 Observer Studies

The clinical utility of HM-MRI to diagnose PCa was validated through two observer studies (18) involving four radiologists with experience levels spanning 18, 10, 5, and 2 years. These studies comprised 61 subjects who underwent mpMRI and HM-MRI scans. In the first study, radiologists interpreted mpMRI first, followed by HM-MRI after a four-week interval and found improved performance using HM-MRI for three out of four readers in a per-lesion analysis (R1, R2, and R3: 78%, 85%, and 82% respectively) compared with mpMRI (68%, 81%, and 73%; $P < .001$ for all). In a per-patient analysis, the least experienced reader exhibited the most significant improvement (with the AUC increasing from 0.46 using mpMRI to 0.64 using HM-MRI). Per-patient specificity was higher for HM-MRI among readers 2–4 compared to mpMRI (specificity for readers 1–4: 48%, 78%, 48%, and 46% vs 37%, 26%, 0%, and 7%; $P = .34$, $p < .001$, $P < .001$, and $p < .001$, respectively). In addition, radiologists reported significantly reduced interpretation time using HM-MRI (73 ± 43 s) vs mpMRI (254 ± 133 s, $p = 0.03$) and greatly improved interobserver agreement: HM-MRI (Cronbach alpha = 0.88) vs mpMRI (0.26, $p = 0.04$).

In the second study comparing mpMRI and HM-MRI + mpMRI (19), the combination of HM-MRI and mpMRI improved specificity (for R3 and R4: 82%, 81% vs 77%, 71%; $p < 0.01$) and accuracy (48% vs 7%, $p < 0.01$) for less-experienced readers compared to conventional mpMRI alone, as well as improved overall inter-reader agreement (Fleiss Kappa = 0.36 for HM-MRI + mpMRI vs 0.17 for mpMRI, $p = 0.009$).

17.4.4 Prospective Clinical Trial

A recently completed prospective single center clinical trial (NCT03585660) evaluated the efficacy of HM-MRI in prospectively identifying areas for PCa biopsy in 91 patients. The findings indicated that HM-MRI was more reliable than random biopsy and/or targets detected by an expert radiologist based on mpMRI/PI-RADS in patients undergoing MR-US fusion biopsy (20).

The diagnostic performance of HM-MRI for diagnosing clinically significant cancers (≥ Gleason 3+4), was either higher than or on par with that of mpMRI. On a per-patient basis, HM-MRI had significantly higher accuracy (55% vs 44%, $p = 0.022$) and specificity (36% vs 14%, $p = 0.002$) than mpMRI. On a per-lesion basis, HM-MRI has significantly higher accuracy compared to mpMRI (58% vs 39%, $p < 0.001$) and PPV (31% vs 22%, $p = 0.003$). A representative example is shown in Figure 17.4. On a per-sextant basis, HM-MRI showed significantly better performance than mpMRI for all metrics, including primary endpoints, AUC (0.760 vs 0.647, $p < 0.001$) and accuracy (83.9% vs 79.0%, $p < 0.001$. This study underscores the potential of HM-MRI to improve MR-US fusion biopsy results by providing more accurate results compared to PIRADS based evaluation by expert radiologists. A multicenter trial (ClinicalTrials.gov ID—NCT05710380) validating the tool is currently underway.

17.5 Discussion

Hybrid multi-dimensional MRI is a new microstructural imaging technique that provides virtual pathology of the prostate. HM-MRI is able to provide noninvasive measurement of prostate tissue composition: fractional volume of stroma, epithelium, and lumen, where presence of cancers and increased aggressiveness/Gleason score are characterized by increased epithelium and reduced lumen (15). These results are similar to results from pathology (13, 14, 22). Prostate tissue composition measured noninvasively using HM-MRI have been validated against reference standards and show excellent correlation and agreement with quantitative histology results (16) and pathologists' measures (17) from whole mount prostatectomy.

HM-MRI has been validated clinically through observer studies and single center prospective trial where HM-MRI was used to guide prostate biopsies. Observer studies revealed improved PCa diagnosis using HM-MRI both as a standalone tool (18) and in combination with mpMRI (19) compared to the current standard—mpMRI/radiologists. HM-MRI was also shown to improved interobserver agreement and reduced interpretation time compared to mpMRI. The imaging technique was tested in a prospective clinical trial to guide and was shown to be was more reliable than random biopsy and biopsy targets chosen by expert radiologists based on mpMRI/PI-RADS in patients undergoing MR-US fusion biopsy (20).

Unlike conventional mpMRI, HM-MRI probes a much wider imaging parameter space by sampling at a series of b-values at multiple TE and is able to exploit the distinct MRI properties of prostate tissue components (23). The three-compartment signal model is able to exploit the unique information provided in this parameter and isolate signal from these tissue components. For example, at high b-signal, MR signal is primarily from cancer epithelial cells. Since we go to much lower TE than conventional DWI (T2 ~80 ms), HM-MRI has more signal contribution from cancers (pure cancerous epithelium has T2 ~50–60 ms) that conventional mpMRI doesn't account for (24). Whereas at high TE, signal comes predominantly from the luminal fluid that is then attenuated at higher b-values due to high diffusivity and T2 relaxation times, which we know from MRI microscopy (23) and clinical studies (25, 26), including those showing draining of prostatic fluid (27). HM-MRI models the signal change with the change in imaging parameters (b, TE) to give the best estimates of the fractional volume, as well as T2 relaxation times and diffusivity of these tissue components: stroma, epithelium, and lumen. While we can actually visualize prostate glandular microstructure using 16T preclinical MRI when imaged over 18 hours to get images at 40 micron resolution (12, 23), it is not feasible to image at such high resolution and for such a long time in a clinical setting. Therefore, these tissue composition measures are potentially the best true biomarkers we can derive from clinical MRI, and they relate very closely to pathology including cancer presence and unlike conventional parameters ADC and T2 that are dependent on imaging parameters used (28, 29), unlike HM-MRI.

There are other microstructural MR imaging techniques, such as luminal water imaging (LWI) (26); vascular, extracellular, and restricted diffusion for cytometry in tumors (VERDICT) (30); restriction spectrum imaging (RSI) (31); time dependent diffusion imaging (32); and diffusion-relaxation correlation spectrum imaging (33), which have been investigated for their applicability to prostate MR imaging. These methods have demonstrated superior performance compared to mpMRI in various studies. RSI, for instance, has undergone clinical evaluation in retrospective observer study and quantitative studies, while VERDICT was validated in the prospective INNOVATE study. It's noteworthy that the decision to conduct biopsies in the INNOVATE study was based on mpMRI Likert scores and PSAD and not on VERDICT MRI findings. However, to our knowledge, HM-MRI stands out as the only automated and quantitative imaging tool validated in a prospective clinical trial for guiding prostate biopsies and has been demonstrated to outperform mpMRI in terms of accuracy and efficacy.

HM-MRI has currently been validated in a prospective single center study. A multicenter clinical trial using HM-MRI to guide prostate biopsies (NCT05710380) is currently underway. The true potential of HM-MRI will be more accurately gauged through large-scale prospective multicenter studies across diverse clinical settings. Such studies are crucial for validating the effectiveness of these techniques in real-world scenarios, particularly in the general screening population. Further testing of these techniques on multiple MRI vendor platforms (34) with robust quality control measures (reproducibility, noise rejection, etc.) is essential to assess their generalizability and are currently undergoing. Additionally, more rigorous correlation between MRI findings and histological results may be vital in enhancing the accuracy of this method. The integration of new quantitative imaging and AI-based tools (35–37) using HM-MRI could potentially enhance diagnostic performance of HM-MRI even further.

HM-MRI has a few limitations. First, it does not account for the vascular component. While the vascular component contributes a very small fraction of the signal, it has been shown to differ between benign and cancer regions, which could potentially improve diagnosis (38). Second, HM-MRI does not account for exchange of water between the compartments/components. Studies have found that alteration of membrane permeability can significantly affect diffusion model parameters (39). Third, diffusion time for different b-values is kept the same, but the model could potentially be improved by exploiting the diffusion time dependence of the signal (40).

In summary, HM-MRI has been shown to improve radiologists' performance and holds promise for improving biopsy outcomes through more reliable detection of prostate cancer, surpassing the evaluation of mpMRI by expert radiologists using PIRADS guidelines. The advantage of HM-MRI is its ability to provide automated, quantitative image interpretation, standardized image acquisition, and reproducible outcomes, which could enhance prostate cancer diagnosis and potentially minimize unnecessary biopsies. The present findings, combined with ongoing advancements in HM-MRI technology, strongly advocate for its integration into routine clinical practice.

17.6 Future Trends

There is increasing interest in noninvasive techniques like HM-MRI to estimate prostate tissue composition to guide better PCa diagnosis. To our knowledge, HM-MRI is the only automated and quantitative tool imaging tool that has been validated in a prospective clinical trial to guide prostate biopsies and shown to have better performance than mpMRI. However, further research is necessary for the implementation of these methods in routine clinical practice. The true potential of HM-MRI will be more accurately gauged through large-scale prospective multicenter studies across diverse clinical settings. Such studies are crucial for validating the effectiveness of these techniques in real-world scenarios, particularly in screening populations.

Further testing of HM-MRI on multiple MRI vendor platforms with robust quality control measures (reproducibility, noise rejection, etc.) is essential to assess their generalizability. Additionally, more rigorous correlation between MRI findings and histological results is vital to enhance the accuracy of the method. The integration of radiomics and artificial intelligence with these MRI techniques could potentially enhance diagnostic precision further, reducing subjectivity and improving interpretation (37, 41, 42). However, more extensive validation of clinical data is needed in order to avoid misinterpretation by AI and its integration into routine clinical practice. Another promising direction is the exploration of feasibility of HM-MRI in low-field and low-gradient strength MRI systems (43), which could make this more accessible and cost-effective, particularly in resource-limited settings. Ensuring these MR methods are affordable and widely available is critical for broader implementation and benefitting a larger population. Most of the HM-MRI work discussed here was focused on diagnosis of cancer, but it can be used to predict and monitor response to therapy similar to conventional MRI (44, 45). In addition, implementation of similar methods on low field magnets can be leveraged for MR-guided radiotherapy and MR-guided in-bore biopsies (43, 46).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN Estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. 2021;71(3):209–49.
2. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*. 2024;74(1):12–49.
3. Turkbey B, Rosenkrantz AB, Haider MA, et al. Prostate imaging reporting and data system version 2.1: 2019 update of prostate imaging reporting and data system version 2. *European Urology*. 2019;76(3):340–51.
4. Kasivisvanathan V, Rannikko AS, Borghi M, et al. MRI-targeted or standard biopsy for prostate-cancer diagnosis. *New England Journal of Medicine*. 2018;378(19):1767–77.
5. Ahmed HU, El-Shater Bosaily A, Brown LC, et al. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. *The Lancet*. 2017;389(10071):815–22.
6. Mohammadian Bajgiran A, Afshari Mirak S, Shakeri S, et al. Characteristics of missed prostate cancer lesions on 3T multiparametric-MRI in 518 patients: based on PI-RADSv2 and using whole-mount histopathology reference. *Abdominal Radiology*. 2019;44(3):1052–61.
7. Niaf E, Lartizien C, Bratan F, et al. Prostate focal peripheral zone lesions: characterization at multiparametric MR imaging—influence of a computer-aided diagnosis system. *Radiology*. 2014;271(3):761–9.
8. Burdette JH, Elster AD, Ricci PE. Acute cerebral infarction: quantification of spin-density and T2 shine-through phenomena on diffusion-weighted MR images. *Radiology*. 1999;212(2):333–9.
9. Stanisz GJ, Henkelman RM. Diffusional anisotropy of T2 components in bovine optic nerve. *Magnetic Resonance in Medicine*. 1998;40(3):405–10.
10. Wang S, Peng Y, Medved M, et al. Hybrid multidimensional T2 and diffusion-weighted MRI for prostate cancer detection. *Journal of Magnetic Resonance Imaging*. 2014;39(4):781–8.
11. Sadinski M, Karczmar G, Peng Y, et al. Pilot study of the use of hybrid multidimensional T2-weighted imaging–DWI for the diagnosis of prostate cancer and evaluation of gleason score. *American Journal of Roentgenology*. 2016;207(3):592–8.
12. Bourne RM, Kurniawan N, Cowin G, et al. Microscopic diffusivity compartmentation in formalin-fixed prostate tissue. *Magnetic Resonance in Medicine*. 2012;68(2):614–20.
13. Langer DL, van der Kwast TH, Evans AJ, et al. Prostate tissue composition and MR measurements: investigating the relationships between ADC, T2, K(trans), v(e), and corresponding histologic features. *Radiology*. 2010;255(2):485–94.
14. Chatterjee A, Watson G, Myint E, Sved P, McEntee M, Bourne R. Changes in epithelium, stroma, and lumen space correlate more strongly with gleason pattern and are stronger predictors of prostate ADC changes than cellularity metrics. *Radiology*. 2015;277(3):751–62.
15. Chatterjee A, Bourne RM, Wang S, et al. Diagnosis of prostate cancer with noninvasive estimation of prostate tissue composition by using hybrid multidimensional MR imaging: a feasibility study. *Radiology*. 2018;287(3):864–73.
16. Chatterjee A, Mercado C, Bourne RM, et al. Validation of prostate tissue composition by using hybrid multidimensional MRI: correlation with histologic findings. *Radiology*. 2022;302(2):368–77.
17. Chatterjee A, Antic T, Gallan AJ, et al. Histological validation of prostate tissue composition measurement using hybrid multi-dimensional MRI: agreement with pathologists' measures. *Abdominal Radiology*. 2022;47(2):801–13.
18. Lee GH, Chatterjee A, Karademir I, et al. Comparing radiologist performance in diagnosing clinically significant prostate cancer with multiparametric versus hybrid multidimensional MRI. *Radiology*. 2022;305(2):399–407.
19. Lee G, Chatterjee A, Harmath C, et al. Improving reader accuracy and specificity with the addition of hybrid multidimensional-MRI to multiparametric-MRI in diagnosing clinically significant prostate cancers. *Abdominal Radiology*. 2023;48(10):3216–28.
20. Chatterjee A, Yousuf AN, Engelmann R, Harmath C, Lee G, Medved M, Jamison EB, Lorente Campos A, Gundogdu B, Gerber G, Reynolds LF, Modi PK, Antic T, Giurcanu M, Eggener S, Karczmar GS, Oto A. Prospective validation of an automated hybrid multidimensional MRI tool for prostate cancer detection using targeted biopsy: Comparison with PI-RADS-based assessment. *Radiology: Imaging Cancer*. 2025;7(1):e240156. doi: 10.1148/rycan.240156. PMID: 39836080; PMCID: PMC11791675.
21. Chatterjee A, Harmath C, Oto A. New prostate MRI techniques and sequences. *Abdominal Radiology*. 2020;45(12):4052–62.
22. Kobus T, Laak JAWMvd, Maas MC, et al. Contribution of histopathologic tissue composition to quantitative MR spectroscopy and diffusion-weighted imaging of the prostate. *Radiology*. 2015;278(3):801–11.

23. Bourne R, Kurniawan N, Cowin G, Sved P, Watson G. 16 T diffusion microimaging of fixed prostate tissue: preliminary findings. *Magnetic Resonanance in Medicine*. 2011;66(1):244–7.
24. Chatterjee A, Kim A, Gundogdu B, et al. Effect of echo times on prostate cancer detection on diffusion weighted images at high b-value and apparent diffusion coefficient maps. *ISMRM & ISMRT Annual Meeting & Exhibition*. Toronto, Canada, 2023.
25. Kjaer L, Thomsen C, Iversen P, Henriksen O. In vivo estimation of relaxation processes in benign hyperplasia and carcinoma of the prostate gland by magnetic resonance imaging. *Magnetic Resonance Imaging*. 1987;5(1):23–30.
26. Sabouri S, Chang SD, Savdie R, et al. Luminal water imaging: a new MR imaging T2 mapping technique for prostate cancer diagnosis. *Radiology.* 2017;284(2):451–9.
27. Medved M, Sammet S, Yousuf A, Oto A. MR imaging of the prostate and adjacent anatomic structures before, during, and after ejaculation: qualitative and quantitative evaluation. *Radiology.* 2014;271(2):452–60.
28. Peng Y, Jiang Y, Antic T, et al. Apparent diffusion coefficient for prostate cancer imaging: impact of B values. *American Journal of Roentgenology.* 2014;202(3):10917.
29. Gilani N, Rosenkrantz AB, Malcolm P, Johnson G. Minimization of errors in biexponential T2 measurements of the prostate. *Journal of Magnetic Resonance Imaging*. 2015;42(4):1072–7.
30. Panagiotaki E, Walker-Samuel S, Siow B, et al. Noninvasive quantification of solid tumor microstructure using VERDICT MRI. *Cancer Research*. 2014;74(7):1902–12.
31. McCammack KC, Schenker-Ahmed NM, White NS, et al. Restriction spectrum imaging improves MRI-based prostate cancer detection. *Abdominal Radiology.* 2016;41(5):946–53.
32. Lemberskiy G, Fieremans E, Veraart J, Deng F-M, Rosenkrantz AB, Novikov DS. Characterization of prostate microstructure using water diffusion and NMR relaxation. *Frontiers in Physics*. 2018;6(91).
33. Zhang Z, Wu HH, Priester A, et al. Prostate microstructure in prostate cancer using 3-T MRI with diffusion-relaxation correlation spectrum imaging: validation with whole-mount digital histopathology. *Radiology.* 2020;296(2):348–55.
34. Chatterjee A, Lee G, Dietz D, Oto A, Karczmar G. Cross vendor validation of Hybrid Multidimensional MRI in the non-invasive measurement of prostate tissue composition. *Society of Abdominal Imaging (SAR) Annual Scientific Meeting*. Maui, USA, 2020; p. 3285826.
35. Chatterjee A, Fan X, Oto A, Karczmar G. Four-quadrant vector mapping of hybrid multidimensional MRI data for the diagnosis of prostate cancer. *Medical Physics*. 2024;51(3):2057–65.
36. Enríquez-Mier-y-Terán FE, Chatterjee A, Antic T, Oto A, Karczmar G, Bourne R. Multi-model sequential analysis of MRI data for microstructure prediction in heterogeneous tissue. *Scientific Reports*. 2023;13(1):16486.
37. Gundogdu B, Chatterjee A, Kalidoss S, Karczmar GS, Oto A. Pathology without a knife: MRI-based non-invasive determination of prostate cancer gradewith physics-informed deep learning. *Proceedings of the International Society for Magnetic Resonance in Medicine,* Singapore, 2024.
38. Panagiotaki E, Chan RW, Dikaios N, et al. Microstructural characterization of normal and malignant human prostate tissue with vascular, extracellular, and restricted diffusion for cytometry in tumours magnetic resonance imaging. *Investigative Radiology.* 2015;50(4):218–27.
39. Bourne R, Panagiotaki E. Limitations and prospects for diffusion-weighted MRI of the prostate. *Diagnostics*. 2016; 6(2):21.
40. Wu D, Jiang K, Li H, et al. Time-dependent diffusion MRI for quantitative microstructural mapping of prostate cancer. *Radiology.* 2022;303(3):578–87.
41. Wei X, Zhu L, Zeng Y, et al. Detection of prostate cancer using diffusion-relaxation correlation spectrum imaging with support vector machine model—a feasibility study. *Cancer Imaging*. 2022;22(1):77.
42. McGarry SD, Bukowy JD, Iczkowski KA, et al. Gleason probability maps: a radiomics tool for mapping prostate cancer likelihood in MRI space. *Tomography.* 2019;5(1):127–34.
43. Tian Y, Nayak KS. New clinical opportunities of low-field MRI: heart, lung, body, and musculoskeletal. *Magnetic Resonance Materials in Physics, Biology and Medicine.* 2024;37(1):1–14.
44. Westin C, Chatterjee A, Ku E, et al. MRI findings following MRI guided focal laser ablation of prostate cancer. *American Journal of Roentgenology.* 2018;211(3):595–604.
45. Chatterjee A, Turchan WT, Fan X, et al. Can pre-treatment quantitative multi-parametric MRI predict the outcome of radiotherapy in patients with prostate cancer? *Academic Radiology.* 2022;29(7):977–85.
46. Satya P, Adams J, Jr, Venkataraman SS, et al. Office-based, single-sided, low-field MRI-guided prostate biopsy. *Cureus*. 2022;14(5).

18

Detection and Characterization of Brain Metastases Using Quantitative Chemical Exchange Saturation Transfer MR Fingerprinting (CEST-MRF)

Ouri Cohen

18.1 Introduction

18.1.1 Current Imaging Protocols Are Poor Predictors of Treatment Response

MRI is the current standard of care to evaluate patients with brain metastases (BM) undergoing treatment. Although it provides excellent soft tissue contrast and sensitivity, MRI lacks specificity and treatment decisions often rely on changes in size of the enhancing lesion. Decreasing size is an accepted surrogate for treatment response, but an increase in size may occur with both metastasis progression and treatment-induced inflammatory changes such as pseudo-progression [1, 2]. Augmenting the imaging protocol with specialized pulse sequences (e.g., diffusion, perfusion) can mitigate the issue at the cost of increased scan time and decreased patient comfort [3, 4], but lesion characterization in many common clinical scenarios remains ambiguous [5] (Figure 18.1). Size is also known to lag metabolic changes, such that some treated metastases may show decreased perfusion or PET uptake but remain stable in size for weeks or months before shrinking.

18.1.2 CEST MRI Is Highly Promising for Cancer Imaging

Chemical exchange saturation transfer (CEST) uses frequency selective radiofrequency (RF) pulses to saturate the magnetization of labile protons on proteins and metabolites [6]. The saturated protons exchange with the water protons and lead to a measurable reduction in the water MRI signal. Unlike MR spectroscopy, CEST is sensitive to metabolite concentrations at imaging resolutions (approximately 1 mm) in clinical scan times. Importantly, the measured CEST signal depends on the chemical exchange rate which is pH-sensitive and can be used to noninvasively map tumor acidosis [7]. Among others, the CEST contrast has been used to distinguish pseudo-progression and radiation necrosis from true progression in brain tumors [8–10], quantify tumor extracellular pH [11], and evaluate the grading and cellularity of gliomas [12].

18.1.3 Clinical Translation of CEST MRI Is Hampered by Technical Challenges

Although preclinical and early clinical CEST studies have demonstrated the potential utility of CEST methods for assessing disease pathologies, disease progression, and therapeutic response, they have not been widely adopted for clinical use due to several challenges encountered in clinical translation:

18.1.3.1 Qualitative CEST Contrast

In a traditional CEST experiment, the RF saturation pulse's frequency offset is stepped across the water resonance to create a CEST Z-spectrum displaying signal intensity relative to saturation frequency offset. The CEST contrast is then determined by comparing signal intensities obtained with negative and positive frequency offsets from water. However, CEST contrast is influenced by various factors such as chemical exchange rate, solute proton volume fraction, water longitudinal relaxation rate, RF saturation time, and RF saturation efficiency. Despite this, quantitative parameters are not typically measured in traditional CEST experiments, leading to the computation of a semi-quantitative CEST contrast. This lack of quantification makes it challenging to compare CEST contrast measurements across different MRI scanners or time points, as variations in field strength, inhomogeneity between scanners, and differences in RF coils can impact the results. Moreover, traditional CEST methods do not allow for the isolation of the contributions of chemical exchange rate, labile proton volume fraction, and water T1 and T2 to the CEST contrast. The simultaneous changes in pH, protein/metabolite concentrations, and water T1 and T2 within tumors add further complexity to accurately measuring pH or protein/metabolite concentration changes.

18.1.3.2 Long Image Acquisition Times

Acquiring a complete CEST Z-spectrum for a single image slice can be time-consuming, taking ten minutes or more. This is because many saturation frequency offsets are typically required, along with long repetition times, to allow

DOI: 10.1201/9781003449942-21

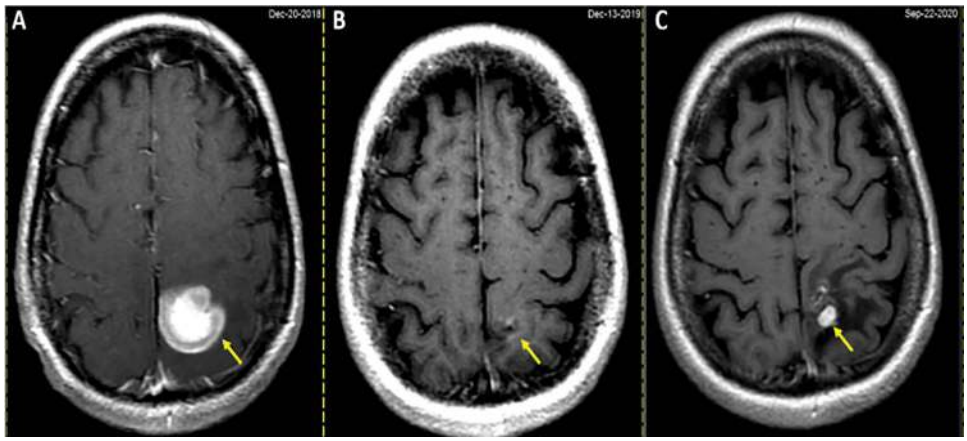


FIGURE 18.1 Example of a 52-year-old female with metastatic breast cancer. The left parietal lobe tumor shown in (A) was successfully treated with stereotactic radiosurgery and nearly disappeared (B) but subsequently resumed growing (C). It is unclear if the growth is due to a tumor or treatment-related changes.

for relaxation between image acquisitions. For multi-slice CEST sequences, acquisition times can be even longer, making it difficult to perform whole brain imaging for clinical applications. In addition to the lengthy Z-spectrum acquisition time, a separate B0 field map and T1 map must also be acquired to correct CEST Z-spectra for field inhomogeneities or T1 effects, adding to the overall scan time.

18.1.3.3 Complex Data Analysis

The analysis of the CEST contrast is further complicated by the presence of several exchangeable proton pools typically found *in vivo*. These include aliphatic protons with chemical shifts between −2 and −5 ppm from water, leading to nuclear Overhauser enhancement (NOE) effects. Additionally, there are amide exchangeable protons 3–4 ppm from water, amine exchangeable protons 2–3 ppm, hydroxyl exchangeable protons 1–2 ppm, and broad magnetization transfer (MT) effects caused by semi-solid, macromolecular exchangeable protons centered around −2.5 ppm. The sensitivity of the CEST contrast to B0 field inhomogeneity requires the acquisition of additional B0 field maps for correcting the Z-spectra [13]. Furthermore, T1 effects also impact the CEST contrast and need to be accounted for [14–16]. The data processing needed to perform these corrections adds further complexity to the clinical translation of the CEST contrast.

18.1.4 CEST Combined with Magnetic Resonance Fingerprinting Offers Fast Quantitative Imaging

This chapter presents a quantitative CEST pulse sequence that utilizes magnetic resonance fingerprinting (MRF) for rapid tissue parameter quantification [17]. Unlike traditional MRI techniques, MRF does not require the magnetization to reach a steady state before imaging, resulting in shorter scan times. By varying acquisition parameters such as flip angle and repetition time, MRF induces dynamic changes in the signal based on tissue properties like T1 and T2 relaxation. The extracted tissue parameters are obtained by comparing acquired signals to a precomputed dictionary or using neural networks like the deep reconstruction network (DRONE) [18]. Our approach combines CEST's sensitivity and specificity with MRF's quick quantitative mapping [19]. We demonstrate the effectiveness of the CEST-MRF method in both healthy individuals and those with brain metastases, showing its ability to differentiate between lesions invisible on conventional imaging.

18.2 Materials

18.2.1 MRI Studies

All experiments described in this chapter were conducted on a 3T GE Signa Premier MRI scanner (GE Healthcare, Waukesha, WI) equipped with a built-in transmit body coil and a head coil array with 48 elements for reception. The MRI data was processed using Python 3.11 and Matlab 2021b (The Mathworks, Natick, MA).

18.2.2 Neural Network Training

The neural networks, CEST-MRF simulations, and schedule optimization code were implemented in PyTorch 2.1.2 [20] using an Nvidia RTX2080 Ti graphics processing unit (GPU) (Nvidia, Santa Clara, CA, USA) with 11 GB of memory.

18.3 Methods

18.3.1 Pulse Sequence

The CEST-MRF pulse sequence depicted in Figure 18.2A illustrates a representative time-point. To saturate the solute protons exchanging with the bulk water protons, an

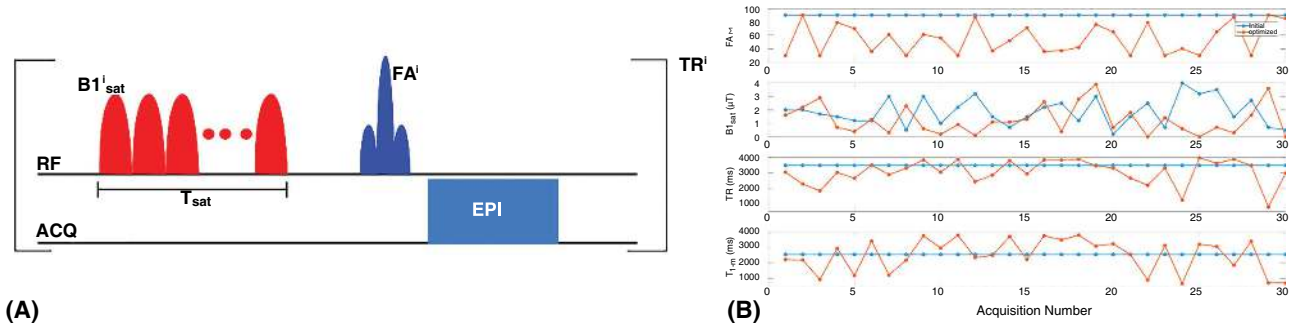


FIGURE 18.2 (A) Clinical CEST-MRF pulse sequence shown for one schedule point. The magnetization is saturated with a Gaussian-shaped pulse train and exchanges with the water. The water signal is then excited and read out with an EPI k-space sampling. (B) Original and optimized MRF schedules of scan parameters used. In the optimized schedule, the $B1_{sat}$, T_{sat}, FA, and TR were varied at each time point to induce dynamic variations in the signal whereas only $B1_{sat}$ was varied in the initial schedule.

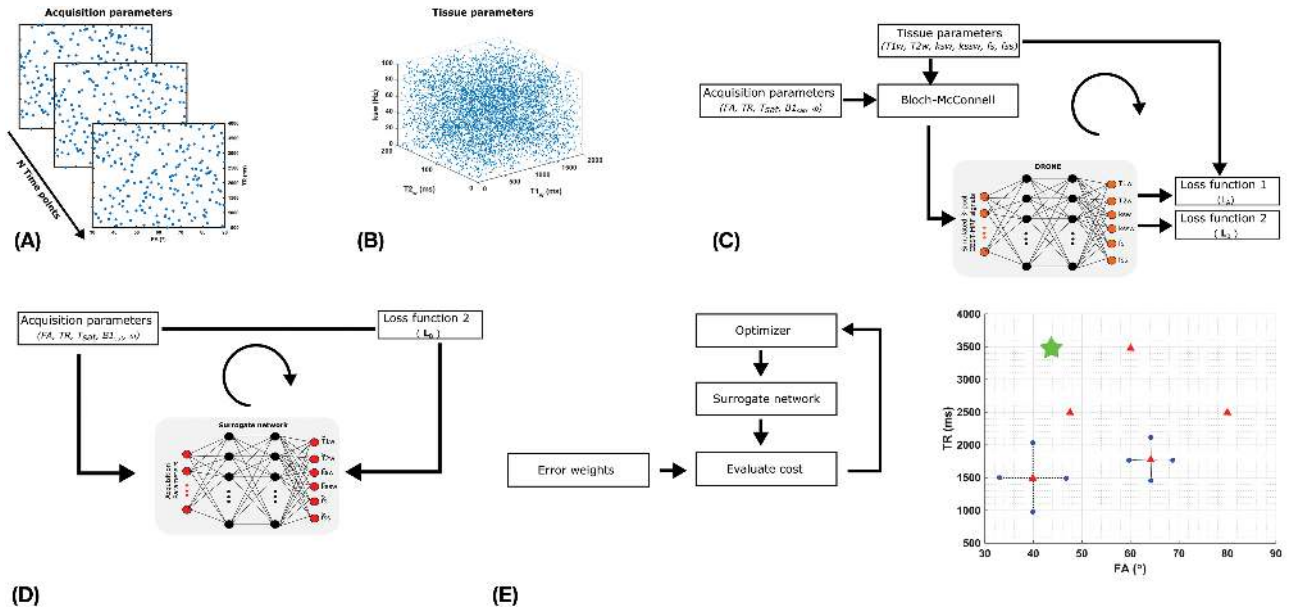


FIGURE 18.3 The deep learning approach to optimize the CEST-MRF acquisition schedule. (A) Sampling of the acquisition parameter space shown for two out of five parameters for visualization purposes. A set of 1000 acquisition schedules was generated by random sampling. (B) Sampling of the tissue parameter space shown for three out of seven parameters. (C) The 5000 tissue parameter samples were used to train a DRONE neural network to convergence for each acquisition schedule by back-propagation of the scalar loss L_A. (D) The surrogate network was used to map the acquisition schedules to the parameter-specific reconstruction errors (L_B) calculated by the DRONE network in (C). Following training, the surrogate network outputs the error in each tissue parameter (denoted with a tilde) given an input schedule. (E) Block diagram (left) and illustration (right) of the optimization process. A pattern-search optimizer is initialized with random candidate schedules (red triangles) and searches for the schedule with the lowest weighted error, as calculated by the surrogate network, by repeated testing of neighboring points (blue circles) using an adaptable grid to reach the global minimum (green star). (Reprinted under Creative Commons CC BY-NC-ND 3.0 from DOI: 10.1002/nbm.4954.)

initial pulse train consisting of non-selective 16 ms subpulses with Gaussian shapes was used. The signal reduction was assessed by exciting the magnetization using an on-resonance RF pulse with a variable flip angle (FA), followed by measurement of the signal using an echo-planar-imaging (EPI) readout. The magnetization was given time to recover during a repetition time (TR). The power of the pulse train ($B1_{sat}$) was adjusted in the initial schedule for each time point, while the optimized schedule, as shown in Figure 18.2B and detailed in Section 18.3.2, involved simultaneous variation of the pulse train duration (T_{sat}), FA, and TR as well as $B1_{sat}$.

18.3.2 Acquisition Schedule Optimization

The schedule of scan parameters was optimized using the deep-learning dual neural network approach shown in Figure 18.3. A randomly sampled set of scan parameters was used to

simulate CEST-MRF acquisitions for a fixed set of tissue parameters. The simulated measurements were reconstructed using a DRONE network and the error quantified. The set of schedules and corresponding reconstruction errors was used to train a surrogate network to enable rapid evaluation of candidate schedules. A pattern-search optimizer was used to explore the space of schedules and find the one the yielded the lowest reconstruction error.

18.3.3 Neural Network Definition and Training

The DRONE neural network used for quantification consisted of a fully connected network with a 30-node input layer, two 300-node hidden layers, and a 6-node output layer. A training dataset consisting of 5000 tissue parameter samples was generated. All signals were normalized to have a l2-norm of 1. A validation set was defined using 20% of the generated data to quantify the quality of the network training with the remainder used for training. The transmit field inhomogeneity (B1) was included in the training dataset generation but excluded from the loss calculation to induce the network to minimize the error in the other parameters instead. The networks were trained with the Adam optimizer [21] with a batch size of 1000 and an adaptive learning rate with weight decay of 10^{-4}. Gaussian noise (zero mean, 1% SD) was added during training to promote robust learning [22], and the network was trained for 1000 epochs. The surrogate network was defined with the same architecture but with a 150-node input layer corresponding to the five acquisition parameters for each of the 30 time points. The surrogate network was trained for 3000 epochs using a dataset comprising 1000 acquisition schedules.

18.3.4 *In Vivo* Human Subject Studies

18.3.4.1 Healthy Subjects

A healthy 34-year-old female volunteer was recruited for this study and provided informed consent in accordance with our institution's IRB protocol. The subject was scanned with the CEST-MRF pulse sequence using the optimized schedule, with the following image acquisition parameters: field of view=280×280mm^2, matrix size=192×192, in-plane resolution = 1.5 × 1.5 mm^2, slice thickness = 5 mm, echo time (TE) = 21 ms, partial Fourier, acceleration factor = 3, number of averages = 1, and bandwidth=250kHz. The measured data from each schedule were reconstructed with an appropriately trained DRONE network.

18.3.4.2 Patients with Brain Metastases

Five patients with brain metastases were recruited and gave informed Institutional Review Board consent. Four of the patients were scanned with the initial acquisition schedule and one with the optimized schedule. The patients were scanned per the consensus standardized brain tumor imaging protocol [23] and institutional standard-of-care that included T1-weighted sequences pre- and post-gadolinium (Gd) contrast injection, FLAIR, diffusion, and perfusion acquisitions. All patients were previously treated with radiation therapy and/or chemotherapy. The CEST-MRF sequence was acquired prior to Gd injection using the same imaging parameters and tissue quantification as for the healthy volunteer scans (Section 18.3.4.1) except for the resolution, which was set to 1 × 1 × 3 mm^3 to enable visualization of small lesions. The tumor was contoured into regions-of-interest (ROIs) comprising a necrotic core, solid core, and edema as well as a contra-lateral healthy region by a trained radiation oncologist. Statistical significance of differences in the reconstructed tissue map values for each ROI in a representative patient (patient 1) was evaluated using a multi-comparison analysis of variance test with Tukey honest significant difference [24] with a significance level set at $P = 0.05$.

18.4 Results

18.4.1 Healthy Subject

The CEST-MRF tissue parameter maps obtained from the healthy volunteer are shown in Figure 18.4. The six parameter maps were the water T1 and T2 relaxation maps (T1w, T2w), amide exchange rate (ksw) and volume fraction (fs) and semi-solid pool exchange rate (kssw) and volume fraction (fss). The CEST-MRF derived T1w and T2w maps were similar to those obtained with alternative methods [25]. The relatively constant ksw map is consistent with a constant pH in the healthy brain [26]. In contrast, the elevated fss in the white matter reflects the increased myelination in that tissue.

18.4.2 Patients with Brain Metastases

The CEST-MRF maps for a representative patient (patient 1) scanned with the initial protocol are shown in Figure 18.5 alongside corresponding images from a standard clinical protocol. Figure 18.6 displays a box and whiskers plot for all patients, revealing prominent trends in the parameter maps. In

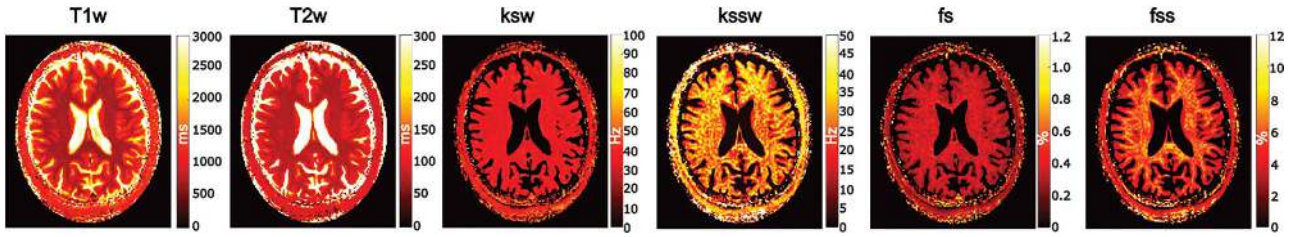


FIGURE 18.4 CEST-MRF tissue parameters maps obtained from a healthy subject using the optimized schedule. Shown are the water T1 and T2 relaxation maps (T1w, T2w), amide exchange rate (ksw) and volume fraction (fs) and semi-solid exchange rate (kssw) and volume fraction (fss).

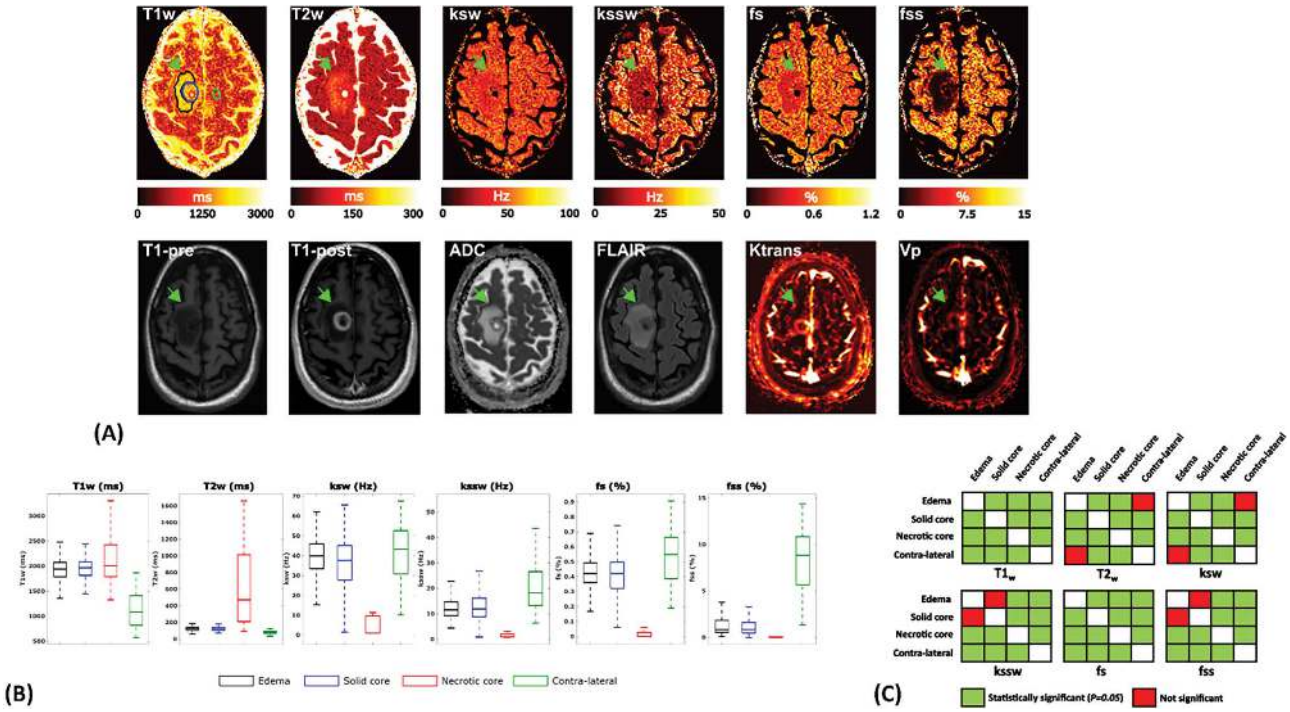


FIGURE 18.5 (A) *In vivo* CEST-MRF maps from a patient with brain metastasis and the corresponding images from a standard clinical protocol for comparison. Green arrows indicate the location of the lesion. The segmented tumor regions are denoted by the colored outlines on the T1w map and include the edema (black), solid core (blue), necrotic core (red), and contra-lateral (green) regions. The T1-pre and T1-post denote the T1-weighted acquisition before and after contrast injection, whereas Ktrans and Vp refer to the perfusion and plasma volume maps. The marked differences in the tissue map values between the lesion and healthy tissues are notable. (B) Box and whiskers plots of the reconstructed tissue maps values for the different ROIs. The distribution of the parameter values along with the median and the first and third quartile ranges are shown. (C) Graphical illustration of the statistical significance of the differences between the various tumor ROI pairs. All regions denoted in green were statistically significantly (P = 0.05) as determined by a multi-comparison analysis of variance test with Tukey honest significant difference. (Reprinted under Creative Commons CC BY-NC-ND 3.0 from DOI 10.1002/mrm.29448.)

the tumor, both the T1w and T2w values were elevated compared to the contra-lateral tissue, especially in the necrotic area. The amide exchange rate in the necrotic region was lower, indicating a potentially lower pH environment. Additionally, the amide volume fraction was reduced in both the necrotic and edema regions. Surprisingly, despite brain metastases from distinct primary tumors typically appearing similar on conventional MRI, noticeable disparities in the CEST-MRF tissue parameter maps were observed among different patients with different primary tumors, as shown in Figure 18.6. Furthermore, the CEST-MRF maps from the patient scanned using the optimized schedule (Figure 18.7) demonstrate heterogeneous increases in the middle frontal gyrus metastasis and decreases in the precentral gyrus metastasis, despite both showing similar enhancement and lower DCE-derived Ktrans and plasma volume.

18.5 Discussion

18.5.1 Advantages of CEST-MRF

The tissue maps provided by CEST-MRF are quantitative and accurately reflect underlying biophysical processes, in contrast to the qualitative signal obtained with traditional MRI scans. The various tissue maps offer a detailed view of complex pathologies like cancer, where tissue parameters can change simultaneously. The additional diagnostic information from the CEST-MRF maps can lead to more personalized and effective treatments. Furthermore, the quick acquisition time of less than two minutes makes it easier for young or frail patients to undergo scans, while also reducing costs. Notably, the neural network reconstruction takes less than a second. Brain metastases from different primary tumors appear distinct on CEST-MRF maps (Figure 18.6), but are often indistinguishable on conventional MRI scans. The enhanced specificity of the CEST-MRF maps improves lesion detection compared to standard imaging methods (Figure 18.7), suggesting a potential role for CEST-MRF in monitoring treatment responses.

18.5.2 Limitations

The current CEST-MRF method implementation utilizes EPI k-space sampling to minimize scan time and enhance the signal-to-noise ratio. However, EPI sampling comes with known limitations such as increased sensitivity to phase dispersion leading to geometric distortions and restrictions on the resolution (due to the resultant longer echo), affecting lesion visibility. As a result, the current method is not ideal for body imaging

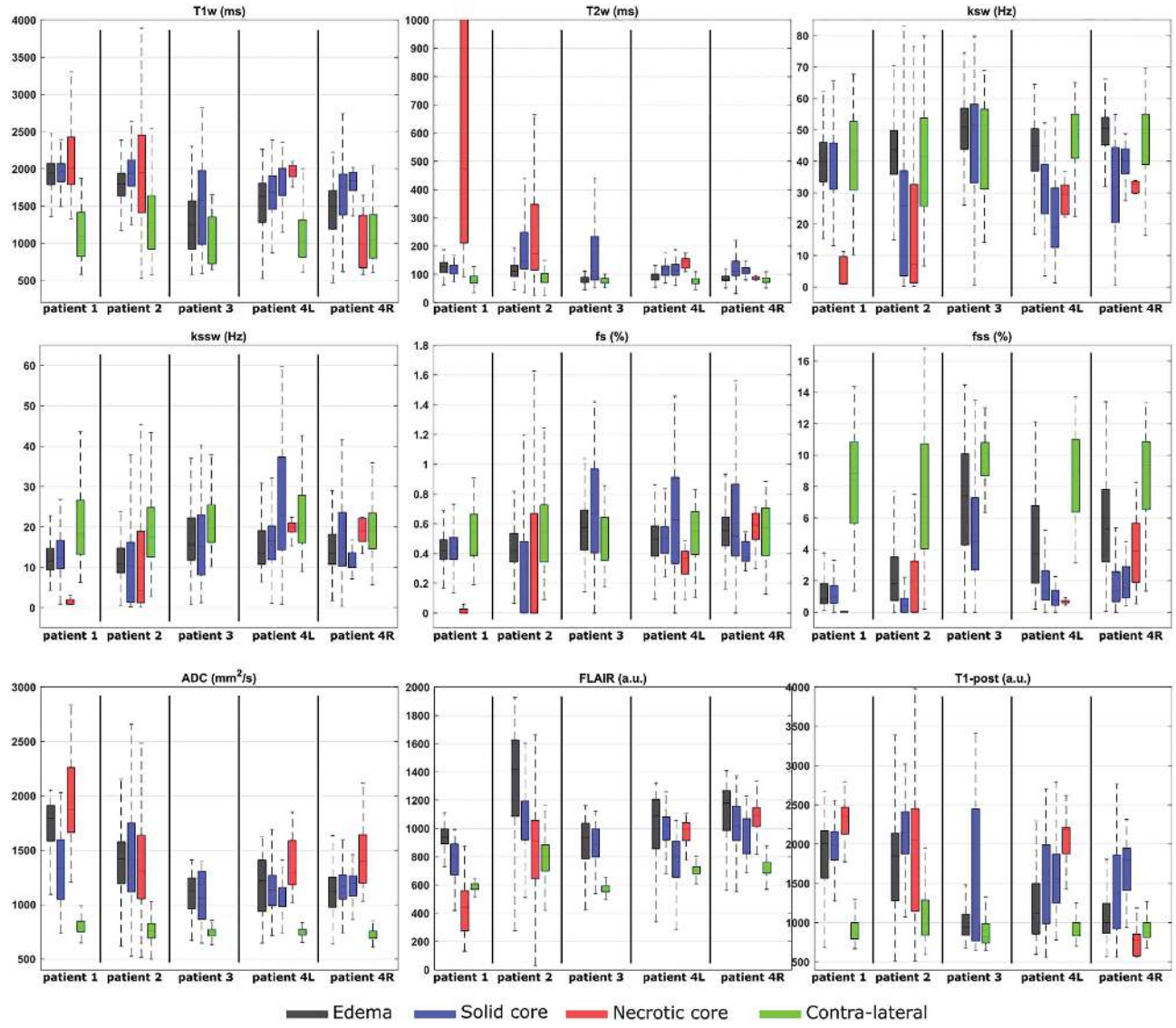


FIGURE 18.6 Box and whiskers plots for CEST-MRF-derived T1, T2, ksw, kssw, fs, and fss as well as conventional sequence values from T1 post Gd contrast, ADC, and FLAIR images. Patient #1: 61-y-old male, non-small-cell lung cancer adenocarcinoma; Patient #2: 65-y-old female, melanoma; Patient #3: 48-y-old female non-small-cell lung cancer adenocarcinoma/small cell; Patient #4: 40-y-old male, melanoma (L: left side with two tumors and surrounding region. R: right side with two tumors and surrounding regions). (Reprinted under Creative Commons CC BY-NC-ND 3.0 from DOI 10.1002/mrm.29448.)

due to significant respiratory and motion-induced phase shifts. We aim to investigate alternative k-space sampling techniques, such as radial sampling [27], to address these challenges and facilitate quick CEST-MRF imaging in body scans.

18.6 Future Trends

Most studies on CEST-MRF have traditionally focused on detecting amide protons, which are abundant *in vivo*, have slow exchange rates, and are sensitive to pH, making them a valuable cancer biomarker [19, 28–31]. However, CEST-MRF can also be used to identify other solutes by adjusting the resonance frequency of the saturation pulse to analyze protons from hydroxyl or amine functional groups. Amine protons, which have lower concentrations and rapid exchange rates, pose a challenge but are important in cancer imaging due to tumors' heightened amino-acid metabolism [32]. Future research efforts to develop a rapid CEST-MRF scan protocol for analyzing various solutes show promise in enhancing the technique's diagnostic capabilities.

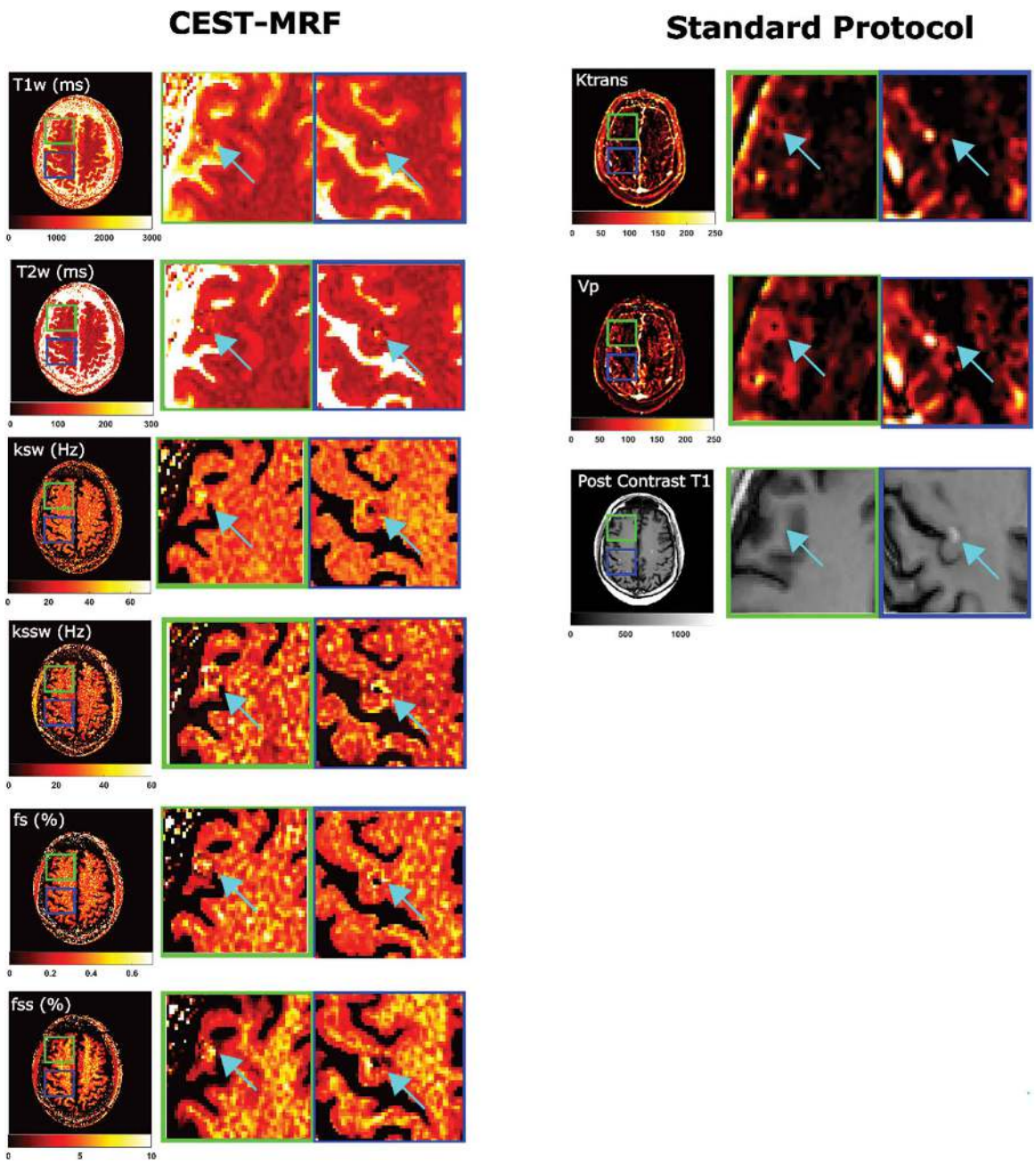


FIGURE 18.7 Patient who has metastatic RCC and irradiated enhancing metastases in the right middle frontal gyrus (green boxes) and right precentral gyrus (blue boxes). The CEST-MRF derived water relaxation (T1w, T2w), amide pool exchange rate (ksw), semi-solid pool exchange rate (kssw), amide pool volume fraction (fs), and semi-solid pool volume fraction (fss) maps show heterogeneous increases in the middle frontal gyrus metastasis and decreases in the precentral gyrus metastasis, although both show similar enhancement and decreased DCE-derived Ktrans and plasma volume.

REFERENCES

1. K. Parvez, A. Parvez, and G. Zadeh, "The diagnosis and treatment of pseudoprogression, radiation necrosis and brain tumor recurrence," *International Journal of Molecular Sciences*, vol. 15, no. 7, pp. 11832–11846, 2014.
2. J. V. Cohen *et al.*, "Melanoma brain metastasis pseudoprogression after pembrolizumab treatment," *Cancer Immunology Research*, vol. 4, no. 3, pp. 179–182, 2016.
3. M.-A. Weber *et al.*, "Assessment of irradiated brain metastases by means of arterial spin-labeling and dynamic susceptibility-weighted contrast-enhanced perfusion MRI: initial results," *Investigative Radiology*, vol. 39, no. 5, pp. 277–287, 2004.
4. K. R. Fink and J. R. Fink, "Imaging of brain metastases," *Surgical Neurology International*, vol. 4, no. Suppl 4, p. S209, 2013.

5. E. L. Chang *et al.*, "The role of tumor size in the radiosurgical management of patients with ambiguous brain metastases," *Neurosurgery*, vol. 53, no. 2, pp. 272–281, 2003.
6. B. Wu *et al.*, "An overview of CEST MRI for non-MR physicists," *EJNMMI Physics*, vol. 3, no. 1, p. 19, 2016.
7. D. L. Longo *et al.*, "In vivo imaging of tumor metabolism and acidosis by combining PET and MRI-CEST pH imaging," *Cancer Research*, vol. 76, no. 22, pp. 6463–6470, 2016.
8. B. Ma *et al.*, "Applying amide proton transfer-weighted MRI to distinguish pseudoprogression from true progression in malignant gliomas," *Journal of Magnetic Resonance Imaging*, vol. 44, no. 2, pp. 456–462, 2016.
9. J. Zhou *et al.*, "Differentiation between glioma and radiation necrosis using molecular magnetic resonance imaging of endogenous proteins and peptides," *Nature Medicine*, vol. 17, no. 1, pp. 130–134, 2011.
10. H. Mehrabian, K. L. Desmond, H. Soliman, A. Sahgal, and G. J. Stanisz, "Differentiation between radiation necrosis and tumor progression using chemical exchange saturation transfer," *Clinical Cancer Research*, vol. 23, no. 14, pp. 3667–3675, 2017.
11. E. A. Randtke, J. C. Granados, C. M. Howison, M. D. Pagel, and J. Cárdenas-Rodríguez, "Multislice CEST MRI improves the spatial assessment of tumor pH," *Magnetic Resonance in Medicine*, vol. 78, no. 1, pp. 97–106, 2017.
12. Y. Bai *et al.*, "Noninvasive amide proton transfer magnetic resonance imaging in evaluating the grading and cellularity of gliomas," *Oncotarget*, vol. 8, no. 4, p. 5834, 2017.
13. M. Kim, J. Gillen, B. A. Landman, J. Zhou, and P. C. van Zijl, "Water saturation shift referencing (WASSR) for chemical exchange saturation transfer (CEST) experiments," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 61, no. 6, pp. 1441–1450, 2009.
14. J. Xu *et al.*, "On the origins of chemical exchange saturation transfer (CEST) contrast in tumors at 9.4 T," *NMR in Biomedicine*, vol. 27, no. 4, pp. 406–416, 2014.
15. P. Z. Sun, Y. Murata, J. Lu, X. Wang, E. H. Lo, and A. G. Sorensen, "Relaxation-compensated fast multislice amide proton transfer (APT) imaging of acute ischemic stroke," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 59, no. 5, pp. 1175–1182, 2008.
16. M. Zaiss *et al.*, "Relaxation-compensated CEST-MRI of the human brain at 7 T: unbiased insight into NOE and amide signal changes in human glioblastoma," *Neuroimage*, vol. 112, pp. 180–188, 2015.
17. D. Ma *et al.*, "Magnetic resonance fingerprinting," *Nature*, vol. 495, no. 7440, pp. 187–192, 2013.
18. O. Cohen, B. Zhu, and M. S. Rosen, "MR fingerprinting deep reconstruction network (DRONE)," *Magnetic Resonance in Medicine*, vol. 80, no. 3, pp. 885–894, 2018.
19. O. Cohen, S. Huang, M. T. McMahon, M. S. Rosen, and C. T. Farrar, "Rapid and quantitative chemical exchange saturation transfer (CEST) imaging with magnetic resonance fingerprinting (MRF)," *Magnetic Resonance in Medicine*, vol. 80, no. 6, pp. 2449–2463, 2018.
20. A. Paszke *et al.*, "Pytorch: an imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, vol. 32, pp. 8026–8037, 2019.
21. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proceedings of International Conference on Learning Representations*, San Diego, CA, USA, May 2015.
22. R. M. Zur, Y. Jiang, L. L. Pesce, and K. Drukker, "Noise injection for training artificial neural networks: a comparison with weight decay and early stopping," *Medical Physics*, vol. 36, no. 10, pp. 4810–4818, 2009.
23. B. M. Ellingson *et al.*, "Consensus recommendations for a standardized brain tumor imaging protocol in clinical trials," *Neuro-oncology*, vol. 17, no. 9, pp. 1188–1198, 2015.
24. J. W. Tukey, "Comparing individual means in the analysis of variance," *Biometrics*, pp. 99–114, 1949.
25. J. Z. Bojorquez, S. Bricq, C. Acquitter, F. Brunotte, P. M. Walker, and A. Lalande, "What are normal relaxation times of tissues at 3 T?," *Magnetic Resonance Imaging*, vol. 35, pp. 69–80, 2017.
26. T. Cadoux-Hudson, B. Rajagopalan, J. Ledingham, and G. Radda, "Rapid communication: response of the human brain to a hypercapnic acid load in vivo," *Clinical Science*, vol. 79, no. 1, pp. 1–3, 1990.
27. O. Cohen and R. Otazo, "Golden-angle radial CEST MR fingerprinting with temporal compressed sensing reconstruction," in *Proceedings of the International Society of Magnetic Resonance in Medicine, Virtual Conference*, May 2021, p. 1486.
28. O. Perlman, K. Herz, M. Zaiss, O. Cohen, M. S. Rosen, and C. T. Farrar, "CEST MR-fingerprinting: practical considerations and insights for acquisition schedule design and improved reconstruction," *Magnetic Resonance in Medicine*, vol. 83, no. 2, pp. 462–478, 2020.
29. O. Perlman *et al.*, "Quantitative imaging of apoptosis following oncolytic virotherapy by magnetic resonance fingerprinting aided by deep learning," *Nature Biomedical Engineering*, vol. 6, no. 5, pp. 648–657, 2022.
30. O. Cohen and R. Otazo, "Global deep learning optimization of CEST MR fingerprinting (CEST-MRF) acquisition schedule," *NMR in Biomedicine*, p. e4954, 2023.
31. O. Cohen *et al.*, "CEST MR fingerprinting (CEST-MRF) for brain tumor quantification using EPI readout and deep learning reconstruction," *Magnetic Resonance in Medicine*, vol. 89, no. 1, pp. 233–249, 2023.
32. N. S. Cho *et al.*, "Amine-weighted chemical exchange saturation transfer magnetic resonance imaging in brain tumors," *NMR in Biomedicine*, vol. 36, no. 6, p. e4785, 2023.

19

Narrow-Beam CT: A Solution to the Limitations in Breast Cancer Screening

Peymon Ghazi

19.1 Introduction

An estimated 40% to 50% of the female population in the United States has breasts with large volumes of mammographically "dense" tissue.[1,2] When these women's breasts are imaged with the current standard of care for breast cancer screening—full field digital mammography (FFDM) or digital breast tomosynthesis (DBT, often called "3D mammography")—their dense breast tissue appears opaque in mammographic images. As a result, potentially cancerous lesions can be hidden or masked from view.[3,4] While the majority of women (50%–60%) may be well served by the standard of care (routine screening mammography), women with dense breasts are at risk for potential cancers going undetected. Exacerbating the problem, breast density is itself an independent elevated risk factor for breast cancer.[5,6] The need for expanded care for women with dense breasts has become such a well-known issue that the US Food and Drug Administration recently announced a mandate requiring clinicians to notify women who have dense breasts of the implications for their health.[7] FFDM and DBT are the only standalone imaging modalities currently available to the general population for breast cancer screening despite their poor sensitivity in dense breast tissue imaging.[3,8] While expanded access to screening with these modalities has saved countless lives, the need for an imaging solution better suited for dense breasts remains.

Aside from these pervasive clinical needs, breast compression, an inherent aspect of mammographic imaging, remains a problem. Many lives have been saved due to increasingly widespread access to and utilization of mammographic screening.[9] However, pre-COVID pandemic figures indicate that 23% of the eligible female population failed to obtain screening.[10] While reasons for screening avoidance are varied and often multifaceted, a common complaint from patients is regarding the pain or discomfort that can arise from breast compression. Indeed, one study that queried specific causes for screening avoidance found that 20% of women refused free screening due to fear of procedure-induced pain.[10,11] A 2013 "failure analysis" study in the Harvard University showed that 71% of women who died from breast cancer were among the 23% who were not participating in screening.[12] Detecting early-stage cancer via screening is the most effective approach for mortality reduction.[13] Providing women with an imaging solution that does not require painful compression could be a critical factor in increasing the utilization of this form of care. A comfortable imaging solution may aid in facilitating earlier detection of cancer in one third of women whose cancers are not diagnosed until they have metastasized beyond the breast.[14]

These limitations are not limited to breast cancer screening. The current paradigm in breast cancer imaging requires long diagnostic paths. Women with dense breasts are up to five times more likely to be recalled for diagnostic exams.[15] Patients often undergo multiple rounds of image, and often repeated rounds using the same imaging modality, on their path to a breast cancer diagnosis.[16] This places more than 30% of women at risk for "financial toxicity," a state in which expenses related to care can become financially catastrophic for their lives.[17]

In current breast cancer radiology, there is no alternative standalone solution to screening mammography. Breast ultrasound, due to its inherent limitations, is used only as a supplement to mammography. Breast MRI is an effective diagnostic tool, but the cost of acquiring and maintaining the technology makes it an unrealistic solution for general breast cancer screening applications. Dedicated breast CT has been developed as a potential solution. The commercialized solutions, though, have been based on traditional renditions of CT: cone-beam CT or spiral CT. As a tomographic (three-dimensional) imaging modality, CT is more effective for dense breast imaging. However, known issues such as (1) unreliable tissue coverage, (2) poor image quality, and (3) high radiation dose have prevented CT from being a viable option for solving this problem.

The technology outlined in this chapter is a new CT image acquisition methodology, with design specificities that make this imaging modality a potential solution to the need for an effective standalone breast cancer imaging detection and diagnosis.

19.2 Methods

19.2.1 Scientific Premise

In x-ray computed tomography (CT), the primary type of photon interaction is Compton scattering.[18] Consequently, a CT image is essentially a map of scatter occurrences, specifically the first scatter events. While scattered photons can undergo secondary

DOI: 10.1201/9781003449942-22

scatter events, these do not contribute to the beam's attenuation profile captured by the imager. Due to the random nature of scatter events, it is impossible to determine the exact location of a scatter event for an acquired scattered photon. Thus, the obtained projections should ideally be free of scattered x-rays. However, in current CT implementations, the raw data includes both scattered and primary photons; however, due to stochastic nature of scatter events, it is practically impossible to deterministically differentiate between the primary and scattered photons.

Lemma—Acquiring x-ray scatter reduces image quality across the entire spatial frequency response of the imaging system.

Proof—For any given sinogram that results from a CT scan, the received signal (I) in the x-ray detector can be modeled as:[19]

$$I(X) = N_o \int (\Phi(E) e^{-p_{(\mu)}} [1 + SPR])\, dE \tag{19.1}$$

where Φ is the x-ray fluence of an x-ray beam with energy E, $P(\mu)$ is the compound attenuation property μ between the x-ray source and the sinogram location X, and SPR is the scatter-to-primary ratio. According to Equation 19.1, if we perform a scan in the absence of an object, the resulting un-attenuated and un-scattered signal is:

$$I_0(X) = N_0 \int \Phi(E) dE \tag{19.2}$$

Carrying out the Beer-Lambert logarithmic subtraction results in:

$$\begin{aligned} \mathrm{In}\left(\frac{I_0(X)}{I(X)}\right) &= \mathrm{In}\frac{e^{p_{(\mu)}}}{\int [1 + SPR] dE} \\ &= p(\mu) - In \int [1 + SPR] dE \end{aligned} \tag{19.3}$$

Simply put, Equation 19.3 states that a scattered signal results in a constant term (the integral containing SPR) being added to the term representing the material attenuation properties and thickness ($P(\mu)$). In other words, the I_0-logged results of an x-ray transmission are no longer a linear operation. This violates the fundamental assumption of all image reconstruction algorithms: that the I_0-logged operation should lead to a linear term solely dependent on the attenuation properties and thickness of the object.

We can study the impact of this nonlinearity. Assume that an object in the x-ray path is a circular object with uniform attenuation of μ and radius of R. The parallel beam projection[18] of this object is $P(\mu) = 2\mu\sqrt{R^2 - X^2}$. Deriving the Taylor expansion of Equation 19.3 for the logarithm containing the $[1 + SPR]$ element yields:

$$In\left(\frac{I_0(X)}{I(X)}\right) = 2\mu\sqrt{R^2 - X^2} - \mathcal{F}(SPR), \tag{19.4}$$

Equation 19.4 defines $\mathcal{F}$ as a function that includes the SPR-dependent terms. According to this equation, if R is significantly larger than X (in other words, when a large object is imaged), or μ is large (i.e., imaging a highly attenuating object), then the relative magnitude of the μ-independent part (F) is small, and the impact of scatter acquisition is reduced. However, when imaging a small and/or low-attenuating object, the SPR-containing term (F) weakens the received signal. As a result, the contrast resolution of the imaging system throughout its frequency response is affected by the acquisition of scattered x-rays. Furthermore, and more directly relevant to breast cancer imaging, the impact of scatter dilution is exacerbated in low-dose (high-noise) imaging protocols, such as breast cancer screening.

The effects of scatter in CT imaging can be reduced through various methods, including the following:

- Anti-scatter grids: Placing an anti-scatter grid between the patient and the detector can reduce the amount of scattered radiation reaching the detector, thus improving image quality. However, this method increases patient dose and can produce grid artifacts in the image.
- Increased x-ray exposure: Increasing the number of x-rays directed at the patient can increase the number of primary photons detected, reducing the relative contribution of scattered photons to the output CT image. However, this also increases patient dose and may not be feasible in all cases.
- Scatter correction algorithms: Applying scatter correction algorithms during post-processing can estimate and subtract the scattered radiation from the acquired data, improving image quality. However, these algorithms can be computationally intensive and may introduce artifacts if not applied correctly. More importantly, as proven in the lemma, scatter impacts the image throughout the frequency response of the imaging system. These algorithms can only address the low-frequency (slow-moving) impacts. But, due to the stochastic nature of the scatter events, the high-frequency impacts, such as diminished visibility of small details, are often not correctable. Therefore, there is always a limit to the effectiveness of these postprocessing algorithms.

In short, each of these methods has its drawbacks. The choice of method depends on the specific imaging requirements and the tradeoffs between image quality and other factors such as dose and the effectiveness of the employed method. The most effective approach to manage the negative consequences of acquiring scatter in CT is to prevent scattered photons from reaching the detector during the scan. We hypothesized that by blocking scattered photons from reaching the x-ray detector, we could enhance the visibility of small and low-contrast details in breast tissue, which are often crucial for diagnosis. Our initial simulation studies supported this

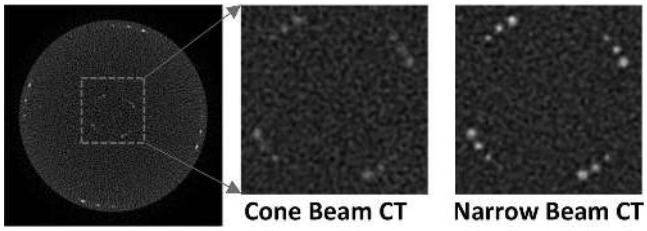


FIGURE 19.1 Qualitative comparison between cone-beam CT (center) and narrow-beam CT (right).[20] We simulated a CT scan through Monte Carlo simulations in two scan modes. Both simulations were done in iso-dose and iso-spatial-resolution conditions. The difference in the visibility of small details is attributed to the difference in acquisition of scattered x-ray photons. In narrow-beam CT, scatter acquisition is minimized to near-zero levels.

hypothesis (refer to Figure 19.1). In the following sections, we will describe the technology that leads to near scatter-free raw data.

19.2.2 Image Acquisition Methodology

The prior lemma and simulation results describe a strong causal relationship between scatter buildup in the sinogram domain and the impairment of visibility of small details in the imaging domain. Here, we introduce narrow-beam CT (NB-CT), which was developed to resolve this problem. NB-CT is designed to avoid the accumulation of x-ray scatter in raw data, leading to near scatter-free imaging. Importantly, the benefits of this approach are highlighted as being independent of, and unrelated to, component hardware selections.

During the initial development of the technology, we discovered a key design choice that results in a spatial modulation of x-ray flux during image acquisition based on the patient's specific breast size. This means that different parts of the breast are exposed to varying numbers of x-rays. This feature is crucial for making CT a viable solution for breast screening applications, as it minimizes dose exposure while maintaining imaging quality. Our simulation studies validated the hypotheses that with this design, the visibility of small details is improved and total dose exposure is reduced. Additionally, this approach increases the coverage of the posterior breast anatomy. These findings not only justified and motivated our technology but also demonstrated its necessity.

NB-CT is centered on a "narrow-beam" image acquisition architecture, which is fundamentally distinct from existing cone-beam or spiral CT design architectures. In Figure 19.2, a typical NB-CT imaging geometry is illustrated. The source, comprising the x-ray tube and a rotational collimator, and the detector are mounted on two mechanically decoupled robotic arms, rotating independently yet synchronized with one another. Assuming a hypothetical 100% quantum detection efficiency, the positioning and moving trajectories of the imaging system are meticulously planned so that all primary transmitted x-ray photons within a narrow beam are captured by the detector panel. In NB-CT, the collimator's opening moves with the same rotational speed and phase as the detector panel (throughout this chapter, the terms "collimator" and "fluence modulator" are used interchangeably).

In NB-CT, the exposure mode used is pulsed fluoroscopy, as illustrated in Figure 19.3. As the x-ray source rotates along its trajectory, multiple exposures are made. Each exposure begins when a clear line of sight is established between the x-ray source, the collimator's opening, the object, and the x-ray detector. The detector's motion during an exposure can cause motion blurring artifacts in the resulting images. To reduce this, the detector operates in time-delay-integration (TDI) mode. In TDI mode, the charge collected in one line of the detector is passed to an adjacent line and added to the charge being collected in that new line. This charge transfer and accumulation process continues across all the lines within the detector's active area. Once the charge has been transferred through all the lines, an output line is formed and transferred to the processor.

19.3 Materials

A fully functional prototype NB-CT system was developed to validate the hypotheses of the study. In NB-CT, a cadmium telluride photon-counting line detector (from Varex Imaging, Salt Lake City, UT) is used instead of a flat-panel detector to further reduce image contamination from scatter to near-zero. This detector selection not only eliminates scatter contamination but also reduces electronic noise, resulting in additional improvements in image quality over those achieved with NB-CT geometry and acquisition methodology alone.

As described earlier, in NB-CT, an important operational aspect is the use of TDI. In TDI mode, the speed of line acquisition can vary based on the imaging system and settings. The concept of TDI is to enable continuous image acquisition along the direction of motion (usually the line-scan direction) without introducing motion blur. The speed of line acquisition is determined by the rate at which charge is transferred from one pixel to the next in the detector.

In industrial applications using TDI, this transfer rate is synchronized with the movement of the object being scanned. This ensures that each line of the image is formed from a continuous integration of signal as the object moves past the detector. However, in medical imaging such as NB-CT where the object of interest (e.g., the patient) remains stationary, a different approach is needed to avoid blurring.

To prevent blurring, the detector in NB-CT moves in the opposite direction to the TDI direction. The speed of the detector's movement must match the speed of TDI charge handoff. A balance must be struck between the TDI speed and the sampling of space. The TDI speed is typically high enough to ensure that there are enough lines of data acquired per unit time to form a complete image without gaps or blurring. The exact speed can vary based on factors like detector resolution, object or detector movement speed, and desired image quality.

In this prototype, we used a pulsed fluoroscopy x-ray source (M-1500, Varex Imaging, Salt Lake City, UT) operated at 60 kVp, with an additional 0.2 mm of copper filtration. To ensure complete imaging of the field of view in each exposure, we calculated the beam ON time. For example, with a detector

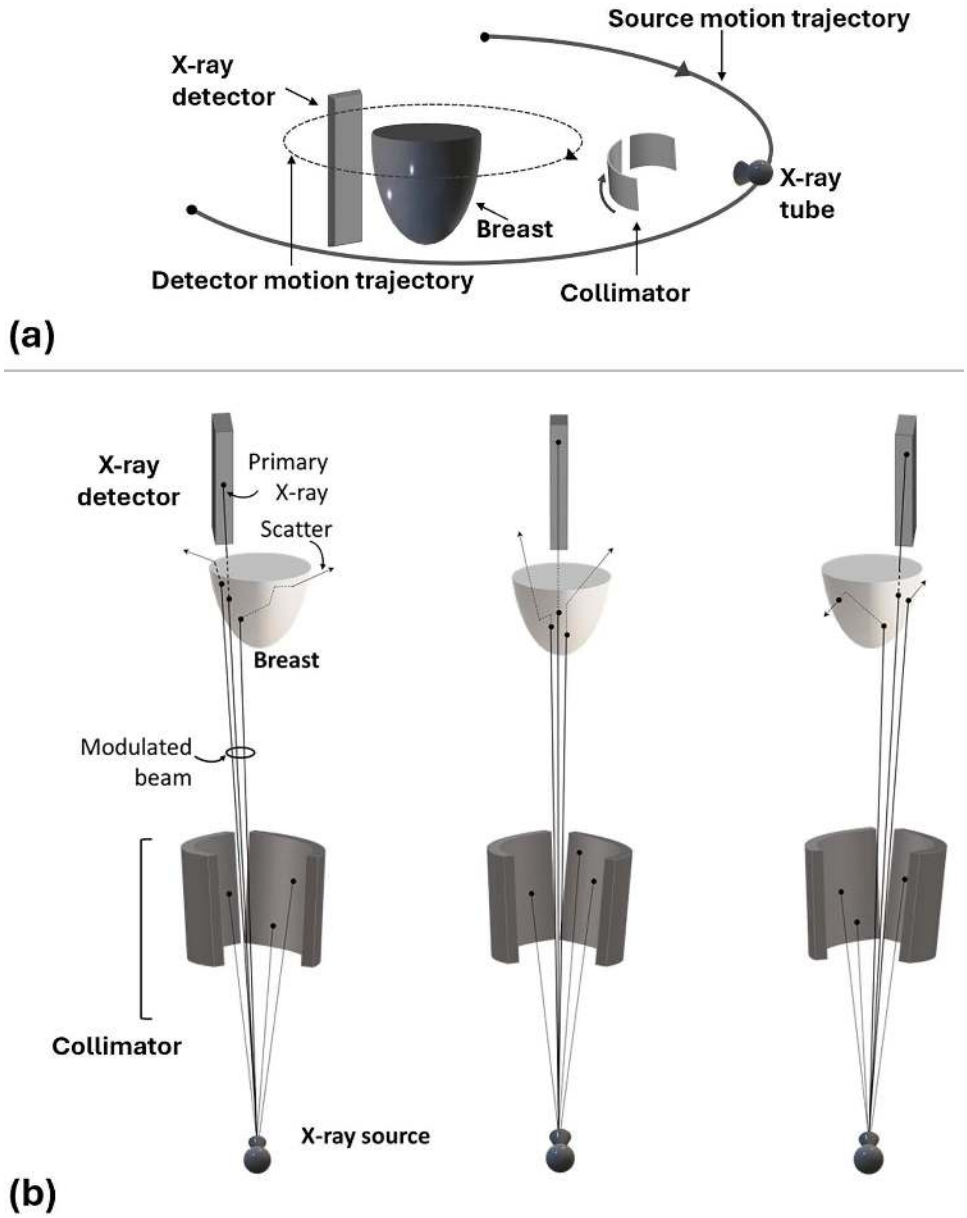


FIGURE 19.2 Narrow-beam apparatus design and method of imaging. (a) A prospective view of the system geometry along with the x-ray generation and detection assemblies' trajectories around the breast, and (b) positioning of the main components of the imaging system at three instances of time.

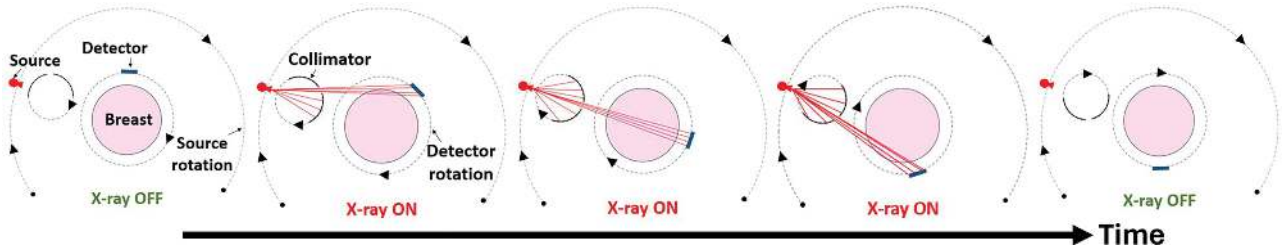


FIGURE 19.3 Narrow-beam breast CT sequencing of image acquisition. As the x-ray source transitions along a path, the breast is measured from different view angles. In each view angle, the detector makes one full rotation. Therefore, in a scan, the x-ray source rotates only once around the breast, but the detector rotates many times. Exposure is in pulsed fluoroscopic mode, therefore, there is no wasted dose.

rotation speed of one rotation per second, we determined that the beam must be turned on for 300 milliseconds to capture the entire field of view.

19.4 Results

A photograph of the developed prototype is shown in Figure 19.4.

19.4.1 Scatter Buildup

Our proof-of-concept studies unequivocally show that NB-CT image acquisition results in nearly scatter-free CT raw data. We conducted a comparison of scatter buildup in a projection in our system with four currently available dedicated breast CT systems under iso-dose and iso-resolution conditions. The Monte Carlo simulations fully implemented the imaging geometry, image acquisition methodologies, and technique factors of these systems. Figure 19.5 illustrates that NB-CT leads to nearly scatter-free projections (raw data).

19.4.2 Radiation Dose

In NB-CT, a narrow beam is generated within the collimator. The collimator consists of two sheets on a revolving disk positioned between the focal spot and the breast. There is a gap or window between the sheets through which x-ray photons pass to the breast. As the disk rotates, the size of this window changes, altering the size of the narrow beam. This variation is intentional, as the center of a pendent breast is thicker than its periphery, requiring a higher number of x-rays for measurement. Figure 19.6c compares the ideal fluence modulation with the resulting fluence from NB-CT's collimator.

As depicted in Figure 19.7, the x-ray fluence modulation in NB-CT has substantial implications for patient dose. In a small breast (10 cm in diameter at the most posterior part), there is a 30% reduction in dose. For a medium-sized breast (14 cm in

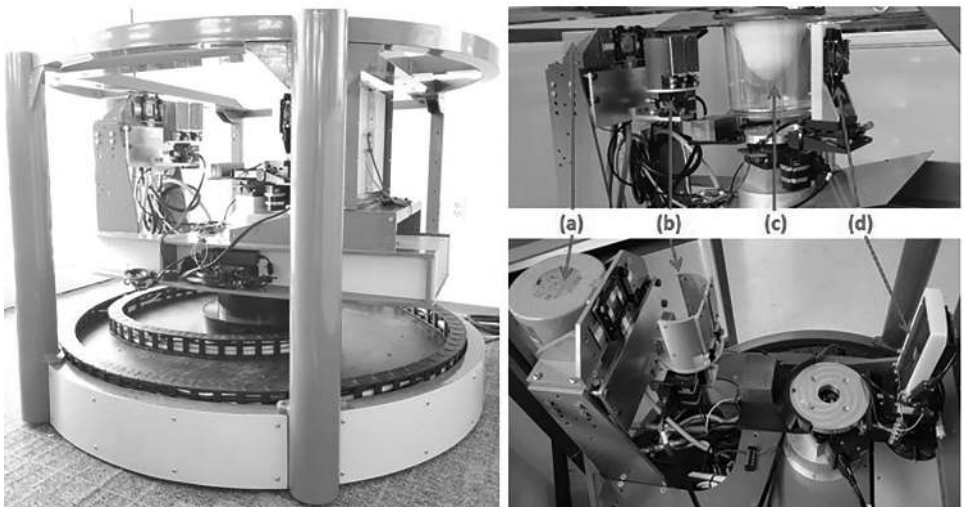


FIGURE 19.4 Close views of prototype narrow-beam breast CT system. A photograph of the entire apparatus is shown on the left. The key components of the CT system from side-view photo (upper) and angled overview (lower) are (a) x-ray source, (b) collimator assembly, (c) breast phantom hanging pendant uncompressed, and (d) photon-counting line detector.

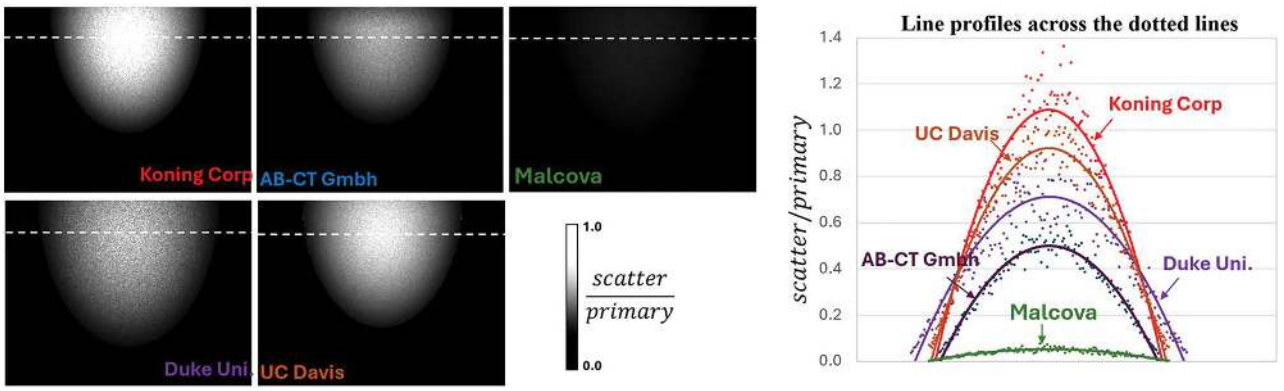


FIGURE 19.5 Scatter buildup in projections measured in scatter-to-primary-ratio compared to the systems developed by Koning Corp.,[21] Advanced Breast CT Gmbh,[22] Duke University,[23] and UC Davis.[24]

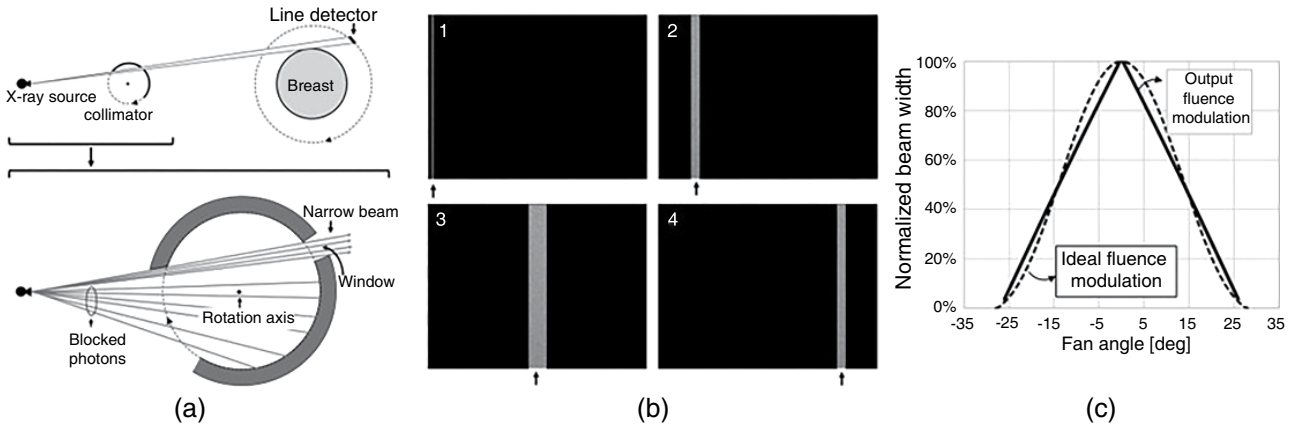


FIGURE 19.6 (a) A planar diagram of the collimator assembly, (b) the resulting recorded beam at four instances of time, and (c) a comparison between the ideal and the resulting fluence modulations. Note that a simple rotation of the collimator around an axis that is positioned in front of the focal spot leads to a spatial modulation of the number x-rays incident on the imaging system's field of view.

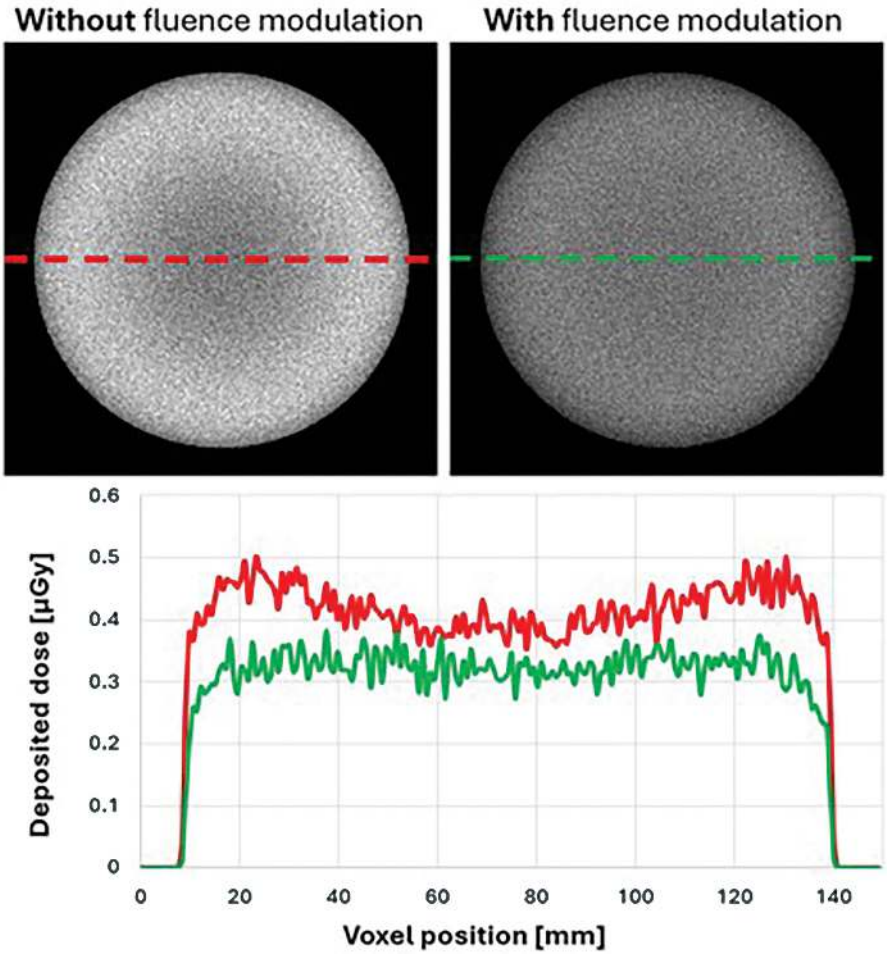


FIGURE 19.7 Spatial distribution of dose in a coronal slice of a medium size breast with and without fluence modulation.[25]

diameter), the reduction is 34%, and for a large breast (18 cm in diameter), it is 40%. This dose reduction is a key factor in making narrow-beam CT a viable solution for breast cancer screening applications.

19.4.3 Spatial Resolution

Figure 19.8 illustrates that NB-CT offers higher system spatial resolution compared to existing 3D imaging solutions like MRI and current dedicated breast CT systems.

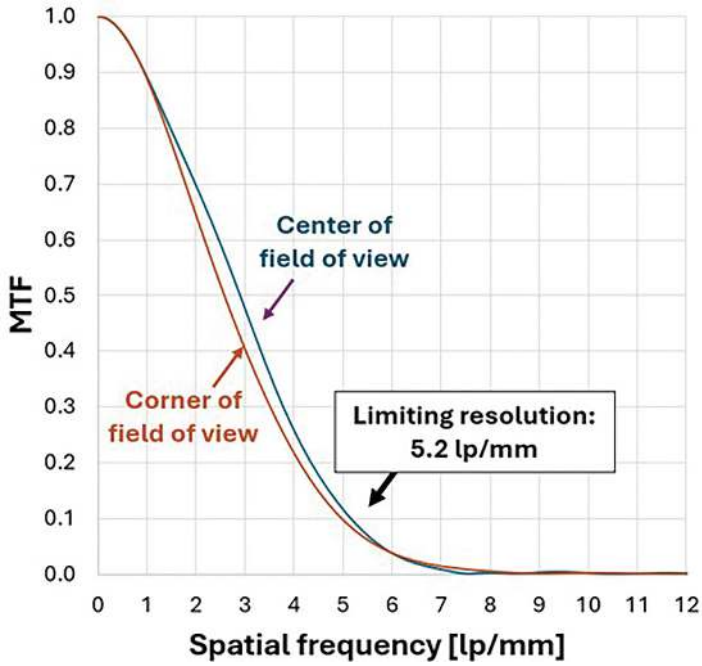


FIGURE 19.8 The measured modulation transfer function (MTF) of the imaging system.

This enhanced resolution, coupled with reduced scatter acquisition, enables the visualization of tiny features crucial for cancer detection and diagnosis. In their publication, the researchers demonstrated that their NB-CT prototype, when imaging the American College of Radiology's Mammography accreditation phantom with the same radiation dose as in screening mammography (5.6 mGy mean glandular dose, below the dose limit in two-view screening mammography of a medium-sized breast), could visualize the smallest features. These features are typically only visible in high-end mammography systems.

19.5 Discussion and Conclusion

Breast cancer is a significant health concern, representing about 30% of all new cancer cases among women annually. In 2024, the US Centers for Disease Control and Prevention estimate that over 300,000 individuals in the US will be diagnosed with aggressive breast cancers. Early detection is crucial for effective treatment, and screening mammography, introduced in the early 1990s, has helped reduce the incidence of late-stage cancers. However, mammography's effectiveness can be limited by breast density, which reduces sensitivity due to tissue overlap.

Dedicated breast CT, a three-dimensional imaging modality, has been considered a potential solution for the past two decades. However, it has not been widely adopted in clinical radiology due to limitations such as low image quality, high radiation dose, and inadequate coverage of the breast posterior region. To address these limitations, narrow-beam CT was developed.

The development of NB-CT was driven by the hypothesis that preventing scattered photon acquisition would enhance the contrast resolution of fine image details. Simulation studies confirmed this hypothesis, leading to the novel method of beam collimation via fluence modulation to create a "narrow beam," which reduces image scatter contamination even with a wide flat-panel detector. To further reduce scatter contamination to near-zero (less than 5% scatter accumulation in the detector), the prototype system uses a "line" detector instead of a flat-panel detector.

A key feature of NB-CT is its rotary collimator design, which varies the number of x-rays at different parts of the field of view, leading to a significant reduction in radiation dose to the patient. This dose reduction is crucial for integrating CT into breast cancer screening. Additionally, the system's spatial resolution is sufficient to resolve details as small as 100 microns, partly due to the high-resolution cadmium telluride photon-counting line detector. This detector not only reduces electronic noise in the raw data but also adds to the substantial image quality improvements achieved with NB-CT geometry and acquisition methodology alone.

In current dedicated breast CT systems, coverage of the breast's posterior anatomy, including the axilla, is inferior to that of breast MRI and even mammography (2D/3D), which also has its own coverage limitations. Our ongoing efforts are focused on addressing this limitation and assessing the clinical feasibility of this innovative design. As noted previously, narrow-beam CT is fundamentally different from cone-beam or spiral CT geometries. Traditional CT approaches rely on a large physical gantry that mechanically couples the x-ray tube and x-ray detector. In contrast, NB-CT eliminates the need for a large gantry because, as shown in Figure 19.9, the x-ray tube

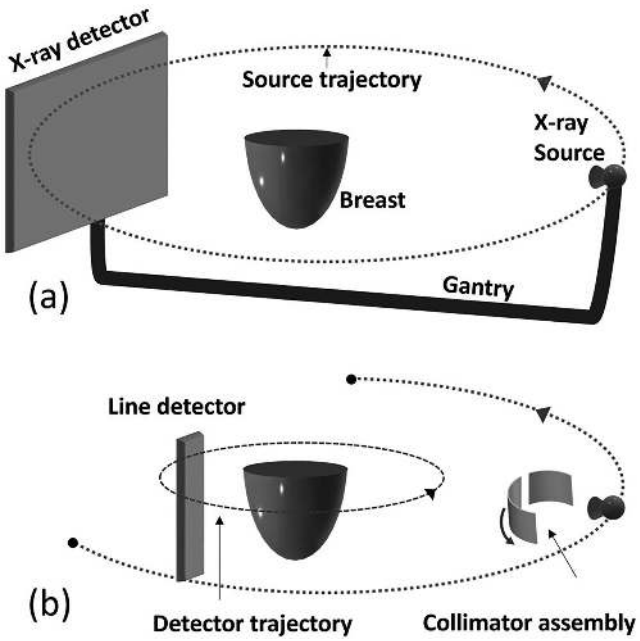


FIGURE 19.9 (a) Structure of cone-beam/spiral CT designs compared to (b) narrow-beam CT design. In narrow-beam CT, the x-ray source and detector are mechanically decoupled, allowing for a gantry-free system design. This design employs a narrow, vertically positioned detector that rotates closely around the breast, while the x-ray source follows a larger path at a different rotational speed.

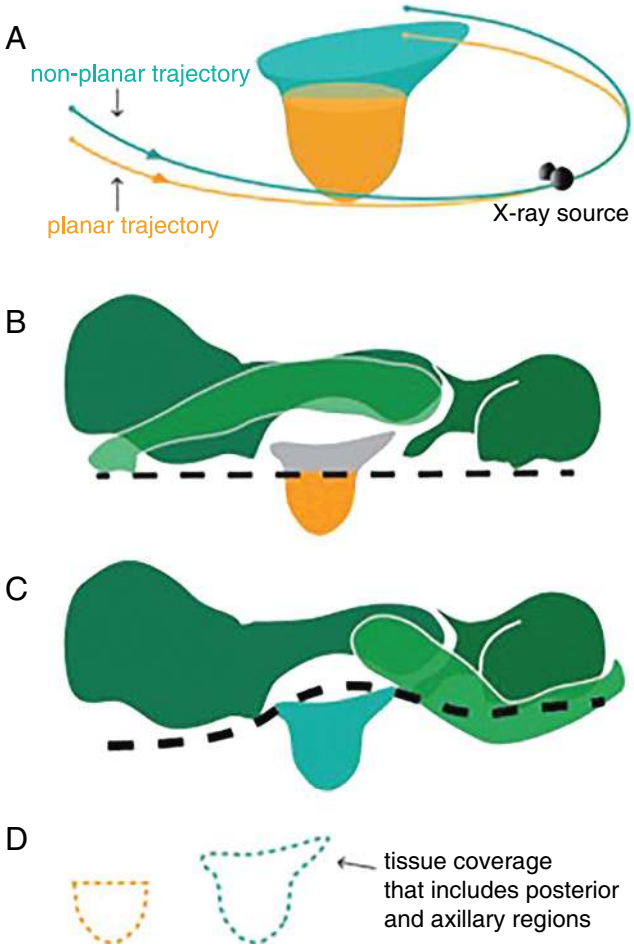


FIGURE 19.10 Objective 1 of this application focuses on updating the trajectory of the x-ray source to maximize the coverage of the breast anatomy. By updating the path that the source takes from planar trajectory (already developed in beta prototype) to non-planar (to be developed in the proposed project), we can reach the entire posterior anatomy of the breast, including the axilla.

and detector are physically decoupled and attached to separate robotic structures, making narrow-beam CT a "gantry-free" design for CT.

Due to the mechanical decoupling in NB-CT, fully tomographic images can be acquired without all x-ray components needing to rotate 360 degrees around an object of interest. In NB-CT, only the detector rotates 360 degrees continuously around the object of interest; the bulky x-ray source does not. This provides a major clinical advantage, especially for imaging the axillary tail of Spence, a tail of breast tissue that extends into the underarm area where breast cancer can develop. Expanding the field of view of the imaging system to include this part of the breast anatomy is a focus of our future work.

In the prototype, the source's movement around the breast was limited to a planar circular orbit, as shown in Figure 19.4. This simplified the design and allowed the team to focus on the fundamental challenges of narrow-beam image acquisition. A new trajectory, as shown in Figure 19.10, significantly improves coverage of the posterior tissue.

The non-planar trajectory of the NB-CT system depends on factors such as the breast diameter and the gap between the focal spot and the x-ray tube's roof. For our product, we are using an M15xx series x-ray tube (Varex Imaging, Salt Lake City, UT), where the distance between the focal spot and the x-ray tube's roof is 1.8 cm. The maximum breast diameter, defined as the 95th percentile breast size in the US female population, is reportedly between 18 to 20 cm. Based on these parameters, our preliminary data suggest that an out-of-plane movement of only 8 cm along a hyperbolic trajectory would meet Tuy's data sufficiency condition for imaging at least 95% of women, providing full sampling needed to measure the posterior regions of the breast. We are focused on these and other relevant considerations in their current and future studies.

19.6 Future Trends

The advancements in narrow-beam CT outlined in this chapter represent a significant potential breakthrough in breast cancer detection and diagnosis, particularly for women with dense breast tissue. By minimizing x-ray scatter through innovative design and image acquisition methodologies, the NB-CT system enhances image quality and spatial resolution while reducing radiation dose, addressing critical limitations of existing breast imaging modalities.

The near scatter-free imaging acquisition, in itself, does not fully resolve the limitations of current CT. Without addressing other shortcomings, we will not completely fulfill the need for a standalone imaging modality that could potentially replace mammography for imaging women with dense breasts. Our ongoing and future work is focused on providing a comprehensive solution.

In NB-CT, the use of a "gantry-free" system allows for non-planar movements of the source, which promise improved anatomical coverage, including challenging regions like the axillary tail, vital for comprehensive breast cancer screening. Additionally, the source's non-planar trajectory opens up new possibilities for patient positioning, thereby increasing clinical throughput and improving the patient experience.

The use of contrast enhancement remains an ongoing area of research in breast cancer radiology. Contrast-enhanced mammography (CEM), for instance, combines traditional mammography with an intravenous contrast agent to highlight areas of increased blood uptake, often associated with cancer. However, the drawbacks of CEM limit its viability. For one, like mammography, CEM requires breast compression, which slows down blood perfusion. Additionally, the long time gap between contrast injection and image acquisition (90 to 120 seconds) can lead to diagnostic errors. These issues, combined, limit CEM's suitability for routine screening and diagnostic use.

We are currently working on spectral CT imaging, which effectively addresses the limitations of CEM. Our system uses a photon-counting detector, allowing us to acquire spectral information in intervals as short as 10 seconds. This enables us to capture diagnostically rich data within the critical first 30 seconds after contrast administration.

These and other topics form the core of our future efforts to develop this important, potentially life-saving technology.

REFERENCES

1. Yaffe MJ. Mammographic density. Measurement of mammographic density. *Breast Cancer Res*. 2008;10(3):209. doi:10.1186/bcr2102
2. Slanetz PJ, Freer PE, Birdwell RL. Breast-density legislation—practical considerations. *N Engl J Med*. 2015;372(7):593–595. doi:10.1056/NEJMp1413728
3. Moshina N, Aase HS, Danielsen AS, et al. Comparing screening outcomes for digital breast tomosynthesis and digital mammography by automated breast density in a randomized controlled trial: results from the to-be trial. *Radiology*. 2020;297(3):522–531. doi:10.1148/radiol.2020201150
4. Carney PA, Miglioretti DL, Yankaskas BC, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med*. 2003. doi:10.7326/0003-4819-138-3-200302040-00008
5. Yaghjyan L, Colditz GA, Collins LC, et al. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to tumor characteristics. *J Natl Cancer Inst*. 2011;103(15):1179–1189. doi:10.1093/jnci/djr225
6. Bertrand KA, Tamimi RM, Scott CG, et al. Mammographic density and risk of breast cancer by age and tumor characteristics. *Breast Cancer Res*. 2013;15(6). doi:10.1186/bcr3570
7. Harris E. FDA updates breast density reporting standards, other mammogram rules. *JAMA*. 2023;329(14). doi:10.1001/jama.2023.4004
8. Förnvik D, Kataoka M, Iima M, et al. The role of breast tomosynthesis in a predominantly dense breast population at a tertiary breast centre: breast density assessment and diagnostic performance in comparison with MRI. *Eur Radiol*. 2018;28(8):3194-3203. doi:10.1007/s00330-017-5297-7
9. Kopans DB. Digital breast tomosynthesis from concept to clinical care. *Am J Roentgenol*. 2014. doi:10.2214/AJR.13.11520
10. Holland K, Sechopoulos I, Mann RM, den Heeten GJ, van Gils CH, Karssemeijer N. Influence of breast compression pressure on the performance of population-based mammography screening. *Breast Cancer Res*. 2017;19(1):1–8. doi:10.1186/s13058-017-0917-3
11. Whelehan P, Evans A, Wells M, MacGillivray S. The effect of mammography pain on repeat participation in breast cancer screening: a systematic review. *The Breast*. 2013;22(4):389–394. doi:10.1016/j.breast.2013.03.003
12. Webb ML, Cady B, Michaelson JS, et al. A failure analysis of invasive breast cancer: most deaths from disease occur in women not regularly screened. *Cancer*. 2014;120(18):2839–2846. doi:10.1002/cncr.28199
13. Onitilo AA, Engel JM, Liang H, et al. Mammography utilization: patient characteristics and breast cancer stage at diagnosis. *Am J Roentgenol*. 2013;201(5). doi:10.2214/AJR.13.10733
14. Monticciolo DL, Newell MS, Moy L, Lee CS, Destounis S V. Breast cancer screening for women at higher-than-average risk: updated recommendations from the ACR. *J Am Coll Radiol*. 2023. doi:10.1016/j.jacr.2023.04.002
15. Kerlikowske K, Zhu W, Tosteson ANA, et al. Identifying women with dense breasts at high risk for interval cancer. *Ann Intern Med*. 2015;162(10):673–681. doi:10.7326/M14–1465
16. Vlahiotis A, Griffin B, Stavros AT, Margolis J. Analysis of utilization patterns and associated costs of the breast imaging and diagnostic procedures after screening mammography. *Clin Econ Outcomes Res*. 2018;10:157–167. doi:10.2147/CEOR.S150260
17. Ehsan AN, Wu CA, Minasian A, et al. Financial toxicity among patients with breast cancer worldwide: a systematic review and meta-analysis. *JAMA Netw Open*. 2023;6(2). doi:10.1001/jamanetworkopen.2022.55388
18. Kak AC, Slaney M, Wang G. Principles of computerized tomographic imaging. *Med Phys*. 2002. doi:10.1118/1.1455742
19. Joseph PM, Spital RD. The effects of scatter in computed tomography. *Med Phys*. 1982. doi:10.1118/1.595111
20. Ghazi P. Reduction of scatter in breast CT yields improved microcalcification visibility. *Phys Med Biol*. 2020;65(23):235047. doi:10.1088/1361–6560/abae07
21. Ning R, Tang X, Conover D. X-ray scatter correction algorithm for cone beam CT imaging. *Med Phys*. 2004;31(5):1195–1202. doi:10.1118/1.1711475

22. Kalender WA, Kolditz D, Steiding C, et al. Technical feasibility proof for high-resolution low-dose photon-counting CT of the breast. *Eur Radiol*. 2017;27(3):1081–1086. doi:10.1007/s00330-016-4459-3
23. Shah JP, Mann SD, Tornai MP. Characterization of X-ray scattering for various phantoms and clinical breast geometries using breast CT on a dedicated hybrid system. *J Xray Sci Technol*. 2017. doi:10.3233/XST-16202
24. Hernandez AM, Boone JM. Average glandular dose coefficients for pendant-geometry breast CT using realistic breast phantoms. *Med Phys*. 2017;44(10):5096–5105. doi:10.1002/mp.12477
25. Ghazi P. A fluence modulation and scatter shielding apparatus for dedicated breast CT: theory of operation. *Med Phys*. 2020;47(4):1590–1608. doi:10.1002/mp.14026

20

Broadband Coherent Anti-Stokes Raman Scattering Microscopy for Metabolic and Phenotypic Imaging in Cancer

Jessica Z. Dixon, Wei-Wen Chen, Rajas Poorna, Xavier Audier, and Marcus T. Ciceronex

20.1 Introduction

The phenotypic and metabolic states of cells within a solid tumor are spatially and temporally complex. These lesions often contain rich micro-economies with a range of metabolically specialized cells which can be critical for tumor function and growth [1, 2]. Additionally, immune infiltrates play a central role in cancer outcomes [3], and locally suppressed immune function has been observed across a large spectrum of human cancers [4–13].

A fuller understanding of these micro-economies and their interactions with immune cells would likely result in increased precision in prognosis and diagnosis, provided there was also an accessible microscopy tool that could fully convey these interactions by directly visualizing tumor metabolic and phenotypic heterogeneity. Raman spectroscopy is an excellent candidate as a contrast-generating mechanism for such a microscopy since spectra can be obtained at submicron resolution [14] and provide highly specific identification of cancer-related metabolites that otherwise cannot be visualized reliably or at all in microscopy [15–26]. Having high dimensionality (many peaks), Raman spectra have ample information capacity to resolve even the subtly different cell phenotypes expected in a tumor environment. Despite all these advantages, spontaneous Raman scattering is not a good candidate for clinical or even routine research imaging due to small Raman scattering cross sections and resulting too-slow signal acquisition [14].

Coherent Raman imaging (CRI) techniques such as coherent anti-Stokes Raman scattering (CARS) and stimulated Raman scattering (SRS) generate Raman scattered light much more efficiently than spontaneous methods [27] and are commonly used for rapid imaging of biological specimens [28–30]. However, most of these CRI methods are largely constrained to the strongly scattering but information-lean "CH-stretch" spectral window. These narrowband CRI methods have proven to be excellent label-free substitutes for H&E staining [30], but are not practical for acquiring broad swaths of spectral information in the weakly scattering but information-rich fingerprint spectral region. Thus, neither spontaneous Raman nor the most common CRI techniques are of practical use for harnessing the rich information available via Raman fingerprint vibrational spectroscopy for histopathology or imaging in cancer research. Broadband coherent anti-Stokes Raman scattering (BCARS) microscopy stands out among CRI methods in that it is optimized to provide fingerprint Raman spectra; it acquires full Raman spectra at each image pixel, but 1000 times faster than spontaneous Raman [29], so enables routine imaging of biological cells and tissues. It remains the only instrument that can acquire full Raman spectra from tissue at such high speed.

Coherent Raman approaches generate signal through a third-order nonlinear effect. BCARS uses three fields (pump, Stokes, and probe) generated by ultrafast pulses to generate this nonlinear signal. An expression for the BCARS signal can be written as:

$$I_{CARS}(\omega) \propto \left| \left\{ \left[E_S(\omega) \star E_p(\omega) \right] \chi^{(3)}(\omega) \right\} * E_{pr}(\omega) \right|^2 \qquad (20.1)$$

where E_S, E_p, and E_{pr} are the electromagnetic field strengths of the Stokes, pump, and probe pulses respectively. BCARS uses a highly efficient impulsive scheme to generate fingerprint signal. One of the BCARS pulses is temporally short (15 fs) and spectrally broad so it provides both the pump and Stokes fields, efficiently driving vibrational coherence in the fingerprint spectral region. The magnitude of this fingerprint coherence is proportional to the third-order nonlinear susceptibility of the sample, $\chi^{(3)}$, which is given by:

$$\chi^{(3)}(\omega) = \chi^{(3)}_{NR}(\omega) + \chi^{(3)}_{R}(\omega) \qquad (20.2)$$

where $\chi^{(3)}_{NR}(\omega)$ is the nonresonant susceptibility originating from electrons directly responding to the pulse fields, and $\chi^{(3)}_{R}(\omega)$ is the resonant susceptibility originating from the nuclear vibrations responding to the electron motion.

In the fingerprint spectral region, $\chi^{(3)}_{NR} >> \chi^{(3)}_{R}$ even with efficient coherence generation, the fingerprint spectra are still *extremely* weak. However, since nonresonant and resonant responses are mechanistically linked, a heterodyne term is also present in the BCARS signal that contains the fingerprint information, but with amplitude found at the geometric mean of the purely nonresonant and purely resonant signals. Efficient fingerprint generation and heterodyne detection together allow fingerprint spectral acquisition on the order of milliseconds, allowing entire images to be generated in a reasonable timeframe. We note that while BCARS remains the state-of-the-art for acquiring fingerprint Raman spectra quickly with high fidelity, BCARS imaging speeds are still orders of magnitude too slow for routine histopathology

DOI: 10.1201/9781003449942-23

imaging. However, in concert with faster imaging modalities, BCARS could be a powerful aid to histopathology and improve diagnostic accuracy.

20.2 Materials

20.2.1 *C. Elegans*

The wild-type *C. elegans* (Bristol N2) worms used in this study were obtained from the Caenorhabditis Genetics Center (CGC), University of Minnesota. Nematodes for measurements were synchronized by the two-generation egg-laying method. Synchronized worms were maintained on nematode growth media (NGM) plates, seeded with *Escherichia coli* strain OP50 as a bacterial food source, and incubated at 20 °C until one-day adults. Sodium azide (Sigma-Aldrich) was used to anesthetize worms for microscopy [31].

20.2.2 Cell Lines

PC-3 and RWPE-1 cell lines were obtained from the American Type Culture Collection (ATCC: CRL-1435, CRL-3607). PC-3 cells were grown in RPMI 1640 medium supplemented with 10% FBS and 5% penicillin-streptomycin. RWPE-1 cells were grown in KSF-M medium supplemented with penicillin-streptomycin, human recombinant epidermal growth factor (rEGF), and bovine pituitary extract (BPE).

20.2.3 Broadband Coherent Anti-Stokes Raman Scattering (BCARS) Microscope

The BCARS instrument described here uses 1040 nm, 100 MHz, 130 fs laser pulses to drive the creation of a broadband supercontinuum (SC) and a narrowband probe pulse (Figure 20.1). A portion of the laser is guided into an optical parametric oscillator (OPO) to produce an 809 nm, 5 ps, 60–90 mW probe pulse. The remaining portion of the laser light is guided into a compression grating and a 10 cm mode area polarization maintaining photonic crystal fiber, yielding a 950–1250 nm, 110 mW SC pulse. The continuum is then compressed with chirp mirrors and a prism pair before being combined with the probe pulse.

The combined light is introduced to 1D line-scan optics, which can be turned off for sample raster-scanning, and then to the sample via a high numerical aperture (NA) objective lens. The signal is collected in transmission, collimated with a second objective lens, and passed through spectral filters to isolate the anti-Stokes signal. This signal is then focused on the slit of a spectrometer and detected with a digital complementary metal-oxide-semiconductor (CMOS) camera (Figure 20.1). The combined light is introduced to 1D line-scan optics and then to the sample via a high NA objective lens. The signal is collected in transmission, collimated with a second objective lens, and passed through two broad bandpass filters to isolate the anti-Stokes signal which is focused on the slit of a spectrometer.

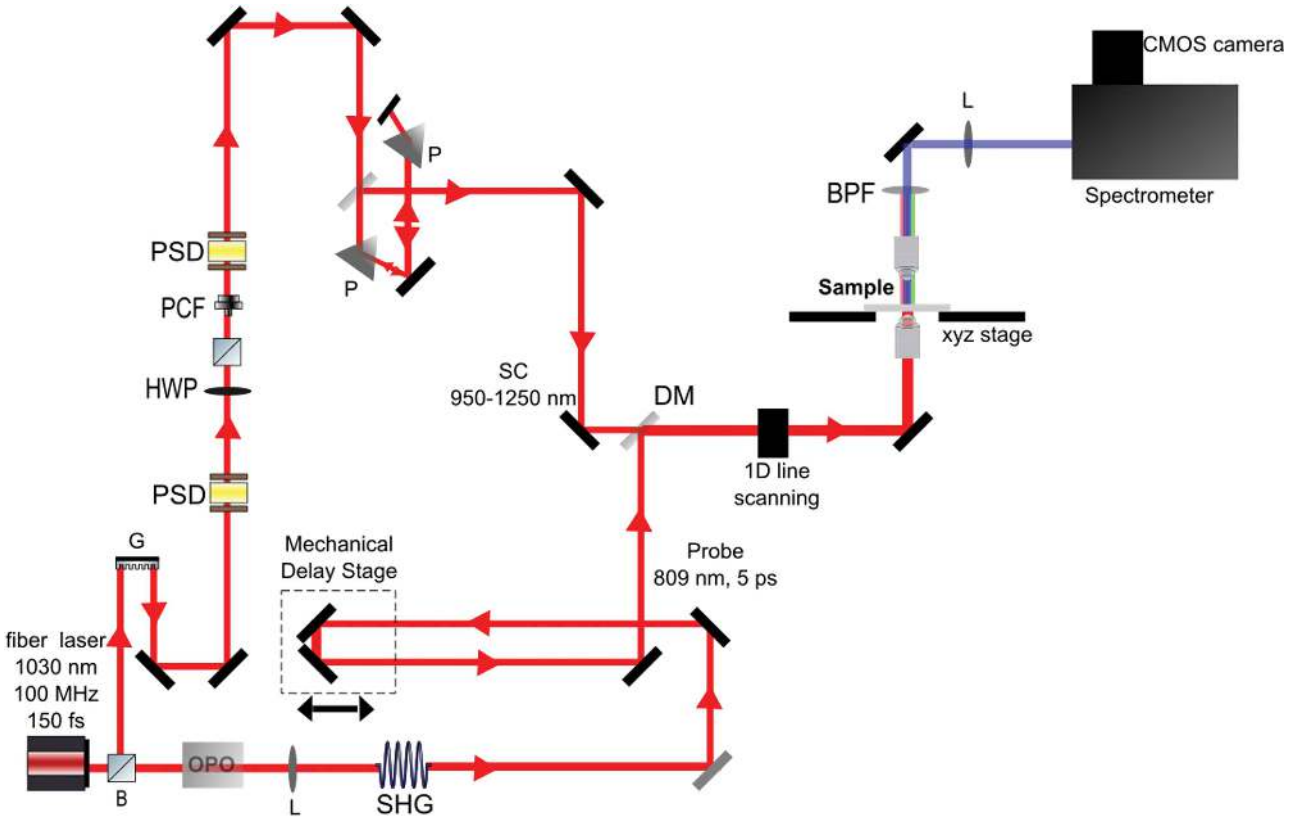


FIGURE 20.1 BCARS experimental setup. B: beamsplitter, OPO: optical parametric oscillator, L: Lens, SHG: second harmonic generator, DM: dichroic mirror, G: grating compressor, HWP: half-wave plate, PCF: photonic crystal fiber, P: prism, PSD: position sensitive detector, BPF: bandpass filter.

The signal is detected with a low-noise scientific CMOS camera (Figure 20.1).

20.2.4 Extracting and Processing the Raman Spectra

The heterodyne BCARS signal is retrieved from the overall signal through a Hilbert transform [32]. This requires the acquisition of non-resonant background (NRB) and baseline (dark) measurements in addition to the BCARS hyperspectral sample image. All three signals are processed using custom Python software, CRIkit2 (https://github.com/CCampJr/CRIkit2) [32].

20.2.5 Spanning-Tree Progression Analysis of Density-Normalized Events (SPADE) Algorithm

SPADE is described in detail elsewhere [33]. As applied here, SPADE calculates the density of points from the BCARS spectrum of each pixel in the hyperspectral BCARS image. Then, the data is downsampled in a density-dependent manner, so that the density of points in the spectral domain is approximately uniform in volume. Where the density of points is lower than a set threshold, the points are treated as outliers and removed. Next, SPADE utilizes either K-means or agglomerative clustering to partition the downsampled data into a number of clusters specified by the user. Then, a minimum spanning tree (MST) is constructed to connect the clusters based on their similarity. Finally, the MST is displayed using a force-directed drawing algorithm [31].

20.3 Methods

20.3.1 BCARS Measurements

Cell images were acquired using the BCARS microscope described earlier in a single-pixel raster-scanning modality with 30 mW of probe pulse power and 50 mW of SC power with a 1.2 NA, 60x water objective lens (UPLSApo60X W IR, Olympus) to collect the signal. The exposure time was set to 10 ms per pixel or 30 ms/μm. The NRB measurement was performed on a region of the sample with no cells and consisted of 2500 time points with 10 ms exposure for each point. The dark measurement was conducted in the same region of the sample as the NRB with the same parameters but with the SC and probe pulses temporally separated so there would be no overlap and no signal generated. The BCARS *C. elegans* image analyzed with SPADE was acquired with a similar BCARS microscope that is described elsewhere [31].

20.3.2 Raman Spectral Processing and Extraction

The initial processing of the data applies a 7.5 cm^{-1} window median filter to the image, NRB, and dark spectra to filter camera noise. For line-scanning images, an intensity correction is applied to correct for the uneven illumination across the line-scanning axis. No additional pre-processing steps are required for images obtained by point scanning.

The remaining processing steps are performed using the CriKIT2 (https://github.com/CCampJr/CRIkit2) software and a visual overview of the method is described in Figure 20.2 [32].

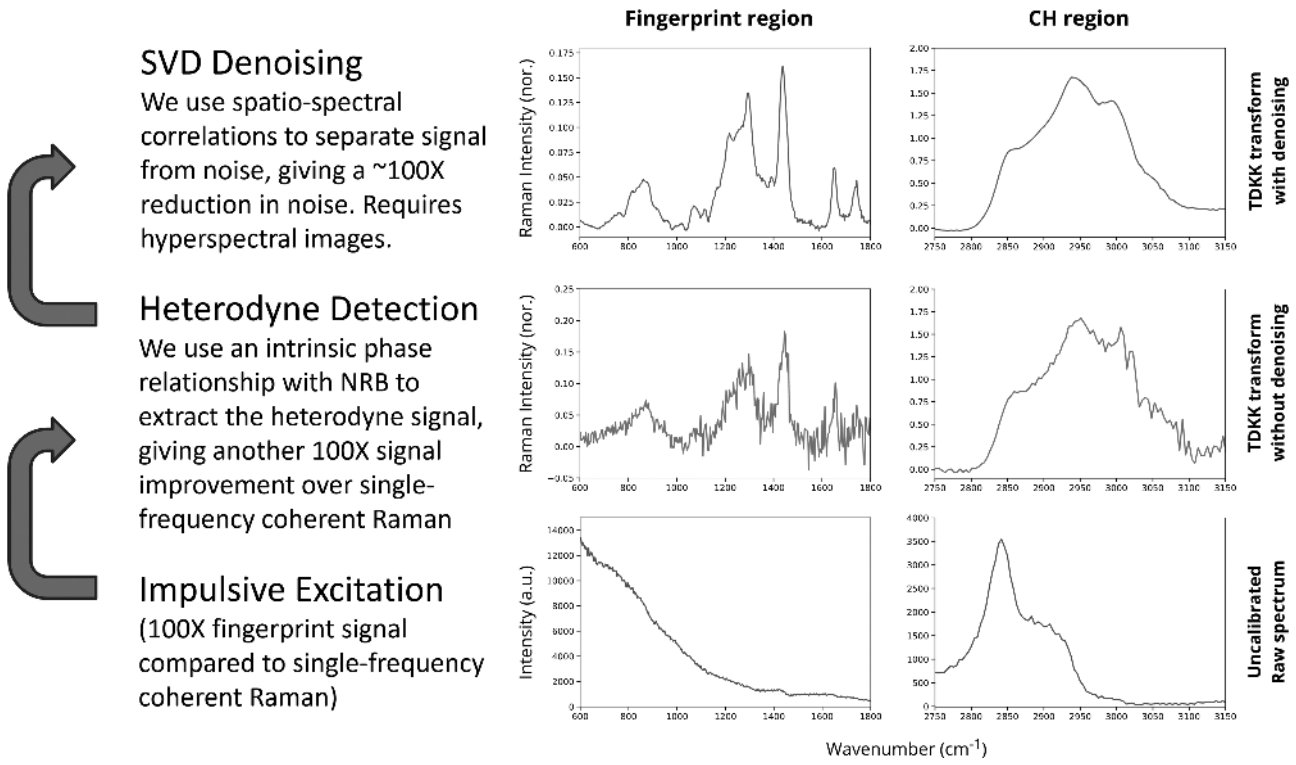


FIGURE 20.2 Summary of the transformation of the measured noisy, BCARS signal with non-resonant background (NRB) to a denoised, Raman spectrum.

First, the pre-processed image, NRB, and dark measurements are uploaded, and the dark is subtracted from both the image and the NRB. Then, all spectra are reduced to the frequency window of approximately 350–3400 cm^{-1}.

Next, the data is denoised in a three-part process that begins with an Anscombe transformation to normalize the variance of the noise, which is a combination of Poisson and Gaussian noise, as is typical for CCD or CMOS cameras by approximating it as purely Gaussian noise [32, 34]. The variance stabilization then allows for the denoising of the image by singular value decomposition (SVD), which factors the 3D hyperspectral image into orthonormal basis vectors that are ordered by the magnitude of their singular value. Each basis vector is represented visually by a 2D spatial representation and a corresponding spectral vector, which are examined for recognizable spatial and spectral features. Only singular values that contain such features are retained, and the remaining are discarded as noise [32]. Following SVD, an inverse Anscombe transformation is applied to return the data to its true mixed noise form [34].

While BCARS generates both a vibrationally resonant and non-resonant response, which are described in Equation 20.2, the Raman spectrum is fully represented by the imaginary portion of the resonant term and must be isolated from the overall nonlinear response ($I_{CARS}(\omega)$) [32, 35, 36]. A windowed Hilbert transform ($\widehat{H_W}$) allows us to retrieve the phase ($\varphi_{CARS}(\omega)$) that isolates the Raman spectrum from the heterodyned BCARS signal [32, 35, 36].

$$\phi_{CARS}(\omega) = \widehat{H_W}\left\{\frac{1}{2} ln I_{CARS}(\omega)\right\} \tag{20.3}$$

If the non-resonant part of the response can be measured or realistically approximated and is fully separate from the Raman part of the response, then Equation 20.3 can be rewritten as:

$$\phi_{CARS/NRB}(\omega) = \widehat{H_W}\left\{\frac{1}{2} ln \frac{I_{CARS}(\omega)}{I_{NRB}(\omega)}\right\} \tag{20.4}$$

which is equivalent to performing a Kramers–Kronig transform on I_{CARS} / I_{NRB} [32]. This then allows the imaginary portion of the spectrum to be retrieved.

The retrieved imaginary component is proportional to the spontaneous Raman signal, but errors in phase and amplitude may still remain. The error in phase estimation is corrected by an automated asymmetric least squares (ALS) algorithm with a Whittaker smoother [37–39]. Amplitude errors are corrected by a broad Savitzky–Golay filter. Lastly, frequency calibration is performed based on the location of a reference peak, typically phenylalanine at 1004 cm^{-1}.

20.3.3 SPADE Algorithm

The user-defined target density for the downsampling step was set to be the third percentile of the local density of all data points while the maximum number of allowable cells in the pooled downsampled data was set to 2×10^5. The resulting downsampled data were clustered using K-means to generate 300 clusters, each of which was represented by one tree node. The position information associated with the pixels that were represented in each tree node was used to highlight the corresponding pixels in the BCARS image (Figure 20.3a–d).

20.4 Results

The described system provides hyperspectral images with a spectral range of 350–3400 cm^{-1} at 2.5 ms/spectrum and spectral resolutions of 10 cm^{-1}. Such characteristics provide the ability to image a variety of biological samples such as tissue slices, *C. elegans*, and single cells. Later we demonstrate the ability of BCARS microscopy to distinguish the unique spatio-chemical makeup of a variety of biological materials. We give examples of correlating BCARS imaging with transcriptomic data of *C. elegans* and identifying metabolic products in human cell lines PC-3 and RWPE−1.

C. elegans is a commonly used experimental model for studies of aging, development, and reproduction with its signaling pathways and genetic makeup very well documented. The gonad, in particular, serves as an excellent example of spatially and temporally discrete critical cellular events such as germline stem cell proliferation and differentiation, meiotic maturation, and meiotic chromosome reorganization, due to the rate at which germline cells progress through it (approximately 6.5 hours) [40]. By comparing spectroscopic BCARS images of a *C. elegans* gonad to the spatially resolved transcriptomic profile of the gonad, we demonstrate the capability for the rich chemical information included in BCARS images to correlate directly with transcriptomic information.

The BCARS hyperspectral image contains a complete Raman spectrum from 500 to 3400 cm^{-1} at each pixel with approximately 50 peaks in each spectrum. For an image with 660 × 660 pixels, this equates to a surplus of information at approximately 10^9 data points. Due to the distinct characteristics of the fingerprint (500–1800 cm^{-1}) and CH stretch (2800–3400 cm^{-1}) regions, we exclusively use the fingerprint region for this analysis reducing the spectral space to 651 points. However, interpreting the subtle differences between spectra at each pixel can still be difficult without dimension reduction and visualization tools.

In biological samples, we expect most of the pixels to contain very similar chemical components consisting of slightly different concentrations of various lipids, proteins, and nucleic acids. Therefore most pixel spectra should cluster distinctly with only some pixels falling into intermediate clusters at the boundary between two clusters with distinct chemical makeups that can be well represented by a network of nodes such as with SPADE [31]. Figure 20.3a–d demonstrates a SPADE tree with four unique clusters highlighted and the corresponding pixels (green) in the BCARS image that are used to compose those nodes in the tree. From the BCARS image it is clear that the pixels represent distinct regions of *C. elegans* that are expected to have unique chemical compositions related to their function. Such differences in chemical components are shown in the corresponding spectra in Figure 20.3e, where

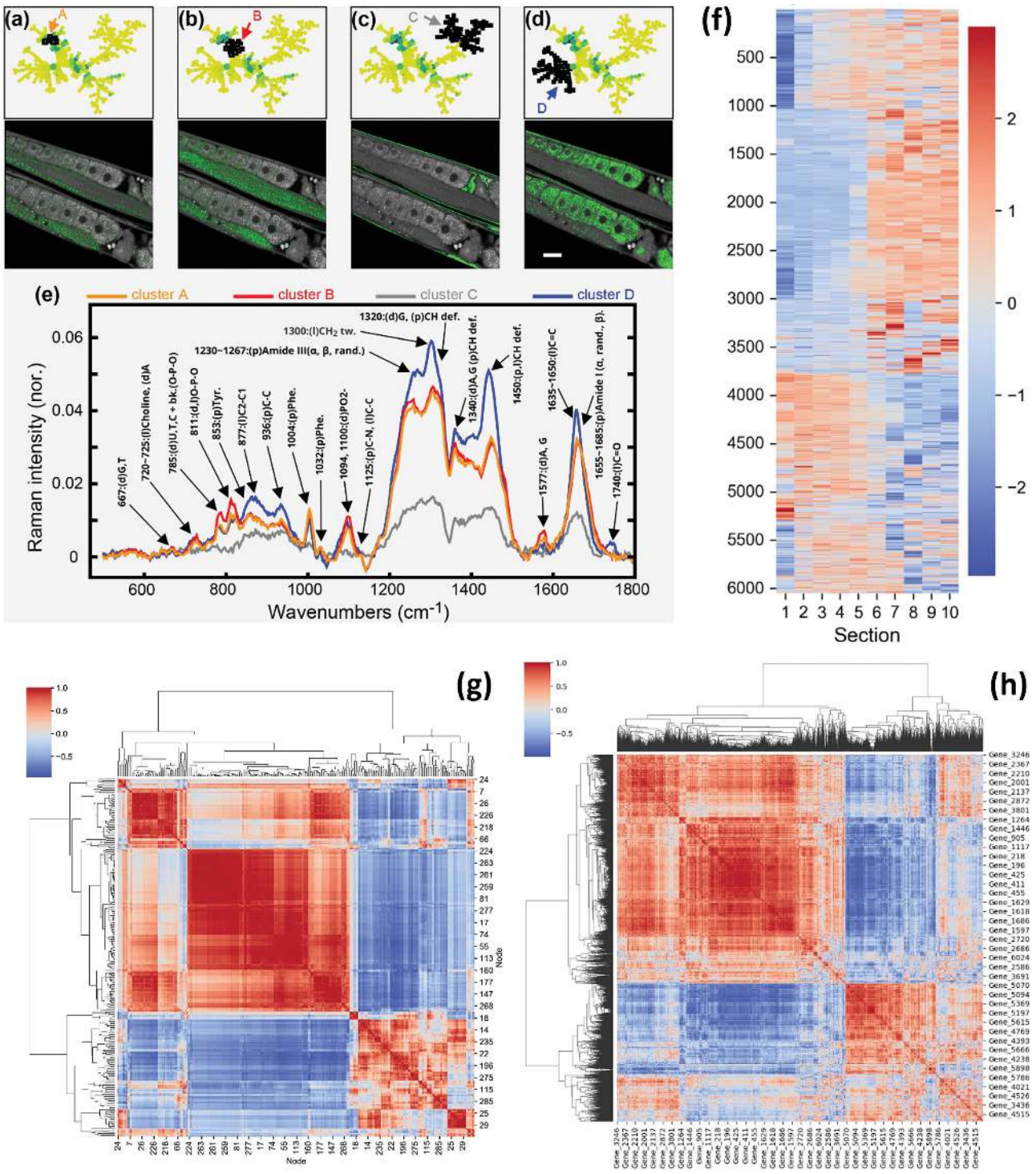


FIGURE 20.3 Clusters of nodes (a–d, black squares) and their spatial distribution (green pixels) in the BCARS image and the corresponding BCARS spectra (e) with Raman peak assignments (d: DNA/RNA or nucleic acids, l: lipid, p: protein). Scale bar: 20 μm. (f) Correlation of standardized gene expression data from Tzur et al. [41] with ten sections of the *C. elegans* gonad. The color indicates the z-score standardized gene frequency. (g) Correlation between each SPADE node. (h) Correlation between genes from the same dynamically expressed gene expression data as in (f) from the same region of the *C. elegans* gonad from Tzur et al. (Figures reproduced with permission from Poorna et al. at https://pubs.acs.org/doi/10.1021/acs.jpcb.3c01446. Further permissions related to the material excerpted should be directed to the American Chemical Society.)

the relative intensities of different Raman bands representing protein (p), lipid (l), and nucleic acid (d) components differ for each cluster.

To demonstrate the relationship of spectral information with transcriptomic activity, we compared the BCARS spectra with gene expression profile data [42] along the hermaphrodite

gonad. Panels (g) and (h) of Figure 20.3 show Seaborn correlation plots of spectral components represented by SPADE nodes (g) and individual gene activity (h). It is clear that the spectral components appear in family groups that precisely mirror the transcriptomic activity groupings. The implication is that BCARS acquires enough information to identify phenotypic state at the level of a 6000-gene transcriptomic read, but much more quickly, noninvasively, and without labels [32]. Information at this level of phenotypic resolution would obviously be useful as a diagnostic tool, especially if it were routinely available. BCARS has the potential to bring this about.

In addition to imaging model organisms such as *C. elegans*, BCARS can be used to analyze human cell lines and tissues. Here we show BCARS data from two human cell lines, RWPE-1, an epithelial prostate cell line, and PC-3, a highly metastatic cell line from the bone metastasis of a prostatic adenocarcinoma (Figure 20.4). The BCARS images in Figure 20.4a and b demonstrate the chemical contrast at 1004 cm^{-1} and show distinct subcellular details such as the nucleus and lipid droplets without the need for any labeling or clustering. The morphology of the two cell lines is clearly distinct, with the PC-3 cells also displaying more heterogeneity.

The average spectral signature of each numbered cell in the two images was calculated along with the overall average spectrum for each cell line (Figure 20.4c). As expected, the average spectra of the two cell types are similar but have important differences in the relative amplitudes of certain Raman peaks indicating differences in chemical composition, likely related to the different metabolic and transcriptomic activity between non-cancer and cancer cells. For example, the PC-3 average spectrum has significantly higher contributions from the vibrational bands at 700 cm^{-1} and 720 cm^{-1}, representing cholesterol and choline, respectively, suggesting higher metabolic activity.

Such differences in the average chemical composition of each cell line are commonly attributed to the Warburg effect [43]. Here, the variation in lipid levels between the two cell lines is likely explained by the increased capacity for *de novo* fatty acid synthesis in many cancers and prostate cancer in particular [41, 44]. The increase in fatty acid synthesis naturally leads to higher accumulation of intracellular lipids in cancer cells and is generally believed to support membrane production for rapidly dividing cells [45] or to play a role in regulating redox balance [46].

While the average spectra of each cell line can be informative about general differences between the behaviors of the cell

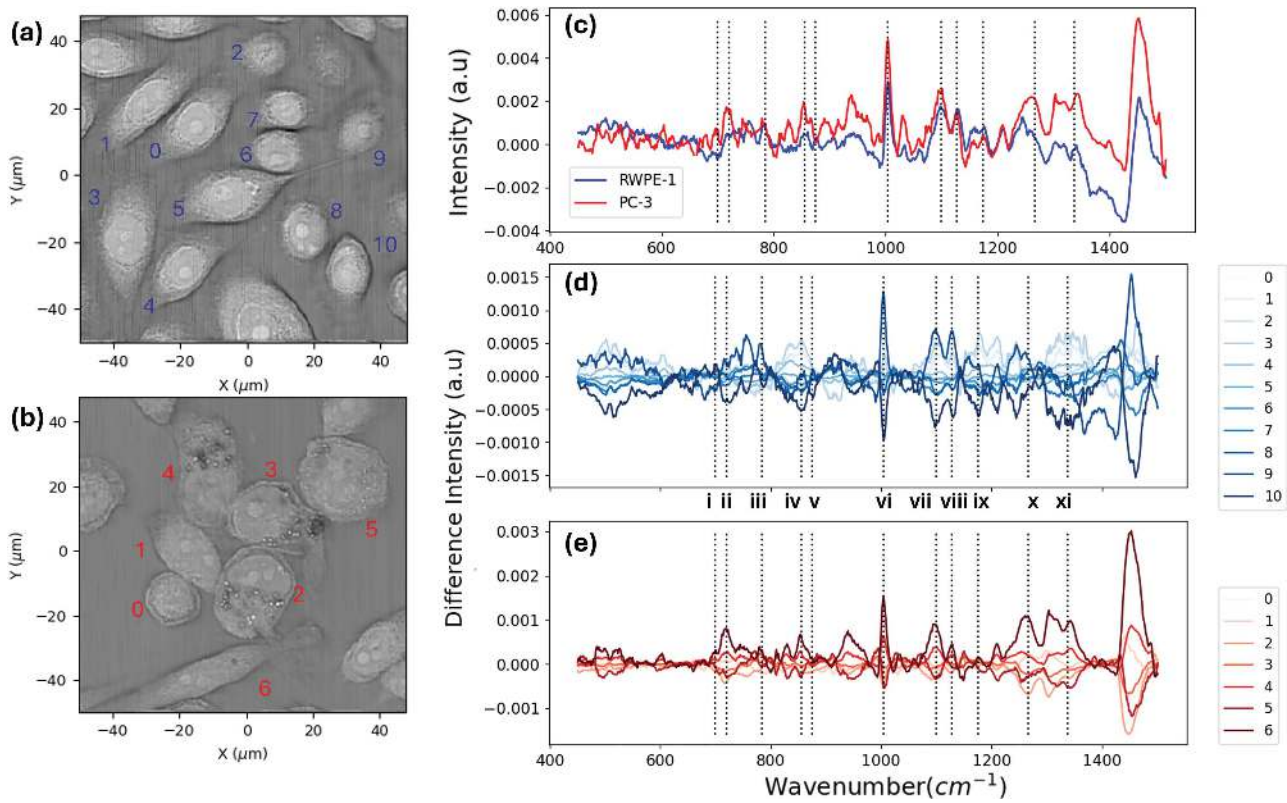


FIGURE 20.4 BCARS images at 1004 cm^{-1} of RWPE-1 prostate epithelial cells (a) and highly metastatic PC-3 cells from a grade IV prostatic adenocarcinoma (b) and the corresponding average Raman spectrum from each of the numbered cells for each cell line (c). The difference spectrum between each individual cell and the average of all cells for RWPE-1 (d) and PC-3 (e). Vertical dashed lines mark specific vibrational bands at i: 700 cm^{-1}, (*l*)Cholesterol; ii: 720 cm^{-1}, (*l*)Choline; iii: 785 cm^{-1}, (*d*)U,T,C + bk(O-P-O); iv: 855 cm^{-1} (*p*)Tyr.; v: 875 cm^{-1}, (*l*)$C_2 - C_1$; vi: 1004 cm^{-1}, (*p*)Phe.; vii: 1100 cm^{-1}, (*d*)PO_2; viii: 1128 cm^{-1} (*p*)C-N, (*l*)C-C; ix: 1175 cm^{-1}, (*p*)Tyr., Phe.; x: 1267 cm^{-1}, (*l*) = CH def.; and xi: 1338 cm^{-1}, (*d*)A,G, (*p*)CH def. *p*: protein related, *d*: DNA/nucleic acid related, *l*: lipid related.

lines, BCARS can provide information related to population heterogeneity. Figure 20.4d–e shows the difference spectrum of each numbered cell in Figure 20.4a–b from the average spectra of the respective cell line in Figure 20.4c. The difference spectra highlight the unique composition of each cell where some cells have more significant contributions of protein, indicated by positive differences of bands such as 855 cm^{-1}, 1004 cm^{-1}, and 1175 cm^{-1}, such as RWPE-1 cell 2 and PC-3 cell 6. In contrast, other cells, such as PC-3 cell 4, show higher nucleic acid content, exhibited by increased contributions of the bands at 785 cm^{-1}, 1100 cm^{-1}, and 1338 cm^{-1}. These variations likely relate to variations in function and state.

20.5 Discussion

A Google Scholar search with "Raman spectroscopy" and "cancer" returns more than 200k citations, with references going back to the 1970s. There is no question that Raman spectroscopy provides signals that can inform on important chemical changes associated with cancer. For example, as mentioned in the introduction, it can provide contrast relevant to metabolic and immune heterogeneity within a solid tumor. However, only in the past decade has it been possible to acquire Raman spectra from tissues with enough speed and fidelity to map tissue regions at sufficient resolution to provide image-encoded information that could be meaningful to a pathologist. As we demonstrate earlier, the information it makes available is roughly equivalent to a deep transcriptomic read at 300 nm resolution. Such tremendously rich information, once interpreted, would often prove tremendously helpful for diagnostic determinations. As required for a diagnostic aid, the BCARS signal reliability is high since BCARS has an intrinsic reference. The so-called nonresonant background is generated with the signal of interest, making it possible to derive Raman signatures with absolute intensity calibration, making measurements repeatable and comparable across systems. BCARS images can also mimic H&E staining by combining the signal from particular vibrational bands, providing a familiar visual for pathologists, while still containing all of the information from the Raman spectrum.

BCARS remains the fastest tissue-imaging microscopy to provide full Raman spectra at each pixel. Even so, it acquires the spatially resolved chemical information in ~1 ms/pixel. This is far too slow for whole-slide imaging but could be employed to acquire highly informative scans of smaller regions of interest.

With all these advantages we feel broadband coherent Raman imaging such as BCARS could be a significant boon to cancer research and diagnosis. Presently, BCARS microscopes must be custom-built and require extensive training for instrument maintenance and imaging. Our research group continues to focus on simplifying the instrument so that it might soon be possible to produce it commercially. Continued development of the experimental setup and techniques as well as the processing methodologies will improve the efficiency, ease of implementation, and affordability of the required technology, making it more accessible and applicable to the biomedical community.

20.6 Future Trends

Given the significant chemical and functional information made available rapidly and at subcellular resolution with BCARS microscopy, we believe it has tremendous potential as a research and diagnostic tool. Without staining, it will be possible to identify metabolic aberrations and other functional variations in tissues and cells, providing valuable supporting information for diagnosis and mechanistic discovery. BCARS will reach its full potential only as it becomes widely used, and data from its application can standardized and shared, as is done with mass spectrometry. Reaching this state will require advances in instrument design and signal processing.

Presently, most BCARS microscopes around the world follow the design we demonstrated in 2014 [29]. This is a highly efficient design from a signal generation perspective, but it requires an 8' × 5'optical table and a full-time optics expert to keep it running optimally. New time-domain designs [47] are easier to implement but significantly less efficient in signal generation, so tissue imaging with this approach is infeasible. BCARS microscopy instrumentation that is both efficient and simple to implement and operate is an unmet need and one that our lab and others are actively working on.

Before biologically relevant data can be extracted from BCARS data, its Raman spectra must be extracted through signal processing. Reliable BCARS signal processing pipelines have been established [32]. However, these require significant user training and intervention, so they are not amenable to routine clinical or biological laboratory use. Machine learning approaches to BCARS signal processing [48] are currently under development and show some promise.

Finally, as simple and reliable instrumentation and signal processing are introduced, categorizing and verifying the connection between BCARS (Raman) signatures and biological systems can begin in earnest. Raman spectra have been correlated with biological structures and function for at least five decades, so much qualitative work has already been done. Over the past decade, quantitative connections have been increasingly deployed between Raman spectra and -omics data, laying an increasingly solid foundation. While exciting, these studies are relatively rare as they are time-consuming and require expertise in Raman spectroscopy and biology. Since BCARS acquires Raman spectra thousands of times faster than spontaneous methods, we anticipate the amount of data connecting Raman spectra with biological function will balloon. The BCARS signal is also absolute and quantitative, so spectra acquired through BCARS will be comparable across instruments and labs, further accelerating the ability to impart robust, actionable meaning to its signatures by reliable analytical methods and interlaboratory comparison. This feature is not available with present Raman methods.

20.7 Funding

This work is supported by the National Institute of Health under Agreement No. 5R21CA240214.

REFERENCES

1. S. Pavlides, D. Whitaker-Menezes, R. Castello-Cros, N. Flomenberg, A. K. Witkiewicz, P. G. Frank, M. C. Casimiro, C. Wang, P. Fortina, S. Addya, R. G. Pestell, U. E. Martinez-Outschoorn, F. Sotgia, and M. P. Lisanti, The reverse Warburg effect: aerobic glycolysis in cancer associated fibroblasts and the tumor stroma, *Cell Cycle* 8 (2014), pp. 3984–4001.
2. M. P. Lisanti, U. E. Martinez-Outschoorn, Z. Lin, S. Pavlides, D. Whitaker-Menezes, R. G. Pestell, A. Howell, and F. Sotgia, Hydrogen peroxide fuels aging, inflammation, cancer metabolism and metastasis, *Cell Cycle* 10 (2014), pp. 2440–2449.
3. M. D. Wellenstein, and K. E. de Visser, Cancer-cell-intrinsic mechanisms shaping the tumor immune landscape, *Immunity* 48 (2018), pp. 399–416.
4. R. Baffa, R. Santoro, F. Bullrich, B. Mandes, H. Ishii, and C. M. Croce, Definition and refinement of chromosome 8p regions of loss of heterozygosity in gastric cancer, *Clin Cancer Res* 6 (2000), p. 1372.
5. K. Wright, P. J. Wilson, J. Kerr, K. Do, T. Hurst, S.-K. Khoo, B. Ward, and G. Chenevix-Trench, Frequent loss of heterozygosity and three critical regions on the short arm of chromosome 8 in ovarian adenocarcinomas, *Oncogene* 17 (1998), pp. 1185–1188.
6. C. L. Wu, L. Roz, P. Sloan, A. P. Read, S. Holland, S. Porter, C. Scully, P. M. Speight, and N. Thakker, Deletion mapping defines three discrete areas of allelic imbalance on chromosome arm 8p in oral and oropharyngeal squamous cell carcinomas, *Gene Chromosome Canc* 20 (1997), pp. 347–353.
7. M. Schoenberg, J. D. Brooks, F. F. Marshall, J. I. Epstein, W. B. Isaacs, P. Cairns, and D. Sidransky, Frequent loss of chromosome arms 8p and 13q in collecting duct carcinoma (CDC) of the kidney, *Gene Chromosome Canc* 12 (1995), pp. 76–80.
8. D. Schullerus, R. von Knobloch, J. Chudek, J. Herbers, and G. Kovacs, Microsatellite analysis reveals deletion of a large region at chromosome 8p in conventional renal cell carcinoma, *Int J Cancer* 80 (1999), pp. 22–24.
9. M. Chang, K. Tsuchiya, R. H. Batchelor, P. S. Rabinovitch, B. G. Kulander, R. C. Haggitt, and G. C. Burmer, Deletion mapping of chromosome 8p in colorectal carcinoma and dysplasia arising in ulcerative colitis, prostatic carcinoma, and malignant fibrous his- tiocytomas, *A m. J Pathol* 1 44 (1994), pp. 1–6.
10. L. A. Takle and M. A. Knowles, Deletion mapping implicates two tumor suppressor genes on chromosome 8p in the development of bladder cancer, *Oncogene* 12 (1996), pp. 1083–1087.
11. R. Anbazhagan, H. Fujii, and E. Gabrielson, Allelic loss of chromosomal arm 8p in breast cancer progression, *Am J Pathol* 152 (1998), pp. 815–819.
12. J. J. König, E. Kamst, A. Hagemeijer, J. C. Romijn, J. Horoszewicz, and F. H. Schröder, Cytogenetic characterization of several androgen responsive and unresponsive sublines of the human prostatic carcinoma cell line LNCaP, *Urol Res* 17 (1998), pp. 79–86.
13. J. Kagan, J. Stein, R. J. Babaian, Y. S. Joe, L. L. Pisters, A. B. Glassman, A. C. von Eschenbach, and P. Troncoso, Homozygous deletions at 8p22 and 8p21 in prostate cancer implicate these regions as the sites for candidate tumor suppressor genes, *Oncogene* 11 (1995), pp. 2121–2126.
14. H. G. Schulze, S. O. Konorov, J. M. Piret, M. W. Blades, and R. F. B. Turner, Label-free imaging of mammalian cell nucleoli by Raman microspectroscopy, *Analyst* 138 (2013), pp. 3416–3423.
15. A. S. Haka, K. E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R. R. Dasari, and M. S. Feld, Diagnosing breast cancer by spectroscopy, *PNAS* 102 (2005), pp. 12371.
16. N. Stone, P. Stavroulaki, C. Kendall, M. Birchall, and H. Barr, Raman spectroscopy for early detection of laryngeal malignancy: preliminary results, *The Laryngoscope* 110 (2000), pp. 1756–1763.
17. Z. Huang, A. McWilliams, H. Lui, D. I. McLean, S. Lam, and H. Zeng, Near-infrared Raman spectroscopy for optical diagnosis of lung cancer, *Int J Cancer* 107 (2003), pp. 1047–1052.
18. S. Kaminaka, H. Yamazaki, T. Ito, E. Kohda, and H. O. Hamaguchi, Near-infrared Raman spectroscopy of human lung tissues: possibility of molecular-level cancer diagnosis, *JRS* 32 (2001), pp. 139–141.
19. A. Nijssen, K. Maquelin, L. F. Santos, P. J. Caspers, T. C. Bakker Schut, J. C. den Hollander, M. H. A. Neumann, and G. J. Puppels, Discriminating basal cell carcinoma from perilesional skin using high wave-number Raman spectroscopy, *J Biomed Opt* 12 (2007), p. 034004.
20. G. R. Lloyd, J. Wood, C. Kendall, T. Cook, N. Shepherd, and N. Stone, Histological imaging of a human colon polyp sample using Raman spectroscopy and self organising maps, *Vib Spectrosc* 60 (2012), pp. 43–49.
21. A. Mizuno, H. Kitajima, K. Kawauchi, S. Muraishi, and Y. Ozaki, Near-infrared Fourier transform Raman spectroscopic study of human brain tissues and tumours, *JRS* 25 (1994), pp. 25–29.
22. P. H. Chen, R. Shimada, S. Yabumoto, H. Okajima, M. Ando, C. T. Chang, L. T. Lee, Y. K. Wong, A. Chiou, and H. O. Hamaguchi, Automatic and objective oral cancer diagnosis by Raman spectroscopic detection of keratin with multivariate curve resolution analysis, *Sci Rep* 6 (2016), p. 20097.
23. S. Rehman, Z. Movasaghi, A. T. Tucker, S. P. Joel, J. A. Darr, A. V. Ruban, and I. U. Rehman, Raman spectroscopic analysis of breast cancer tissues: identifying differences between normal, invasive ductal carcinoma and ductal carcinoma in situ of the breast tissue, *JRS* 38 (2007), pp. 1345–1351.
24. S. Koljenović, L. Choo-Smith, T. C. Bakker Schut, J. M. Kros, H. J. van den Berge, and G. J. Puppels, Discriminating vital tumor from necrotic tissue in human glioblastoma tissue samples by Raman spectroscopy, *Lab Invest* 82 (2002), pp. 1265–1277.
25. T. C. Bakker Schut, K. Maquelin, T. van der Kwast, C. H. Bangma, D. J. Kok, and G. J. Puppels, Discrimination between nontumor bladder tissue and tumor by Raman spectroscopy, *Anal Chem* 22 (2006), pp. 7761–7769.

26. K. Klein, A. Gigler, T. Aschenbrenner, R. Monetti, W. Bunk, F. Jamitzky, G. Morfill, R. Stark, and J. Schlegel, Label-free live-cell imaging with confocal Raman microscopy, *Biophys J* 102 (2012), pp. 360–368.
27. W. Min and X. Gao, Fundamental detectability of Raman scattering: a unified diagrammatic approach, *J Chem Phys* 160 (2024), p. 094110.
28. C. W. Freudiger, W. Min, B. G. Saar, S. Lu, G. R. Holtom, C. W. He, J. C. Tsai, J. X. Kang, and X. S. Xie, Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy, *Science* (2008) pp. 1857–1861.
29. C. H. Camp Jr., Y. J. Lee, J. M. Heddleston, C. M. Hartshorn, A. R. H. Walker, J. N. Rich, J. D. Lathia, and M. T. Cicerone, High-speed coherent Raman fingerprint imaging of biological tissues, *Nat Photon* 8 (2014) pp. 627–634.
30. T. C. Hollon, B. Pandian, A. R. Adapa, et al., Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks, *Nat Med* 26 (2020), pp. 52–58.
31. L. H. Higgins, H. G. Withers, A. Garbens, H. D. Love, L. Magnoni, S. W. Hayward, and C. D. Moyes, Hypoxia and the metabolic phenotype of prostate cancer cells, *Biochim Biophys Acta Bioenerg* 1787 (2009), pp. 1433–1443.
32. J. Lu, M. Tan, and Q. Cai, The Warburg effect in tumor progression: mitochondrial oxidative metabolism as an anti-metastasis mechanism, *Cancer Lett* 2 (2015), pp. 156–164.
33. R. Poorna, W. Chen, P. Qiu and M. T. Cicerone, Toward gene-correlated spatially re- solved metabolomics with fingerprint coherent Raman imaging, *JPC B* 127 (2023), pp. 5576–5587.
34. C. H. Camp Jr., Y. J. Lee, and M. T. Cicerone, Quantitative, comparable coherent anti- Stokes Raman scattering (CARS) spectroscopy: correcting errors in phase retrieval, *J Raman Spectrosc* 47 (2016), pp. 408–415.
35. P. Qiu, E. F. Simonds, K. D. Gibbs Jr., R. V. Bruggner, M. D. Linderman, K. Sach, G. P. Nolan, and S. K. Plevritis, Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE, *Nat Biotechnol* 29 (2011), pp. 886–891.
36. M. Makitalo and A. Foi, Optimal inversion of the generalized anscombe transformation for poisson-gaussian noise, *IEEE Trans Image Proc* 22 (2013), pp. 91–103.
37. M. T. Cicerone, A. K. Aamer, and Y. J. Lee, and E. Vartiainen, Maximum entropy and time-domain Kramers–Kronig phase retrieval approaches are functionally equivalent for CARS microspectroscopy, *J. Raman Spectrosc* 43 (2012), pp. 637–643.
38. Y. Liu, Y. J. Lee, and M. T. Cicerone, Broadband CARS spectral phase retrieval using a time-domain Kramers–Kronig transform, *Opt Lett* 9 (2009), pp. 1363–1365.
39. P. H. C. Eilers, A perfect smoother, *Anal Chem* 75 (2003), pp. 3631–3636.
40. P. H. C. Eilers, Unimodal smoothing, *J Chemom* 19 (2005), pp. 317–328.
41. A. A. Urbas and S. J. Choquette, Automated spectral smoothing with spatially adaptive penalized least squares, *Appl Spec* 65 (2011), pp. 665–677.
43. G. A. Lemieux and K. Ashrafi, Insights and challenges in using C. elegans for investigation of fat metabolism, *Crit Rev Biochem Mol* 50 (2015), pp. 69–84.
42. Y. B. Tzur, E. Winter, J. Gao, T. Hashimshony, I. Yanai, and M. P. Colaiácovo, Spa-tiotemporal gene expression analysis of the *Caenorhabditis elegans* Germline uncovers a syncytial expression switch, *Genetics* 210 (2018), pp. 587–605.
44. G. Medes, A. Thomas, and S. Weinhouse, Metabolism of neoplastic tissue. IV. A study of lipid synthesis in neoplastic tissue slices in vitro, *Cancer Res* 13 (1953), pp. 27–29.
45. J. V. Swinnen, P. P Van Veldhoven, L. Timmermans, E. De Schrijver, K. Brusselmans, F. Vanderhoydonc, T. Van de Sande, H. Heemers, W. Heyns, and G. Verhoeven, Fattyacid synthase drives the synthesis of phospholipids partitioning into detergent-resistant membrane microdomains, *BBRC* 302 (2003), pp. 898–903.
46. P. W. Hochachka, J. L. Rupert, L. Goldenberg, M. Gleave, and P. Kozlowski, Going ma-lignant: the hypoxia-cancer connection in the prostate, *BioEssays* 24 (2002), pp. 749–757.
47. K. Hiramatsu, T. Ideguchi, Y. Yonamine, S. Lee, Y. Luo, K. Hashimoto, T. Ito, M. Hase, J. W. Park, Y. Kasai, S. Sakuma, T. Hayakawa, F. Arai, Y. Hoshino, and K. Goda, High-throughput label-free molecular fingerprinting flow cytometry, *Sci Adv* 5 (2019), p. eaau0241.
48. Vernuccio, F., Broggio, E., S. Sorrentino, et al., Non-resonant background removal in broadband CARS microscopy using deep-learning algorithms, *Sci Rep* 14 (2024), p. 23903.

21

Spatial Multiplexed Immunofluorescence Imaging Exploring The Tumor Immune Microenvironment of Melanoma

Jun Jiang, Shankar Suman, Caitlin Ward, Raymond M. Moore, Zachary C. Fogarty, Alexey Leontovich, Wendy K. Nevala, Chathu L. Atherton, Jill M. Schimke, and Svetomir N. Markovic

21.1 Introduction

The tumor-immune microenvironment (TIME) refers to the complex interactions between tumor cells and the immune cells within the tumor's local environment. It profoundly influences tumor development, growth, and metastasis. The interactions of tumor cells with T cells, B cells, macrophages, and dendritic cells can either promote or inhibit tumor progression, as well as shape the tumor's behavior against immunotherapy. Research on characterizing and understanding the TIME holds significant promise for identifying new therapeutic targets and refining existing treatment approaches, thereby improving patient outcomes. For many years, microscopy was used to study cellular structures and the function of a few visual characteristics that were directly correlated to tumor spatial organization rather than dealing with homogenized tumor samples. Multiplexed immunofluorescence (MxIF) spatial imaging has emerged as a valuable tool in melanoma immuno-oncology research in recent decades.

Although conventional chromogenic or fluorescent method for antigen detection is a widely used diagnostic technique in tissue pathology, it is insufficient to reveal the details in the TIME due to its limited capacity to label only one marker per tissue section [1]. However, MxIF allows the simultaneous visualization and quantification of multiple biomarkers (proteins or cell markers) within a single tissue section and, therefore, facilitates comprehensive characterization of the TIME by simultaneously assessing multiple immune cell types, their activation status, distribution patterns, and interactions within the tumor. The general data acquisition and analysis workflow for MxIF typically involves the following steps (Figure 21.1): (1) In the sample preparation step, tissues are collected, fixed, and embedded in paraffin or frozen sections. Then, antigen retrieval methods such as heat-induced epitope retrieval or enzymatic digestion are commonly used to expose antigens for antibody binding. (2) In the antibody preparation step, multiple antibody clones are selected according to the study design. Different fluorophores are conjugated to the antibody, which enables visualization. (3) Raw images are obtained through cyclic multiplex staining, scanning, and bleaching iterations (also known as imaging rounds). In each iteration, the antibody of interest is conjugated with fluorescent tags, scanned, and then subjected to dye inactivation (bleaching) to remove the fluorescent signal. High-resolution imaging systems, such as fluorescence and bright-field microscopes equipped with filter sets for each fluorophore or chromogen, capture images of the stained tissue sections. (4) Autofluorescence (AF) is removed from the obtained raw images, and image channels are assembled into standard file formats, such as OME-TIFF. (5) Image quality is assessed using manual spot-checking or automatic methods to ensure no apparent tissue damage or imaging artifacts [2]. (6) Within the processed images, cell boundaries are identified using automatic cell segmentation methods, such as DeepCell, Stardist, and CellPose [3–5]. (7) Based on cell segmentations, cells are quantified by measuring the protein marker signal level within different components of cells (nuclei and cytoplasm). The quantified cells are saved as a tabular dataset, in which each row is a cell, and each column is a statistic metric of a marker. (8) Supervised or unsupervised methods are used to classify cells into several subtypes. (9) The classified cells are analyzed to extract meaningful information about biomarker expression patterns, cellular interactions, and other relevant parameters such as patient outcome and recurrence. Statistical analysis and data visualization techniques are employed to aid biology interpretation.

Given the complexity of MxIF images, there are many challenges during routine slide preparation and digital processing that may affect the accuracy and reliability of the results. One of the significant challenges is autofluorescence from tissue components, such as lipofuscin or collagen, which can interfere with specific signal detection, reducing the signal-to-noise ratio and affecting the accuracy of target identification. Although there are postprocessing steps to calibrate this imaging problem, it can cause truncated protein signals [6, 7]. Preparing the right antibody concentration and applying it on slides are also crucial, as nonspecific binding or cross-reactivity of antibodies with unintended targets can lead to false-positive signals through variations in antibody penetration, staining efficiency, or antigen expression levels within the tissue, which can cause variability in signal intensity across different regions. From computational aspects, it is challenging to handle large amounts of high-resolution image data generated during MxIF experiments, as it requires efficient storage, management, and processing. Additionally, the downstream analysis, such as cell segmentation, classification, and neighborhood analysis, could be confounded by the artifacts and noise introduced in previous

DOI: 10.1201/9781003449942-24

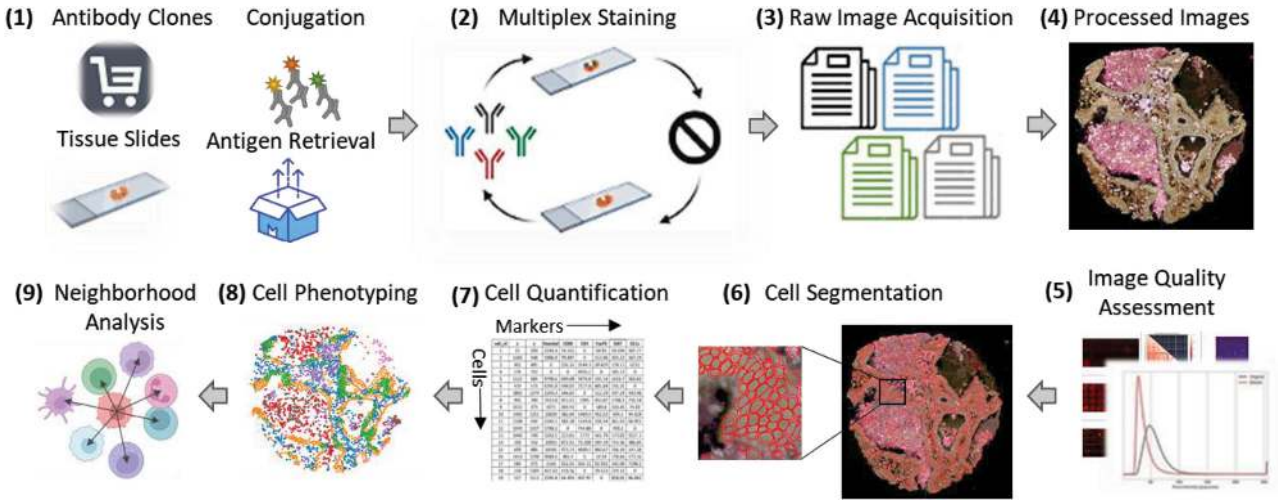


FIGURE 21.1 Workflow of multiplexed immunofluorescence data acquisition and processing.

steps. To mitigate these technical difficulties, combinational strategies need to be employed for experimental optimization, image acquisition techniques, computational methods, and rigorous validation protocols to ensure the accuracy and reliability of MxIF image analysis.

Our group progresses in spatial immunology, emphasizing understanding the heterogenic tumor immune milieu, mainly through tumor-centric cellular neighborhood (TCCN) analysis. We focused on sharing our experience in investigating the TIME using MxIF. Specifically, studies on lymph node biopsies from patients with regionally metastatic melanoma who underwent subsequent anti-PD1 therapy obtained MxIF images with 45 protein markers. We achieved accurate cell segmentation and classification results as the basis for quantitative assessments to investigate spatial interactions between the immune system and melanoma tumors. Through cell neighborhood and community analysis, we found dendritic cell enrichment can be positively related to patient treatment response and non-recurrence [8].

21.2 Materials

21.2.1 Study Cohort

Out of an initial group of 33 untreated patients diagnosed with Stage III melanoma and eligible for anti-PD1 immunotherapy, 25 patients were chosen for analysis using MxIF in this study. All patient biospecimens and clinical data were collected following the principles of the Declaration of Helsinki, with approval granted by the Institutional Review Board at the Mayo Clinic.

21.2.2 Slide Preparation

A singular 5 μm section of formalin-fixed paraffin-embedded (FFPE) tissue was prepared from each patient for the MxIF workflow. The slide preparation process involved encompassing deparaffinization, rehydration, and permeabilization, following previously outlined procedures [9].

21.2.3 Antibody Panel Design

We designed a specialized panel consisting of 45 antibodies to characterize distinct subsets of tumor, immune, and stromal cells, ultimately categorizing them into ten unique classification phenotypes [9].

21.2.4 Cyclic MxIF Staining

Cyclic MxIF methods were used to evaluate multiple antibodies on a single tissue section. Within each iterative round, tissues were stained with fluorescent antibodies, digitally scanned, and dye-inactivated. During staining, slides were placed in a humidified dark chamber at room temperature for one hour. Subsequently, slides were washed in PBS and coverslipped for image scanning. After imaging, slides were decoverslipped in PBS and subjected to dye inactivation using a sodium bicarbonate solution, following previously established procedures [10]. Post-dye inactivation, slides were rewashed in PBS and coverslipped for imaging.

21.3 Methods

21.3.1 Image Acquisition and Preprocessing

The INCell Analyzer 2500HS (Leica Microsystems) was utilized to capture images. To match the scanner settings, fields of view (FOVs) were selected from a virtual H&E image by a board-certified pathologist, encompassing areas of tumor core, tumor-immune interface, and lymphoid tissue. The annotated FOVs were sequentially imaged at the 20X objective [9]. The AF removed images were standardized based on the minimum exposure time. The processed image channels were joined into a standard OME.TIFF file format [11].

21.3.2 Image Quality Checking

MxIF images were obtained through multiple staining, scanning, and bleaching iterations. In each iteration, the antibody of the marker was conjugated with fluorescent tags, scanned, and then washed away. Considering that the DAPI stain in each imaging iteration provides references for tissue localization and auto-focusing for the imaging system, we developed a DAPI channel-referenced evaluation method for automatic MxIF image quality checking (QC) [2]. By comparing DAPI signals in all the imaging iterations, our [automated or programmatic] QC modular can detect tissue damage and halo-like artifacts to distill the dataset for downstream analysis.

21.3.3 Cell Segmentation and Quantification

We adapted the pixel classification-based workflow to achieve both nucleus and cytoplasm segmentation. Two major tools were employed during cell segmentation: Ilastik [12] and CellProfiler [13]. Ilastik was used to encode individual pixels within standard OME.TIFF from the previous workflow into the probability of three classes to differentiate nuclear, cytoplasm, and background [12, 14]. Experts' annotations were used to train the supervised model, which predicts the uncertainty of a single-pixel belonging to the nucleus, cytoplasm, or background within unlabeled image crops. The trained model was applied to every FOV in the dataset, generating pixel classification probabilities, which were then input into CellProfiler for algorithmically constrained object propagation and the generation of labeled masks for whole cells and nuclei. The obtained cell segmentation masks were imported into QuPath [15] with the full panel OME.TIFF MxIF images, in which the morphological and protein marker signal features for each cell were calculated and exported for further cell subtyping.

21.3.4 Cell Subtyping

Depending on the downstream analysis tasks, our cell subtyping approaches were divided into two tasks: phenotypic classification and functional classification. The phenotypic classification method aims to follow the hierarchical nature of cell development. In contrast, functional classification focuses on capturing the functional state of a cell determined by the known biomarker, such as PD1. To provide training data for supervised learning, human experts were invited to annotate 11 types of cells (tumor cell, macrophage, dendritic cell, stromal cell, B cell, cytotoxic T cell, helper T cell, neutrophil, unclassified immune, regulatory T cell, and monocyte) within QuPath for FOVs from tumor core [8]. The obtained annotations were used to develop models for phenotypic and functional classification. More details can be found in our previous publications [8, 9].

21.3.5 Cell Neighborhood Analysis

To investigate cell-cell interactions within the TIME, cells and their neighborhoods were clustered into communities based on both cell-cell proximity and their phenotypes. Thus, two parameters were defined to implement method: the K number of cells to aggregate, and C number of clusters to cluster the cells. Cell communities were assessed at various resolutions (k = 5, 25, 50, 75, 100) mimicking different ranges of cell neighborhoods in distinct macro-architectures within the tumor tissue, such as primarily tumor-dominant communities vs. tumor communities with either lymphocyte or myeloid dominant immune cell subsets.

21.3.6 Statistical Models for Spatial Image Analysis of MxIF Data

Spatial generalized linear mixed models were used to test for significant differences in cell-cell interactions within the TIME by response to immunotherapy while accounting for within-subject correlation, batch effects resulting in image-to-image variability, and spatial correlation between nearby cells [16]. Models were implemented at the tumor cell level and considered colocalization outcomes, such as the proportion of CD8+ cells in each tumor's microenvironment.

21.4 Results

21.4.1 DAPI-Referenced QC Identifies Low Qualities That Harm Accurate Cell Phenotyping

To visually check how the quality issues within MxIF impact the cell distribution and downstream analysis, we used UMAP [17] to project all the cells from high-dimensional feature space to two-dimensional points. By mapping the cells back to the original location and overlaying them with the image, we found that MxIF quality issues have a limited impact on differentiating tumor and stroma cells. Still, for the FOVs passed vs. failed in our QC (denoted as HQ and LQ, respectively), the LQ group presented dramatically higher (more than 20%) stroma cell density than the HQ group in a local region of the projection map [2].

21.4.2 Cyclic MxIF Enables Cell Detection and Phenotyping within Melanoma TIME

Within metastatic lymph node resected tissue, the delineated tumor, immune, and stromal climate within this specific TIME were compared with our automatic detection and phenotyping methods. For visual assessment, classification labels for both cellular phenotypes and functional expressions were returned to QuPath and merged with the spatial coordinates of each cell. The tumor core spatial landscape was differentiated at various levels of resolution, including the tissue macro-architecture and the phenotypic and functional status of diverse cell subsets (Figure 21.2A–B). For quantitative phenotype evaluation, a confusion matrix was used to summarize the performance by comparing model-predicted labels to the true labels. Our latest results demonstrated that the XGBoost model was outperformed in cell phenotyping (Figure 21.2C). The model has reached higher than 90% classification accuracy for most cell types. Dendritic cells and monocytes were the two subtypes that were most likely to be misclassified. Monocytes were likely to be misclassified as macrophages.

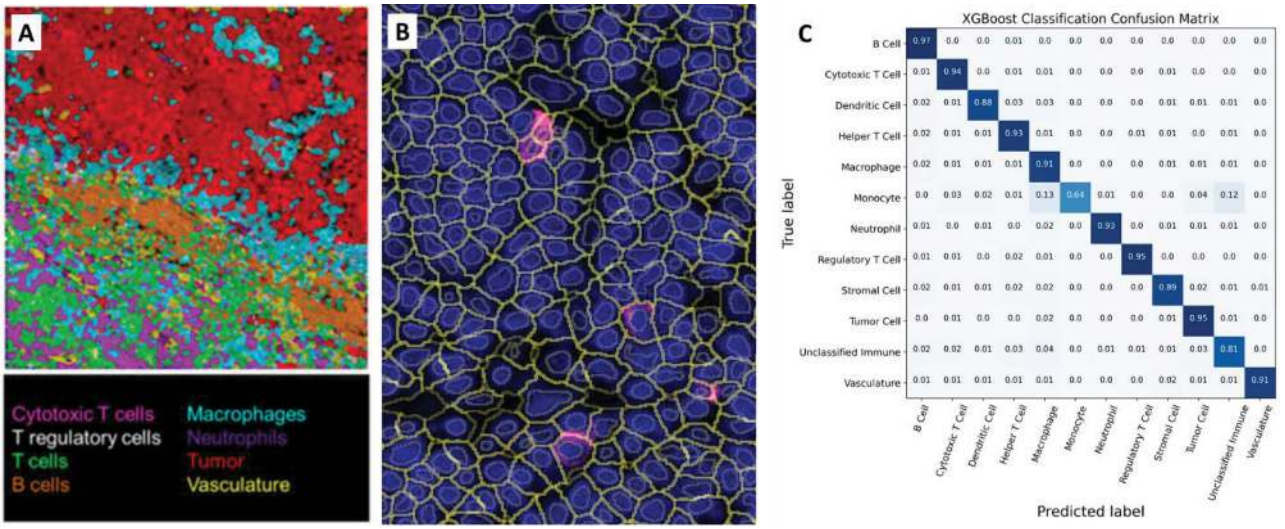


FIGURE 21.2 Cell detection, phenotyping, and cellular composition within the TIME. (A) Visualization of cell detection and phenotyping. Different cell types were colored with a specific color. (B) Zoom-in view of visualization. (C) Confusion matrix of cell classification.

21.4.3 Phenotypic and Spatial Paracrine Interactions within Cell-Centered Neighborhoods

To interrogate paracrine interactions between individual cells and their nearest spatial neighbors, we adapted the method of cell neighborhood analysis by modeling a micro-niche surrounding each tumor cell in an FOV, termed a "tumor cell centric neighborhood," or TCCN. Two criteria were defined for a valid TCCN: (i) the center (index) cell must be classified as a tumor cell, and (ii) the neighborhood must be occupied by at least one additional non-index cell [9]). By these two criteria, the optimal TCCN size (~25 μm) was obtained by investigating the distribution of TCCN counts that satisfy both criteria across various radius sizes [9].

Within TCNN, we used the percentage of other types of cells as a biologically derived spatial metric to describe the regional characteristics of cancer tissue. Neighborhood composition was compared between patient cohorts by evaluating the percentage of neighborhoods populated with immune cell phenotypes classified per patient. In TCCN, patients who did not experience recurrence had a higher proportion of TCCN populated with most immune cell subsets (CTL or cytotoxic CD8+ T cell, dendritic cell, helper T cells, and Tregs) than patients who did not recur. Notably, B cells and macrophages had an increased representation in neighborhoods of patients who experienced recurrence [8]. Similar to TCCN, we also characterized the "cytotoxic T cell-centric neighborhood" (CTCN), with CTL at the center of the cellular neighborhood [8].

21.4.4 Evaluation of Spatial Tumor-Centric Community to Comprehend Tumor Cell Heterogeneity

One challenge of failure for cancer therapy is that tumor cells are heterogeneous, which we can define with spatial analysis with MxIF. Intra-tumoral heterogeneity, or genetic variation among cancer cells within a tumor and their spatial interactions with surrounding stromal and/or immune cells, may be key to unlocking long-standing mysteries in cancer biology and improving the identification and management of particular cancer subtypes. Recent developments in technology have demonstrated the dynamic and molecular diversity of both malignant and non-malignant cells, including stromal and immune cells, with functional ramifications related to the features of cancer.

21.4.4.1 Macro-Architecture Assessment of the Tumor Core Region Using Cellular Communities

The variety of cell types that were spatially grouped surrounding each cell, producing a relational structure that defines the environment of a cell, is referred to as a "community." The number of cells closest to one another determines the relationship between those cells (Figure 21.3A). The majority of multiplex imaging investigations have used tissue microarray analysis to resolve clinical heterogeneity across a sizable patient cohort. We decided to examine many FOVs inside a single, spatially preserved lymph node biopsy. We drew on this pioneering work in spatial mapping to sample the various locations of the tumor core, margin, and lymphoid tissue. A more comprehensive examination of the TIME composition concerning the tumor core, margin, and lymphoid tissue regions was obtained by sampling FOVs in metastatic lymph nodes (Figure 21.3). The single-cell composition in the tumor core revealed proven predictive associations between cytotoxic T cells and clinical response to immunotherapy.

21.4.4.2 Immunosuppression in the Context of the TCCN

The histological evaluation of PD1 and PDL1 expression in tumors are considered important parameters for

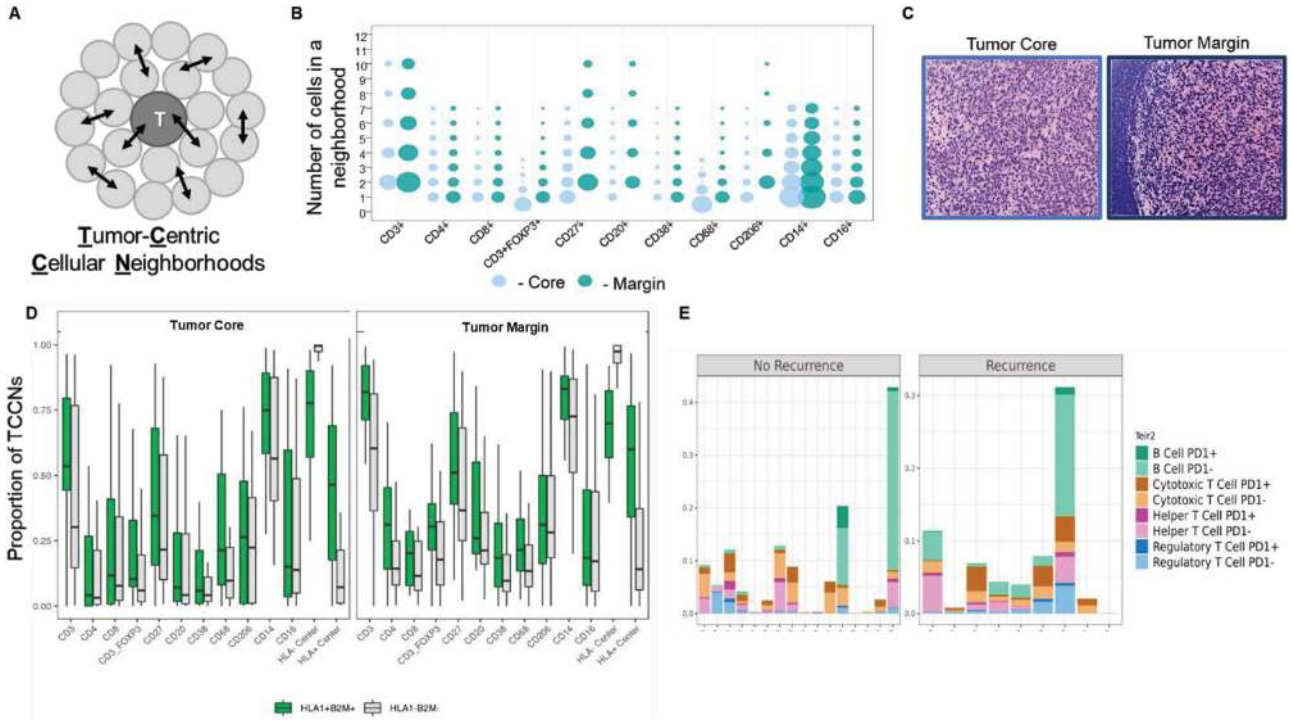


FIGURE 21.3 The tumor-immune climate's TCCN contextualizes the functional activities for immune subversion. (A) Tumor neighborhood analysis by TCCN approach. (B) The MxIF overlay shows the expression of different immune cells in the same FOVs as well as the tumor-immune landscape. (C) Tumor core and tumor margin analysis with MxIF. (D) Box plot comparing tumor core and tumor margin in proportion of TCCN. (E) Immune cell distribution in recurrence and no recurrence as neighbors in the TCCN analysis.

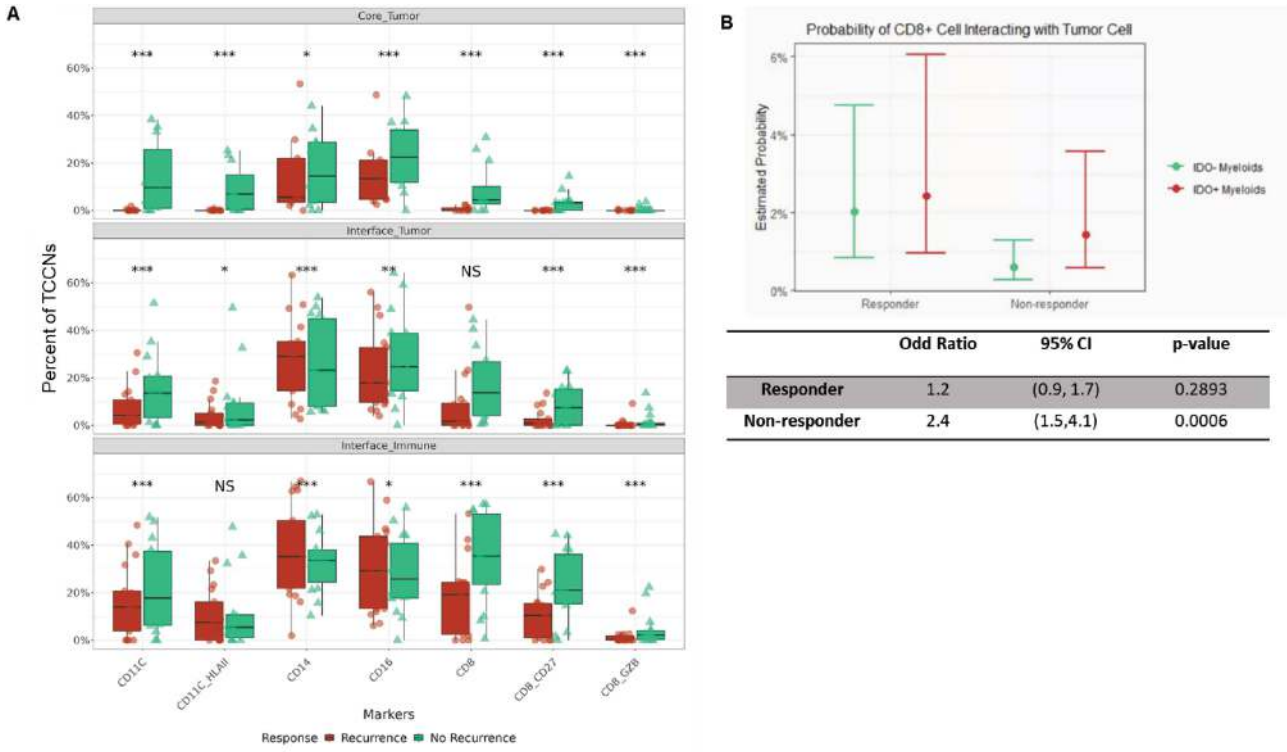
immunotherapy. We observed that tumors expressing PDL1 were only expressed in a small percentage of cells. It was observed in a small subset of tumor FOVs obtained from patients regardless of the recurrence status. TCCNs were imposed onto the example FOV map, each consisting of at least one tumor-infiltrating lymphocyte (TIL) expressing PD1. B cells with PD1 expression, T helper cells, and Tregs were more common in the TCCN of melanoma patients. In the TCCN of patients who had B cells with PD1 expression, T helper cells and Tregs had an enriched frequency of recurrence. Those patients also had high PD1-expressing CTLs in their TCCNs. Moreover, our finding also suggests that the PD1 expression on TIL alone is insufficient to predict clinical responses to anti-PD1 therapy.

21.4.5 Spatial Evaluation of Tumor-Immune Interactions by the Impact of Immunotherapy in Metastatic Melanoma Lymph Nodes

To examine two clinically divergent outcomes and the related tissue landscapes, we chose patients in our study who had either demonstrated a clinical response to immunotherapy or had not responded well. We assessed the potential use of TCCN immunological signatures as a predictive factor in the context of pretreatment. Using the TCCN as the analytic unit of paracrine interaction, our work offers a framework for physiologically relevant resolution of geographical, cellular, and clinical heterogeneity. Our findings indicate that the tumor cells' anatomical area and HLA-I status affect the paracrine immune signatures within TCCNs. These further refine the spatial landscape of the TIME beyond single-cell measurement. The prognostic importance of macrophages and cytotoxic T cells in the tumor core and new spatially relevant interactions between immune cell subsets and malignancies were investigated. This may influence future clinical responses to immunotherapy by demonstrating the MxIF INCell's feasibility as a single-cell resolving bioimaging platform.

21.4.6 Myeloid Cells Impact the Functionality of T Cells in the Tumor Microenvironment

Melanoma-derived factors were responsible for developing premetastatic niche (PMN) in sentinel lymph nodes (SLN). This mechanism creates a favorable environment for nodal metastasis by inducing an immunological suppressive mechanism. The molecular features of PMN were not yet fully understood, but it has been found that dysfunctions of myeloid cells were crucial components for immune subversion in SLN. Melanoma secretory factors govern the immune escape mechanism by inducing immunosuppressive macrophages [18, 19]. The PMN development is a critical in suppressing anti-tumor immunity [20–23] and directly transforming monocytes to acquire immunosuppressive macrophage features within the PMN [24]. Mounting evidence shows that melanoma-derived factors upregulate PD-L1 and indoleamine 2,3- dioxygenase (IDO)-1 on myeloid cells to transform them into myeloid-derived suppressive cells



	Odd Ratio	95% CI	p-value
Responder	1.2	(0.9, 1.7)	0.2893
Non-responder	2.4	(1.5,4.1)	0.0006

FIGURE 21.4 (A) The pilot study analysis with MxIF data in 11 patients indicates significant deregulation of myeloid cells (CD11c, CD11c+ HLAII, CD14, CD16) and CD8-T cells in the SLNs of non-responders or recurred melanoma patients compared to responder cases with immunotherapy. (B) In non-responders, the odds of CD8+/tumor cell interactions were 2.4 times higher when IDO+ myeloid cells present in the TIME of the responder.

(MDSC) and inhibit T-cell function [25, 26]. Spatial analysis on SLNs of responder and non-responder of immunotherapy revealed upregulation of the CD14+ interface of tumor and downregulated CD8+T cells in the tumor core and interface region. Our spatial analysis in SLNs also indicates that IDO with myeloid cells in the tumor cell vicinity significantly increases the odds of tumor-CTL cell interactions in non-responder patients. However, this was not found to be significant in responder patients (Figure 21.4B).

Our research group has previously shown the dendritic cells tolerization by the loss of maturation and differentiation factors [27]. We also analyzed the role of the tumor-promoting macrophages in the immune subversion through CTL-macrophage interaction [8]. In our study, we employed TCCN and CTCN approaches to evaluate myeloid markers like CD14, CD15, CD16, CD11b, CD68, CD11c, and other more classified myeloid markers. Our findings indicate a significant difference in the myeloid panel between responders and non-responders, which are known to manifest T cell activity.

21.5 Discussion

Compared to the study within tissue FOVs, we plan to scale our analysis to the whole slide level, which enables us to evaluate the TIME without human biases. Although we have extended the analysis to cell neighborhoods within different anatomic subregions, most of our results were based on cell-centric neighborhood analysis, such as TCCN and CTCN. This method maximizes the biomedical interpretation of the statistical results, though it was heuristic in parameter selection. Increasingly, studies were leveraging the computational graph-based method to interrogate cell neighborhoods. In this cell, spatial patterns were formulated into a cell graph with each node as a cell and each edge defined by the proximity between cells. Then, graph theories and graph neural networks (GNNs) can be used to improve cell phenotyping [28, 29] and recognize spatial patterns within the TIME [30–32]. According to our literature survey, deep GNNs carry scalable AI potential to make breakthrough discoveries in cancer spatial biology.

Our study demonstrates that MxIF is an important platform for resolving patterns of paracrine interactions within the TIME of the metastatic melanoma lymph node about the clinical response to PD1 immunotherapy. Our findings indicate that the anatomical area and indexed tumor cells affect the paracrine immune signatures within TCCNs, and they further refine the spatial landscape of the TIME beyond single-cell measurement. Proving the MxIF INCell's feasibility as a single-cell resolving bioimaging platform may offer a paradigm for physiologically relevant resolution of geographical, cellular, and clinical heterogeneity using the TCCN approach. We used multiplex imaging methods to examine the geographical dynamics and phenotypic variety of TIL subsets in the context of metastatic tumors. Using our 45-antibody panel design in conjunction with a machine-based modeling technique, we achieved high accuracy in single-cell segmentation and a 10-class phenotypic categorization model. To assess the single-cell TIME at different resolutions, we also modified the

spatial tools that were already in place. To situate immunotherapy in this dynamic tissue environment, it is imperative to investigate cellular network interactions beyond pairwise connections. Our novel findings show that the myeloid compartment may more directly modulate the CTL-tumor interaction in the LN TIME. Even though we evaluated 416 tumor core FOVs that pathologists classified for a group of 25 patients, our dataset is not yet experimental. To more accurately reflect traditional TIL evaluations in the context of primary tumors, we decided to concentrate on the tumor-enriched areas of the metastatic LN. Nevertheless, this dramatically lowers the number of immune cell observations per FOV, and additional research will be required to confirm these results in tumor regions similar to ours and other histopathologically significant areas of the TIME. In line with earlier findings [33], nine individuals (or 36%) in this sample had a melanoma recurrence within a year after starting anti-PD1 medication. Our data demonstrate the variety of cellular interactions that contribute to heterogeneous responses to immunotherapy in addition to the intended effector-target mechanism of action, which goes beyond the established associations between immune-infiltrated tumors and favorable responses to immune checkpoint inhibitors [34].

Our findings are significant because they provide a more thorough assessment of TILs and spatial immunological markers in the metastatic context and information on the likelihood of recurrence after immunotherapy. Further research is needed to confirm our results in broader cohorts. In human cutaneous melanoma, micrometastasis at tumor-draining lymph nodes promotes distant metastasis and mortality. Melanoma-cell-derived mediators contribute to immunological tumor escape and support metastatic tumor colonization in the premetastatic niche of the tumor-draining lymph nodes [20]. The MxIF platform is a robust technology that can recognize spatial alteration even at the molecular level. Using multiplexed bioimaging technologies is crucial in improving our understanding of the spatial TIME. However, they have limitations. One major obstacle is still cell classification. Cell types not quantified in the abundance plot for a given sample may need to be more represented. Furthermore, optimization of the antibody panel design in subsequent research, particularly the augmentation of phenotypic indicators and the integration of functional markers such as PDL1, would facilitate fine-tuning immune cell and tumor subgroup classification.

21.6 Future Trends

With the rapid development of MxIF, increasing numbers of platforms are available to capture cellular profiles. There is a trend that combines the data from multiple platforms to obtain enriched information from cancers, as each platform has its unique strength in a particular aspect. For example, CellDive provides more than 60 protein biomarkers, while NanoString CosMx is advanced in quantifying RNA levels in cancer tissues.

We recently investigated the feasibility of integrating datasets from CosMx and CellDive platforms to obtain the full spectrum of cell profiles. To this end, cell location correspondence was established by aligning DAPI images from both platforms using human-labeled landmarks as reference. Regional signal concordance was evaluated using a sliding window approach to iterate through the aligned images from CellDive and CosMx. The correlations of cellular biomarkers and expression were quantified using averaged mutual information (AMI) of signal distributions within the sliding windows. We found that immune-related proteomics and transcriptomics from CosMx and CellDive presented regional-level concordance, especially in the immune-cell enriched subregions.

REFERENCES

1. Tan, W.C.C., et al., Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Communications*, 2020. 40(4): p. 135–153.
2. Jiang, J., et al., Multiplex immunofluorescence image quality checking using DAPI channel–referenced evaluation. *Journal of Histochemistry & Cytochemistry*, 2023. 71(3): p. 121–130.
3. Bannon, D., et al., DeepCell Kiosk: scaling deep learning–enabled cellular image analysis with Kubernetes. *Nature Methods*, 2021. 18(1): p. 43–45.
4. Stevens, M., et al., StarDist image segmentation improves circulating tumor cell detection. *Cancers*, 2022. 14(12): p. 2916.
5. Stringer, C., et al., Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, 2021. 18(1): p. 100–106.
6. Neumann, M. and D. Gabel, Simple method for reduction of autofluorescence in fluorescence microscopy. *Journal of Histochemistry & Cytochemistry*, 2002. 50(3): p. 437–439.
7. Baharlou, H., et al., Digital removal of autofluorescence from microscopy images. *bioRxiv*, 2019: p. 566315.
8. Maus, R.L., et al., Quantitative spatial evaluation of tumor-immune interactions in the immunotherapy setting of metastatic melanoma lymph nodes. *Frontiers in Immunology*, 2022. 13: p. 1024039.
9. Maus, R.L., et al., Resolving the heterogeneous tumor-centric cellular neighborhood through multiplexed, spatial paracrine interactions in the setting of immune checkpoint blockade. *Cancer Research Communications*, 2022. 2(2): p. 78–89.
10. Gerdes, M.J., et al., Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of Sciences*, 2013. 110(29): p. 11982–11987.
11. Leigh, R., et al., OME files-an open source reference library for the OME-XML metadata model and the OME-TIFF file format. *BioRxiv*, 2016: p. 088740.
12. Berg, S., et al., Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods*, 2019. 16(12): p. 1226–1232.
13. McQuin, C., et al., CellProfiler 3.0: next-generation image processing for biology. *PLoS Biology*, 2018. 16(7): p. e2005970.
14. Bai, X., F. Zhou, and B. Xue, Image enhancement using multi scale image features extracted by top-hat transform. *Optics & Laser Technology*, 2012. 44(2): p. 328–336.
15. Bankhead, P., et al., QuPath: open source software for digital pathology image analysis. *Scientific Reports*, 2017. 7(1): p. 1–7.

16. Kondo, A., et al., Highly multiplexed image analysis of intestinal tissue sections in patients with inflammatory bowel disease. *Gastroenterology*, 2021. 161(6): p. 1940–1952.
17. McInnes, L., J. Healy, and J. Melville, Umap: uniform manifold approximation and projection for dimension reduction. *arXiv Preprint* arXiv:1802.03426, 2018.
18. Wang, T., et al., Melanoma-derived conditioned media efficiently induce the differentiation of monocytes to macrophages that display a highly invasive gene signature. *Pigment Cell & Melanoma Research*, 2012. 25(4): p. 493–505.
19. Wang, H., et al., Pro-tumor activities of macrophages in the progression of melanoma. *Human Vaccines & Immunotherapeutics*, 2017. 13(7): p. 1556–1562.
20. Suman, S. and S.N. Markovic, Melanoma-derived mediators can foster the premetastatic niche: crossroad to lymphatic metastasis. *Trends Immunology*, 2023. 44(9): p. 724–743.
21. Garcia-Silva, S., et al., Melanoma-derived small extracellular vesicles induce lymphangiogenesis and metastasis through an NGFR-dependent mechanism. *Nature Cancer*, 2021. 2(12): p. 1387–1405.
22. Rincon-Riveros, A., et al., Regulation of antitumor immune responses by exosomes derived from tumor and immune cells. *Cancers (Basel)*, 2021. 13(4).
23. Leary, N., et al., Melanoma-derived extracellular vesicles mediate lymphatic remodelling and impair tumour immunity in draining lymph nodes. *Journal of Extracellular Vesicles*, 2022. 11(2): p. e12197.
24. Morrissey, S.M., et al., Tumor-derived exosomes drive immunosuppressive macrophages in a pre-metastatic niche through glycolytic dominant metabolic reprogramming. *Cell Metabolism*, 2021. 33(10): p. 2040–2058 e10.
25. Fleming, V., et al., Melanoma extracellular vesicles generate immunosuppressive myeloid cells by upregulating PD-L1 via TLR4 signaling. *Cancer Research*, 2019. 79(18): p. 4715–4728.
26. Arkhypov, I., et al., HSP90alpha induces immunosuppressive myeloid cells in melanoma via TLR4 signaling. *The Journal for Immunotherapy of Cancer*, 2022. 10(9).
27. Maus, R., Jakub, J., Nevala, W., Christensen, T., Noble-Orcutt, K., Sachs, Z., Hieken, T., Markovic, S. Human melanoma-derived extracellular vesicles regulate dendritic cell maturation. *Frontiers in Immunology*, 2017. 8: p. 358. 10.3389/fimmu.2017.00358.
28. Brbić, M., et al., Annotation of spatially resolved single-cell data with STELLAR. *Nature Methods*, 2022. 19(11): p. 1411–1418.
29. Kim, J., et al., Unsupervised discovery of tissue architecture in multiplexed imaging. *Nature Methods*, 2022. 19(12): p. 1653–1661.
30. Cang, Z., et al. Scan-IT: domain segmentation of spatial transcriptomics images by graph neural network, in *BMVC: proceedings of the British Machine Vision Conference. British Machine Vision Conference*. 2021. NIH Public Access.
31. Yuan, Y. and Z. Bar-Joseph, GCNG: graph convolutional networks for inferring gene interaction from spatial transcriptomics data. *Genome Biology*, 2020. 21(1): p. 1–16.
32. Wen, H., et al. Graph neural networks for multimodal single-cell data integration, in *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22)*. 2022. Association for Computing Machinery, New York, NY, USA. p. 4153–4163. https://doi.org/10.1145/3534678.3539213
33. Weber, J., et al., Adjuvant nivolumab versus ipilimumab in resected stage III or IV melanoma. *New England Journal of Medicine*, 2017. 377(19): p. 1824–1835.
34. Bruni, D., H.K. Angell, and J. Galon, The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy. *Nature Reviews Cancer*, 2020. 20(11): p. 662–680.

22

Spectroscopic Optical Coherence Tomography for In Situ *Analysis of Colonic Epithelium*

Jatin Roper, Wesley Kendall, and Adam Wax

22.1 Introduction

Colorectal cancer (CRC) is the second leading cause of cancer-related mortality, with a lifetime risk of approximately 4% to 5% [1], despite the widespread use of screening colonoscopy and other colorectal cancer screening modalities that have reduced mortality rates [2, 3]. Colonoscopy aims to remove precancerous polyps (i.e., adenomas) by visual identification, and thus prevent CRC. However, less than three quarters of eligible adults receive regular screening [4] and the rate of incidence has not diminished further in the past decade among adults ages 50–64 [5]. Even more concerning is an apparent rise in CRC incidence among those aged 49 years or less [5]. In addition, adenoma detection by colonoscopy is limited by several factors, including the lack of quantitative biomarkers and lack of real-time feedback aside from visual guidance. Additionally, the limited capacity to meet growing demand and the cost of endoscopic procedures can result in reduced access in underserved populations. Lower cost alternative methods such as fecal immunochemical testing (FIT), multi-DNA stool testing, and newly available blood-based testing offer alternatives to colonoscopy but with reduced sensitivity [6].

There is also a need for surveillance of colorectal tissues in patients with increased risk of CRC. Diseases of the colon such as inflammatory bowel disease (IBD), including ulcerative colitis (UC) and Crohn's disease (CD), require regular surveillance using endoscopy with up to 40–50 random or targeted biopsies per procedure. Patients with IBD are at an increased risk of developing CRC, up to six times compared to the general population [7, 8]. The extent of involvement and duration of symptoms both correlate with increased CRC risk [9] and endoscopic surveillance at one- to two-year intervals is recommended [10]. In spite of this increased risk, there is significant debate on the benefit of surveillance due to the cost and relatively low likelihood of discovering dysplasia, a precancerous condition that is the primary marker used to determine the need for colectomy [11].

Once dysplastic tissues are found, there is an additional need to improve imaging of colonic polyps to enhance identification of lesion margins and to facilitate planning of therapy (i.e., biopsy, choice of polypectomy technique such as endoscopic mucosal resection or endoscopic submucosal dissection), or referral for surgery [12].

Recent innovations in endoscopic examination aim to improve efficiency by decreasing procedure time and its associated risks, as well as minimizing pathology costs by potentially leaving diminutive polyps (< 5mm) in place. Such small lesions may account for up to 80% of polyps resected during colonoscopies [13]. If a leave-in-place strategy were implemented, it could realize cost savings of millions of dollars [14]. Initially, strategies have focused on adopting novel imaging approaches [15]. More current efforts seek to employ deep learning strategies to identify low-risk colon polyps in imaging data [16, 17]. The concept is that such lesions could either be left in place or resected and simply discarded instead of incurring the expense of histopathology processing.

Here we describe our efforts to develop a new diagnostic modality within colonoscopy, by using the spectroscopic capabilities of optical coherence tomography (OCT). Since its introduction over 30 years ago [18], OCT has seen extensive clinical use for retinal imaging [19]. However, the high cost of instrumentation is often a limiting factor in expanding OCT to new clinical applications. A research OCT system may cost upwards of $100,000 and require extensive modification to adapt it to a clinical need such as colonoscopy. With the recent introduction of the first low-cost OCT system in developmental work at Duke University [20], the cost of exploring new applications has dropped significantly.

22.2 Materials

22.2.1 Animal Studies

The studies described later used the azoxymethane (AOM) rat carcinogenesis model, a well characterized and established model for colon cancer research and drug development [21], and two types of genetically modified mice [22, 23], each representing different stages in the development of colorectal cancer. Animals in the AOM study received intraperitoneal (IP) injections of AOM > 90% pure with a molar concentration of 13.4 M (Sigma, St. Louis MO) at a dose level of 15 mg/kg body weight, once per week, for two consecutive weeks (two doses per animal). The remaining ten animals received saline by IP and served as the control group. At 4, 8, and 12 weeks after the completion of the dosing regimen, the animals (ten AOM-treated and three or four saline-treated rats per time point) were sacrificed by CO_2 asphyxiation. The colon tissues were harvested, opened longitudinally, and washed with saline. Then, the tissues were split into four to five different segments, each with a length of 3–4 cm, and imaged using SOCT.

DOI: 10.1201/9781003449942-25

The two mouse models were produced using genetic modification to inactivate the *Apc* tumor suppressor gene (hereby referred to as APC mice), while the other (hereby referred to as AKP mice) also has additional mutations to activate the *KrasG12D* oncogene and inactivate the *Trp53* tumor suppressor gene to represent a more advanced adenocarcinoma. Localized tumors were induced with injection of 4-hydroxytamoxifen into the colon of each mouse. After allowing tumor growth for two to four weeks, the colons were extracted and the tissue imaged with OCT *ex vivo*. Three biopsies were taken from each classification of colon tissue: three benign tissues from control *Apc*-floxed mice, three adenomas from APC mice, and three adenocarcinomas from AKP mice.

22.2.2 Human Tissue Study

The human clinical study described later was performed at the Duke University Hospital Endoscopy Unit with approval from the Duke University School of Medicine IRB. The study was designed to include a diverse set of adenomas with different pathologies, in order to characterize the range of optical coefficients expected for different adenoma types. OCT imaging data were collected from 19 patients, consisting of volumetric scans of freshly excised adenomas. The protocol was modified after patient 11 to include normal mucosa tissue samples from the colonic epithelium for comparison.

Inclusion criteria included those who were 18 years or older, able to provide consent, and referred for colonoscopy at the Duke University Hospital Division of Gastroenterology for resection of previously identified colon polyps or multiple polyps. Exclusion criteria included a history of coagulopathy, high international normalized ratio (INR) > 2, and low platelet count (less than 50×10^9/L), and lack of "lift sign" indicating the presence of a cancerous lesion that must be removed via alternative procedures. Patients meeting the inclusion criteria and enrolled in the study underwent colonoscopy where their polyps were resected using endoscopic mucosal resection (EMR) or endoscopic submucosal dissection (ESD). After the polyps were removed, they were immediately imaged *ex vivo* using the SOCT instrument described later, followed by fixation in formalin for standard of care histopathological analysis. Additionally, six biopsies of normal-appearing colonic epithelium, located approximately 20 cm from resected polyps, were also removed using forceps. These normal tissue samples were also immediately imaged using the SOCT instrumentation.

22.3 Methods

22.3.1 Spectroscopic Optical Coherence Tomography

To date, OCT has been applied for examining colonic tissue using imaging, but has been less useful for tissue *diagnostics*, in part because it only provides structural information and lacks functional information. Several functional extensions have been developed for OCT [24]. Most relevant to this work is spectroscopic OCT (SOCT), where a short time Fourier transform (STFT)—a windowed transform—is used to analyze the wavelength dependence of OCT signals [25, 26]. This approach has a well-known drawback of exchanging depth resolution for spectral information.

The SOCT system used in these studies is shown in Figure 22.1. Two superluminescent diodes (Superlum, Ireland) were used to provide a broad bandwidth. The combined source provided a center wavelength of λ_o = 903 nm, with a FWHM bandwidth of 180 nm, producing an axial resolution of 1.4 microns. The spectrometer contains a 2048-pixel CCD line array camera (OctoPlus, Teledyne e2v), with a line rate of up to 80,000 per second. The lateral resolution was set at 5 microns by using a 10X microscope objective attached to the end of the scanner. The system includes a motorized reference arm to control its optical pathlength. Custom algorithms

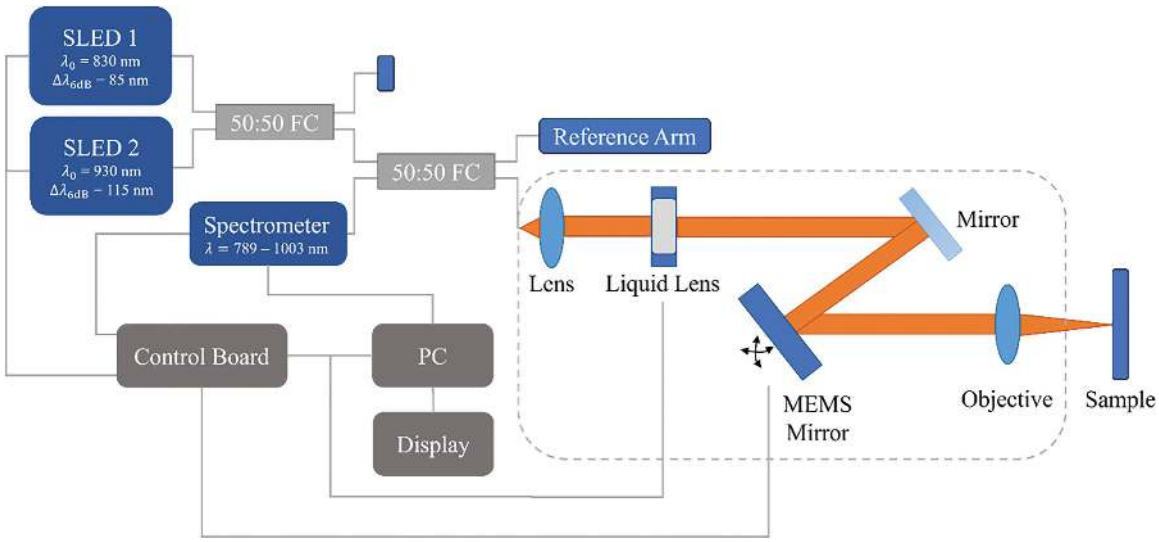


FIGURE 22.1 Block diagram of SOCT system. Setup includes two superluminescent diode modules (SLED), 50:50 fiber couplers (FC), collimating lens, liquid lens, mirror, MEMS mirror, objective, and spectrometer. (Adapted from [27].)

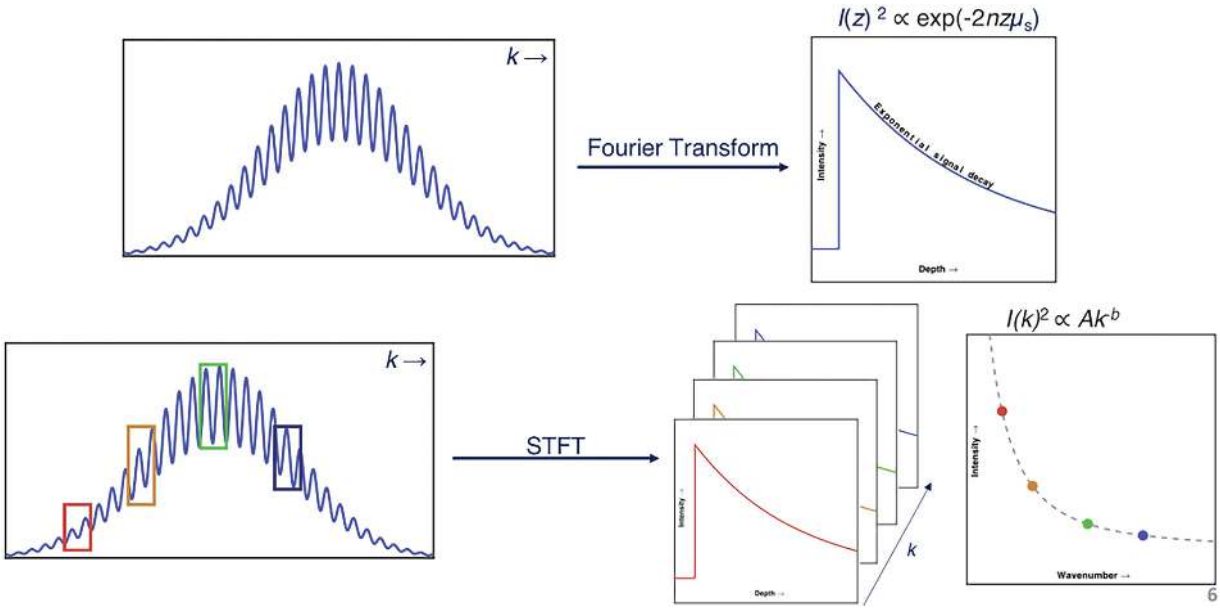


FIGURE 22.2 Illustration of OCT signal analysis scheme. (Top) OCT data are acquired in the spectral domain as a function of wavenumber, k, and Fourier transformed to produce a depth-resolved reflection profile for each pixel. Attenuation information is determined from analyzing this depth profile. (Bottom) In SOCT, a windowed Fourier transform is used to examine the signal as a function of k. (Adapted from [35].)

(Matlab) were used to analyze the signals and obtain spectroscopic information.

Figure 22.2 shows the approach for analyzing SOCT images [28]. OCT images are acquired by obtaining a spectral domain measurement for each pixel (x,y). A depth-resolved reflection profile is obtained by Fourier transform. The depth profile is then analyzed to determine the optical attenuation in the axial direction (z) by fitting the data to an exponential function to extract the scattering coefficient μ_s. SOCT is then implemented by analyzing the data using a windowed Fourier transform to provide a depth profile as a function of wavelength λ. In this dimension the signal is proportional to $A\,k^b$ where A is the scattering power, k is the wavenumber ($k = 2\pi/\lambda$) and b is the spectral exponent, a measure of the scatterer distribution within the tissue. Figure 22.1 illustrates the extraction of tissue optical properties from OCT data.

Several recent results have shown that SOCT can detect cancerous changes in colonic epithelium. We first showed that SOCT could be used to detect cancer development in the AOM rat carcinogenesis model [29]. In this study, SOCT was used to analyze tissue optical properties and determine the size of cell nuclei as a function of depth into the tissue (Figure 22.3). The results showed evidence of a field effect of carcinogenesis where enlarged cell nuclei were observed across the tissue that foreshadowed the appearance of aberrant crypt foci (ACF), the traditional biomarker used in the AOM model. Another approach, termed inverse spectroscopic OCT (ISOCT), uses a rigorous model of tissue scattering to determine optical properties from SOCT data [30, 31]. The ISOCT approach was used to discriminate advanced adenoma in colorectal epithelium from biopsy samples. Most recently, the Zhu group has found that colorectal epithelium could be evaluated using maps of scattering coefficient [32, 33], an optical property measured with OCT. Extending their approach using deep learning also showed improved diagnostic capacity [34].

22.4 Results

22.4.1 SOCT Imaging of Mouse Colon Polyps

We have conducted initial studies to use SOCT to evaluate tissue type in a mouse model of colon polyps [35], demonstrating high accuracy and the ability to distinguish polyp type by malignant potential. Previous efforts have shown SOCT measurements of tissue optical properties [29, 36, 37]. Following the measurement scheme in Figure 22.1 [35], OCT images are acquired and the optical attenuation in the axial direction (z) is fit to an exponential function to extract the scattering coefficient μ_s. SOCT is then used to examine the tissue properties as a function of wavelength. In this dimension the signal is proportional to k^{4-D} where k is the wavenumber ($k = 2\pi/\lambda$) and D is the fractal dimension, a measure of the scatterer distribution within the tissue.

To show the feasibility of this approach, we analyzed OCT measurements from mouse models of CRC in which 4-hydroxytamoxifen is delivered to the mouse colon of genetically engineered mice via colonoscopy guided injection [38, 39]. We have developed this model to model carcinogenesis in the colonic epithelium. Images were acquired from control mice, mice bearing an *Apc*-null precancerous adenoma (APC), and mice with an Apc-null, Kras-mutant, Trp53-null advanced cancer (AKP, which represents an invasive tumor phenotype).

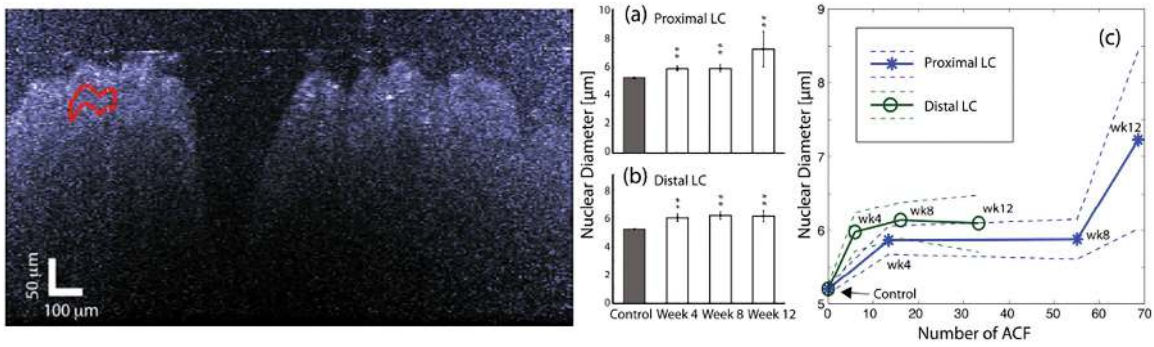


FIGURE 22.3 (Top) OCT image of an *ex vivo* rat colon sample. The red line delineates an example region that is averaged across to determine the nuclear diameter. (Bottom) Nuclear diameter measured for distal and proximal colon. Highly statistical differences (p-value $<10^{-4}$ **) were observed between the control group and treated groups. (c) Measured cell nuclear diameter as a function of the number of ACF. (From [29].)

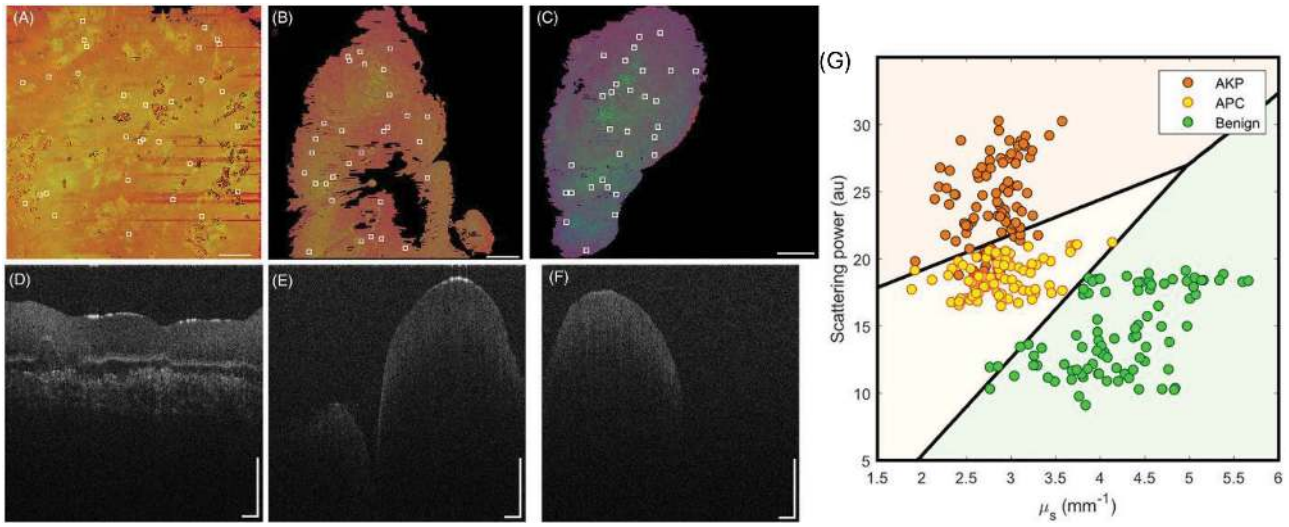


FIGURE 22.4 Examples of imaging data. (A–C) En face depictions of each tissue sample image, overlaid with false color corresponding to values of μ_s and SP. The white boxes show the location of the 30 randomly selected biopsy sites used for analysis. Scale bars for (A–C) are 1 mm. (D–F) Example B-scans taken from the volumetric images in (A–C). (A) and (D) show benign tissue, (B) and (E) show an APC tumor, and (C) and (F) show an AKP tumor. Scale bars for (D–F) are 500 µm. (G) Classification results from spectroscopic and scattering analysis of tissue. Plot shows scattering attenuation coefficient vs scattering power with regions defined by linear discriminant analysis. (Adapted from [35].)

Optical biopsy sites were simulated from 90 sites from tissues from each type, producing 270 points of comparison. Facile comparison was enabled by using the optical property maps as a visual guide, as shown in Figure 22.4. The 30 biopsy locations on each volumetric image were randomly selected 10×10 pixel regions with no black pixels present. The optical properties for each biopsy location were averaged across the 100 pixels to form each data point.

Figure 22.4G shows that these tissue types can be separated with 91.5% accuracy by using both parameters. Using OCT alone would not discriminate AKP from AKC, and would result in an uncertain determination of invasive tumors.

22.4.2 SOCT Imaging of Human Colon Polyps

A clinical study was performed to evaluate the optical properties from a range of adenoma types in the human colonic epithelium [27]. Figure 22.5 shows typical imaging data for the SOCT colon polyp study. Each resected polyp was photographed and 3D imaged with OCT. In addition to traditional OCT images, the depth scans were analyzed to determine attenuation coefficients and spectroscopic parameters.

To assess the potential diagnostic capacity of SOCT image analysis, a deep learning classifier was constructed using the optical coefficients derived from the SOCT data. The novel deep learning architecture builds upon the prior work developing a biLSTM approach for classifying mouse colon tissue data [35], discussed earlier. In the mouse tissue, the scatter power was found to offer good discrimination and the biLSTM approach was able to provide good accuracy in identifying tissue type in the human colon imaging study of 93.2%. To improve upon this, we initially used a *k*-nearest neighbor (kNN) approach. Although it did not provide clear diagnostics that could improve discrimination, it did enable a minimum redundancy maximum

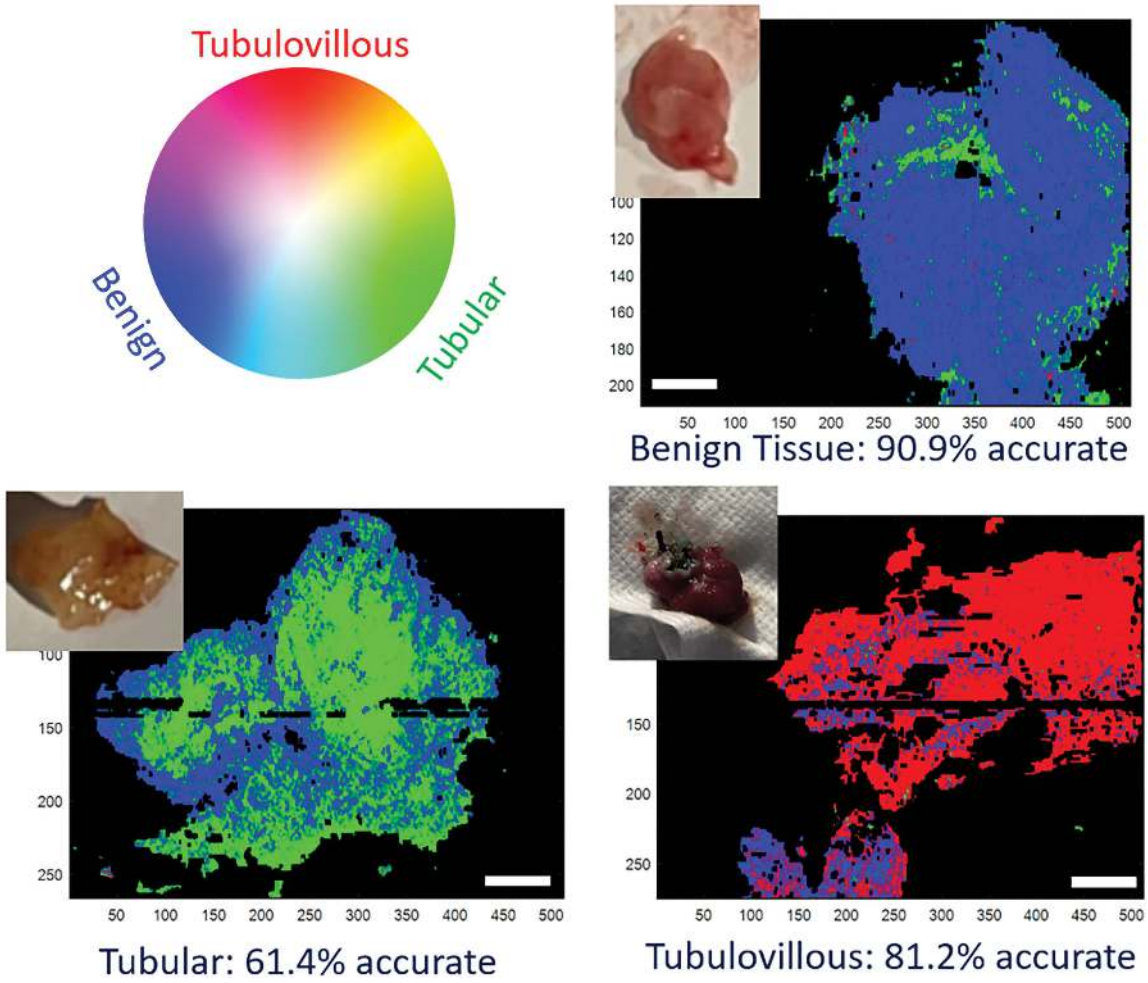


FIGURE 22.5 False color imaging of OCT data based on deep learning algorithm. Each pixel in the field of view was classified based on the predicted tissue type. While not every pixel agrees with the classification, the predominance of a particular class can identify suspected tissues based on majority voting. (Adapted from [36].)

relevance (MRMR) feature selection algorithm, which identified 12 distinct features related to scattering attenuation (e.g., mean μs, RMSE, R^2 of linear fit) from raw spectral data. To use both the scattering power (spectral data) and scattering attenuation, we combined both of these approaches into a deep learning approach featuring two parallel and fully connected networks. Each fully connected network was uniquely optimized to the physical properties of scattering attenuation coefficient and power law values. The networks' outputs were combined via a concatenation layer, transforming inputs into a unified tensor. This tensor was then processed through a sequence of convolutional layers, yielding a final decision of benign, tubulovillous, or tubular classification. Our dataset included human polyp data from 11 patients, partitioned into 90% for training, 5% for validation, and 5% for testing. The total number of A-scans used in this study were tubular: 783,353, tubulovillous: 757,468, normal mucosa: 764,481, and total 2,307,302.

The approach of using optical tissue properties significantly reduced data requirements compared to using complete depth scans (512-vector in z) for each x,y pixel. Instead, we utilized key parameters as inputs for the classifier, resulting in a concise 25-entry vector. This design enabled the construction of a robust classifier with rapid evaluation. Classification of over 100,000 optical biopsies was accomplished in 23.56 seconds or approximately 0.2 milliseconds each. Overall, the classifier performed quite well, producing an overall accuracy of 97.9% with a sensitivity of 98.2% and a specificity of 95.2%. The 98.6% positive predictive value points to a good clinical potential in identifying malignant polyps. Meanwhile, the negative predictive value (NPV) demonstrated solid performance at 96.3%. This level of accuracy is clinically acceptable, given the outcome of incorrect classification would only involve acquiring a biopsy from a potentially benign tissue site. It's worth noting that current colonoscopy procedures commonly extract tissue samples from primarily benign tissues anyway.

To demonstrate how OCT data can be used in colonoscopy procedures, we constructed false color images, based on the output of the deep learning classifier (Figure 22.5). In this example, the classifier was applied to each pixel in the OCT image and used to color code the image based on the color wheel shown in the top left. The benign tissue was correctly identified in 90.9% of the pixels in the image. The tubular

adenoma was less accurate at 61.4%, but as can be seen, the remaining pixels were identified as benign tissue. This is an expected outcome since not necessarily every pixel in the image would necessarily present as tubular adenoma and instead benign tissue would be mixed in. This type of classification would be clinically acceptable, particularly because both of the tissue types displayed are viewed as possessing lower malignant potential. In comparison, the image for the tubulovillous adenoma polyp shows a clear dominance of classification (81.2%) and would be a sufficient red flag marker to induce the endoscopist to resect this polyp.

22.5 Discussion

We have evaluated the use of SOCT as a potential method for tissue assessment during screening and diagnostic colonoscopies. The preliminary studies presented here show feasibility of discriminating tissue type based on SOCT measurements of optical properties in mouse and human *ex vivo* colonic epithelial tissues. The accuracy of spectroscopic OCT in classifying colon tissue shows promise for clinical applications. Specifically, the high negative predictive value indicates potential utility in implementing a leave-in-place strategy, where doctors may decide to not remove a biopsy if it is confirmed to be low risk, producing potential benefits in reduced procedure time, fewer complications, and lower costs of tissue processing and subsequent pathological analysis.

Several factors need to be considered based on the limitations of the models used here. The mouse studies may have limited applicability to human studies as the differences in tissue morphology and route for inducing carcinogenesis may yield different optical properties than those found in human tissues. Further, differences in SOCT instrumentation may provide different spectral information that may not allow direct mapping of this information. However, the human tissue studies provide solid confirmation of the feasibility of the approach. The data obtained from these samples were able to provide discrimination of tissue type even though the optical properties and spectral windows varied from the mouse model. This was enabled in part by the adoption of novel deep learning architectures. This information was used to produce a false coloring scheme as a means to visually identify tissue type. This aspect could prove to be an essential feature for encouraging clinical adoption in the future.

22.6 Future Trends

Further development is planned to implement a low-cost, compact OCT system that can be applied as an adjunct modality during screening colonoscopy to improve adenoma and cancer diagnosis. Improved imaging information will save overall costs through earlier and more comprehensive detection of CRC while also reducing pathology costs by selective targeting of potentially precancerous polyps. The potential clinical utility of this SOCT approach may be during colonoscopy procedures for CRC screening where real-time definition of adenomatous tissue would provide the opportunity for immediate diagnosis. This could improve screening procedures while also potentially reducing expense and time cost associated with further procedures. Adoption for clinical use will require larger studies that characterized the optical properties of a wider range of adenoma and normal mucosa tissue morphologies, followed by clinical trials in screening colonoscopy patient cohorts.

Acknowledgments

The authors wish to acknowledge funding from the National Institutes of Health: 1R41EB032693.

REFERENCES

1. Seer Database. https://seer.cancer.gov/statfacts/html/colorect.html.
2. Selby, J., G. Friedman, C. Quesenberry, and N. Weiss, A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *New England Journal of Medicine,* 1992. 326(10): p. 653–657.
3. Mandel, J.S., T.R. Church, J.H. Bond, F. Ederer, M.S. Geisser, S.J. Mongin, D.C. Snover, and L.M. Schuman, The effect of fecal occult-blood screening on the incidence of colorectal cancer. *New England Journal of Medicine*, 2000. 343(22): p. 1603–1607.
4. Meissner, H.I., N. Breen, C.N. Klabunde, and S.W. Vernon, Patterns of colorectal cancer screening uptake among men and women in the United States. *Cancer Epidemiology Biomarkers & Prevention,* 2006. 15(2): p. 389–394.
5. Siegel, R.L., K.D. Miller, and A. Jemal, Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, 2020. 70(1): p. 7–30.
6. Sullivan, B.A., M. Noujaim, and J. Roper, Cause, epidemiology, and histology of polyps and pathways to colorectal cancer. *Gastrointestinal Endoscopy Clinics of North America*, 2022. 32(2): p. 177–194.
7. Ekbom, A., C. Helmick, M. Zack, and H.-O. Adami, Ulcerative colitis and colorectal cancer. *New England Journal of Medicine*, 1990. 323(18): p. 1228–1233.
8. Bernstein, C.N., J.F. Blanchard, E. Kliewer, and A. Wajda, Cancer risk in patients with inflammatory bowel disease. *Cancer*, 2001. 91(4): p. 854–862.
9. Lukas, M., Inflammatory bowel disease as a risk factor for colorectal cancer. *Digestive Diseases*, 2010. 28(4–5): p. 619–624.
10. Farraye, F.A., R.D. Odze, J. Eaden, and S.H. Itzkowitz, AGA medical position statement on the diagnosis and management of colorectal neoplasia in inflammatory bowel disease. *Gastroenterology*, 2010. 138(2): p. 738–745.
11. Provenzale, D., K.V. Kowdley, S. Arora, and J.B. Wong, Prophylactic colectomy or surveillance for chronic ulcerative colitis? A decision analysis. *Gastroenterology*, 1995. 109(4): p. 1188–1196.
12. Lee, A. and N. Tutticci, Enhancing polyp detection: Technological advances in colonoscopy imaging. *Translational Gastroenterology and Hepatology*, 2021. 6: p. 61.

13. Gupta, N., A. Bansal, D. Rao, D.S. Early, S. Jonnalagadda, S.B. Wani, S.A. Edmundowicz, P. Sharma, and A. Rastogi, Prevalence of advanced histological features in diminutive and small colon polyps. *Gastrointestinal Endoscopy*, 2012. 75(5): p. 1022–1030.
14. Hassan, C., P.J. Pickhardt, and D.K. Rex, A resect and discard strategy would improve cost-effectiveness of colorectal cancer screening. *Clinical Gastroenterology and Hepatology*, 2010. 8(10): p. 865–869. e3.
15. Rex, D.K., C. Kahi, M. O'Brien, T. Levin, H. Pohl, A. Rastogi, L. Burgart, T. Imperiale, U. Ladabaum, and J. Cohen, The American Society for Gastrointestinal Endoscopy PIVI (Preservation and Incorporation of Valuable Endoscopic Innovations) on real-time endoscopic assessment of the histology of diminutive colorectal polyps. *Gastrointestinal Endoscopy*, 2011. 73(3): p. 419–422.
16. Hassan, C., G. Balsamo, R. Lorenzetti, A. Zullo, and G. Antonelli, Artificial intelligence allows leaving-in-situ colorectal polyps. *Clinical Gastroenterology and Hepatology,* 2022. 20(11): p. 2505–2513. e4.
17. Mori, Y., S.-e. Kudo, M. Misawa, Y. Saito, H. Ikematsu, K. Hotta, K. Ohtsuka, F. Urushibara, S. Kataoka, and Y. Ogawa, Real-time use of artificial intelligence in identification of diminutive polyps during colonoscopy: a prospective study. *Annals of Internal Medicine*, 2018. 169(6): p. 357–366.
18. Huang, D., E.A. Swanson, C.P. Lin, J.S. Schuman, W.G. Stinson, W. Chang, M.R. Hee, T. Flotte, K. Gregory, and C.A. Puliafito, Optical coherence tomography. *Science (New York, NY)*, 1991. 254(5035): p. 1178.
19. Drexler, W. and J.G. Fujimoto, State-of-the-art retinal optical coherence tomography. *Progress in Retinal and Eye Research*, 2008. 27(1): p. 45–88.
20. Kim, S., M. Crose, W.J. Eldridge, B. Cox, W.J. Brown, and A. Wax, Design and implementation of a low-cost, portable OCT system. *Biomedical Optics Express*, 2018. 9(3): p. 1232–1243.
21. Reddy, B.S., Studies with the azoxymethane-rat preclinical model for assessing colon tumor development and chemoprevention. *Environmental and Molecular Mutagenesis*, 2004. 44(1): p. 26–35.
22. Roper, J., T. Tammela, N.M. Cetinbas, A. Akkad, A. Roghanian, S. Rickelt, M. Almeqdadi, K. Wu, M.A. Oberli, F.J. Sánchez-Rivera, Y.K. Park, X. Liang, G. Eng, M.S. Taylor, R. Azimi, D. Kedrin, R. Neupane, S. Beyaz, E.T. Sicinska, Y. Suarez, J. Yoo, L. Chen, L. Zukerberg, P. Katajisto, V. Deshpande, A.J. Bass, P.N. Tsichlis, J. Lees, R. Langer, R.O. Hynes, J. Chen, A. Bhutkar, T. Jacks, and H. Yilmaz Ö, In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. *Nature Biotechnology*, 2017. 35(6): p. 569–576.
23. Roper, J., T. Tammela, A. Akkad, M. Almeqdadi, S.B. Santos, T. Jacks, and H. Yilmaz Ö, Colonoscopy-based colorectal cancer modeling in mice with CRISPR-Cas9 genome editing and organoid transplantation. *Nature Protocols*, 2018. 13(2): p. 217–234.
24. Kim, J., W. Brown, J.R. Maher, H. Levinson, and A. Wax, Functional optical coherence tomography: principles and progress. *Physics in Medicine and Biology*, 2015. 60(10): p. R211–237.
25. Leitgeb, R., M. Wojtkowski, A. Kowalczyk, C. Hitzenberger, M. Sticker, and A. Fercher, Spectral measurement of absorption by spectroscopic frequency-domain optical coherence tomography. *Optics Letters*, 2000. 25(11): p. 820–822.
26. Xu, C., C. Vinegoni, T.S. Ralston, W. Luo, W. Tan, and S.A. Boppart, Spectroscopic spectral-domain optical coherence microscopy. *Optics Letters*, 2006. 31(8): p. 1079–1081.
27. Kendall, W., Q. Tian, S. Zhao, S. Mirminachi, A. Joseph, D. Dufault, C. Shi, J. Roper, and A. Wax, Deep learning classification of ex vivo human colon tissues using spectroscopic OCT. *Journal of Biophotonics*, 2024 17(9): p. e20240008
28. Yi, J. and V. Backman, Imaging a full set of optical scattering properties of biological tissue by inverse spectroscopic optical coherence tomography. *Optics Letters*, 2012. 37(21): p. 4443–4445.
29. Robles, F.E., Y. Zhu, J. Lee, S. Sharma, and A. Wax, Detection of early colorectal cancer development in the azoxymethane rat carcinogenesis model with Fourier domain low coherence interferometry. *Biomedical Optics Express*, 2010. 1(2): p. 736–745.
30. Winkelmann, J.A., A. Eid, T.-Q. Nguyen, T. Bui, J. Yi, and V. Backman, In vivo broadband visible light optical coherence tomography probe enables inverse spectroscopic analysis. *Optics Letters*, 2018. 43(3): p. 619–622.
31. Ji, Y., J.R. Andrew, S.-C. Yolanda, N.M. Nikhil, M.A. Samira, H. Elizabeth, J.G.M.D. Michael, B. Laura, B. Shailesh, K.R. Hemant, and B. Vadim, Spatially resolved optical and ultrastructural properties of colorectal and pancreatic field carcinogenesis observed by inverse spectroscopic optical coherence tomography. *Journal of Biomedical Optics*, 2014. 19(3): p. 1–17.
32. Zeng, Y., W.C. Chapman Jr, Y. Lin, S. Li, M. Mutch, and Q. Zhu, Diagnosing colorectal abnormalities using scattering coefficient maps acquired from optical coherence tomography. *Journal of Biophotonics*, 2021. 14(1): p. e202000276.
33. Zeng, Y., B. Rao, W.C. Chapman, S. Nandy, R. Rais, I. González, D. Chatterjee, M. Mutch, and Q. Zhu, The angular spectrum of the scattering coefficient map reveals subsurface colorectal cancer. *Scientific Reports*, 2019. 9(1): p. 2998.
34. Zeng, Y., S. Xu, W.C. Chapman, Jr., S. Li, Z. Alipour, H. Abdelal, D. Chatterjee, M. Mutch, and Q. Zhu, Real-time colorectal cancer diagnosis using PR-OCT with deep learning. *Theranostics*, 2020. 10(6): p. 2587–2596.
35. Kendall, W.Y., J. Bordas, S. Mirminachi, A. Joseph, J. Roper, and A. Wax, Spectroscopic optical coherence tomography for classification of colorectal cancer in a mouse model. *Journal of Biophotonics*, 2022. 15(7): p. e202100387.
36. Graf, R.N., F.E. Robles, X. Chen, and A.P. Wax, Detecting precancerous lesions in the hamster cheek pouch using spectroscopic white-light optical coherence tomography to assess nuclear morphology via spectral oscillations. *Journal of Biomedical Optics*, 2009. 14(6): p. 064030.
37. Robles, F.E., C. Wilson, G. Grant, and A. Wax, Molecular imaging true-colour spectroscopic optical coherence tomography. *Nature Photonics*, 2011. 5(12): p. 744–747.
38. Roper, J., T. Tammela, N.M. Cetinbas, A. Akkad, A. Roghanian, S. Rickelt, M. Almeqdadi, K. Wu, M.A. Oberli, F.J. Sánchez-Rivera, Y.K. Park, X. Liang, G. Eng, M.S. Taylor, R. Azimi,

D. Kedrin, R. Neupane, S. Beyaz, E.T. Sicinska, Y. Suarez, J. Yoo, L. Chen, L. Zukerberg, P. Katajisto, V. Deshpande, A.J. Bass, P.N. Tsichlis, J. Lees, R. Langer, R.O. Hynes, J. Chen, A. Bhutkar, T. Jacks, and Ö.H. Yilmaz, In vivo genome editing and organoid transplantation models of colorectal cancer and metastasis. *Nature Biotechnology*, 2017. 35(6): p. 569–576.

39. Roper, J., T. Tammela, A. Akkad, M. Almeqdadi, S.B. Santos, T. Jacks, and Ö.H. Yilmaz, Colonoscopy-based colorectal cancer modeling in mice with CRISPR–Cas9 genome editing and organoid transplantation. *Nature Protocols*, 2018. 13(2): p. 217–234.

23

Analysis of Circulating Tumor DNA (ctDNA) by Digital Droplet PCR (ddPCR)

Bryant L. McCue and George Miles

23.1 Introduction

Digital polymerase chain reaction (PCR) allows for ultrasensitive measurement of mutated and wild-type copies of DNA over tens of thousands of individual partitions and is best suited for applications requiring quantitation of low-level targets. This technology spans a diverse range of applications, including rare mutant detection in cancer, pathogen detection and environmental surveillance, quantitating gene therapy/gene editing events, and noninvasive prenatal testing, to name a few. Various microfluidic technologies exist for effective molecular partitioning, including the use of nanoliter-sized prefabricated chambers and via generation of nanoliter-sized droplets in a water-oil emulsion. Individual chambers or droplets are quantitated based on discernment of emission spectra of different fluorescent probe dyes or differing fluorescent intensities. We illustrate methodology that focuses specifically on competing duplex digital droplet PCR using the Bio-Rad QX200 platform with HEX and FAM-labeled primer probes. An automated water-emulsion droplet generator (QX200DG, Bio-Rad) creates tens of thousands of individual droplets that partition analyte molecules and detection probes into separate PCR reactions. A typical starting sample of 20 µL should yield up to 20,000 droplets. The QX200 can process and analyze up to 96 samples per run, yielding and counting approximately 1.5 million droplets per 96-well plate. Such microfluidics-based sample partitioning and statistical analyses of up to millions of droplets provide absolute quantification of target DNA or RNA molecules for probe-based detection at high sensitivity without the need for standard curves. Completed droplet reactions are streamed single file through a specialized reader for fluorescence analysis using a multi-pixel photon counter. Positive droplets exhibit increased fluorescence of target (mutant) probes compared to negative droplets (wild type), and the fraction of PCR-positive droplets are quantitated according to Poisson distribution. Thus, the fractional abundance of the mutant allele can be calculated by the ratio of the number of mutant-FAM probe copies per droplet over the total DNA concentration of the reaction. In this chapter we demonstrate basic assay methodology that can be used to detect circulating tumor DNA isolated from blood plasma.

For ddPCR we must know the mutations we want to detect *a priori*. Unlike next-generation sequencing panels that serve largely for discovery, digital PCR is specific to confirming the presence or absence of a mutation at specific locus. Qualifying mutations for ctDNA analysis by digital PCR are selected from sequenced surgical tissue samples with mutations prioritized by likely clonal/driver status, or from a sequenced liquid biopsy sample from the patient. For each mutant allele, a custom ddPCR probe set is designed. Digital droplet construction is performed on the Bio-Rad AutoDG QX200 (automatic droplet generator) ddPCR system with assays utilizing TaqMan hydrolysis chemistry with custom and/or commercially available primer-probe sets (e.g., Bio-Rad, Integrated DNA Technologies, etc.). The partitioned PCR droplet assays are subjected to amplification using a C1000 Touch™ thermal cycler and fluorometrically assessed on the QX200 droplet reader (Bio-Rad) along with patient-specific positive controls. The fraction of positive droplets will be fitted to a Poisson distribution using available QuantaSoft Pro software (Bio-Rad) to calculate the number of mutant and wild-type (WT) copies of DNA in the sample.

23.2 Materials

23.2.1 Reagents

ctDNA Blood Collection Tubes: Streck Cell-Free DNA BCT blood collection tubes (Streck, cat #230470) were used to collect 20 ml whole blood (2 × 10 ml tubes).

gBlocks: gBlocks were obtained from Integrated DNA Technologies (IDT) and are constructed to specified length.

Control Diluent: SeraCare, Basematrix Negative Diluent, cat# 1805–0076 was used as a diluent to spike in known copies of gBlock for extraction and PCR process controls.

Cell-Free DNA Extraction kits: Applied Biosystems by Thermo Scientific, MagMax Cell-Free DNA Isolation Kit (cat# A29319), was used for the cfDNA extraction via a magnetic bead automated extraction process.

Supermix: Bio-Rad, ddPCR Supermix for Probes (No dUTP), cat # 186–3023 is a ready to go reaction mixture containing all the needed elements to conduct digital droplet DNA PCR reactions.

DOI: 10.1201/9781003449942-26

ddPCR Assays: Dual color ddPCR assays (probe/primer mix) were obtained from Bio-Rad available as either ready-made or custom made to a sequence of interest.

23.2.2 Laboratory Layout

To reduce the chance of sample contamination, the molecular laboratory floor plan incorporates a unidirectional work flow consisting of independent but adjacent zones and rooms: (1) sample preparation including DNA extraction, (2) PCR reaction setup including digital droplet generation, and (3) PCR amplification. Whole blood samples are received into the initial preparation zone for record of receipt and labeling. The whole blood tubes are then moved to an adjacent sample preparation room where centrifugation and plasma separation occurs. Plasma samples are then transferred to our DNA extraction zone, which can be a manual or automated process (e.g., Hamilton Nimbus-Presto®). Extracted DNA is sent to the PCR setup room, which is a separate "clean room" under positive pressure and ceiling HEPA air filtration. Here, templates and assays are assembled manually or with dedicated robotic liquid handling systems (e.g., Eppendorf epMotion®) and subjected to droplet formation (Bio-Rad® QX200DG). Sealed plates are then placed inside the cleanroom pass-through window and transferred to the amplification room for PCR and droplet reading. Note, in addition to the unidirectional lab layout, an isolated room completely detached from the main molecular laboratory is used for handling, diluting, and storing all positive controls, which may include cell lines and synthetic DNA fragments. Plasmid DNA is expressly forbidden in our laboratory.

23.2.3 Good Laboratory Practices

Good laboratory practices are utilized throughout the entire workflow, including the use of clean or disposable lab coats and frequent glove changes between sections and critical assay points. The workflow is a stepwise process so that each step can be separate with definitive stopping points managed by different technicians, creating natural breaks for any potential of carry over or contamination events. All workstations have lined disposal bins that are sealed prior to removal. Each workstation has dedicated supplies, consumables, reagents, and equipment (pipettors, centrifuges, cabinets, etc.). Our lab uses only filtered, low-binding PCR grade tips.

All work surfaces, equipment, cabinets, and handling devices are cleaned before and after use with 70% ETOH followed by a commercially available DNA degrading solution (i.e., DNA Zap by Thermo Fisher). Desktop PCR cabinets have UV cleaning capability.

23.3 Methods

23.3.1 cfDNA Blood Sample Collection

Whole blood samples are collected and stabilized into two Streck ctDNA BCT tubes and transported at room temperature. Plasma is isolated from the whole blood the same or next day following collection; however, cell-free DNA is stabilized for up to seven days at ambient temperature. One tube completely filled with whole blood can yield approximately 4 ml of plasma.

23.3.2 Plasma Isolation

Whole blood tubes are centrifuged for 15 minutes at 10 °C at 1200 xg, then carefully removed from the centrifuge to avoid disturbing the buffy coat. The plasma is then removed using a transfer pipette, leaving the last 3–5 mm level of plasma to avoid picking up cells from the buffy coat. Plasma from both Streck tubes is transferred into a single low-binding 15 ml conical centrifuge tube. The plasma is gently mixed by two to three inversions. The combined plasma in the 15 ml conical centrifuge tube is centrifuged for 15 minutes at 10 °C at 3200 xg. The plasma can then be aliquoted into cryovials for storage or pipetted into an extraction tube for DNA extraction. Any cell pellet appearing at the bottom of the 15 ml tube is avoided during this transfer to mitigate contamination with genomic DNA. A total of 4 ml plasma is used per run. Plasma can be stored at 4 °C for up to 24 hours or at −80 °C. Previously frozen plasma is thawed at 4 °C for up to 24 hours prior to use.

23.3.3 Control Materials

Synthetic DNA fragments or cell lines are utilized as positive controls for each unique target. GBlocks (Integrated DNA Technologies) or similar constructs can be made and mixed in varying ratios or through a series of dilutions to mimic observed mutant abundances. These control materials can be used as spike-in extraction controls to monitor ctDNA extraction efficiency as well as for individual ddPCR assays.

23.3.4 ctDNA Extraction

23.3.4.1 Pre-Extraction Preparation

Fresh plasma or previously frozen thawed plasma is aliquoted (4 ml) into low-binding 15 ml conical centrifuge tubes. If the sample had been previous frozen/thawed and cryoprecipitate has formed, the tube can be vortexed for 30 seconds to suspend. The plasma (4 ml) is treated with 60 µl of 20 mg/ml proteinase K (Invitrogen, AM2546) and 200 µl of 20% SDS (Invitrogen, AM9820). Mix slowly by three to four inversions and incubate at 60 °C for 20 minutes followed by cooling at 4 °C for 5 minutes.

23.3.4.2 Automated DNA Extraction

The cfDNA is extracted using a Hamilton Nimbus liquid handler integrated with a Thermo Presto sample preparation instrument and Thermo MagMax Isolation kits, which utilizes magnetic bead technology for separation. This combination provides optimized, consistent, and automated hands-free cfDNA extraction. Immediately following the pre-extraction preparation, the 15 ml conical tubes of plasma are placed on the Hamilton deck along with the extraction reagents

(beads, elution buffer, 80% ETOH, binding, and wash buffer) and plastic consumables (conductive pipette tips, deepwell plates, and magnetic tip combs) needed for the extraction. The deck arrangement and materials are specific for the Thermo MagMax cfDNA extraction process and workflow programmed accordingly. The final elution volume is 50 µl and the samples are eluted into on deck 1.5 ml snap cap tubes. Following the completion of the extraction, the elution tubes should be immediately removed from the instrument and closed.

23.3.5 DNA Quantification and QC

DNA quantity is measured using a Qubit 4 Fluorometer and Qubit 1x dsDNA HS Assay Kit (Q33231). This is to determine absolute quantity of DNA extracted. Quality and any presence of high molecular weight DNA is measured using an Agilent 4200 Tape station with Agilent Cell-Free Screentapes (95067–5630) and Reagents (5067–5631). These measurements are made immediately following extraction.

23.3.6 ddPCR

A 96-well worksheet/map is created to assign sample well placement within the plate. Following quantity and/or quality measurements of the extracted ccfDNA, the entire volume of the eluate for each sample is carefully transferred to a corresponding well of a 96-well ddPCR plate using a P50 pipettor and the elution is carefully dispensed at the bottom of its designated well. It is crucial that the sample be placed following the map and that care is taken to not cross contaminate wells. Automated liquid handlers may be considered for throughput. The eluates are dried down in the 96-well ddPCR plate by placing plate in a vacufuge set to 60 °C for 1.5 hours. After 1.5 hours, a visual check is made to ensure complete drying has occurred. Additional time can be added as needed.

23.3.6.1 Reaction Plate Setup for Droplet Generation

A total of 22 µl reaction mix is required per well (assayed amount is 20 µl with ~9% overage to ensure enough volume is available for the instrument). To each well containing the dried samples the following are added: 11 µl Bio-Rad supermix, 1.1 µl probe, and 9.9 µl DNase free water. A master mix can be created containing supermix, probe, and water and then 22 µl added to each well as appropriate. In some instances, a sample may be a direct "wet" add (vs dried), then the quantity of water can be adjusted to volume. The plate is sealed with pierceable foil heat seal, then vortexed (set to high) for 10 seconds on each corner of the plate and the 10 seconds on each side. Plate is centrifuged for 1 minute at 2000 rpm and visually checked to make sure no bubbles exist, and the liquid is situated at the bottom of each well. As noted in prior steps, to aid with consistency and to reduce opportunities for handling errors, an Eppendorf EP Motion® 5073 or similar robotic device can be used to automate and facilitate high throughput reaction setups. The epMotion® liquid handler can be programmed to add any combinations of reaction mix, water, probe, or master mix to the plate. Programming steps are documented, and a "virtual" run is conducted prior to actual use.

TABLE 23.1

Typical Amplification Conditions for ddPCR Assays

Cycling Step	Temp, °C	Time	Ramp Rate	# of Cycles
Enzyme activation	95	10 min	2 C/sec	1
Denaturation	94	30 sec		40
Annealing/extension	60	1 min		40
Enzyme deactivation	98	10 min		1
Hold (optional)	4	Infinite		1

23.3.6.2 Droplet Generation

The plate is loaded on to the Bio-Rad automated droplet generator instrument along with droplet cartridges, oil, tips, and 96-well droplet plate. The instrument will automatically transfer samples to the droplet generation cartridges, add oil, create droplet emulsions, and then transfer droplets to a 96-well plate. The plate is removed from the droplet generator and sealed with a pierceable foil heat seal and placed in the transfer door to the post-PCR processing room. The plate is sealed and placed on the thermocycler within 30 minutes of loading (Table 23.1).

23.3.7 PCR

The 96-well ddPCR plate containing the droplets is placed on a Bio-Rad C1000® Touch Thermo Cycler set to the following conditions using a heated lid at 105°C and 40 µl sample volumes:

Critical note: PCR primer-probe sets and reaction conditions should be optimized for each target assay. A helpful site for primer-probe design and *in silico* analysis is Primer3web (https://primer3.ut.ee). Following the PCR reaction, the plate is incubated on the thermocycler for 30 minutes before being transferred to the QX200 droplet reader.

23.3.8 Droplet Reader and Data Analysis

The plate is placed in the QX 200 droplet reader, which will stream the droplets for reading. QX Manager 1.2 Standard Edition software is used to run the instrument and to collect and analyze the data. The software can be set to "auto analyze" data using software predetermined parameters, or the data can be manually analyzed with manual control of the parameters.

23.4 Results

23.4.1 Circulating Tumor DNA Integrity

Representative cell-free DNA extraction analysis of various plasma samples is illustrated in Figure 23.1. Inadequate presence of short-sized DNA (representing putative ccfDNA/ctDNA) (16) or the presence of any larger DNA fragment carryover (> 400bp) can result in diminished detection and performance of downstream ctDNA analyses. While Qubit analysis can be used to quantitate extracted DNA from plasma, Agilent TapeStation® or a similar fragment analysis approach is highly recommended as a pre-analytic step to ascertain the integrity of extracted circulating cell-free DNA (ccfDNA) and

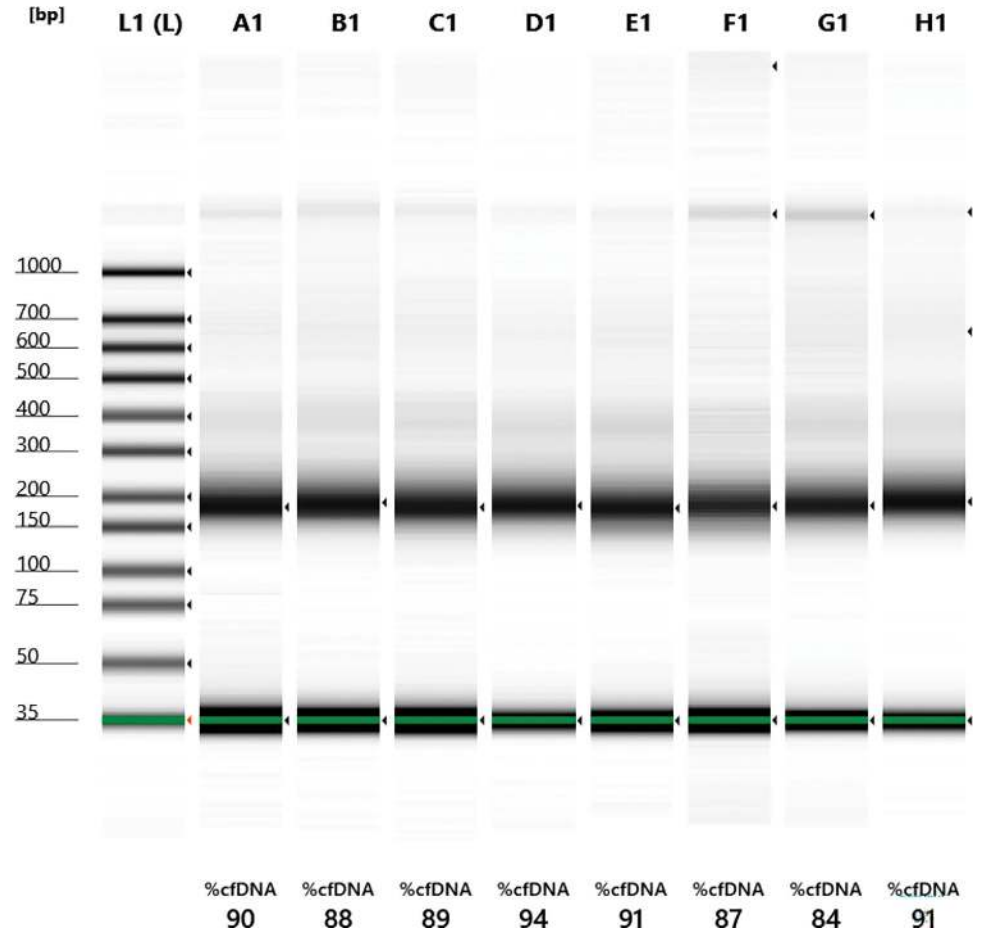


FIGURE 23.1 *ctDNA Tapestation Analysis*. A representative digital electropherogram (Agilent) showing size and percent purity of extracted circulating cell-free DNA in samples.

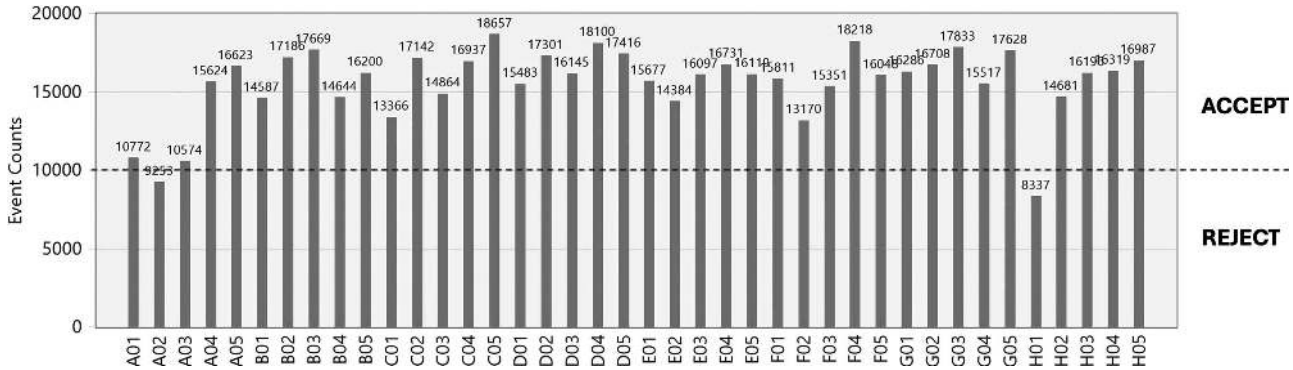


FIGURE 23.2 *Droplet Event Counts*. This chart illustrates the number of individual droplet events from a 96-well plate as detected on the Bio-Rad QX 200 Reader. In this instance, a minimum quality threshold of 10,000 droplets (dashed line) is required for further analysis. Samples A02 and H01 are below this cutoff and would be rejected.

to discern levels and purity of circulating tumor (ctDNA) in a sample prior to running any digital PCR assay. This in turn will help decrease test failure rates.

23.4.2 Digital Droplet PCR

Basic guidelines for droplet quality control should be implemented for every run that meets laboratory requirements. For example, in our lab individual sample wells with less than 10,000 accepted "events" are rejected from further analysis (Figure 23.2). The detection and quantitation of tumor-specific mutations in ctDNA for each patient was determined using bespoke digital droplet PCR assays constructed from baseline NGC ctDNA testing. Additionally, droplet clusters should be examined under both 1D and 2D fluorescence amplitude plots to assess individual assays for consistency and appropriately defined and classified droplet clusters. Further assay optimization and analysis are reviewed elsewhere (17, 18).

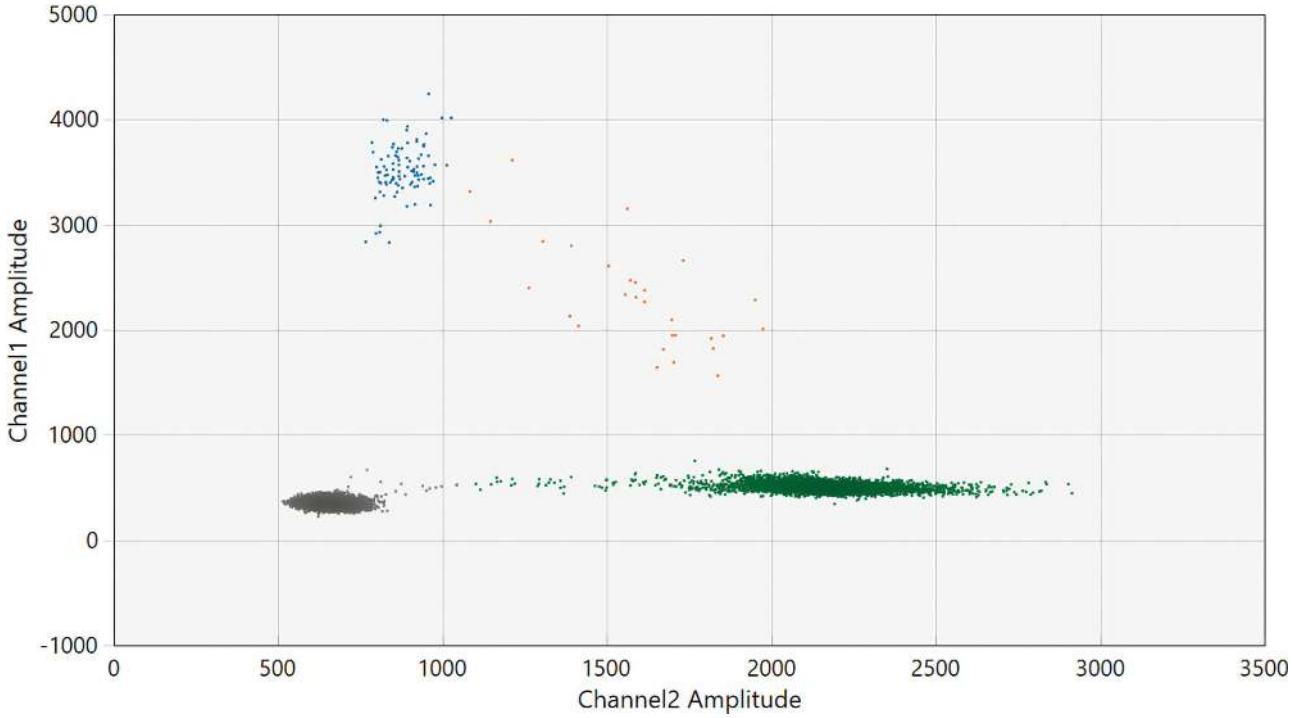


FIGURE 23.3 *ddPCR 2-D Cluster Map.* This diagram represents a 2D fluorescence amplitude view of clustered droplets in a typical dual-color ddPCR assay. Positive detection of the mutant (MT) is identified by the cluster of blue droplets in the upper left quadrant (FAM channel, y axis), whereas the green droplets (HEX channel, x axis) indicate abundance of wild type only. The orange spread represents double positive (mix of both MT and WT), and the gray droplets lack any detectable fluorescence. Quadrant thresholds that discriminate probe signals were calculated using the Bio-Rad QuantaSoft® software package. The fractional abundance or allelic frequency of this particular mutant (HER2 S310F) is 2.09%.

Figure 23.3 plots results from a ddPCR assay with the typical four clusters observed under two channels. Assay validations should be designed that allow for reliable detection of fractional mutant abundances over the anticipated dynamic range of the specimens being tested. No-template controls (NTCs) should be carefully examined to identify workflow issues and potential contamination. Positive controls should reflect performance requirements of the assay underway and are most effectively sourced from synthetic or cell line controls constructed in the lab or obtained commercially. We include contrived control materials assayed at varying levels of known abundance with our unknown specimens to ensure reliable positive mutant calls over a range of detection (Figure 23.4).

23.5 Discussion

Genomic characterization of cancer has revolutionized our ability to decipher the complexities of tumor biology and promote more precise cancer treatments. Advances in DNA sequencing have enabled the detection of mutations in the tumor and now in tumor DNA that circulates in the blood (ctDNA). Cell-free circulating tumor DNA has been proposed as a surrogate biological sample to define the genetic change(s) of a primary tumor and/or metastatic disease in a cancer patient and to serve as a biomarker for diagnosis, prognostication, and monitoring of response to therapy. Clinically, digital droplet PCR technology affords the opportunity for rapid molecular surveillance for known targets of interest and with high sensitivity. This platform is experiencing rapid growth in the oncology market as a complement to next-generation sequencing (NGS). Results can be generated quickly and frequently as an economical means for monitoring disease progression or response to therapy. Frequent time sampling will allow for earlier detection of recurrence or response compared to standard of care imaging.

23.6 Future Trends

The use of digital droplet PCR to detect and monitor ctDNA is continuing to evolve. Multiplex digital PCR that can simultaneously and precisely detect multiple mutations in a single assay is a promising advance. These developments should be particularly useful in the monitoring space, for example in screening resistance mutations acquired during treatment. Panels consisting of common mutations can be utilized, or more so, bespoke panels can be more effectively developed to guide patient management. As the technology evolves, the detection of ctDNA in other sample types (urine, saliva, CSF, etc.) are being explored for both their proximal and their distal solid tumor monitoring potential (19).

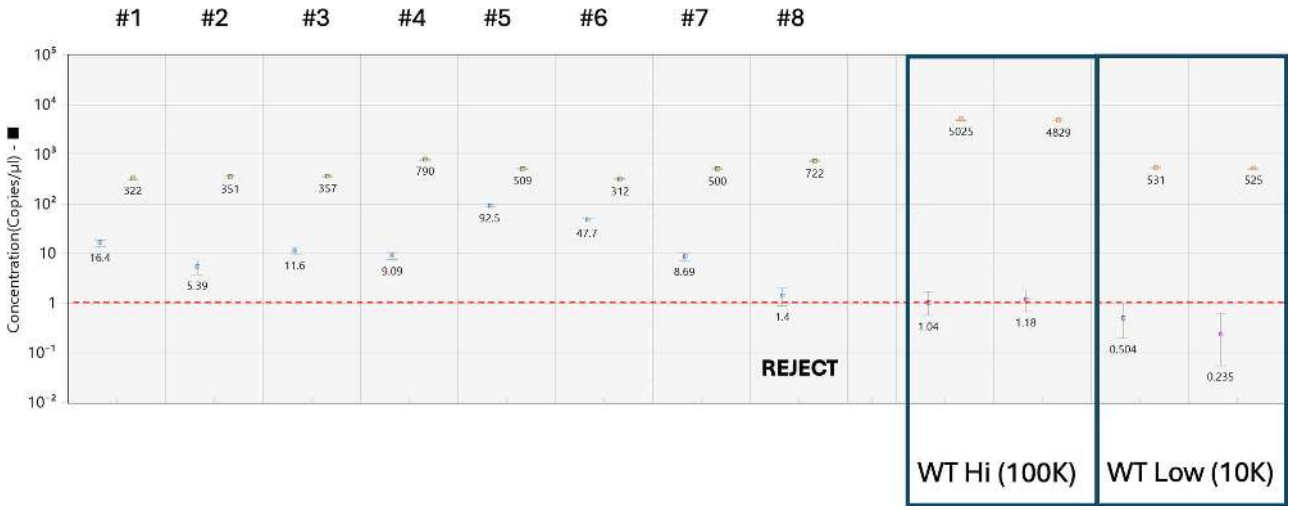


FIGURE 23.4 *Establishing Confidence in Positive Mutant Calls*. This plot illustrates the concentration of detected mutant DNA (blue boxes) in eight different patient samples (copies/mL) bounded by 95% confidence intervals (CI) as calculated with the Bio-Rad QX Manager® software. On the right hand side are WT-only positive controls (no mutant) run in duplicate at two different concentrations (enclosed columns). Sample 8 shows a detected mutant concentration that overlaps the 95% CI bar observed in the similarly concentrated WT-*only* low control (see red dashed line), and therefore cannot be reliably called a positive mutant.

REFERENCES

1. Bidard FC, Weigelt B, Reis-Filho JS. Going with the Flow: From Circulating Tumor Cells to DNA. *Sci Transl Med*. 2013;5(207):207ps14. Epub 2013/10/18. doi: 10.1126/scitranslmed.3006305. PubMed PMID: 24132635.
2. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N. Analysis of Circulating Tumor DNA to Monitor Metastatic Breast Cancer. *N Engl J Med*. 2013;368(13):1199–209. Epub 2013/03/15. doi: 10.1056/NEJMoa1213261. PubMed PMID: 23484797.
3. De Mattos-Arruda L, Cortes J, Santarpia L, Vivancos A, Tabernero J, Reis-Filho JS, Seoane J. Circulating Tumour Cells and Cell-Free DNA as Tools for Managing Breast Cancer. *Nat Rev Clin Oncol*. 2013;10(7):377–89. Epub 2013/05/29. doi: 10.1038/nrclinonc.2013.80. PubMed PMID: 23712187.
4. Forshew T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, Dawson SJ, Piskorz AM, Jimenez-Linan M, Bentley D, Hadfield J, May AP, Caldas C, Brenton JD, Rosenfeld N. Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA. *Sci Transl Med*. 2012;4(136):136ra68. Epub 2012/06/01. doi: 10.1126/scitranslmed.3003726. PubMed PMID: 22649089.
5. Murtaza M, Dawson SJ, Tsui DW, Gale D, Forshew T, Piskorz AM, Parkinson C, Chin SF, Kingsbury Z, Wong AS, Marass F, Humphray S, Hadfield J, Bentley D, Chin TM, Brenton JD, Caldas C, Rosenfeld N. Non-Invasive Analysis of Acquired Resistance to Cancer Therapy by Sequencing of Plasma DNA. *Nature*. 2013;497(7447):108–12. Epub 2013/04/09. doi: 10.1038/nature12065. PubMed PMID: 23563269.
6. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, Kinzler KW, Vogelstein B, Diaz LA, Jr. Circulating Mutant DNA to Assess Tumor Dynamics. *Nat Med*. 2008;14(9):985–90. Epub 2008/08/02. doi: 10.1038/nm.1789. PubMed PMID: 18670422; PMCID: PMC2820391.
7. Figg WD, 2nd, Reid J. Monitor Tumor Burden with Circulating Tumor DNA. *Cancer Biol Ther*. 2013;14(8):697–8. Epub 2013/06/25. doi: 10.4161/cbt.25361. PubMed PMID: 23792566; PMCID: PMC3841207.
8. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW, Jr, Alizadeh AA, Diehn M. An Ultrasensitive Method for Quantitating Circulating Tumor DNA with Broad Patient Coverage. *Nat Med*. 2014;20(5):548–54. Epub 2014/04/08. doi: 10.1038/nm.3519. PubMed PMID: 24705333; PMCID: PMC4016134.
9. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, Gandara DR, Mack PC, Odegaard JI, Nagy RJ, Baca AM, Eltoukhy H, Chudova DI, Lanman RB, Talasaz A. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. *Clin Cancer Res*. 2018;24(15):3528–38. Epub 2018/05/20. doi: 10.1158/1078–0432.Ccr-17–3837. PubMed PMID: 29776953.
10. Xi L, Pham TH, Payabyab EC, Sherry RM, Rosenberg SA, Raffeld M. Circulating Tumor DNA as an Early Indicator of Response to T-Cell Transfer Immunotherapy in Metastatic Melanoma. *Clin Cancer Res*. 2016;22(22):5480–6. Epub 2016/08/03. doi: 10.1158/1078–0432.Ccr-16–0613. PubMed PMID: 27482033.
11. O'Leary B, Hrebien S, Morden JP, Beaney M, Fribbens C, Huang X, Liu Y, Bartlett CH, Koehler M, Cristofanilli M, Garcia-Murillas I, Bliss JM, Turner NC. Early Circulating Tumor DNA Dynamics and Clonal Selection

with Palbociclib and Fulvestrant for Breast Cancer. *Nat Commun*. 2018;9(1):896. Epub 2018/03/03. doi: 10.1038/s41467-018-03215-x. PubMed PMID: 29497091; PMCID: PMC5832789.

12. Riva F, Bidard FC, Houy A, Saliou A, Madic J, Rampanou A, Hego C, Milder M, Cottu P, Sablin MP, Vincent-Salomon A, Lantz O, Stern MH, Proudhon C, Pierga JY. Patient-Specific Circulating Tumor DNA Detection During Neoadjuvant Chemotherapy in Triple-Negative Breast Cancer. *Clin Chem*. 2017;63(3):691–9. Epub 2017/01/12. doi: 10.1373/clinchem.2016.262337. PubMed PMID: 28073896.
13. Garcia-Murillas I, Schiavon G, Weigelt B, Ng C, Hrebien S, Cutts RJ, Cheang M, Osin P, Nerurkar A, Kozarewa I, Garrido JA, Dowsett M, Reis-Filho JS, Smith IE, Turner NC. Mutation Tracking in Circulating Tumor DNA Predicts Relapse in Early Breast Cancer. *Sci Transl Med*. 2015;7(302):302ra133. Epub 2015/08/28. doi: 10.1126/scitranslmed.aab0021. PubMed PMID: 26311728.
14. Olsson E, Winter C, George A, Chen Y, Howlin J, Tang MH, Dahlgren M, Schulz R, Grabau D, van Westen D, Ferno M, Ingvar C, Rose C, Bendahl PO, Ryden L, Borg A, Gruvberger-Saal SK, Jernstrom H, Saal LH. Serial Monitoring of Circulating Tumor DNA in Patients with Primary Breast Cancer for Detection of Occult Metastatic Disease. *EMBO Mol Med*. 2015;7(8):1034–47. Epub 2015/05/20. doi: 10.15252/emmm.201404913. PubMed PMID: 25987569; PMCID: PMC4551342.
15. McDonald BR, Contente-Cuomo T, Sammut SJ, Odenheimer-Bergman A, Ernst B, Perdigones N, Chin SF, Farooq M, Mejia R, Cronin PA, Anderson KS, Kosiorek HE, Northfelt DW, McCullough AE, Patel BK, Weitzel JN, Slavin TP, Caldas C, Pockaj BA, Murtaza M. Personalized Circulating Tumor DNA Analysis to Detect Residual Disease After Neoadjuvant Therapy in Breast Cancer. *Sci Transl Med*. 2019;11(504). Epub 2019/08/09. doi: 10.1126/scitranslmed.aax7392. PubMed PMID: 31391323.
16. Underhill HR, Kitzman JO, Hellwig S, Welker NC, Daza R, Baker DN, Gligorich KM, Rostomily RC, Bronner MP, Shendure J. Fragment Length of Circulating Tumor DNA. *PLoS Genet*. 2016;12(7):e1006162. Epub 2016/07/18. doi: 10.1371/journal.pgen.1006162. PubMed PMID: 27428049; PMCID: PMC4948782.
17. Bio-Rad. Rare mutation detection best practices guidelines; Bulletin 6628.
18. Devereaux KA, Souers RJ, Merker JD, Lindeman NI, Graham RP, Hameed MR, Vasalos P, Moncur JT, Lockwood CM, Xian RR. Clinical Testing for Tumor Cell-Free DNA: College of American Pathologists Proficiency Programs Reveal Practice Trends. *Arch Pathol Lab Med*. 2023;147(4):425–33. doi: 10.5858/arpa.2021–0585-CP. PubMed PMID: 35687785; PMCID: PMC10189890.
19. Herzog H, Dogan S, Aktas B, Nel I. Targeted Sequencing of Plasma-Derived vs. Urinary cfDNA from Patients with Triple-Negative Breast Cancer. *Cancers (Basel)*. 2022 Aug 24;14(17):4101. doi: 10.3390/cancers14174101. PMID: 36077638; PMCID: PMC9454533.

24

Noninvasively Monitoring the Delivery of DNA Alkylating Agents and Cellular Responses to These Agents Using Chemical Exchange Saturation Transfer Magnetic Resonance Imaging

Ethel J. Ngen

24.1 Introduction

Central nervous system tumors (CNSTs) are the most common solid tumors and the second most common cancers diagnosed in children in the United States [1, 2]. Although treatment regimens vary depending on the tumor type, tumor grade, and patient age at diagnosis, standard treatment options for CNSTs include total-resection surgery followed by cranial or craniospinal radiotherapy and adjuvant chemotherapy [3, 4]. Given the greater vulnerability of the developing brain to ionizing radiation, radiotherapy is usually delayed in pediatric patients below the age of three or avoided completely in some subsets of pediatric patients, to prevent chronic neurological deficits and cognitive impairment associated with radiotherapy [3–5]. Thus, chemotherapy after total resection surgery is the preferred treatment option for this group of patients [3–5]. Chemotherapy is also the preferred treatment option for pediatric patients with tumors considered dispersed and inoperable [3–5].

24.1.1 Challenges of delivering drugs to brain tumors

A major impediment to the successful treatment of CNSTs has been a difficulty in specifically delivering high doses of chemotherapeutic agents to brain tumors, while sparing surrounding healthy tissue [6, 7]. Intravenously administered chemotherapeutic agents usually result in severe systemic side effects, which set a threshold for the maximum drug dose that can be used [6, 7]. One of the challenges of delivering high drug doses to brain tumors is the presence of the blood brain barrier (BBB), which restricts drug delivery into the brain [6, 7]. Consequently, several minimally invasive neurosurgical interventional procedures are currently being developed and evaluated to enhance the delivery of high drug doses across the BBB and into brain tumors, while also alleviating systemic side effects associated with high-dose chemotherapy [6, 7]. These include procedures such as convection-enhanced drug delivery, high-intensity focused ultrasound-assisted drug delivery, and intra-arterial drug delivery [8, 9]. Additionally, targeted drug delivery systems such as nanoparticles are also being developed and evaluated for drug delivery to brain tumors, with reduced systemic side effects [8–10]. The precision and efficacy of these targeted drug delivery systems can be tremendously enhanced by imaging approaches capable of noninvasively monitoring the drug delivery process [11–13]. Also critical is the need for imaging techniques capable of reporting on early tumor response to therapy [14, 15].

Current strategies to image drug delivery rely on using imaging agents as drug surrogates [11–13]. However, directly imaging drugs would be beneficial for monitoring their pharmacokinetics and biodistribution longitudinally. Although labeling drugs with radioactive isotopes provides a means to visualize drugs, this involves the use of ionizing radiation, which could be detrimental to pediatric patients [16]. Additionally, radiolabeling drugs could alter their physicochemical properties and subsequently their therapeutic efficacies [17]. Furthermore, standard methods to monitor tumor response to therapy rely on measuring either changes in bulk tumor volume or changes in tumor metabolism. However, these methods do not report on the efficiency of the drug delivery process [15, 18]. Additionally, tumor size reduction and treatment-induced metabolic changes usually require several weeks to occur [15, 18]. Thus, developing molecular imaging strategies capable of directly monitoring clinically approved drugs (without chemically modifying the drugs), and also capable of detecting early molecular changes associated with tumor response to therapy, such as pH changes, will be beneficial for the management of chemotherapeutic regimens in pediatric brain tumor patients. In this chapter, we describe a previously reported molecular imaging approach that enables the direct visualization of some clinically approved chemotherapeutic agents, using magnetic resonance imaging (MRI) [19].

24.1.2 Imaging DNA Alkylating Agents with Chemical Exchange Saturation Transfer Magnetic Resonance Imaging (CEST MRI)

Diamagnetic chemical exchange saturation transfer (CEST) magnetic resonance imaging (MRI), is an advanced MRI technique that can be used to detect organic molecules, unlike traditional MRI techniques that rely on metal-based contrast agents [19–21]. Although CEST MRI is based on a magnetic resonance spectroscopic (MRS) technique, it is several orders of magnitude more sensitive (~1000 times) than MRS [19–22].

DOI: 10.1201/9781003449942-27

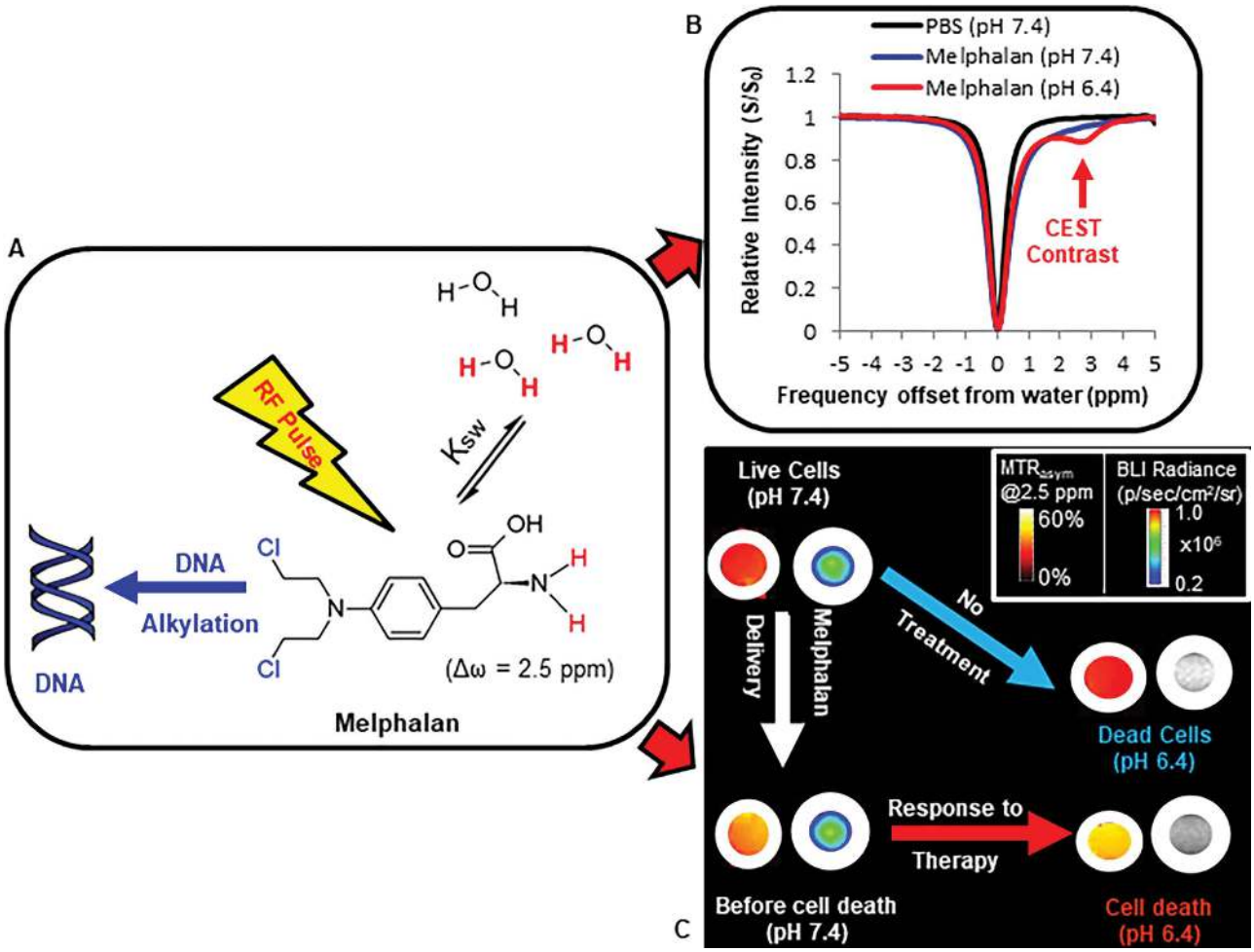


FIGURE 24.1 *Detection of the bi-functional DNA alkylating agent melphalan by CEST MRI.* (A) Mechanism of CEST MRI detection of melphalan. This shows simultaneous CEST imaging and DNA alkylation at different terminals of the molecule. (B) pH-dependent CEST MRI signals of melphalan (Z-spectra) at pH 6.4 and 7.4, and of phosphate buffered saline (PBS) at 7.4. (C) Schematic representing the detection of melphalan delivery to cancer cells and cellular responses to melphalan therapy.

This greater sensitivity has been attributed to the chemical exchange phenomenon between a proton on a molecule of interest and its surrounding water protons [19–21]. The exchangeable protons (amine, amide, and hydroxyl protons) of some organic molecules such as chemotherapeutic agents can generate a pH-dependent CEST MRI signal, depending on the molecular structure of the agent [19–21]. This pH-dependence of CEST MRI can be used to simultaneously detect drug delivery and early responses to therapy (Figure 24.1) [19–21].

DNA alkylating agents are integral components of all brain tumor treatment plans. Here we discuss a previously reported proof-of-concept study which showed that CEST MRI could be used to image 4-[bis(2-chloroethyl)amino]-l-phenylalanine (melphalan), a clinically used DNA alkylating agent, without the need for chemical modification of the drug and, subsequently, without compromising its therapeutic effectiveness (Figure 24.1) [19]. This study also demonstrated that the pH-dependent CEST MRI signal of melphalan could be used to image early molecular changes such as pH changes, associated with early cellular responses to melphalan therapy (Figure 24.1) [19].

24.2 Materials

24.2.1 Chemicals and Reagents

All chemicals and reagents were obtained from Sigma-Aldrich (St. Louis, MO) and Thermo Fisher Scientific (Waltham, MA), unless otherwise stated. D-luciferin potassium salt was obtained from PerkinElmer Inc. (Shelton, CT). The anti-melphalan mouse monoclonal antibody was obtained from Abcam (Waltham, MA), and the Alexa Fluor^{TM594} rabbit anti-mouse secondary antibody from Invitrogen (Waltham, MA).

24.2.2 Cell Culture

All cell culture reagents were obtained from Thermo Fisher Scientific (Waltham, MA), unless otherwise stated. Human medulloblastoma cells (DAOY HTB-186^{TM} cells), available at the American Type Culture Collection (Manassas, VA), were cultured in improved minimum essential media (Richter's modification), supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin [19].

24.3 Methods

24.3.1 Elucidating the pH-Dependent CEST MRI Signal of Melphalan

All CEST MRIs were acquired on an 11.7 T vertical Bruker Avance system (Billerica, MA) equipped with a 20 mm radio-frequency (RF) coil, at 37 °C. A high throughput screening sample holder, which permits the imaging of a large number of samples, was used to image the drug at different pH values and drug concentrations [23]. Briefly, 10 mM melphalan solutions were prepared in 10 mM PBS and the pH of the solutions were adjusted to 6.7, 7.7, and 8.7 respectively. Additional melphalan solutions were also prepared and imaged at lower concentrations (1.5, 3, and 6 mM) and at different pH values, ranging from 6.0 to 7.4. 80 μL of the respective solutions were then transferred and sealed in 1 mm micro-capillaries. The micro-capillaries were next placed in a micro-capillary holder and imaged.

To acquire CEST MRIs a modified spin echo pulse sequence, rapid acquisition with relaxation enhancement (RARE), including a magnetization transfer (MT) module, was used. Acquisition parameters included echo time (TE) = 5 ms; effective echo time (ETE) = 20 ms; repetition time (TR) = 6000 ms; RARE factor = 16; slice thickness (ST) = 1 mm; number of slices (NSL) = 1; field of view (FOV) = 17 × 17 mm; matrix size (MS) = 128 × 64; resolution = 0.13 × 0.26 mm^2; number of averages (NA) = 2; saturation power (B_1) = 4.6 μT; saturation module time (t_{sat}) = 4000 ms; initial saturation transfer frequency offset (Δω) = 5 ppm; final saturation transfer frequency offset (−Δω) = −5 ppm; saturation offset steps = 0.2 ppm; pulse shape = hermite; encoding order = linear; and acquisition time = 62.4 min [24].

The absolute water resonance frequency shift (B_0) was measured using a water saturation shift reference (WASSR) method modified with Lorentzian analysis, as previously described [23]. The same parameters as those used for the CEST MRI acquisitioning stated above were used for WASSR acquisitioning, except for the following modifications: TR = 1500 ms; t_{sat} = 500 ms, B_1 = 0.5 μT; Δω = 1.5 ppm; −Δω = −1.5 ppm; and saturation offset steps = 0.1 ppm. The acquired data was then processed using a custom-written MATLAB (Mathworks) scripts. The CEST contrast was next computed by calculating the magnetization transfer asymmetry ratio, as previously described [23, 24].

$$\mathbf{MTR_{Asym} = [(S_{-\Delta\omega} - S_{\Delta\omega})/S_0] \times 100\%} \quad \mathbf{(1)}$$

where S_0 represents the MRI signal intensity without RF saturation and $S_{-\Delta\omega}$ and $S_{+\Delta\omega}$ represent the MRI signal intensity with RF saturation at the frequency offsets of −Δω and +Δω relative to water (set at 0.0 ppm), respectively.

24.3.2 Imaging Melphalan Delivery *In Vitro* by CEST MRI

Luciferase-transduced DAOY cells were plated in T150 flasks (3×10^6 cells/flask) and cultured in improved minimum essential medium (Richter's modification), without serum for 18 h. The cells were then treated with melphalan at its determined IC_{80} (100 μM) for 6 h, then rinsed thrice with PBS, harvested, and 6×10^6 cells were counted using a cellometer (Nexcelom Bioscience). Next, the cells were pelleted in 4-inch nuclear magnetic resonance (NMR) spectroscopy tubes and the pH of the supernatant was measured. The melphalan-treated luciferase-transduced cell pellets were then imaged at pH 7.4, in comparison to control non-treated luciferase-transduced live DAOY cells at pH 7.4. CEST MRI was performed as described in the prior section, with the following modifications: ST = 2 mm; FOV = 20 × 20 mm; resolution = 0.16 × 0.31 mm^2; Δω = 10 ppm; −Δω = −10 ppm. Data processing and B_0 correction were also performed as described in the earlier section [19].

Bioluminescence imaging (BLI) of the cell pellets was then performed after CEST MRI. All BLI experiments were performed using a Xenogen IVIS 200 Series optical imaging system from PerkinElmer Inc. (Shelton, CT). Briefly, the cell pellets were incubated with 10 μL of a 150 μg/mL solution of D-luciferin potassium salt for 10 min, then imaged. The images were next processed using the Xenogen Living Imaging 4.2 software, and the BLI radiance from the samples was used to determine cell viability [19].

24.3.3 Imaging Therapeutic Responses to Melphalan Therapy

Luciferase-transduced DAOY cells were treated with melphalan at its determined IC_{80} (100 μM) for 6 h, then rinsed thrice with PBS, harvested, and 6×10^6 cells were counted. The cells were then pelleted in 4-inch NMR tubes and stored at room temperature. The pH of the supernatant was then monitored until it decreased to 6.4. The cell pellets were next imaged at pH 6.4, using both CEST MRI and BLI, in comparison to control non-treated luciferase-transduced dead DAOY cells at pH 6.4, as described in the prior section [19].

24.3.4 Immunocytochemistry of Melphalan-Treated Cells

Non-transduced DAOY cells were plated in 8-welled chamber slides (1×10^6 cells/well) and cultured without serum for 24 h. 100 μM of melphalan was next added to the cells for 6 h. Then the cells were rinsed thrice with PBS, fixed in acetone for 10 min at room temperature, air-dried for 10 min at room temperature, and stored overnight at −20 °C. The frozen chamber slides were next brought to room temperature, rinsed thrice with PBS for 5 min each time and then treated with 0.1 M of sodium hydroxide for 5 min. Melphalan–DNA adducts were then stained using an anti-melphalan mouse monoclonal antibody (Abcam, ab97554), prepared at a 1:100 dilution in PBS containing 1% (w/v) bovine serum albumin and 10% (v/v) horse serum. Next, the cells were incubated at 4 °C for 30 min with an Alexa Fluor594 rabbit anti-mouse secondary antibody (Invitrogen), prepared at a 1: 500 dilution in PBS containing 1% (w/v) bovine serum albumin and 10% (v/v) horse serum. Following this, the cells were rinsed thrice with PBS (5 min each time) and counter stained with DAPI. Images were next acquired with a Zeiss microscope equipped with the Zen Pro software [19].

24.4 Results

24.4.1 Elucidating the Signature pH-Dependent CEST MRI Signal of a DNA Alkylating Agent for Image-Guided Drug Delivery and Response Assessment

The exchangeable amine protons of the drug melphalan generated a CEST MRI signal at 2.5 ppm that was dependent on the pH of the milieu (Figure 24.2A and B). Generally, the CEST MRI signal was higher between pH 6.2 and 6.4 and lower between pH 7.0 and 7.4 (Figure 24.2). A maximum CEST MRI signal was detected at pH 6.2 (Figure 24.2C) [19].

This pH-dependent CEST MRI trend was detected at all the drug concentrations evaluated (Figure 24.2D). Additionally, a linear correlation was detected between the CEST MRI signal at 2.5 ppm and the drug concentration at all pH values. However, this linear correlation was greatest at pH values between 6.0 and 6.4. This suggested the feasibility of using the CEST MRI signal from the drug to quantify its delivery to tumor cells.

Furthermore, since the optimum pH for imaging the drug is close to the pH of the clinical formulation of the drug (pH 6.5), this suggested the feasibility of directly utilizing the drug as a theranostic agent.

24.4.2 Utilizing the Signature pH-Dependent CEST MRI Signal of a DNA Alkylating Agent for Image-Guided Drug Delivery and Response Assessment

Next, we evaluated the feasibility of using the signature pH-dependent CEST MRI signal of the drug melphalan to detect its delivery to cancer cells. We also evaluated the feasibility of using the drug's signature pH-dependent CEST MRI signal to detect pH changes, associated with cell death in response to therapy. Luciferase-transduced cancer cells with a transduction efficacy (≥ 99%) were used in this study to enable the validation of the CEST MRI results with bioluminescence imaging (BLI) [19].

Briefly, the CEST MRI contrast generated by melphalan could be detected *in vitro* following drug administration to the cells at

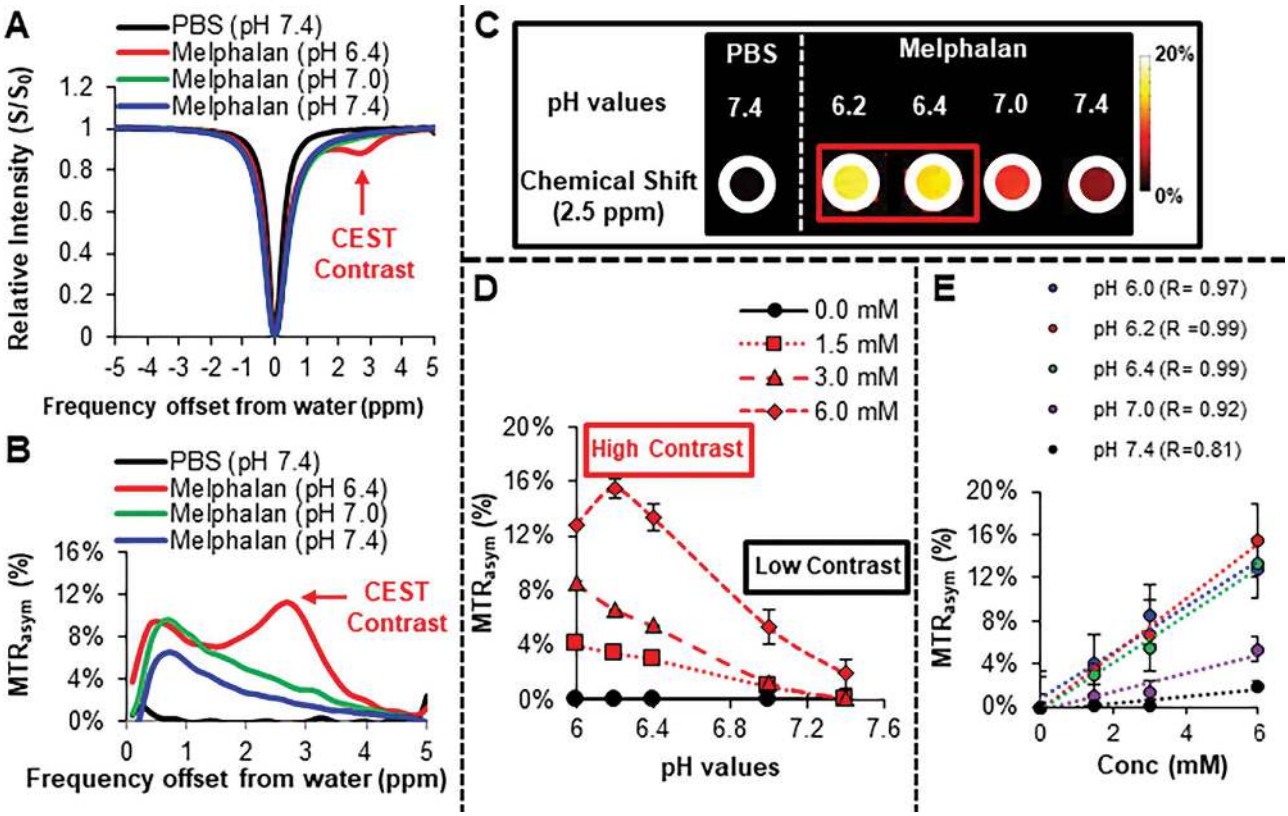


FIGURE 24.2 *pH-dependence of the CEST MRI signal of the bi-functional DNA alkylating agent melphalan.* (A) The CEST MRI Z-spectra of 10 mM melphalan solutions at different pH values (pH 6.0, 6.4, and 7.4); and 10 mM PBS at pH 7.4. The Z-spectra indicates a high CEST signal from melphalan at pH values close to 6.4. (B) CEST MRI magnetization transfer asymmetry ratio (MTR_{asym}) plots of 10 mM melphalan solutions at different pH values (pH 6.0, 6.4, and 7.4); and 10 mM PBS at pH 7.4. The MTR_{asym} plots indicate a high CEST contrast from melphalan close to pH 6.4. (C) CEST images (MTR_{asym} maps) of 10 mM melphalan solutions at different pH values (pH 6.2, 6.4, 7.0, and 7.4) and 10 mM PBS at pH 7.4. A higher CEST contrast was detected between pH 6.0 and 6.4; and a lower contrast between pH 7.0 and 7.4. (D) A graph of the pH-dependence of the CEST signal (MTR_{asym}) of three melphalan solutions at different concentrations (1.5, 3.0, and 6 mM). A higher CEST contrast was detected between pH 6.0 and 6.4; and a lower contrast between pH 7.0 and 7.4, in all three solutions. (E) A graph of the effects of drug concentration on CEST signal at different pH values indicates a higher linear correlation at pH 6.0 to pH 6.4. This linear correlation could be used to quantify the delivered drug concentration and also to assess the effects of drug dose on treatment response. All measurements were done at 37 °C.

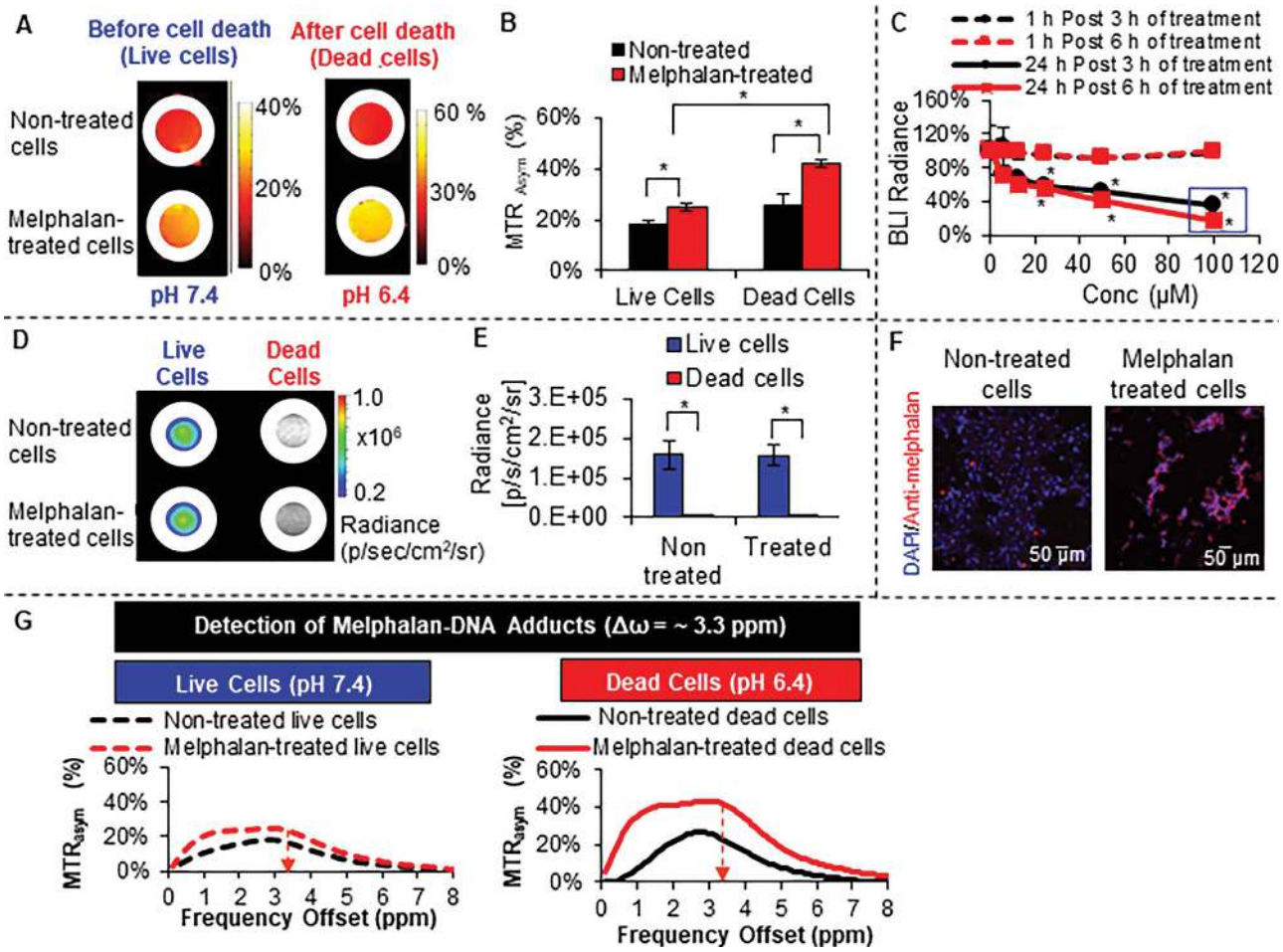


FIGURE 24.3 *Detecting cellular responses to melphalan therapy using melphalan's pH-dependent CEST MRI signal.* (A) CEST images (MTR_{asym} maps) of luciferase-transduced DAOY cells treated with 100 μM of melphalan and non-treated control luciferase-transduced DAOY cells; before cell death (pH 7.4) and after cell death (pH 6.4), respectively. (B) Quantitative CEST MRI signal differences between melphalan-treated and non-treated control luciferase-transduced DAOY cells before cell death (pH 7.4) and after cell death (pH 6.4), respectively. Upon cell death, a decrease in the cellular pH from pH 7.4 to 6.4 caused an amplification of the CEST signal in melphalan-treated dead cells. (C) Melphalan cytotoxicity of luciferase-transduced DAOY cells treated with various concentrations of melphalan for 3 h and 6 h, respectively; and assessed with bioluminescence imaging immediately and 24 h post-treatment, respectively. An IC_{80} of 100 μM was determined from the cytotoxicity and used in the CEST MRI cellular studies. (D) Bioluminescence images (BLIs) of live and dead melphalan-treated and non-treated control DAOY cells, respectively. (E) Quantitative BLI signal differences between melphalan-treated live versus dead cells and non-treated control live versus dead cells, respectively. (F) Immunocytochemistry of DAOY cells treated with 100 μM of melphalan for 6 h and non-treated control cells. Both cell groups were stained with a melphalan-specific antibody (red) and counter-stained with DAPI (blue). Melphalan was detected only in the treated cells (as expected). (G) CEST MRI (MTR_{asym} plots) of non-treated and melphalan-treated live DAOY cell pellets at pH 7.4, and dead DAOY cell pellets at pH 6.4, respectively. An amplification of the CEST MRI signal was detected in melphalan-treated dead cells following cell death and a decrease in the cellular pH from 7.4 to 6.4.

the drug's IC_{80} and at pH 7.4 (Figure 24.3A–C). In melphalan-treated luciferase-transduced live cells, a significant difference in the CEST MRI contrast was detected one-hour post-treatment at pH 7.4, compared to the non-treated control luciferase-transduced live cells (Figure 24.3). Immunocytochemistry, using a melphalan-specific antibody, also confirmed the presence of melphalan in the melphalan-treated cells (Figure 24.3F). Interestingly, a shift in the detection frequency offset maxima from 2.5 ppm to 3.1 ppm suggested the detection of melphalan bound to DNA (melphan-DNA adducts) in the melphalan-treated cells and not the free drug (Figure 24.3G) [19].

In addition to imaging the delivery of melphalan *in vitro*, we evaluated the feasibility of using melphalan's amplified CEST MRI contrast between pH 6.2 and 6.4, to monitor cellular responses to melphalan therapy. A significant difference in the CEST MRI contrast was detected in the melphalan-treated luciferase-transduced dead cells at pH 6.4, compared to the non-treated control luciferase-transduced dead cells imaged at pH 6.4 (Figure 24.3) [19].

Additionally, the CEST contrast generated from melphalan after cell death at pH 6.4 was significantly greater than that generated from melphalan in live cells at pH 7.4, before cell death (Figure 24.3A and B). Upon cell death, a decrease in the cellular pH from 7.4 to 6.4 caused an amplification of the melphalan-DNA adduct CEST signal in melphalan-treated dead cells [19].

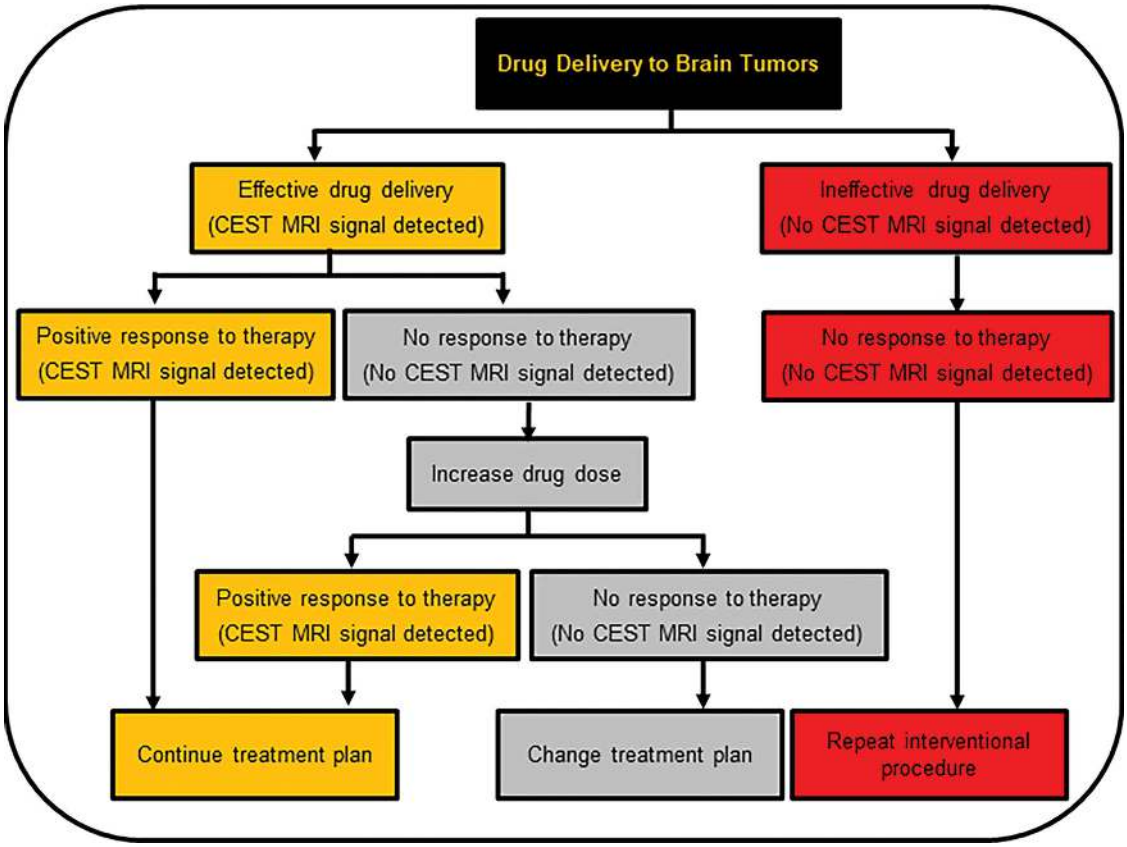


FIGURE 24.4 Schematic of patient stratification and decision-making based on a drug's CEST MRI signal.

24.5 Discussion

Overall, this study shows the feasibility of using the signature pH-dependent CEST MRI signal of a DNA alkylating agent melphalan to noninvasively monitor its delivery to cancer cells and also cellular responses to melphalan therapy.

24.6 Future Trends

This imaging method could be used to noninvasively monitor drug delivery to brain tumors, especially inoperable tumors, during interventional neurosurgical procedures and could also enable the monitoring of early responses to therapy. Furthermore, since several clinically approved drugs used in the treatment of brain tumors possess exchangeable protons that could be imaged by CEST MRI (depending on their chemical structure), this method could be extended to other drugs to develop an imaging library of clinically used drugs for image-guided drug delivery.

Ultimately, this could enable more effective patient management by permitting the rapid selection of the most effective drug and drug dose. This could also permit the early stratification of patients into responding and non-responding patients and guide decision-making on future treatment plans (Figure 24.4).

These efforts could bring personalized therapeutic regimens one step closer to realization for CNST patients and also enable the development of more effective and less toxic therapeutic regimens for pediatric brain tumor patients, who are more susceptible to brain injury from cancer treatment.

Acknowledgement

Reprinted (adapted) with permission from Ngen EJ, Bar-Shir A, Jablonska A, Liu G, Song X, Ansari R, Bulte JW, Janowski M, Pearl M, Walczak P, Gilad AA: Imaging the DNA Alkylator Melphalan by CEST MRI: An Advanced Approach to Theranostics. *Molecular Pharmaceutics* 2016, 13(9):3043–3053. Copyright 2016 American Chemical Society. This research was sponsored by the National Institutes of Health under Awards Numbers R21HD097357 and R01CA262887, respectively.

REFERENCES

1. Miller KD, Ostrom QT, Kruchko C, Patil N, Tihan T, Cioffi G, Fuchs HE, Waite KA, Jemal A, Siegel RL *et al.*: Brain and other central nervous system tumor statistics, 2021. *CA: A Cancer Journal for Clinicians* 2021, 71(5):381–406.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A: Cancer statistics, 2021. *CA: A Cancer Journal for Clinicians* 2021, 71(1):7–33.

3. Pollack IF, Agnihotri S, Broniscer A: Childhood brain tumors: Current management, biological insights, and future directions: JNSPG 75th anniversary invited review. *Article Journal of Neurosurgery: Pediatrics PED* 2019, 23(3):261–273.
4. Plant-Fox AS, O'Halloran K, Goldman S: Pediatric brain tumors: The era of molecular diagnostics, targeted and immune-based therapeutics, and a focus on long term neurologic sequelae. *Current Problems in Cancer* 2021, 45(4):100777.
5. Hollingworth MA, Chattopadhyay A, Walker D, Paine S, Kirkman MA: What you should you know about pediatric brain tumours in 2022. *Pediatrics and Child Health* 2022, 32(9):339–350.
6. Nailor A, Walker DA, Jacques TS, Warren KE, Brem H, Kearns PR, Greenwood J, Penny JI, Pilkington GJ, Carcaboso AM *et al.*: Highlights of children with cancer UK's workshop on drug delivery in pediatric brain tumours. *Ecancermedicalscience* 2016, 10:630.
7. Warren KE: Beyond the blood: Brain barrier: The importance of central nervous system (CNS) pharmacokinetics for the treatment of CNS tumors, including diffuse intrinsic pontine glioma. *Frontiers in Oncology* 2018, 8:239.
8. Mitusova K, Peltek OO, Karpov TE, Muslimov AR, Zyuzin MV, Timin AS: Overcoming the blood–brain barrier for the therapy of malignant brain tumor: Current status and prospects of drug delivery approaches. *Journal of Nanobiotechnology* 2022, 20(1):412.
9. Noorani I, de la Rosa J: Breaking barriers for glioblastoma with a path to enhanced drug delivery. *Nature Communications* 2023, 14(1):5909.
10. Miao Y-B, Zhao W, Renchi G, Gong Y, Shi Y: Customizing delivery nano-vehicles for precise brain tumor therapy. *Journal of Nanobiotechnology* 2023, 21(1):32.
11. Ishida J, Alli S, Bondoc A, Golbourn B, Sabha N, Mikloska K, Krumholtz S, Srikanthan D, Fujita N, Luck A *et al.*: MRI-guided focused ultrasound enhances drug delivery in experimental diffuse intrinsic pontine glioma. *Journal of Controlled Release* 2021, 330:1034–1045.
12. Bell M, Rooks CP, Agrahari V: Drug delivery approaches and imaging techniques for brain tumor. In: *Nanotherapy for Brain Tumor Drug Delivery.* Edited by Agrahari V, Kim A, Agrahari V. New York, NY: Springer US; 2021: 109–126.
13. Priester MI, ten Hagen TLM: Image-guided drug delivery in nanosystem-based cancer therapies. *Advanced Drug Delivery Reviews* 2023, 192:114621.
14. Bhatia A, Sabin ND, Fisher MJ, Poussaint TY: Review of imaging recommendations from response assessment in pediatric neuro-oncology (RAPNO). *Pediatric Radiology* 2023, 53(13):2723–2741.
15. Segtnan EA, Hess S, Grupe P, Høilund-Carlsen PF: 18F-Fluorodeoxyglucose PET/computed tomography for primary brain tumors. *PET Clinics* 2015, 10(1):59–73.
16. Vaidyanathan S, Reed A: Pipeline impact of radiolabeled compounds in drug discovery and development. *ACS Medicinal Chemistry Letters* 2022, 13(10):1564–1567.
17. Vallabhajosula S, Killeen RP, Osborne JR: Altered biodistribution of radiopharmaceuticals: Role of radiochemical/pharmaceutical purity, physiological, and pharmacologic factors. *Seminars in Nuclear Medicine* 2010, 40(4):220–241.
18. Warren KE, Poussaint TY, Vezina G, Hargrave D, Packer RJ, Goldman S, Wen PY, Pollack IF, Zurakowski D, Kun LE *et al.*: Challenges with defining response to antitumor agents in pediatric neuro-oncology: A report from the response assessment in pediatric neuro-oncology (RAPNO) working group. *Pediatric Blood & Cancer* 2013, 60(9):1397–1401.
19. Ngen EJ, Bar-Shir A, Jablonska A, Liu G, Song X, Ansari R, Bulte JW, Janowski M, Pearl M, Walczak P *et al.*: Imaging the DNA alkylator melphalan by CEST MRI: An advanced approach to theranostics. *Molecular Pharmaceutics* 2016, 13(9):3043–3053.
20. van Zijl PCM, Yadav NN: Chemical exchange saturation transfer (CEST): What is in a name and what isn't? *Magnetic Resonance in Medicine* 2011, 65(4):927–948.
21. Li Y, Chen H, Xu J, Yadav NN, Chan KW, Luo L, McMahon MT, Vogelstein B, van Zijl PC, Zhou S *et al.*: CEST theranostics: Label-free MR imaging of anticancer drugs. *Oncotarget* 2016, 7(6):6369–6378.
22. Kogan F, Hariharan H, Reddy R: Chemical exchange saturation transfer (CEST) imaging: Description of technique and potential clinical applications. *Current Radiology Reports* 2013, 1(2):102–114.
23. Liu G, Gilad AA, Bulte JW, van Zijl PC, McMahon MT: High-throughput screening of chemical exchange saturation transfer MR contrast agents. *Contrast Media & Molecular Imaging* 2010, 5(3):162–170.
24. Bar-Shir A, Liu G, Liang Y, Yadav NN, McMahon MT, Walczak P, Nimmagadda S, Pomper MG, Tallman KA, Greenberg MM *et al.*: Transforming thymidine into a magnetic resonance imaging probe for monitoring gene expression. *Journal of the American Chemical Society* 2013, 135(4):1617–1624.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Section 3

Nucleic Acid Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

25

Clonal Analysis of Cancer by Mitochondrial DNA Barcoding

Jonathan D. Good, Ksenia R. Safina, and Peter van Galen

25.1 Introduction

Cancer cell heterogeneity is a key obstacle to understanding cancer development and discovering new treatments [1]. Genetic heterogeneity is characterized by the sequential acquisition of DNA mutations driving clonal expansion. Cell-intrinsic and -extrinsic factors contribute to the heterogeneous landscape by driving cell state transitions and differentiation hierarchies [2, 3]. The relationship between cancer cell differentiation states and outcomes in patients with acute myeloid leukemia (AML) demonstrates the importance of understanding these layers of heterogeneity [4, 5]. Further work highlights that drug sensitivity is impacted by both genetic heterogeneity and cancer cell differentiation states [6–11]. These findings emphasize the need to characterize clonal evolution across diverse transcriptional and epigenetic states to understand cancer cell transformation better and improve treatment options.

Single-cell sequencing tools are essential for illuminating complex tumor ecosystems. This field was pioneered by single-cell RNA sequencing (scRNA-seq) technologies, which give a snapshot of the transcriptional landscape [12, 13]. In addition, multimodal analyses with DNA mutations, chromatin accessibility, and protein levels are gaining popularity [5, 14–16]. Tracing clonal lineages while capturing these single-cell dynamics can build a more comprehensive picture of cellular transformation. Traditional cell lineage tracing methods use nuclear genetic variation and rely on costly tools like whole-genome sequencing (WGS) [17–20]. Additionally, these methods capture lineage information separately from scRNA-seq. The demonstration that mitochondrial genomic variants worked as endogenous cell barcodes opened the possibility of developing a cost-effective tool that combines scRNA-seq and clonal lineage information (Figure 25.1) [21]. This technology has the potential to generate detailed profiles of rare clonal cell populations that drive disease initiation, progression, and recurrence. Understanding the properties of these cells can help identify vulnerabilities and target them more effectively.

To effectively capture mtDNA variants from scRNA-seq cDNA libraries, we developed mitochondrial alteration enrichment from single-cell transcriptomes to establish relatedness (MAESTER). It was initially tested on a CML cell line (K562) mixed with a brain tumor cell line (BT142) and was validated on a human bone marrow sample with clonal hematopoiesis [22]. Further optimization was done using human bone marrow samples from subjects without hematological abnormalities. Along with work from other groups discussed later, these examples underscore MAESTER's capability to accurately assign clonal relationships between cells and demonstrate its effectiveness in diverse experimental settings.

25.2 Materials

25.2.1 Starting cDNA and PCR Primers

The starting material for mitochondrial transcript enrichment is cDNA prepared using a 10x Genomics 3′ Single Cell Gene Expression v3 chemistry kit (PN: 1000268). All PCR primers are custom DNA oligos from Integrated DNA Technologies. Building on original MAESTER primers [22], we recently published an updated list of primers that facilitate experimental setup and sequencing [24].

25.2.2 PCR Reaction

All PCR reactions used KAPA HiFi Hotstart ReadyMix (Thermo Fisher Scientific, KK2602) and are run using an Eppendorf Mastercycler X50 PCR thermocycler. Ambion nuclease-free water (ddH_2O) is used to dilute the PCR reactions and ethanol (Life Technologies, AM9937). PCR product cleaning used AMPure XP beads (Beckman Coulter, A63881). Ethanol from Decon Labs is used for bead washes (#2716). DNA is eluted in TE buffer from Fisher Bioreagents (BP2473-100).

25.2.3 PCR Product Quantification and Quality Assessment

DNA is quantified using a Qubit 4 Fluorometer (Thermo Fisher Scientific) with the 1X dsDNA HS Assay kit (Q33231) and size assessment is performed using a Tapestation automated electrophoresis system (Agilent 4150).

25.2.4 Sequencing and Analysis

Final mitochondrial libraries were submitted for sequencing using an Illumina NovaSeq SP 300-cycle kit aiming for 150 million reads per sample. Data processing was performed using the maegatk package (https://github.com/caleblareau/maegatk) as detailed in Section 25.3.7.

DOI: 10.1201/9781003449942-29

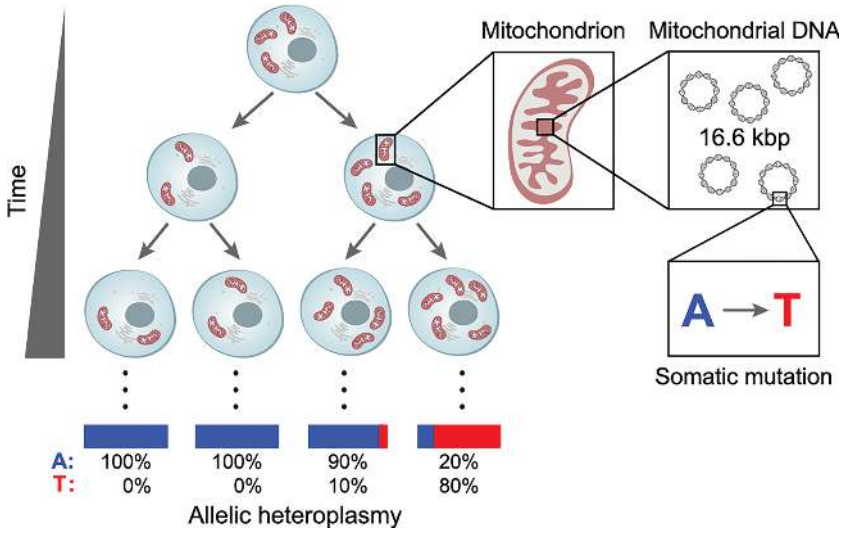


FIGURE 25.1 *Diagram of mtDNA variant acquisition and propagation*. Cancer cells contain ~10–1000 mtDNA copies [23] that can acquire somatic variants over time. Some variants rise to a heteroplasmy level that enables their use for inference of clonal relationships in a snapshot analysis (bottom row). (The figure is adapted with permission from Ludwig et al., 2019 [21].)

TABLE 25.1

Primer Mix Volumes for PCR1

Primer Mix	Forward Primer	Forward Primer Volume (µl)	Number of Reverse Primers (100 µM each)	Reverse Primer Volume (µl)	Add TE (µl)	Total Volume (µl)
Mix 1	PvG1200-PartialRead1 (100 µM stock)	10	11	1 µl each	79	100
Mix 2	PvG1200-PartialRead1 (100 µM stock)	10	11	1 µl each	79	100
Mix 3	PvG1200-PartialRead1 (100 µM stock)	10	10	1 µl each	80	100
Mix 4	PvG1200-PartialRead1 (100 µM stock)	10	8	1 µl each	82	100
Mix 5	PvG1200-PartialRead1 (100 µM stock)	10	5	1 µl each	85	100
Mix 6	PvG1200-PartialRead1 (100 µM stock)	10	5	1 µl each	85	100
Mix 7	PvG1200-PartialRead1 (100 µM stock)	10	4	1 µl each	86	100
Mix 8	PvG1200-PartialRead1 (100 µM stock)	10	2	1 µl each	88	100
Mix R1	PvG1200-PartialRead1 (100 µM stock)	10	3	1 µl each	87	100
Mix R2	PvG1200-PartialRead1 (100 µM stock)	10	2	1 µl each	88	100
Mix R3	PvG1200-PartialRead1 (100 µM stock)	10	2	1 µl each	88	100
Mix R4	PvG1200-PartialRead1 (100 µM stock)	10	2	1 µl each	88	100

25.3 Methods

25.3.1 cDNA Preparation

The cDNA libraries for MAESTER's initial development are described in Miller and colleagues [22]. Libraries for optimization and further development were prepared from our sternum bone marrow (SBM) biobank. The samples are from Mass General Brigham patients undergoing sternotomy for various cardiac conditions who consented to all study procedures. We selected patients with or without cancer-associated mutations in alignment with a separate project's parameters. We constructed the scRNA libraries using 10x Genomics 3′ Single Cell Gene Expression kits.

25.3.2 Primer Plates

Custom oligo primers were ordered from IDT in dry tubes and were resuspended at 100 uM TE per the manufacturer's directions. We made primer mixes for both PCR1 and PCR2 in 96-well plates to streamline reaction setup. They were stored at −20 °C between uses. PCR1 has 65 individual mtDNA primers combined to make 12 final mixes. Mixes 1–8 provide overlapping coverage of the entire human mitochondrial transcriptome and Mixes R1–R4 cover the most abundant mtDNA transcripts. The 12 final PCR primer mixes were prepared following Table 25.1 and added to a row of a 96-well plate for easy multichannel pipetting. For the PCR2 primer plate, we added 90 µl of TE buffer to a row of wells. Then we added 5 µl of the 100 uM XV-P5-i5-BCxx forward primer and 5 µl of the 100 µM XV-P7-i7-BCxx reverse primers in each well, ensuring a single barcode per well.

25.3.3 cDNA Amplification

This step is not required but helps by increasing the amount of starting cDNA. We took 10 ng of cDNA and added 1.5 µl each of the PartialRead primer 1 (10 µM) and the PartialTSO

primer (10 µM). Next, we added 25 µl of the KAPA polymerase on ice. Finally, we added ddH_2O to make a total reaction volume of 50 µl. We programmed a thermocycler with the initial denaturation at 95 °C for 3 minutes. This was followed by eight cycles of a 20-second denaturation at 98 °C, 30 seconds of annealing at 67 °C, and a 3-minute extension at 72 °C. The program was finished with a 3-minute final extension at 72 °C and holding at 4 °C. The amplified product was cleaned using a 0.6 times ratio of AMPure XP beads according to the manufacturer's directions. The final elution was in 40 µl of TE buffer.

25.3.4 PCR1 Reaction and Purification

PCR1 uses the suite of locus-specific mitochondrial primers to enrich the mtDNA from the full-length cDNA. Twelve reactions were set up per cDNA sample using the 12 primer mixes prepared in Section 25.3.2. We made a master mix for 13 reactions to account for pipetting loss. We added 20 ng of cDNA per reaction (260 ng total) to 20 µl of KAPA ReadyMix on ice per reaction (260µl total). We factored in the 4 µl of primers that will be added later and added ddH_2O to make a 40 µl total volume per reaction (520 µl total). In a PCR-rated plate, we added 36 µl of this master mix to each well for 12 reactions per cDNA sample. Finally, we added 4 µl of the primer mixes separately to each reaction and programmed a thermocycler with the initial denaturation at 95 °C for 3 minutes. This was followed by six cycles of a 20-second denaturation at 98 °C, 15 seconds of annealing at 67 °C, and a 3-minute extension at 72 °C. The program was finished with a 3-minute final extension at 72 °C and holding at 4 °C. The PCR products were combined according to Table 25.2. These volumes were determined using trial and error to equalize coverage. We cleaned the combined product with a 0.8 times ratio of AMPure XP beads and eluted in 20 µl of TE buffer. Qubit and Tapestation were used to measure the product (Figure 25.2A).

25.3.5 PCR2 Reaction and Purification

PCR2 adds library indices to enable sample pooling and to make the libraries compatible with the Illumina platform. To the remaining volume from PCR1 (~18 µl), we added 2 µl of the PCR2 primer mix (see Section 25.3.2) and 25 µl of the KAPA polymerase on ice. Finally, we added ddH_2O to a total volume of 50 µl. We programmed a thermocycler with the initial denaturation at 95 °C for 3 minutes. This was followed by six cycles of a 20-second denaturation at 98 °C, 15 seconds of annealing at 67 °C, and a 3-minute extension at 72 °C. The program was finished with a 3-minute final extension at 72 °C and holding at 4 °C. The product was cleaned using a 0.8 times ratio of AMPure XP beads and eluted in 20 µl of TE buffer. Qubit and Tapestation were used to measure the product (Figure 25.2B).

TABLE 25.2

Directions for Combining PCR1 Products

PCR Primer Mix	µl from Rxn
1	32
2	40
3	40
4	40
5	40
6	40
7	40
8	40
R1	8
R2	8
R3	8
R4	16
Total	352

25.3.6 Sequencing

Each sample's final library has a unique barcode that allows sample pooling when sequencing. Pooling volume was calculated to reach 5 nM of DNA per sample. We submitted pools of four samples (100 µl total at 5 nM each) to be run on an Illumina NovaSeq SP 300 cycle kit, with a target minimum of 150 million reads per sample. The run parameters were 28 cycles for Read 1,252 cycles for Read 2, and 10 cycles each for Index 1 (i7) and Index 2 (i5).

25.3.7 Analysis

To generate variant count matrices from raw MAESTER reads, we developed a pipeline that includes the maegatk package [22]. This python-based command line interface processes .bam files with mitochondrial reads and generates high-quality heteroplasmy estimation from sequencing data. A key feature is the consensus base inference by collapsing reads with identical insert positions and UMIs. This allows maegatk to produce robust, error-corrected genotype calls in single cells. These calls are summarized in a RangedSummarizedExperiment object saved as an .rds file readily available for analysis in R. The object is composed of 16 matrices which report the number of supporting consensus reads and the mean sequencing quality of every possible variant in every captured cell in both forward and reverse directions (A, T, G, C × (forward read count, forward base quality, reverse read count, reverse base quality)).

From these matrices, our approach was first to select informative variants [22]. The total number of possible variants is $3 \times 16{,}569 + 1 = 49{,}708$, because the mitochondrial genome (NC_012920) is 16,569 bp with three possible variants each, except base 3107, which has four possible variants (A, C, T, and G), because the reference is N. To select informative variants, we generated a table with information on every possible variant: the mean heteroplasmy level, mean coverage, mean quality score, and the number of cells exceeding a chosen heteroplasmy level. Then, we applied filters to select variants with high coverage, high quality, and an informative heteroplasmy distribution.

25.4 Results

We calculated the average amplification across 17 samples for both PCRs. PCR1 averaged 8.4±3.4 fold amplification with an average fragment size of 1138 base pairs. PCR2

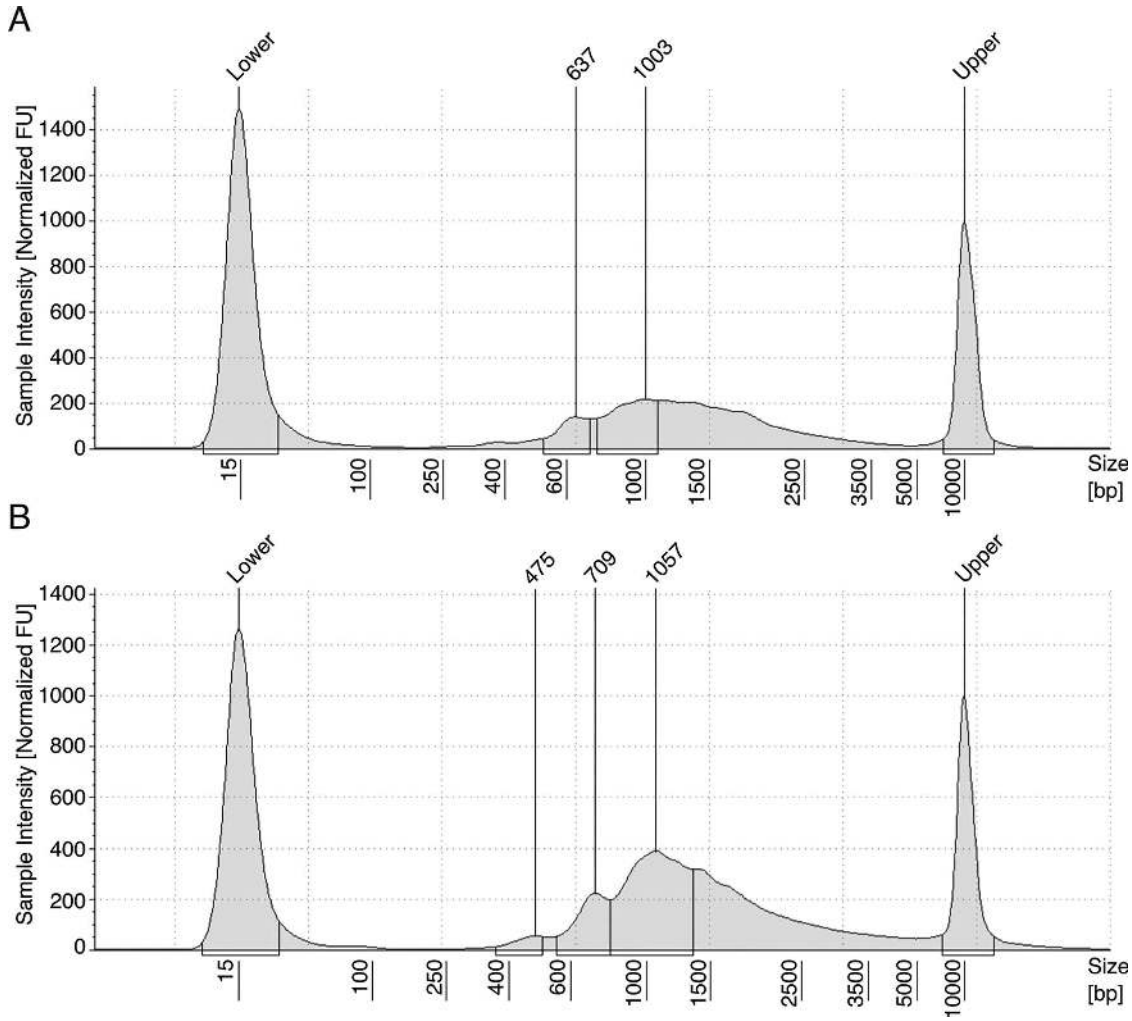


FIGURE 25.2 *Representative Tapestation analysis.* These are Tapestation results from MAESTER on a human bone marrow sample. (A) PCR1 results. (B) PCR2 results.

averaged 2.9±1.8 fold amplification with an average fragment size of 1267 base pairs. See Figure 25.2 for a representative Tapestation analysis.

Miller and colleagues demonstrated MAESTER efficiency using K562 and BT142 cell lines [22]. MAESTER resulted in a hundred-fold increase in the mean coverage (number of UMIs) per cell and a more uniform coverage along the mitochondrial genome compared to scRNA-seq data (Figure 25.3A). They validated these findings using a bone marrow aspirate from a clonal hematopoiesis patient where MAESTER demonstrated a cell-type-dependent coverage (Figure 25.3B). The difference in coverage between cell types was highly correlated to the cells' total mRNA content, and mitochondrial gene expression was very similar between different blood cell types [22]. To investigate clonal structures in patients without any hematological abnormalities, we applied MAESTER to samples from our SBM biobank and selected 5–25 informative variants per sample. We selected cells with any of these informative variants, using a heteroplasmy threshold of ≥ 1% to consider a cell positive. We visualized clonal structures by depicting variants (rows) and cells (columns) colored by the heteroplasmy level (Figure 25.4A). Subsequently, we colored cells in a UMAP dimensionality reduction by heteroplasmy levels to visually evaluate the potential enrichment of variants in certain cell populations. Some variants showed lineage skewing (11456_G>A and 9966_G>A), whereas others did not (15383_T>C, Figure 25.4B). Interestingly, 11456_G>A was indicative of an NK cell expansion, consistent with the recent discovery of clonal persistence of NK cells with heritable epigenetic features [25].

25.5 Discussion

We have shown MAESTER's utility in identifying cell lineage relationships and clonal expansion in the context of normal and malignant hematopoiesis. Other researchers are using similar methods of mtDNA-mediated clonal tracing to understand leukemic transformation from healthy HSCs or to trace variants in the context of chronic lymphocytic leukemia [26,27]. In addition to hematological disease states, MAESTER has been applied to study clonal expansions in

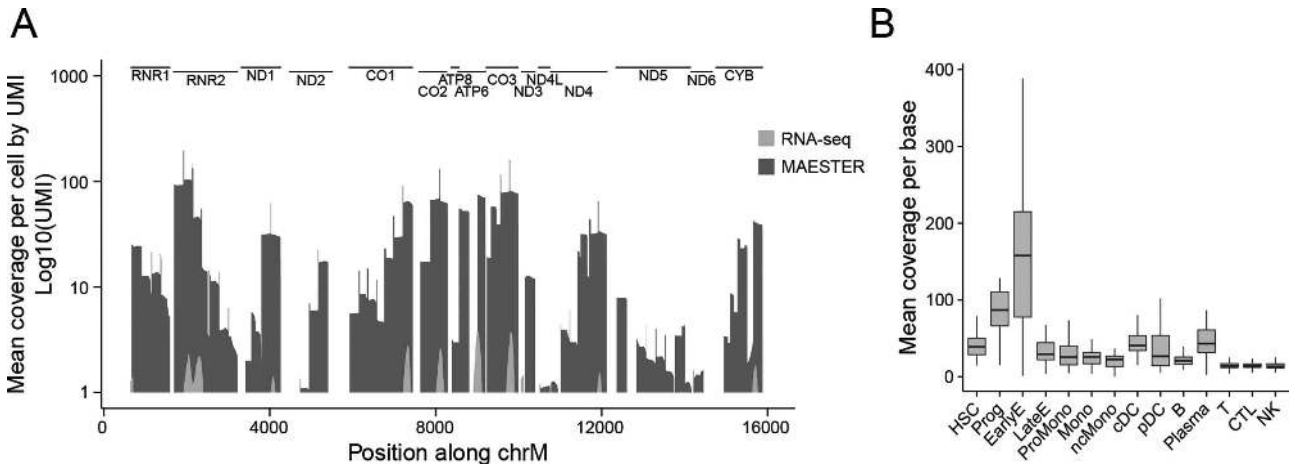


FIGURE 25.3 *Typical mitochondrial genome coverage from MAESTER.* (A) Barplot shows mean mitochondrial genome coverage in 9346 cells from a typical bone marrow sample with and without amplification from 10x libraries using MAESTER. Each of the transcripts (UMIs) was sequenced ≥ three times. (B) Boxplot shows the mean MAESTER coverage per base along the mitochondrial genome per cell (y-axis) split by cell types (on the x-axis). Boxplots show the median, 25th and 75th percentiles, and smallest/largest values (geom_boxplot defaults). HSC: hematopoietic stem cell, Prog: progenitor, E: erythroid, ProMono: promonocyte, Mono: classical monocyte, ncMono: non-classical monocyte, cDC: conventional dendritic cell, pDC: plasmacytoid dendritic cell, B: B cell, Plasma: plasma cell, T: T cell, CTL: cytotoxic T lymphocyte, NK: natural killer. (The figure is adapted with permission from Miller et al., 2022 [22].)

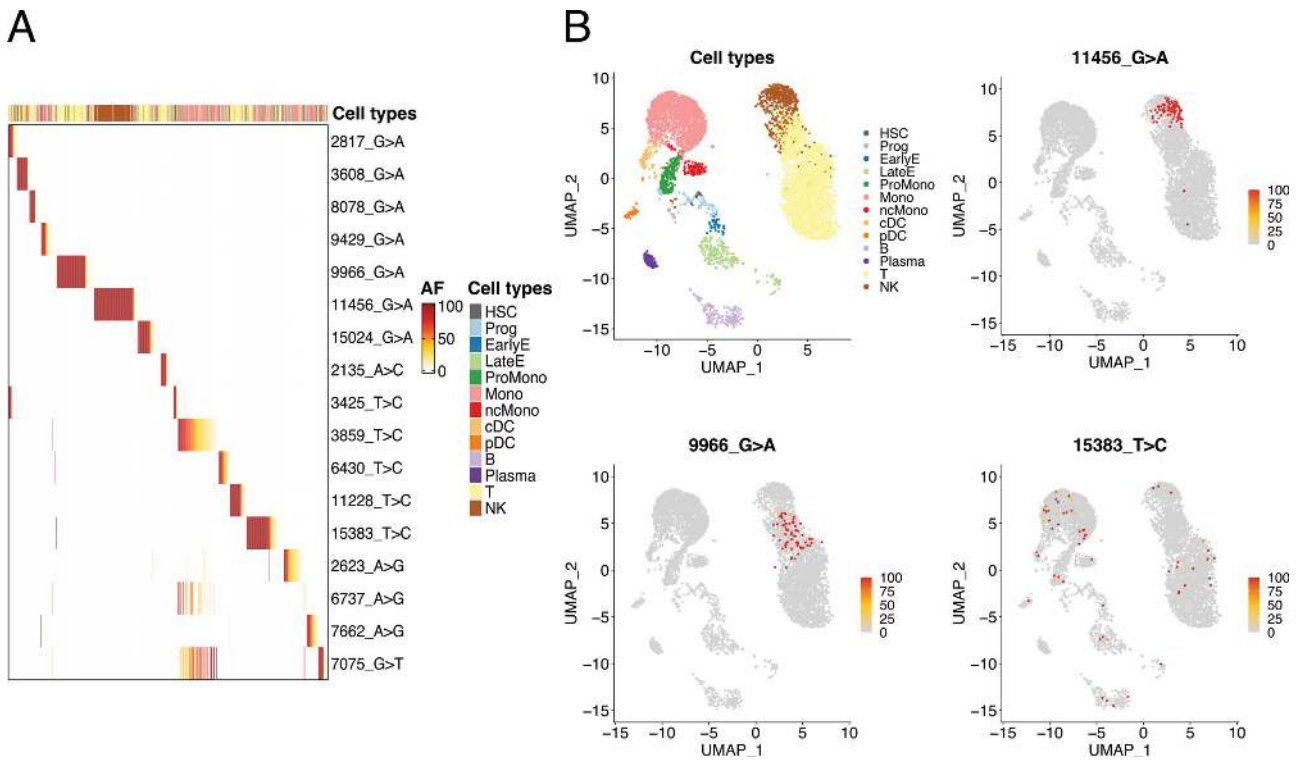


FIGURE 25.4 *A mitochondrial variant indicates a clonal expansion of NK cells in a bone marrow sample.* (A) Heatmap shows the heteroplasmy level and cell type assignment (top bar) for 17 informative variants detected in a bone marrow sample from a 63-year-old White male undergoing mitral valve repair at Mass General Brigham. Note that most variants are transitions, as expected [21]. (B) UMAPs illustrate the distribution of cells carrying variants 11456_G>A, 9966_G>A, and 15383_T>C across the hematopoietic system. AF: allele frequency (heteroplasmy level).

non-hematologic conditions. For example, it was used to show that Barrett's esophagus is polyclonal, but single molecularly aberrant stem cells expand into precancerous foci that transform into esophageal adenocarcinoma [28]. These findings highlight MAESTER's utility in identifying clonal dynamics in other systems relying on stem cell differentiation such as epithelial tissues.

Given the ability of MAESTER to infer clonal relationships, it has been used to study the origins of microglia (macrophages that reside in the central nervous system, CNS). While

these cells are classically considered to derive from erythro-myeloid progenitors in the yolk sac during early embryogenesis, recent research indicates that some microglia-like cells originate from the bone marrow [29, 30]. Thus, conditions like clonal hematopoiesis could have important implications for CNS disease states [31]. Tracking mitochondrial variants could contribute to this question by identifying common variants between microglia and clonal blood cell populations in the bone marrow or peripheral blood. Indeed, MAESTER analysis on peripheral blood and brain tumor cells confirmed that brain-resident myeloid cells can be derived from the blood but adopt gene expression programs that are indistinguishable from microglia [32].

25.6 Future Trends

In addition to MAESTER, technologies that combine scRNA-seq with mtDNA-mediated lineage tracing include CloneTracer [26] and nanoranger [33]. ATAC-based assays such as scmtATAC-seq have the advantage of uniform coverage along the mitochondrial genome because they use mtDNA (instead of RNA) as input [34, 35]. The depth of variant coverage is critical for robust inference of clonal relationships. Recently, the Sankaran and Weissman groups obtained deep coverage of the mitochondrial genome in combination with single-cell gene expression and chromatin accessibility analysis [36]. ReDeeM customizes the 10x Genomics multiome kit to preserve the mitochondrial genome rather than starting with prepared cell nuclei. It then uses custom mtDNA probes to purify mitochondrial genome fragments from the ATAC libraries. While ReDeeM provides the most comprehensive coverage, MAESTER remains more cost-effective (~$3,000 per sample compared to ~$6,000 for ReDeeM) and is retroactively compatible with 10x Genomics 3′ barcoded cDNA libraries. In addition to assay improvements, analytical innovations are essential for reliable inference of clonal relationships based on mtDNA variants. For example, clonal cell populations may be defined based on the similarity of many detected variants rather than the presence of a single variant [36]. Further bioinformatics development will provide greater analytical sophistication such as the reconstruction of phylogenies, differentiation hierarchies, and cancer cell evolution over time.

The impact of obtaining deep mitochondrial variant coverage concurrent with cell states lies in research applications. Other technologies to analyze clonal expansions in native human tissues have different limitations, such as a high dropout rate of nuclear DNA mutations, reliance on variants in highly expressed transcripts, or an inability to concurrently analyze cell states [20, 35, 37]. Although it can be challenging to identify informative mtDNA variants that allow reliable lineage inference, they uniquely enable concurrent assessment of gene expression and chromatin accessibility in primary human tissues while avoiding the inefficiency of capturing nuclear DNA mutations. As such, we expect mtDNA variant tracking to contribute significant insight by deeply profiling clonal lineages over time. For example, in premalignant clonal hematopoiesis, we expect that the ability to measure the clonal output of stem cell populations will deepen our understanding of stem cell dysfunction, akin to our recent work in aging hematopoiesis [36]. After malignant transformation, we expect that paired analysis of diagnosis and remission leukemia samples will identify properties of persistent leukemia cells. This should inform specific vulnerabilities that can be exploited to eradicate malignant cells more comprehensively and prevent disease recurrence.

REFERENCES

1. Kreso A, Dick JE. Evolution of the cancer stem cell model. *Cell Stem Cell*. 2014 Mar 6;14(3):275–91.
2. Trumpp A, Haas S. Cancer stem cells: The adventurous journey from hematopoietic to leukemic stem cells. *Cell*. 2022 Apr 14;185(8):1266–70.
3. Batlle E, Clevers H. Cancer stem cells revisited. *Nat Med*. 2017 Oct 6;23(10):1124–34.
4. Eppert K, Takenaka K, Lechman ER, Waldron L, Nilsson B, van Galen P, Metzeler KH, Poeppl A, Ling V, Beyene J, Canty AJ, Danska JS, Bohlander SK, Buske C, Minden MD, Golub TR, Jurisica I, Ebert BL, Dick JE. Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat Med*. 2011 Aug 28;17(9):1086–93.
5. van Galen P, Hovestadt V, Wadsworth MH, Hughes TK, Griffin GK, Battaglia S, Verga JA, Stephansky J, Pastika TJ, Lombardi Story J, Pinkus GS, Pozdnyakova O, Galinsky I, Stone RM, Graubert TA, Shalek AK, Aster JC, Lane AA, Bernstein BE. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. 2019 Mar 7;176(6):1265–81.e24.
6. Zeng AGX, Bansal S, Jin L, Mitchell A, Chen WC, Abbas HA, Chan-Seng-Yue M, Voisin V, van Galen P, Tierens A, Cheok M, Preudhomme C, Dombret H, Daver N, Futreal PA, Minden MD, Kennedy JA, Wang JCY, Dick JE. A cellular hierarchy framework for understanding heterogeneity and predicting drug response in acute myeloid leukemia. *Nat Med*. 2022 Jun;28(6):1212–23.
7. Pei S, Pollyea DA, Gustafson A, Stevens BM, Minhajuddin M, Fu R, Riemondy KA, Gillen AE, Sheridan RM, Kim J, Costello JC, Amaya ML, Inguva A, Winters A, Ye H, Krug A, Jones CL, Adane B, Khan N, Ponder J, Schowinsky J, Abbott D, Hammes A, Myers JR, Ashton JM, Nemkov T, D'Alessandro A, Gutman JA, Ramsey HE, Savona MR, Smith CA, Jordan CT. Monocytic subclones confer resistance to venetoclax-based therapy in patients with acute myeloid leukemia. *Cancer Discov*. 2020 Apr;10(4):536–51.
8. Shlush LI, Mitchell A, Heisler L, Abelson S, Ng SWK, Trotman-Grant A, Medeiros JJF, Rao-Bhatia A, Jaciw-Zurakowsky I, Marke R, McLeod JL, Doedens M, Bader G, Voisin V, Xu C, McPherson JD, Hudson TJ, Wang JCY, Minden MD, Dick JE. Tracing the origins of relapse in acute myeloid leukemia to stem cells. *Nature*. 2017 Jul 6;547(7661):104–8.
9. Schwenger E, Steidl U. An evolutionary approach to clonally complex hematologic disorders. *Blood Cancer Discov*. 2021 May;2(3):201–15.
10. Kreso A, O'Brien CA, van Galen P, Gan OI, Notta F, Brown AMK, Ng K, Ma J, Wienholds E, Dunant C, Pollett A, Gallinger S, McPherson J, Mullighan CG, Shibata D, Dick

JE. Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science*. 2013 Feb 1;339(6119):543–8.
11. Eide CA, Kurtz SE, Kaempf A, Long N, Joshi SK, Nechiporuk T, Huang A, Dibb CA, Taylor A, Bottomly D, McWeeney SK, Minnier J, Lachowiez CA, Saultz JN, Swords RT, Agarwal A, Chang BH, Druker BJ, Tyner JW. Clinical correlates of venetoclax-based combination sensitivities to augment acute myeloid leukemia therapy. *Blood Cancer Discov*. 2023 Nov 1;4(6):452–67.
12. Acosta J, Ssozi D, van Galen P. Single-cell RNA sequencing to disentangle the blood system. *Arterioscler Thromb Vasc Biol*. 2021 Mar;41(3):1012–8.
13. Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc*. 2018 Apr;13(4):599–604.
14. Miles LA, Bowman RL, Merlinsky TR, Csete IS, Ooi AT, Durruthy-Durruthy R, Bowman M, Famulare C, Patel MA, Mendez P, Ainali C, Demaree B, Delley CL, Abate AR, Manivannan M, Sahu S, Goldberg AD, Bolton KL, Zehir A, Rampal R, Carroll MP, Meyer SE, Viny AD, Levine RL. Single-cell mutation analysis of clonal evolution in myeloid malignancies. *Nature*. 2020 Nov;587(7834):477–82.
15. Granja JM, Klemm S, McGinnis LM, Kathiria AS, Mezger A, Corces MR, Parks B, Gars E, Liedtke M, Zheng GXY, Chang HY, Majeti R, Greenleaf WJ. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat Biotechnol*. 2019 Dec;37(12):1458–65.
16. Montefiori LE, Bendig S, Gu Z, Chen X, Pölönen P, Ma X, Murison A, Zeng A, Garcia-Prat L, Dickerson K, Iacobucci I, Abdelhamed S, Hiltenbrand R, Mead PE, Mehr CM, Xu B, Cheng Z, Chang TC, Westover T, Ma J, Stengel A, Kimura S, Qu C, Valentine MB, Rashkovan M, Luger S, Litzow MR, Rowe JM, den Boer ML, Wang V, Yin J, Kornblau SM, Hunger SP, Loh ML, Pui CH, Yang W, Crews KR, Roberts KG, Yang JJ, Relling MV, Evans WE, Stock W, Paietta EM, Ferrando AA, Zhang J, Kern W, Haferlach T, Wu G, Dick JE, Klco JM, Haferlach C, Mullighan CG. Enhancer hijacking drives oncogenic BCL11B expression in lineage-ambiguous stem cell leukemia. *Cancer Discov*. 2021 Nov;11(11):2846–67.
17. Woodworth MB, Girskis KM, Walsh CA. Building a lineage from single cells: Genetic techniques for cell lineage tracking. *Nat Rev Genet*. 2017 Jan 23;18(4):230–44.
18. Van Egeren D, Escabi J, Nguyen M, Liu S, Reilly CR, Patel S, Kamaz B, Kalyva M, DeAngelo DJ, Galinsky I, Wadleigh M, Winer ES, Luskin MR, Stone RM, Garcia JS, Hobbs GS, Camargo FD, Michor F, Mullally A, Cortes-Ciriano I, Hormoz S. Reconstructing the lineage histories and differentiation trajectories of individual cancer cells in myeloproliferative neoplasms. *Cell Stem Cell*. 2021 Mar 4;28(3):514–23.e9.
19. Lee-Six H, Øbro NF, Shepherd MS, Grossmann S, Dawson K, Belmonte M, Osborne RJ, Huntly BJP, Martincorena I, Anderson E, O'Neill L, Stratton MR, Laurenti E, Green AR, Kent DG, Campbell PJ. Population dynamics of normal human blood inferred from somatic mutations. *Nature*. 2018 Sep;561(7724):473–8.
20. Mitchell E, Spencer Chapman M, Williams N, Dawson KJ, Mende N, Calderbank EF, Jung H, Mitchell T, Coorens THH, Spencer DH, Machado H, Lee-Six H, Davies M, Hayler D, Fabre MA, Mahbubani K, Abascal F, Cagan A, Vassiliou GS, Baxter J, Martincorena I, Stratton MR, Kent DG, Chatterjee K, Parsy KS, Green AR, Nangalia J, Laurenti E, Campbell PJ. Clonal dynamics of haematopoiesis across the human lifespan. *Nature*. 2022 Jun;606(7913):343–50.
21. Ludwig LS, Lareau CA, Ulirsch JC, Christian E, Muus C, Li LH, Pelka K, Ge W, Oren Y, Brack A, Law T, Rodman C, Chen JH, Boland GM, Hacohen N, Rozenblatt-Rosen O, Aryee MJ, Buenrostro JD, Regev A, Sankaran VG. Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics. *Cell*. 2019 Mar 7;176(6):1325–39.e22.
22. Miller TE, Lareau CA, Verga JA, DePasquale EAK, Liu V, Ssozi D, Sandor K, Yin Y, Ludwig LS, El Farran CA, Morgan DM, Satpathy AT, Griffin GK, Lane AA, Love JC, Bernstein BE, Sankaran VG, van Galen P. Mitochondrial variant enrichment from high-throughput single-cell RNA sequencing resolves clonal populations. *Nat Biotechnol*. 2022 Feb 24;1–5.
23. Yuan Y, Ju YS, Kim Y, Li J, Wang Y, Yoon CJ, Yang Y, Martincorena I, Creighton CJ, Weinstein JN, Xu Y, Han L, Kim HL, Nakagawa H, Park K, Campbell PJ, Liang H, PCAWG Consortium. Comprehensive molecular characterization of mitochondrial genomes in human cancers. *Nat Genet*. 2020 Mar;52(3):342–52.
24. Good JD, Safina KR, Miller TE, van Galen P. Protocol for mitochondrial variant enrichment from single-cell RNA sequencing using MAESTER. *STAR Protoc*. 2025 Mar 21;6(1):103564.
25. Rückert T, Lareau CA, Mashreghi MF, Ludwig LS, Romagnani C. Clonal expansion and epigenetic inheritance of long-lasting NK cell memory. *Nat Immunol*. 2022 Nov;23(11):1551–63.
26. Beneyto-Calabuig S, Merbach AK, Kniffka JA, Antes M, Szu-Tu C, Rohde C, Waclawiczek A, Stelmach P, Gräßle S, Pervan P, Janssen M, Landry JJM, Benes V, Jauch A, Brough M, Bauer M, Besenbeck B, Felden J, Bäumer S, Hundemer M, Sauer T, Pabst C, Wickenhauser C, Angenendt L, Schliemann C, Trumpp A, Haas S, Scherer M, Raffel S, Müller-Tidow C, Velten L. Clonally resolved single-cell multi-omics identifies routes of cellular differentiation in acute myeloid leukemia. *Cell Stem Cell*. 2023 May 4;30(5):706–21.e8.
27. Penter L, Gohil SH, Lareau C, Ludwig LS, Parry EM, Huang T, Li S, Zhang W, Livitz D, Leshchiner I, Parida L, Getz G, Rassenti LZ, Kipps TJ, Brown JR, Davids MS, Neuberg DS, Livak KJ, Sankaran VG, Wu CJ. Longitudinal single-cell dynamics of chromatin accessibility and mitochondrial mutations in chronic lymphocytic leukemia mirror disease history. *Cancer Discov*. 2021 Dec 1;11(12):3048–63.
28. Gier RA, Hueros RAR, Rong J, DeMarshall M, Karakasheva TA, Muir AB, Falk GW, Zhang NR, Shaffer SM. Clonal cell states link Barrett's esophagus and esophageal adenocarcinoma. *bioRxiv [Internet]*. 2023 Feb 6. http://dx.doi.org/10.1101/2023.01.26.525564.
29. Loeb AM, Pattwell SS, Meshinchi S, Bedalov A, Loeb KR. Donor bone marrow-derived macrophage engraftment into the central nervous system of patients following allogeneic transplantation. *Blood Adv*. 2023 Oct 10;7(19):5851–9.
30. Sankowski R, Süß P, Benkendorff A, Böttcher C, Fernandez-Zapata C, Chhatbar C, Cahueau J, Monaco G, Gasull AD, Khavaran A, Grauvogel J, Scheiwe C, Shah MJ,

Heiland DH, Schnell O, Markfeld-Erol F, Kunze M, Zeiser R, Priller J, Prinz M. Multiomic spatial landscape of innate immune cells at human central nervous system borders. *Nat Med*. 2024 Jan;30(1):186–98.

31. Zhou Y. Clonal derivation and dynamics in rhesus macaques following hematopoietic stem cell transplantation and the impact of clonal hematopoiesis. In: *65th ASH Annual Meeting & Exposition [Internet]*. ASH; 2023 [cited 2024 Mar 25]. https://ash.confex.com/ash/2023/webprogram/Paper180863.html.
32. Miller TE, El Farran CA, Couturier CP, Chen Z, D'Antonio JP, Verga J, Villanueva MA, Castro LNG, Tong YE, Saadi TA, Chiocca AN, Fischer DS, Heiland DH, Guerriero JL, Petrecca K, Suva ML, Shalek AK, Bernstein BE. Programs, origins, and niches of immunomodulatory myeloid cells in gliomas. *Nature*. 2025 Feb 26. doi: 10.1038/s41586-025-08633-8. Epub ahead of print. PMID: 40011771. https://pubmed.ncbi.nlm.nih.gov/40011771
33. Penter L, Borji M, Nagler A, Lyu H, Lu WS, Cieri N, Maurer K, Oliveira G, Al'Khafaji AM, Garimella KV, Li S, Neuberg DS, Ritz J, Soiffer RJ, Garcia JS, Livak KJ, Wu CJ. Integrative genotyping of cancer and immune phenotypes by long-read sequencing. *Nat Commun*. 2024 Jan 2;15(1):32.
34. Lareau CA, Ludwig LS, Muus C, Gohil SH, Zhao T, Chiang Z, Pelka K, Verboon JM, Luo W, Christian E, Rosebrock D, Getz G, Boland GM, Chen F, Buenrostro JD, Hacohen N, Wu CJ, Aryee MJ, Regev A, Sankaran VG. Massively parallel single-cell mitochondrial DNA genotyping and chromatin profiling. *Nat Biotechnol*. 2021 Apr;39(4):451–61.
35. Izzo F, Myers RM, Ganesan S, Mekerishvili L, Kottapalli S, Prieto T, Eton EO, Botella T, Dunbar AJ, Bowman RL, Sotelo J, Potenski C, Mimitou EP, Stahl M, El Ghaity-Beckley S, Arandela J, Raviram R, Choi DC, Hoffman R, Chaligné R, Abdel-Wahab O, Smibert P, Ghobrial IM, Scandura JM, Marcellino B, Levine RL, Landau DA. Mapping genotypes to chromatin accessibility profiles in single cells. *Nature*. 2024 May;629(8014):1149–57.
36. Weng C, Yu F, Yang D, Poeschla M, Liggett LA, Jones MG, Qiu X, Wahlster L, Caulier A, Hussmann JA, Schnell A, Yost KE, Koblan L, Martin-Rufino JD, Min J, Hammond A, Ssozi D, Bueno R, Mallidi H, Kreso A, Escabi J, Rideout WM, Jacks T, Hormoz S, van Galen P, Weissman JS, Sankaran VG. Deciphering cell states and genealogies of human hematopoiesis. *Nature*. 2024 Jan 22;1–3.
37. Scherer M, Singh I, Braun M, Szu-Tu C, Kardorff M, Rühle J, Frömel R, Beneyto-Calabuig S, Raffel S, Rodriguez-Fraticelli A, Velten L. Somatic epimutations enable single-cell lineage tracing in native hematopoiesis across the murine and human lifespan. *bioRxiv [Internet]*. 2024 Apr 1. http://dx.doi.org/10.1101/2024.04.01.587514.

26

Implementing Clinical Risk Assessment for Genetic Susceptibility to Cancer: Challenges and Successes in a Statewide Initiative

Sarah Austin, Liliana Arida-Moody, Erika Hanson, Shayna Weiner, Matthew Demerath, Emerson Delacroix, Erika Koeppe, Ken Resnicow, Jennifer J. Griggs, and Elena M. Stoffel

26.1 Introduction

Approximately one in three individuals will be diagnosed with cancer during their lifetime [1]. Certain factors can increase an individual's lifetime cancer risk, including environmental exposures, tobacco use, diet, and genetic susceptibility. Early identification of individuals who have elevated risk for cancer provides opportunities for interventions, including specialized screening/surveillance, and pharmacological and behavioral interventions, which decrease cancer morbidity and mortality.

Inherited germline variants associated with genetic susceptibility to cancer are implicated in 5%–20% of cancer diagnoses and influence recommendations for cancer screening, prevention, and treatment for affected individuals and their relatives. Germline genetic testing has been available for nearly three decades, and indications for genetic testing continue to expand [2, 3]. Innovations in next-generation sequencing have increased access to genetic testing, now available at minimal or no cost to patients. Unfortunately, only a fraction of people at risk for hereditary cancer syndromes undergo genetic testing, with a recent study showing that 6.8% of patients with cancer undergo genetic testing, and identifying people who meet clinical guidelines for cancer genetic testing remains a challenge [4].

Family history is a cornerstone of cancer risk assessment and is an integral step in identifying high-risk patients who are candidates for genetic counseling [5]. National guidelines and professional societies recommend clinicians elicit and evaluate all cancer diagnoses in first- and second-degree relatives, including type and age at diagnosis [5, 6]. Despite these recommendations, family history documentation in patient medical records is often limited or inaccurate [7–12]. Wood and colleagues found that only one in three patients with breast cancer and one in four patients with colorectal cancer had a complete family history documented in health records in practices participating in quality improvement initiatives [11].

Barriers to collecting a complete family history exist at the patient, provider, practice, and health system levels. Patient and provider barriers include lack of genetics knowledge and awareness of the importance of family history and limited access to genetics specialty care [9, 13, 14]. In most settings, collecting and documenting family history relies primarily on clinicians; however, lack of time and inefficient approaches to eliciting detailed information about cancers contribute to missing data [14]. While most electronic health record (EHR) applications include family history fields for clinicians to update, the format and structure of these fields are highly variable and often cumbersome and inefficient for clinical decision-making.

Several digital tools have been developed to facilitate collection of family history information. In 2005, the surgeon general launched My Family Health Portrait [15], a free and public-facing web-based tool that allows individuals to input family history information into a digital platform that can be modified and updated over time [My Family Health Portrait (https://cbiit.github.io/FHH/html/index.html)]. Other tools include clinical decision support for patients and clinicians, such as MeTree (MeTree | Duke Department of Medicine), first developed in 2004 by Duke Center for Precision Health. MeTree is a web-based patient-facing tool that has evolved to collect personal and family history information on 48 health conditions and incorporate validated cancer risk models [16–18]. The utility of family history data depends on clinicians having easy access to this information, and integrating this clinical information into the EHR in a systematic and accessible way across health systems is still a formidable challenge [16].

There remains an unmet need for efficient procedures to accurately collect family history and integrate this information into clinical workflows. Funded through the National Cancer Institute (NCI) Cancer Moonshot initiative to improve identification and management of individuals with genetic susceptibility to cancer (Moonshot Blue Ribbon Panel—NCI [cancer.gov]), our UO1 project sought to implement a statewide initiative promoting family history risk assessment to identify individuals at risk for hereditary cancer syndromes. Here we describe the development, implementation, adaptations, and outcomes in deploying our digital family health history tool in clinical practices across the state of Michigan.

26.2 Materials

26.2.1 Development and Technical Implementation of FHHT

The Family Health History Tool (FHHT) is an electronic, patient-facing web-based survey developed by the University of Michigan Cancer Genetics Clinic and Rogel Cancer Center's Informatics Team in response to two unmet clinical needs: (1) systematic and efficient collection of accurate family history and (2) clinical decision support to classify patients' cancer risk at

DOI: 10.1201/9781003449942-30

point of care. To be eligible to complete the FHHT, participants had to be age 18 or older, have access to the Internet (smartphone, laptop, or tablet), and be able to read and speak English.

The FHHT survey is accessible to individuals/patients on smartphone, tablet, or laptop devices. The survey elicits information about cancer types and ages at diagnosis for all first- and second-degree relatives with sufficient detail for construction of a three-generation pedigree. Along with the pedigree and cancer summary, the FHHT report includes a calculated PREMM5 score, a validated tool that uses family cancer diagnoses to predict risk for Lynch syndrome, a common hereditary cancer syndrome [19, 20]. The FHHT does not provide a recommendation regarding indication for genetic testing but provides information to support discussions between patients and clinicians about how family history impacts cancer risk.

At the heart of the FHHT is a custom adaptive survey software that generates questions using the previous answers to guide the output. This allows the FHHT to dynamically alter the survey based on each individual's responses. A user who has a large family with multiple relatives with cancer will be presented a much longer survey than a user with a relatively small family history with few cancer diagnoses. For example, users are asked how many siblings they have. Based on their response, users would be asked to provide specific details (initials, current age or age at death, cancer diagnosis, and age at diagnosis) for each of those siblings. Individuals that report they do not have siblings would not be shown subsequent questions on those relatives. The FHHT is not a monolithic software application but a collection of smaller micro services and user interfaces (e.g., PREMM5 score calculation, pedigree image generation, cancer summary outputs) that work together to provide the needed functionality. Each of these applications run in the Google Cloud Platform (GCP), chosen because of its support of HIPAA compliance. All communications are processed over secure socket layer (SSL) encryption ensuring privacy, authentication, and data integrity in Internet communication.

All data including survey logic, configuration, administrative, and survey response are stored in a secure PostgreSQL relational database. The data are processed by task-specific microservices responsible for risk analysis, pedigree drawing, email, and SPS communication, transferring data to patient information systems and report generation. The survey and administration user interfaces were written in Angular. This application framework is cross-platform compatible (Android, iOS) to support mobile, tablet, and desktop viewing. Microservices that read and write to the database are all written in Java with the Quarkus framework. Angular and Quarkus are both open-source software and have active communities that continually add new features and address security issues.

26.3 Methods

26.3.1 Implementation of the Family Health History Tool in Clinical and Non-Clinical Settings: The MiGHT Study

Funded in 2020 through the NCI Cancer Moonshot Inherited Cancer Syndrome Collaborative (ICSC), we launched the Michigan Genetic Hereditary Testing (MiGHT) study (NIH/NCI U01CA 232827) to improve identification of individuals at risk for hereditary cancer syndromes through provider-, practice-, and patient-based interventions [21]. Our practice-based quality improvement intervention proposed to use the FHHT to increase the proportion of oncology patients who have a complete family history documented in their medical record. We leveraged existing partnerships with the Michigan Department of Health and Human Services (MDHHS), Blue Cross Blue Shield (BCBS) of Michigan, and Michigan Oncology Quality Consortium (MOQC) to launch a statewide effort to improve family cancer history ascertainment in clinical settings. MOQC is a diverse network of nearly all community and academic medical and gynecologic oncology practices across the state of Michigan with extensive experience with quality improvement projects requiring data collection, audit and feedback, data sharing, and sharing of best practices [22].

Our intervention consisted of clinician education on hereditary cancer syndromes and the importance of collecting a complete family cancer history. Our team provided educational materials and instructions for how to refer eligible patients for genetic evaluation (developed as part of the MDHHS CDC Cancer Genomics Best Practices Cooperative Agreement [U58DP 005357]). In collaboration with the MOQC measures committee, we provided orientation regarding the family history quality metric. Practices were offered the opportunity to deploy the FHHT (see Section 26.3.2) along with a $1,00 startup incentive to offset the burden of integrating the FHHT into their workflow. As an additional incentive, MOQC offered practices a value-based reimbursement, such that each region of practices that met specific family history documentation metrics were rewarded with at least a 3% value-based reimbursement for all BCBS Preferred Provider Organization patients.

Analyses included comparison of rates of completion of the FHHT by clinical setting and descriptive characteristics of individuals who completed the FHHT.

26.3.2 Deployment of the Family Health History Tool: Community Oncology Practices

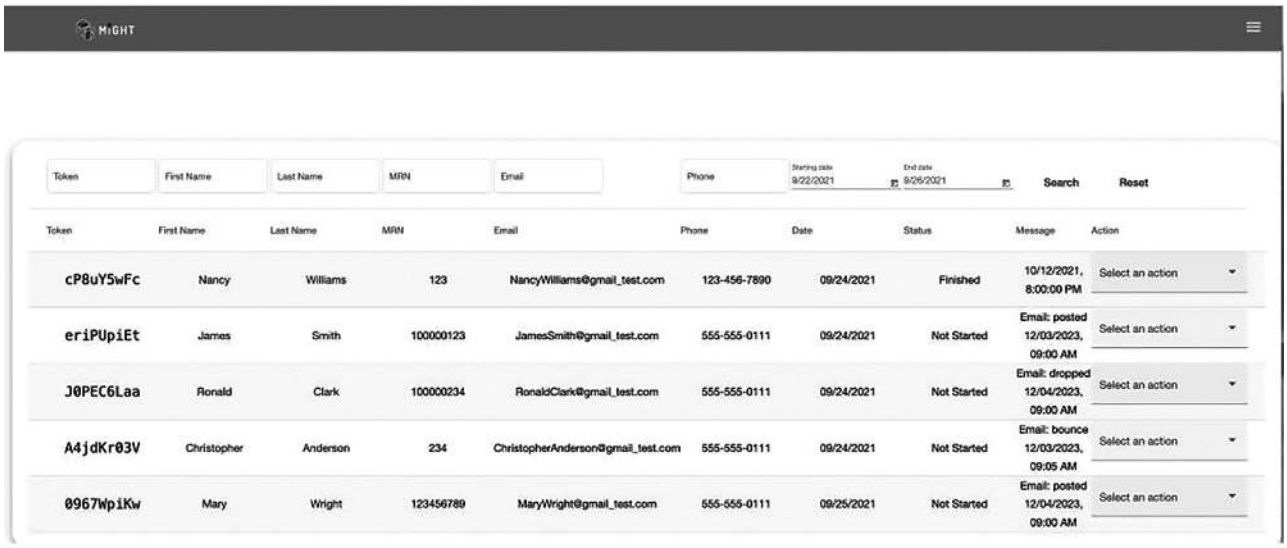
In spring 2021, 63 MOQC practices were introduced to the Family Health History Quality Improvement initiative at a consortium meeting in which collaborative-wide and practice-level information about performance on this measure were presented. Project team members discussed implementation of the FHHT and legal and information technology (IT) requirements, as well as practice-specific workflows as follows. MOQC sites were asked to identify an office champion (e.g., physician, nurse practitioner, or practice manager) to promote the use of the FHHT in their practice, as well as an operations lead (e.g., front desk staff member, practice/office manager, or administrative staff) to manage implementation of the tool. Additional meetings between project team staff, operations leads, and IT representatives from each practice (estimating two hours in total time) were conducted as needed to pilot the tool and provide education and training to office staff.

Practices were offered access to a FHHT study dashboard, through which practices could send their patients email and/

or SMS text messages inviting them to complete the FHHT survey (Figure 26.1). Patients would receive a letter of introduction explaining the importance of family history, with an electronic link to the FHHT. The practice leads were instructed on processes for exporting and uploading CSV files containing appointment schedules containing patient demographics (first and last name, date of birth, medical record number, appointment date, appointment time, email, and phone number) into the dashboard. From the dashboard, practices could view survey completion status, and download results, including the patient's family cancer summary and three-generation pedigree.

Patients could access the FHHT via a personalized link sent by their clinic/practice by email or text in advance on their clinical

1. Family health history tool: Patient Dashboard



2. Family health history tool: Family History Section

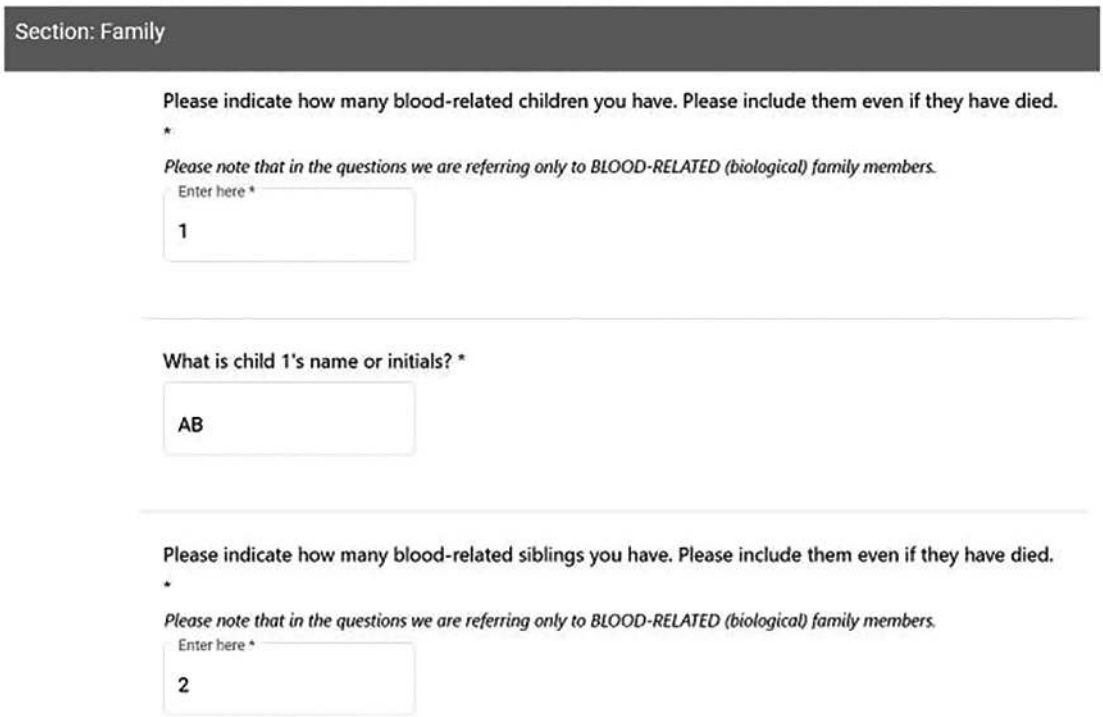


FIGURE 26.1 Image showing practice view of the FHHT study dashboard and patient view of the FHHT. Names in this image are not those of real participants and were randomly generated from https://dlptest.com/sample-data/namedobemail/.

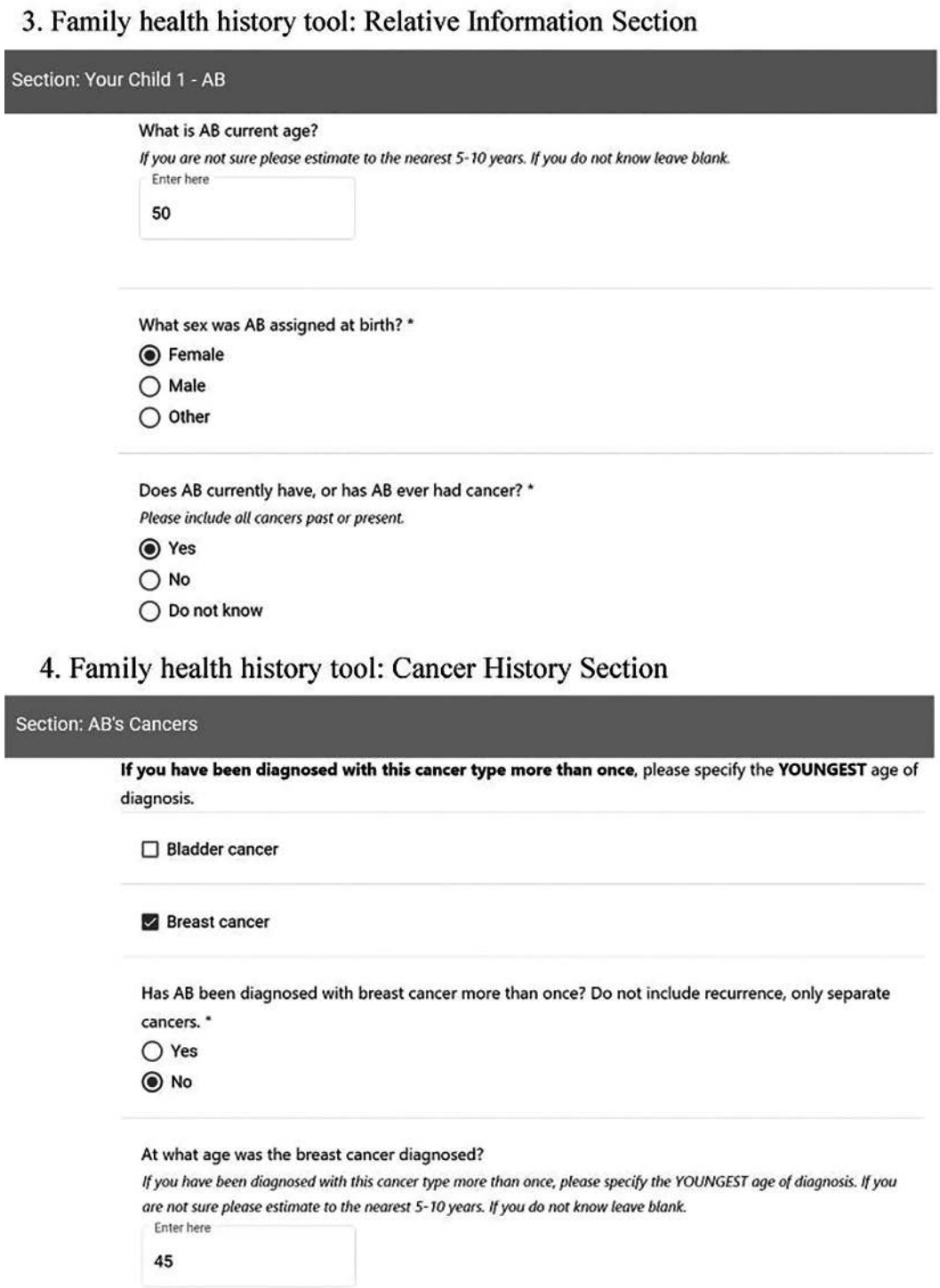


FIGURE 26.1 (Continued)

appointment. The introductory page of the FHHT included an abbreviated consent for HIPAA authorization. Patients were encouraged to complete the FHHT before their upcoming clinic visit. Practices had the option to include the FHHT survey link with their new patient appointment paperwork and could tailor the survey dashboard, choosing how often to upload patient lists, mode of contact (SMS, email, or both), and timing of reminders relative to each new patient appointment.

When patients checked in for an appointment, the practice staff were asked to confirm with the patient that they had completed the FHHT. Patients who had not completed the FHHT were given the opportunity to do so on site (using a smart phone, tablet device, or computer). Upon survey completion, office staff could print a real-time report that includes a summary of the patient's family history of cancer, pedigree, and PREMM5 risk model scores. After completing the FHHT the

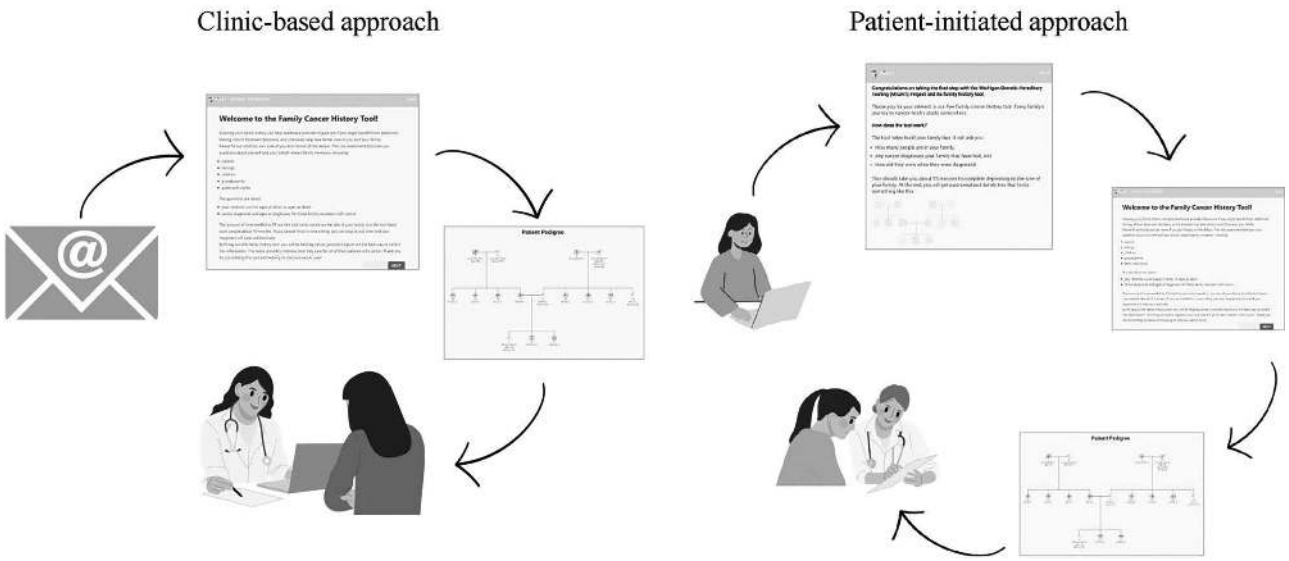


FIGURE 26.2 A comparison of clinic-based and individual-initiated approaches for accessing the family health history tool.

patient was provided a printable summary page listing the cancer diagnoses in their family along with their three-generation pedigree. Patients were prompted to share this summary with their medical providers and were given access to review and reprint the pedigree as needed.

One of the MOQC practice sites (University of Michigan Rogel Cancer Center) was able to work with their information technology team to automate upload of the FHHT summary page into the EHR (EPIC) for clinician review. Other sites developed their own processes for incorporating the FHHT summary into their clinical workflows.

26.3.3 Deployment of the Family Health History Tool: Other Clinical and Non-Clinical Settings

As we encountered lower than expected uptake of the FHHT among MOQC practices, we adopted alternative strategies for disseminating the FHHT. We added non-oncology outpatient clinics affiliated with University of Michigan/Michigan Medicine, including primary care, outpatient gastroenterology, and gastrointestinal endoscopy. Individuals who had previously received their cancer care through U-M Rogel Cancer Center were identified through the Rogel Cancer Registry and were sent an email inviting them to complete the FHHT. In an effort to expand recruitment beyond University of Michigan/Michigan Medicine, we launched population-based recruitment through social media, print media, and television advertisements. We created different QR codes and URL links for each recruitment site, with virtual and paper flyers disseminated through social media campaigns, newspaper articles, health fairs, research study websites, and news/radio programs. In this individual-initiated self-registration model, subjects accessed an online registration page, entered their contact information, and were sent an automated email and/or SMS message containing their personal link to the FHHT. A comparison between clinic-based and individual-initiated models are presented in Figure 26.2.

26.4 Results

26.4.1 Practice Participation with the Family Health History Quality Improvement Initiative

Of the original 63 MOQC practices, 43 were formally invited to participate in the Family Health History Initiative (20 practices had either merged with another practice or were no longer participating in MOQC activities). Of note, this initiative was launched in 2021 during the COVID-19 pandemic and consequently many practices reported logistical barriers to implementing the FHHT that proved insurmountable, and ultimately 32 sites declined participation, with 14 sites having not responded, and the remaining declining for reasons that included having limited resources/bandwidth for a new project (n = 8), bad timing (n = 3), legal concerns (n = 3), overlap with another tool currently in use (n = 1), and not interested (n = 4). Ultimately, 11 MOQC practices (11/43, 25.6%) agreed to participate. Each practice underwent onboarding, during which they met with study team members to review their existing workflows of family history collection (e.g., elicited by clinician at office visit, through pre-visit paperwork, or not collected) and discuss strategies to best incorporate the FHHT into their practice. Onboarding included demonstration of how to upload the practice schedule to the FHHT, how to send FHHT invitations to patients from the project dashboard, review of common notifications and errors, and pilot testing of the FHHT with practice data. This training and onboarding took approximately three to four hours over several meetings and email, with reference materials and training guides provided. Practice managers, MOQC representatives, study team members, and IT representatives from each practice were invited to partake in the onboarding meetings. Of the practices that underwent

onboarding training, seven (7/43, 16.2%) actively deployed the FHHT with their patients. The four out of 11 sites that received training but did not use the FHHT in practice reported encountering additional technology barriers and staffing shortages in the setting of the COVID-19 pandemic that influenced their decision not to implement the FHHT.

26.4.2 Rates of Uptake and Completion Rates for FHHT

Across all recruitment methods, > 100,00 FHHT invitations have been distributed with > 14,000 FHHT surveys completed to date (Figure 26.3). Of participants who completed the FHHT, most were women (59.5%) and in the 60–79 age group (50.9%). Approximately one in five (20.3%, n = 2884) individuals who completed the FHHT met criteria for genetic testing based on their personal and/or family history of cancer (Figure 26.4). Of individuals who met testing criteria, 20.7% had no personal history of cancer, and 79.3% had a personal history of cancer (27.5% breast cancer, 12.0% prostate cancer, 8.0% colorectal cancer, 4.5% pancreatic cancer, 4.7% uterine/endometrial cancer, 5.4% ovarian cancer, 37.9% other cancer). Individuals who met criteria for genetic testing and had not previously undergone genetic testing were invited to partake in Aim 2 of the MiGHT study, a randomized controlled trial aimed to improve uptake of genetic testing for at-risk patients (ClinicalTrials.gov ID NCT05162846). To date, 645 individuals have enrolled in the trial.

26.4.3 Technical, Knowledge, and Regulatory Barriers to Implementation and Dissemination

Dissemination and implementation of the FHHT encountered barriers at practice, provider, and patient levels. Through additional follow-up and self-report, many practices indicated they were unable to integrate the FHHT due to the aforementioned logistical challenges. Practices also encountered complications with implementing the technology, including exporting clinic schedules in the required file format. Additionally, while clinics affiliated with University of Michigan (the institution where the tool was developed) were able to automatically import a PDF of the FHHT report and pedigree into the media tab of the patient's EHR, IT departments at other sites determined that importing the FHHT report into the EHR was not possible due to technical difficulties and/or security concerns. Practices were hesitant to use the tool given challenges in determining how best to make the FHHT report accessible for clinician review. In some clinics, patients were responsible for sharing a digital or print copy of their family history with their providers.

At the provider level, some clinicians were hesitant to use the FHHT out of concern that reviewing, interpreting, and taking action based on FHHT results would generate additional work. Some clinicians reported they did not feel comfortable with diagnosis and/or management of genetic conditions and requested additional training. In response, our team developed clinical materials, including primers on recognizing hereditary cancer syndromes, indications for cancer genetic counseling referrals, and answers to frequently asked questions and made continuing medical education online courses about cancer genetics available to MOQC clinicians.

Anecdotally, sites in which staff members were more engaged with distributing the FHHT had higher completion rates. The MOQC site with the highest FHHT completion rates (50.4%) routinely discussed the goals and purpose of the FHHT with patients at the time the appointment was scheduled. In comparison, MOQC sites with lower completion rates (18.9%) routinely sent the FHHT survey to new patients without any prior introduction/discussion of the purpose of the tool.

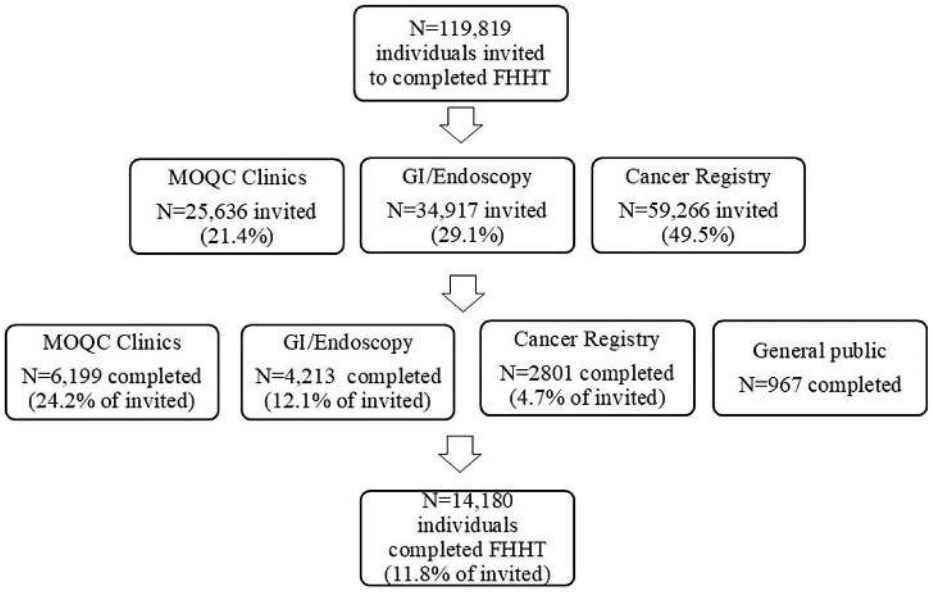


FIGURE 26.3 Flow diagram of completion rate of the family health history tool among oncology clinics, GI/endoscopy clinics, cancer registry, research registry, and general public recruitment sites.

Proband Personal Cancer History	Additional Criteria
Breast Cancer	Proband: Diagnosed age under 50, or has Ashkenazi Jewish ancestry or male (any age) Proband or FDR/SDR: diagnosis of triple-negative breast cancer FDR/SDR: history of ovarian cancer, pancreatic cancer, breast cancer under 50, or male breast cancer
Prostate Cancer	Proband: Diagnosed age under 50 or having Ashkenazi Jewish ancestry FDR/SDR: history of ovarian cancer, pancreatic cancer, breast cancer under 50, or male breast cancer
Colorectal Cancer	Diagnosed age under 50
Endometrial or Uterine Cancer	Diagnosed age under 50
Ovarian Cancer	No additional criteria
Pancreatic cancer	No additional criteria
Personal History of Any Cancer or No Cancer	FDR: history of pancreatic cancer, male breast cancer, colon cancer diagnosed age under 50, endometrial or uterine cancer under 50, or breast cancer under 50 FDR/SDR: history of ovarian cancer, or history of breast cancer at any age ONLY if the proband has Ashkenazi Jewish ancestry Proband: PREMM5 score >= 2.5

FIGURE 26.4 Clinical criteria used to identify individuals meeting criteria for germline genetic testing for cancer susceptibility. (FDR = first degree relative, SDR = second degree relative).

At the patient level, some individuals provided feedback that they were concerned that the emails and SMS texts presenting the link to access the FHHT were spam or scams, prompting them to contact their clinic to confirm the messages were legitimate.

26.5 Discussion

The overall goal of the Family Health History Initiative was to help clinicians identify individuals at risk for hereditary cancer syndromes through facilitating collection of robust family cancer history. Despite consensus among clinical practices, researchers, payers, and the state health department about the importance of family history information for clinical care, dissemination and implementation of the FHHT in clinical practice settings was difficult, and adoption by practices and uptake by patients were lower than anticipated. Logistical barriers to integrating this technology with existing EHR systems and scheduling platforms were major deterrents for many practices who decided not to participate. In addition, FHHT completion rates varied widely (10%–50%) depending on the practice from which individuals were ascertained. Despite lower-than-expected uptake of the FHHT, we identified a cohort of > 600 individuals that met criteria for genetic testing and enrolled in Aim 2 of the MiGHT study, a randomized controlled trial aimed to improve the uptake of testing for high-risk individuals (ClinicalTrials.gov ID NCT05162846). These individuals qualified for genetic testing based their personal and/or family history of cancer (Figure 26.4) and were identified because they completed the FHHT.

26.5.1 Lessons Learned: Dissemination and Implementation

There are several take-aways from our experience in the development, implementation, and use of a digital family history tool in clinical practice. First, our existing relationships with clinical practices and the state health department and our ability to employ a policy lever to reward high-quality care through partnership with payers were crucial in developing this project. Having buy-in from a multidisciplinary team allowed for idea sharing, support, and infrastructure that would not have been possible otherwise and enabled us to adapt recruitment strategies to reach individuals receiving medical care in settings beyond the original MOQC practices.

Unfortunately, the timing of this initiative (grant funded in June 2020 at the start of the COVID-19 pandemic) had a profound negative impact on participation of clinic practices that were struggling with staffing shortages, staff and patient burnout, and challenges with implementing models for delivering virtual care. With these major stressors, the promises of financial compensation through bonuses and value-based reimbursement for a quality improvement project were

deemed insufficient by many practices. At the time of grant submission, MOQC sites were enthusiastic and optimistic about the project, however due to the aforementioned barriers many practices were not able to incorporate FHHT as they had hoped. This suggests a gap between intention and adoption and should be an important consideration in projects requiring multi-site participation.

While the FHHT integration went smoothly for the clinical practices affiliated with the study's home academic institution, lack of integration with the EHR and scheduling systems was a major barrier for implementing the FHHT at other sites. Ideally, the FHHT (and other clinical tools) would be easily integrated across multiple EHR systems, with patient-reported information populating fields to trigger best practice advisories and assist clinicians with medical management recommendations for specialized cancer screening and/or genetic testing.

26.5.2 Future Directions

The success of dissemination and implementation projects depends on the level of engagement of stakeholders as well as the level of investments in technology and personnel. In healthcare, new technologies must integrate easily across a wide variety of clinical systems without additional expertise or funding and must minimize concerns about security of both personal and institutional data networks. Electronic tools should strive to achieve "plug and play" functionality, based on technology that is both free standing and able to integrate with other systems. Although dissemination and implementation projects are proposed as "real world," having a dedicated champion at each clinical site and engaging early with local clinic and IT teams could increase the success of similar initiatives.

REFERENCES

1. Society, A.C. *What is cancer?* 2022. https://www.cancer.org/cancer/understanding-cancer/what-is-cancer.html.
2. Daly, M.B., et al., Genetic/familial high-risk assessment: Breast, ovarian, and pancreatic, version 2.2021, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*, 2021. 19(1): p. 77–102.
3. Weiss, J.M., et al., NCCN Guidelines(R) insights: Genetic/familial high-risk assessment: Colorectal, version 1.2021. *J Natl Compr Canc Netw*, 2021. 19(10): p. 1122–1132.
4. Kurian, A.W., et al., Germline genetic testing after cancer diagnosis. *JAMA*, 2023. 330(1): p. 43–51.
5. Lu, K.H., et al., American society of clinical oncology expert statement: Collection and use of a cancer family history for oncology providers. *J Clin Oncol*, 2014. 32(8): p. 833–40.
6. Syngal, S., et al., ACG clinical guideline: Genetic testing and management of hereditary gastrointestinal cancer syndromes. *Am J Gastroenterol*, 2015. 110(2): p. 223–62; quiz 263.
7. Annoni, A.M. and C. Longhini, Investigating men's motivations to engage in genetic screening for BRCA1 and BRCA2 mutations. *PLoS One*, 2022. 17(3): p. e0265387.
8. Murff, H.J., D. Byrne, and S. Syngal, Cancer risk assessment: Quality and impact of the family history interview. *Am J Prev Med*, 2004. 27(3): p. 239–45.
9. Murff, H.J., R.A. Greevy, and S. Syngal, The comprehensiveness of family cancer history assessments in primary care. *Community Genet*, 2007. 10(3): p. 174–80.
10. Qureshi, N., et al., Collection and use of cancer family history in primary care. *Evid Rep Technol Assess (Full Rep)*, 2007(159): p. 1–84.
11. Wood, M.E., et al., Quality of cancer family history and referral for genetic counseling and testing among oncology practices: A pilot test of quality measures as part of the American society of clinical oncology quality oncology practice initiative. *J Clin Oncol*, 2014. 32(8): p. 824–9.
12. Lin, J., et al., Room for improvement in capturing cancer family history in a gynecologic oncology outpatient setting. *Gynecol Oncol Rep*, 2022. 40: p. 100941.
13. Armstrong, J., et al., Utilization and outcomes of BRCA genetic testing and counseling in a national commercially insured population: The ABOUT study. *JAMA Oncol*, 2015. 1(9): p. 1251–60.
14. Sussner, K.M., L. Jandorf, and H.B. Valdimarsdottir, Educational needs about cancer family history and genetic counseling for cancer risk among frontline healthcare clinicians in New York City. *Genet Med*, 2011. 13(9): p. 785–93.
15. My Family Health Portrait. [cited 2024 Mar 27]. https://cbiit.github.io/FHH/html/index.html#.
16. Orlando, L.A., et al., Development and validation of a primary care-based family health history and decision support program (MeTree). *N C Med J*, 2013. 74(4): p. 287–96.
17. Orlando, L.A., et al., Protocol for implementation of family health history collection and decision support into primary care using a computerized family health history system. *BMC Health Serv Res*, 2011. 11: p. 264.
18. Wu, R.R., et al., Patient and primary care provider experience using a family health history collection, risk stratification, and clinical decision support tool: A type 2 hybrid controlled implementation-effectiveness trial. *BMC Fam Pract*, 2013. 14: p. 111.
19. Kastrinos, F., et al., Development and validation of the PREMM(5) model for comprehensive risk assessment of lynch syndrome. *J Clin Oncol*, 2017. 35(19): p. 2165–2172.
20. Mercado, R.C., et al., Performance of PREMM(1,2,6), MMRpredict, and MMRpro in detecting Lynch syndrome among endometrial cancer cases. *Genet Med*, 2012. 14(7): p. 670–80.
21. Gerido, L.H., et al., The Michigan genetic hereditary testing (MiGHT) study's innovative approaches to promote uptake of clinical genetic testing among cancer patients: A study protocol for a 3-arm randomized controlled trial. *Trials*, 2023. 24(1): p. 105.
22. Michigan Oncology Quality Consortium (MOQC). https://moqc.org/moqc/members/.

27

Assessing Tumor Tissue for the Alternative Lengthening of Telomeres (ALT) Phenotype

C. Patrick Reynolds, Trevor A. Burrow, and Balakrishna Koneru

27.1 Introduction

27.1.1 Telomeres and Cancer

Telomeres are nucleoprotein structures of canonical (TTAGGG) DNA repeats and shelterin proteins located at the end of eukaryotic chromosomes. Their primary function is to safeguard coding DNA sequences located near the ends of the chromosome from the DNA damage repair machinery [1]. Typically spanning 4–10 kilobases in length, telomeres buffer the erosion of terminal coding DNA that arises from the "end-replication problem" in somatic (i.e., non-germ line) cells; that is, the inability of DNA polymerase to fully replicate the chromosome end sequence [2]. When a somatic cell undergoes division, DNA polymerase is unable to replicate the terminal end of the lagging DNA strand where the last RNA primer was present; thus the terminal DNA sequence shortens with every cell division [3]. Extensive telomere shortening (generally after ~40–50 cell divisions) triggers p53-induced cellular senescence and/or cell death to guard against the erosion of coding DNA sequences that would compromise genomic integrity (i.e., "The Hayflick Limit") [2, 4].

In order to achieve replicative immortality, cancer cells need to maintain their telomere length so that they do not shorten upon each cell division. The ability to enable replicative immortality by activating a telomere maintenance mechanism (TMM) is one of the hallmarks of cancer [5]. Approximately 85% of cancer cells use the ribonucleoprotein, telomerase, containing a catalytic subunit (TERT) and an RNA template (hTR) in order to add sequential TTAGGG telomeric repeats to the ends of chromosomes during division [6, 7]. TERT mRNA (the telomerase catalytic subunit) expression is repressed in most type of non-malignant cells. Cancer cells activate telomerase by acquiring *TERT* promoter mutations, *TERT* rearrangements, epigenetic alterations, or transcriptional dysregulation [8–11]. We, as well as other groups, have reported that increased telomerase activity is associated with poor clinical outcomes in cancer patients [11–16].

27.1.2 Alternative Lengthening of Telomeres (ALT)

A subset of cancer cells (10%–15%) use a telomerase-independent mechanism to maintain their telomeres, termed the "alternative lengthening of telomeres (ALT)". Potential biomarkers of ALT cancers are (1) C-circles which are extrachromosomal partially double-stranded (self-priming) circular telomeric DNA sequences [17], (2) longer than normal telomeres and of heterogeneous lengths [18, 19], (3) the presence of ALT-associated, promyelocytic leukemia (PML) nuclear bodies (i.e., APBs) that contain proteins involved in DNA repair, recombination and telomeric-associated proteins [20–22], and (4) a high frequency of inactivating genomic alteration in genes encoding for the alpha thalassemia X-linked intellectual disability syndrome gene (ATRX) and death domain-associated (DAXX) proteins [23]. Unlike cancer cells that use telomerase enzyme to directly lengthen their telomeres, ALT cells rely on a homology-directed DNA repair (HDR) mechanism for telomere elongation and extension [24–26]. While still incompletely understood, it is believed that activation of the ALT mechanism requires the transitioning of telomeres into a "recombination-permissive" state. This process involves changes to telomeric chromatin composition/structure and increased levels of "DNA replication stress." Genomic alterations in *ATRX*, *DAXX*, and histone gene *H3.3* have been shown to cause changes in telomeric chromatin [27–30]; some ALT cells modify their telomeric chromatin via less-defined mechanisms. The telomeres of cells that utilize the ALT pathway (i.e., "ALT telomeres") exhibit increased levels of both DNA damage and replication stress [31, 32], primarily due to aberrant telomeric chromatin and altered protein binding. Unresolved replication forks lead to double strand breaks and promote the recruitment of DNA repair factors and the engagement of HDR mechanisms that extend the telomeres in ALT cells [33].

27.1.3 Biomarkers of Telomere Maintenance Mechanisms (TMM)

27.1.3.1 Biomarkers for Detection of Telomerase

Early studies on telomere maintenance focused on assaying telomerase activity, but increasingly it has been accepted that quantifying TERT (catalytic subunit of telomerase) mRNA expression provides a robust method for defining telomerase-positive vs telomerase-negative cells [8–11, 16]. Detection of telomerase and TERT mRNA requires high-quality protein and RNA samples, which are not always available for all patients. Therefore, several groups use genomic aberrations like *TERT* promoter mutations, *TERT* rearrangements, and *MYCN* amplification as surrogate markers for telomerase

DOI: 10.1201/9781003449942-31

activation [9–11, 34]. However, these markers are not sufficient to detect all tumors that activate telomerase in certain histologies, such as neuroblastoma [16, 35].

27.1.3.2 Biomarkers for Detection of Alternative Lengthening of Telomeres (ALT) Activation

ALT cancers have been identified using high telomere content, but this approach is not sensitive and is plagued with both false-positive and false-negative results. Genomic alterations in *ATRX* and/or *DAXX* genes are associated with ALT, and virtually all, if not all, tumors with inactivating genomic alterations of *ATRX* or *DAXX* are of the ALT phenotype, but ≥ 40% of ALT tumors in some histologies lack such mutations [16, 36]. Staining FFPE sections for ALT-associated PML bodies is a more specific assay for ALT, but as we recently demonstrated the APB assay can miss some low-TERT expressing tumors that are C-circle-positive and have clinical outcomes indicative of ALT [16].

Our data and that of others indicate that the telomeric DNA C-circle assay can detect ALT cancers accurately with a small amount of tumor DNA (~50 ng) and has the potential to be applied as a clinical biomarker for ALT cancers [16, 17]. The C-circle assay consists of two stages: (1) selectively amplifying the partially double stranded telomeric C-circles with *Bacillus subtilis* phage Phi-29 DNA polymerase (Phi-29), a high-fidelity enzyme capable of isothermal processive strand displacement DNA synthesis [17, 37]. (2) C-circle enriched telomeric signal is detected by real-time PCR, normalized to the single copy gene VAV2, and compared against the controls to qualitatively detect the presence of C-circles [16, 38] (Figure 27.1). VAV2 is a validated single copy gene for some tumor types (neuroblastoma, osteosarcoma), but it may not be suitable for all cancer histologies due to possible copy number losses or gain at VAV2 genomic coordinates. Other single copy genes commonly used for the C-circle assay are 36B4 and HBB [37]. While tumors with a clearly positive telomeric DNA C-circle assay are ALT-positive, some ALT tumors (likely due to spatial tumor heterogeneity) can have a low C-circle content but can be alternatively identified by assessing tumor sections for ultrabright telomere foci (UTF) using telomere specific fluorescence *in situ* hybridization (FISH). Figure 27.2 shows an example of a tumor positive for ALT by the UTF assay.

The current approach for identifying ALT-positive tumors in our laboratory is assaying for telomeric C-circles from tumor DNA extracted from fresh or snap-frozen tumor tissue. For C-circle negative tumors that have low TERT-mRNA expression (by RT-PCR) absence of a TMM is confirmed by assessing tumor FFPE sections for UTF by telomere FISH [39]. This approach limits the need to conduct the labor-intensive UTF assay to a small number of patients. The UTF assay can also be used when only formalin-fixed paraffin embedded (FFPE) tumor tissue is available. Employing the C-circle assay together with assessing TERT mRNA expression by real-time PCR and using the UTF assay to confirm TMM has enabled identifying three cohorts of high-risk neuroblastoma patients (telomerase-positive, ALT-positive, and TMM-negative), all with different clinical outcomes [16, 40]. Although not yet done, applying the same approach to other cancer types has the potential for identifying cohorts of patients with varying clinical outcomes by classifying the patients based on TMM phenotype.

Beyond classifying tumors as to TMM phenotype to use in assessing prognosis, defining tumors as ALT-positive has potential for use as companion diagnostics for therapies with high activity for ALT-positive cancers [41–44]. Here we present the two methods used to identify ALT-positive tumors, the C-circle assay and the UTF assay.

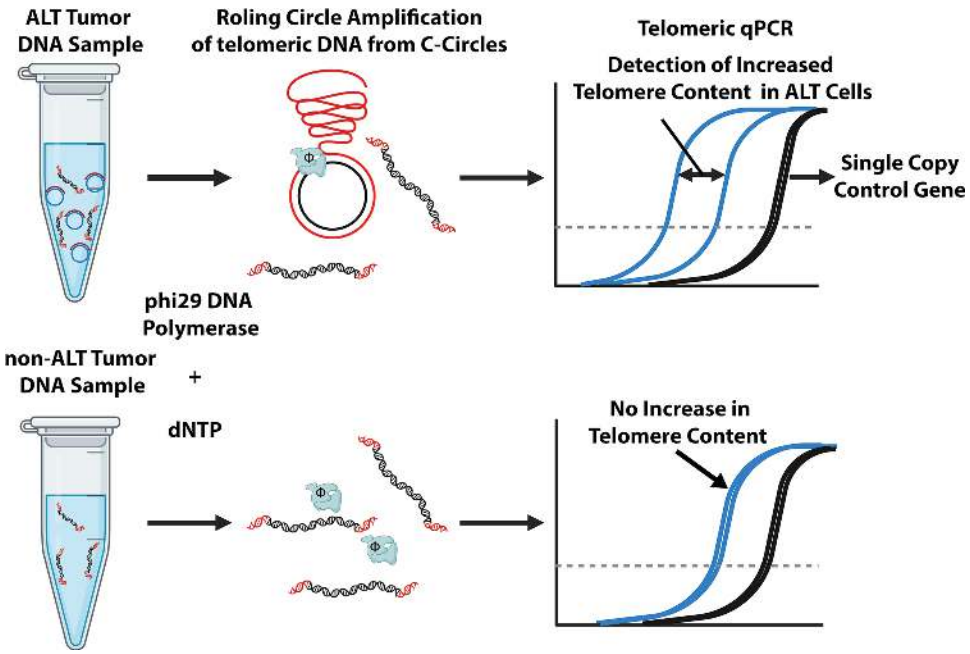


FIGURE 27.1 Illustration of the C-circle assay.

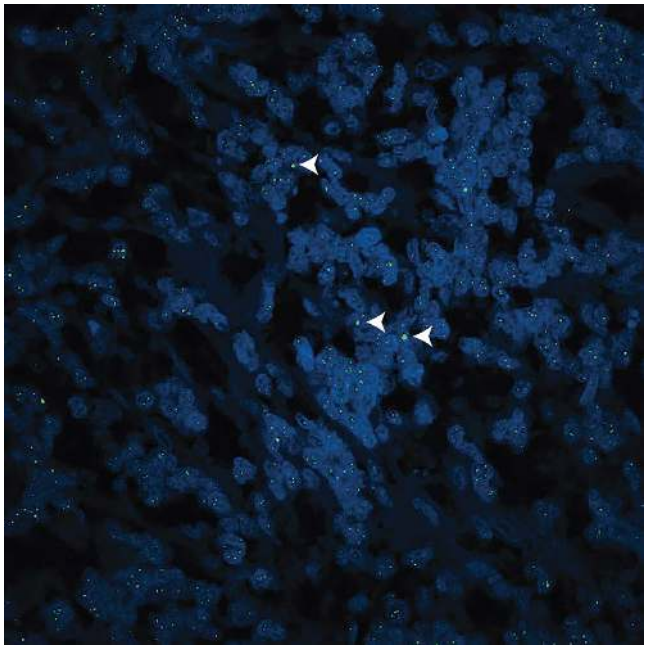


FIGURE 27.2 Example of a FFPE tumor section stained by telomere FISH showing the ultrabright telomere foci (UTF) that identify the tumor as ALT-positive.

27.2 Materials

27.2.1 Telomeric DNA C-Circle Assay

27.2.1.1 Reagents for C-Circle Assay

Controls used are the following. Positive control: cell line DNA from ALT-positive neuroblastoma cell line CHLA–90. Negative control: cell line DNA from telomerase-positive, ALT-negative neuroblastoma cell line CHLA–20. The assay requires 10 mM dNTPs (New England Biolabs Cat. No. N0447L), Phi-29 DNA polymerase and buffer (New England Biolabs Cat. No. M0269L), nuclease-free water (not DEPC treated) (Ambion Cat. No. AM 9932), QuantiTect SYBR Green PCR Kit (1000) (Qiagen Cat. No. 204145), IDTE 1X TE Solution pH 7.5 (Integrated DNA Technologies Cat. No.11-05-01-04).

Primers are: 1 μM Telb Forward Primer (Integrated DNA Technologies):

5′-CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT-3′

9 μM Telb Reverse Primer (Integrated DNA Technologies):

5′-GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT-3′

2 μM VAV2 Forward Primer (Integrated DNA Technologies):

5′-TGGGCATGACTGAAGATGAC-3′

2 μM VAV2 Reverse Primer (Integrated DNA Technologies)

5′-ATCTGCCCTCACCTTCTCAA-3′

27.2.1.2 Equipment for C-circle Assay

Freezer, < −20 °C, refrigerator, 2–8 °C, benchtop Cooler −20 °C (Thermo-Fisher Cat. No. 355501), calibrated pipettes capable of liquid delivery ranges of 100–200 μL, 20–200 μL, 0.5–10 μL, Vortexer, Thermocycler, 2x MicroAmp™ Splash-Free 96-Well Base, PCR Plate Spinner (500 xg), QuantStudio™ 3 Real-Time PCR System 96-well (Thermo-Fisher Cat. No. A28567) or similar qPCR machine.

27.2.1.3 Supplies for C-Circle Assay

Pipette tips 1000 μL, 200 μL, 10 μL pipette tips, 1% Tween, 2 μg/μL Bovine Serum Albumin (BSA), 100μM DTT, DMSO, PCR Reaction Strips (with Attached Caps) (VWR Cat. No. 89133–922), RNase/DNase Free 1.7 mL Microcentrifuge tubes (VWR Cat. No. 87003–294), Blue RNase/DNase Free 1.7 mL Microcentrifuge tubes (VWR Cat. No. 87004–256), Red RNase/DNase Free 1.7 mL Microcentrifuge tubes (VWR Cat. No. 87004–254), Argos Technologies™ Microcentrifuge Tubes—5 mL (Fisher Scientific Cat. No. 03-391–276), MicroAmp™ Optical 96-Well Reaction Plates (Thermo-Fisher Cat. No. N8010560), MicroAmp™ Optical Adhesive Film (Thermo-Fisher Cat. No. 4311971), aluminum foil, sealable plastic bags, 70% ethanol, 0.1% sodium hypochlorite/bleach.

27.2.2 Ultrabright Telomere Foci (UTF) Assay

27.2.2.1 Reagents for the UTF Assay

A blocking solution (10% stock solution) made with 10% blocking reagent (Roche 11096176001) stock in 100 mM maleic acid 150 mM NaCl. Adjust pH to 7.5 (20 °C) with NaOH. Store at −80 °C. A hybridizing solution of 10 mM Tris-HCl pH 7.2, 70% deionized formamide. Use 0.5% blocking reagent (from 10% stock). Hybridization wash buffer 1 is 10 mM Tris-HCl pH 7.2, 70% formamide. Hybridization wash buffer 2 is 50 mM Tris-HCl pH 7.2, 150 mM NaCl, 0.05% Tween–20. DAPI stock solution (3000x) made by dissolving 0.5 mg/ml 4′,6-diamino-2-phenylindole (Sigma D9542) in H_2O. (Store at −30 °C.) The PNA probe stock solution is made by resuspending lyophilized TelC-Cy3 telomere probe (PNA Bio # F1002) in deionized formamide to achieve a final concentration of 50 μM (stock solution can be stored at −80 °C). Other reagents used are 100%, 95% and 70% ethanol, citrate buffer pH 6.0, xylene, and pro-Long Gold Antifade Reagent (Cell Signaling #9071).

27.2.2.2 Equipment and Supplies for the UTF Assay

Humidified slide hybridization chamber, forceps, slide warmer, slide steamer, hybridization oven capable of consistently maintaining 83 °C, glass staining jars, epifluorescence microscope (capable of imaging FITC and Cy3). Microscope cover glass (Corning # 2975–225).

27.3 Methods

27.3.1 Method for C-Circle Assay

Specimens for the assay are fluorometrically quantified DNA extracted from fresh or snap frozen tissue via DNeasy Blood &

Tissue Kit (Qiagen Cat. No.69506). *Storage and Handling Requirements for Reagents*: QuantiTect SYBR Green: Store at –20 °C or colder for long-term storage (performance is guaranteed until the expiration date). Once thawed, QuantiTect SYBR Green can be stored, protected from light, at 2–8 °C for up to six months, depending on the expiration date. Primers stored (dry or eluted in TE) at –20 °C, or colder, are viable for 12 months. Working stocks eluted in TE stored at 2–8 °C are viable for 12 months. Minimize freeze thaws and exposure to UV. Controls stored at –20 °C or colder (in a freezer without freeze thaw cycles) are stable for 12 months. Once thawed controls can be stored at 2–8 °C for up to one month. Avoid/prevent nuclease activity, freeze-thaw cycles, acid hydrolysis, and shearing. dNTPs stored at –20 °C, or colder, are viable for 12 months (aliquot to avoid excessive freeze thaws). Phi-29 DNA Polymerase stored at –20 °C is stable for 12 months. Minimize freeze-thaw cycles to no more than ten cycles and avoid vortexing by flick mixing. Phi-29 DNA Polymerase buffer stored at –20 °C, or colder, is viable for 12 months. Nuclease-free water (not DEPC treated) is viable for 12 months at room temperature (15–30 °C). IDTE 1X TE is viable for 12 months at room temperature (15–30 °C). USP Grade190200 Proof Ethanol (95%–100%, not denatured): Stored at 15–30 °C. Unopened, it is stable for three years. Opened (but resealed tightly) ethanol is stable for 12 months from opening.

TABLE 27.1

C-circle Assay Master Mix Reagent Volumes

Phi Reaction		No Phi Reaction	
2 μg/μl BSA	2.4 μL	2 μg/μl BSA	2.4 μL
1% Tween	2.4 μL	1% Tween	2.4 μL
10 mm dNTPs	2.4 μL	10 mm dNTPs	2.4 μL
Phi-29 Buffer	2.4 μL	Phi-29 Buffer	2.4 μL
Phi-29 (10 U/μl)	0.96 μL	DTT (100 μM)	0.96 μL
DTT (100 μM)	0.96 μL		

TABLE 27.2

Teolmere PCR Reagent Volumes

Telomere Reaction (Telb)		Single-Copy Gene Reaction (VAV2)	
SYBR Green	15 μL	SYBR Green	15 μL
1 μM Telb FWD primer	3 μL	2 μM VAV2 FWD primer	3 μL
9 μM Telb REV primer	3 μL	2 μM VAV2 REV primer	3 μL
DTT (100 μM)	1.2 μL	DTT (100 μM)	1.2 μL
DMSO	0.6 μL	DMSO	0.6 μL
Nuclease-Free Water (not DEPC treated)	1.2 μL	Nuclease-Free Water (not DEPC treated)	1.2 μL

27.3.1.1 Isothermal PCR Preparation and Run for C-Circle Assay

Remove 2 μg/μL BSA, 1% Tween, 10 mM dNTPs, Phi-29 Buffer, and DTT from the –20 °C and place in a bench top rack to thaw—*do not remove* Phi-29 Polymerase from –20 °C. Remove positive and negative control aliquots from –20 °C (or from 4 °C if there is a current working stock). Set up red and blue PCR reaction strips with individually attached caps in a PCR tube rack. (Each sample, including controls, will require one tube on each colored strip—eight samples will require one red and one blue PCR strip—scale accordingly). Add fluorometrically quantified DNA template 32 ng per control/sample per Phi (blue PCR reaction strips) and No Phi reactions (red PCR reaction strips). The CCA has been shown to have a linear range of 1–32 ng for input template. QS Phi controls/samples with nuclease-free water (not DEPC treated) to a total volume of 10.4 μL. QS samples No Phi sample with nuclease-free water (not DEPC treated) to a total volume of 11.2 μL. Obtain one blue and one red RNase/DNase Free 1.7 mL microcentrifuge tubes and label the blue tube "Phi" and the red tube "No Phi"(use 5 mL tubes to scale for larger master mix preps). Remove the benchtop cooler containing Phi-29 Polymerase from –20 °C. Master mix regent volumes, per control/sample, and per Phi/No Phi reaction are shown in Table 27.1: (Scale for number of samples + controls for the run accordingly). Do not mix reagent lots.

Vortex the Phi master mix and the no Phi Master Mix for 2–3 seconds. Aliquot 9.6 μLof Phi master mix to each Phi reaction to reach a final reaction volume of 20 μL. Aliquot 8.8 μL of master mix omitting the Phi polymerase to each no Phi reaction to reach a final reaction volume of 20 μL. Close caps on PCR reaction strips and ensure a proper seal. Vortex PCR reaction strips for 2–3 seconds. Centrifuge PCR reaction strips in a tabletop spinner (approximately 3500 xg) 5–10 seconds. Place PCR reaction strips in thermocycler and check to make sure caps are sealed properly before closing thermocycler lid. Program the following run conditions: 30 °C hold for 480 min, 65 °C hold for 20 min, 4 °C hold for ∞ min. Start the run and leave samples overnight.

27.3.1.2 Real-Time PCR Preparation and Run

Remove PCR reaction strips from thermocycler and turn off the thermocycler centrifuge PCR reaction strips in a tabletop spinner at approximately 3500 xg for 5–10 seconds. Add 20 μL of nuclease-free water (not DEPC treated) to each control/sample. Set up two MicroAmp™ Optical 96-well reaction plates in their respective MicroAmp™ Splash-Free 96-Well Bases. Using a VWR® Lab Marker*, label one plate "Telb" and label the other plate "VAV2" on the rim of the plate. Do not use a sharpie or other writing instrument. Do not make any other markings on the plate. Other writing instruments and labeling locations can potentially interfere with the optical imaging signal. Vortex PCR reaction strips for 2–3 seconds. Centrifuge PCR reaction strips in a table top spinner at approximately 3500 xg for 5–10 seconds. Pipette 5 μL of each control/sample in triplicate into the Telb real-time plate. Pipette 5 μL of each control/sample in triplicate into the VAV2 real-time plate. Prepare Telb and VAV2 real-time PCR master mixes. *Protect master mixes from light*. Regent volumes, per control/sample, are shown in in Table 27.2. Scale for number of samples + controls for the run accordingly). Do not mix reagent lots.

Pipette 20 μL of Telb master mix into each control/sample well on the Telb plate. Apply a MicroAmp™ Optical Adhesive Film to the Telb plate and ensure a proper seal by using the

MicroAmp™ Adhesive Film Applicator. Centrifuge the Telb plate in a PCR Plate Spinner at approximately 500 xg for 30 seconds. Select "Load Experiment" on the QuantStudio™. Select "USB". *Real-time PCR conditions are as follows: 95 °C hold for 15 min, 30 cycles of: 95 °C for 15 sec and 56 °C for 2 min. Transfer the Telb real-time from the centrifuge to the QuantStudio™ 3 tray and ensure the plate is oriented correctly. (Real-time plate well A1 should be in the top left hand corner well of the QuantStudio™ 3 heatblock). Select "Start Run" in the bottom right hand corner. Pipette 20 µL of VAV2 master mix into each control/sample well on the VAV2 plate. Apply a MicroAmp™ Optical Adhesive Film to the VAV2 plate and ensure a proper seal by using the MicroAmp™ Adhesive Film Applicator. Cover VAV2 plate in aluminum foil, to protect from light, and place the covered VAV2 plate in a refrigerator (2–8 °C). Once the Telb plate run has completed, remove the Telb plate from the QuantStudio™ 3 tray and run the VAV2 plate. Real-time PCR conditions for VAV2 plate are as follows: 95 °C hold for 15 min, 40 cycles of: 95 °C for 15 sec and 57 °C for 30 sec, then 72 °C for 1 min. Retrieve the VAV2 plate from the refrigerator (2–8 °C) and remove the aluminum foil. Centrifuge the VAV2 plate in a PCR Plate Spinner at approximately 500 xg for 30 seconds. Transfer the VAV2 plate from the centrifuge to the QuantStudio™ and run the assay. Calculations are done using a TelB Excel workbook file and a VAV2 Excel workbook file available in the Data Archive of www.COGcell.org.

27.3.1.3 Interpretation of C-Circle Assay

The real-time PCR C-Circle Assay is a qualitative assay. A minimum of one each of positive control, negative control, non-template negative control, and samples are required for each run. The homology of the Telb primers can result in primer self-amplification at CTs > 25; thus, any real-time Telb signal past CT = 20 is disregarded. Additionally, control/sample CVs should be < 10%. Arbitrary unit values of samples with a C-circle level value < 5% of the positive control are interpreted as "Negative."

27.3.2 Method for the UTF Assay

Specimens for UTF assay are unstained 4 to 5 µm tissue sections from archival formalin-fixed paraffin-embedded tissue blocks, preferably on positively charged slides. Prior to hybridization, dilute the stock solution 1:250 with the hybridizing solution. Heat the dilution for 5 minutes at 60 °C before use. To melt the paraffin, place the tissue slides on a slide warmer and heat at 60 °C for 10 minutes. Wash the tissue slides three times in xylene, each wash lasting 5 minutes (note: all wash steps are performed using glass slide staining jars). Rehydrate the tissue by washing in a graded ethanol series (100%, 95%, and then 70% ethanol) for 5 minutes each. Wash the tissue slides in PBS for 3 minutes, followed by a wash in 1% tween 20 in PBS for 1 minute.

Incubate the slides in a plastic slide holder with citrate buffer pH 6.0 in a preheated (boiling) kitchen steamer for 30 minutes. After incubation, remove the container from the steamer, open the lid, and let the slides cool at room temperature for 10 minutes. Wash the slides in deionized distilled water for 1 minute and then dehydrate them in a graded ethanol series (70%, 95%, and 100% ethanol) for 3 minutes each. Allow the slides to air dry thoroughly.

Add 50 µl of a mixture of diluted TelC-FAM telomere probe and diluted CENT-Cy3 PNA probe (mixed 1:1) to each slide. The CENT-Cy3 probe is included as a control for hybridization efficiency. Gently apply a coverslip to each tissue slide, avoiding air bubbles. Denature the tissue slides at 83 °C for 5 minutes in a hybridization oven, avoiding light exposure for all subsequent steps. Hybridize the tissue slides for four hours at room temperature in a humidified slide hybridization chamber, or overnight at 4 °C.

After hybridization, carefully remove the coverslip from the tissue slides with forceps. Wash the slides twice in hybridization wash buffer 1 for 15 minutes each with gentle agitation. Then wash the slides three times for 5 minutes each in hybridization wash buffer 2, adding 3000 x DAPI stock solution to the second wash. Wash the slides three times in deionized distilled water for 5 minutes each. Drain the tissue slides without allowing them to dry completely.

Apply 50 µl of ProLong antifade mounting medium to each tissue slide slowly, avoiding air bubbles. Carefully apply a coverslip to each tissue slide. Store the slides at room temperature in the dark for 24 hours to allow the ProLong to harden. Subsequently, store the slides at 4 °C for at least one week before imaging; slides can be stored at −20 °C for up to one year.

Visualize and image the slides using a fluorescence microscope equipped with an appropriate fluorescence filter set and a 40x or 63x objective. Telomere signal intensity is assessed using Telometer plugin in ImageJ.

27.3.2.1 UTF Assay Interpretation

Tumors are classified as ALT-positive if they meet the following criteria: (1) the presence of both heterogeneous and ultrabright nuclear foci of telomere FISH signals. (2) The presence of ultra-bright telomere foci, with the integrated total signal intensities for individual foci being more than ten times the mean integrated signal intensities of all telomeric signals in individual stromal cells within the same case. (3) At least 1% of cancer cells display ALT-associated telomeric DNA foci. Tumor samples that lack ALT-associated ultrabright telomeric foci, with at least 2000 cells assessed, are considered ALT-negative.

27.4 Results

Examples of use of the C-circle assay can be found in two studies our laboratory performed to define telomere maintenance mechanism for high-risk neuroblastoma patients. The first study assessed DNA from 107 patients, of which 25 (23%) were found to be ALT+ by the C-circle assay. For this initial study we employed RNA sequencing data (quantitation of TERT mRNA) to define tumors with positive telomerase, of which 54 (51%) were telomerase-positive. We observed 28 tumors had no apparent TMM mechanism, but analyzing

APBs using FFPE tumor sections of those tumors reclassified three as ALT-positive.

The second set of clinical samples we analyzed were DNA and RNA samples from 204 patients enrolled on the COG ANBL0532 phase III study. For this study the C-circle assay was performed as for the prior study but real-time RT-PCR was used to quantify TERT mRNA expression to identify telomerase-positive tumors. In the ANBL 0532 study we observed that the APB assay can miss some low-TERT NB that are both C-circle-positive and have clinical outcomes indicative of ALT, but we did not observe this with the UTF assay [40]. Therefore, for the ANBL 0532 study we employed the UTF assay to confirm TMM negative status of tumors with low TERT expression and a low C-circle content. Analysis of the samples from the ANBL 0532 confirmed that for all patients on the study with ALT-positive tumors had a lower overall survival [40].

27.5 Discussion

While most cancers have reactivated telomerase as a means to maintain telomeres and ensure replicative immortality, as many as 15% of cancers lack telomerase expression and maintain telomeres by an alternative lengthening of telomeres mechanism. Defining a cancer as being ALT-positive can be useful for determining prognosis and risk stratification. Also, as investigational drugs have been identified that have particularly high activity against ALT cancers in preclinical studies [41–44], defining cancers as ALT-positive is needed to carry out clinical trials of such agents and potentially as companion diagnostics for ALT-targeting drugs.

Here we have presented methods for two robust assays to detect ALT in tumor tissue. The telomeric DNA C-circle assay, which employs real-time PCR, is sensitive and specific for detecting ALT when DNA from fresh or snap-frozen tumors is available to use in the assay. The ultrabright telomere foci (UTF) assay utilizes FFPE tumor sections, and thus can be used when DNA from fresh or frozen tumor is not available. As false negative results can be obtained with the C-circle assay, likely due to spatial tumor heterogeneity, the UTF assay is useful for confirming tumors that lack TERT expression and have a negative C-circle assay are indeed TMM-negative.

Disadvantages of the UTF assay are that it's labor-intensive and requires operator subjective interpretation. By contrast, the C-circle assay can be done rapidly on a large number of specimens and provides quantitative data that is interpreted qualitatively with an objectively defined cutoff. One disadvantage of the C-circle assay is the final step is a telomere PCR reaction, and thus the assay relies for a positive assay on a significant difference between the background total telomere PCR signal and that of the C-circles that are amplified selectively by Phi-29 polymerase (Phi-29). Defining cut-offs for a positive sample relies on having robust positive and negative control samples and an appropriate non-telomere reference gene. In spite of the disadvantages of the assays, they can be applied with clinical samples to obtain robust and reproducible data [16, 38].

27.6 Future Trends

The C-circle assay and detection of ultra-bright telomere foci remain the most validated methods for identifying ALT in patient tumor tissue sections. However, several new ALT detection methods, such as ALT-FISH [39, 45], have shown promise as specific, quantifiable markers of ALT at single-cell resolution. These approaches could be valuable for screening ALT-specific drugs and for exploring the spatial heterogeneity of ALT activity. Whether ALT-FISH can replace the C-circle and ultra-bright telomere foci assays requires validation in a large cohort of patient samples.

The exploration of circulating tumor DNA (ctDNA) as a noninvasive approach for cancer diagnosis and disease monitoring represents an exciting advancement in ALT detection. ctDNA allows for real-time tracking of tumor dynamics and potential treatment responses through blood-based sampling, offering a promising route for longitudinal monitoring of ALT-positive tumors. Recent findings from our group and others suggest that C-circles, a key marker of ALT, can be detected in plasma or serum samples with potentially high sensitivity and specificity [46, 47]. However, these assays are still in the early stages of validation, having only been tested in small patient cohorts. A thorough evaluation of the sensitivity and specificity of the C-circle assay in larger, diverse cohorts, and across different cancer types, will be necessary to determine the clinical utility of the cfDNA C-circle assay.

Additionally, recent studies indicate that the ALT pathway produces telomere fusions detectable in the blood of cancer patients, introducing another possible marker for noninvasive ALT detection [48]. While this method could broaden the scope of ALT monitoring in clinical settings, its sensitivity across various ALT-positive cancer types needs to be assessed, particularly in histologies with a high prevalence of the ALT phenotype.

Looking forward, the integration of these novel approaches into multimodal detection platforms, combining tissue- and blood-based assays, could significantly enhance our ability to diagnose ALT activity, monitor disease progression, and evaluate responses to therapies targeting ALT cancers. Developing robust, validated methods that are assessed across larger and more diverse patient cohorts will be necessary to translate these promising detection strategies into clinical practice.

REFERENCES

1. Blackburn EH. Structure and function of telomeres. *Nature* 1991;350(6319):569–73.
2. Levy MZ, Allsopp RC, Futcher AB, *et al*. Telomere end-replication problem and cell aging. *J Mol Biol* 1992; 225(4):951–60.
3. Wellinger RJ. In the end, what's the problem? *Mol Cell* 2014;53(6):855–6.
4. Collado M, Gil J, Efeyan A, *et al*. Tumour biology: Senescence in premalignant tumours. *Nature* 2005; 436(7051):642.
5. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell* 2011;144(5):646–74.

6. Broccoli D, Young JW, de Lange T. Telomerase activity in normal and malignant hematopoietic cells. *Proc Natl Acad Sci U S A* 1995;92(20):9082–6.
7. Nugent CI, Lundblad V. The telomerase reverse transcriptase: Components and regulation. *Genes Dev* 1998;12(8): 1073–85.
8. Huang FW, Hodis E, Xu MJ, *et al*. Highly recurrent TERT promoter mutations in human melanoma. *Science* 2013;339(6122):957–9.
9. Killela PJ, Reitman ZJ, Jiao Y, *et al*. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. *Proc Natl Acad Sci U S A* 2013;110(15):6021–6.
10. Landa I, Ganly I, Chan TA, *et al*. Frequent somatic TERT promoter mutations in thyroid cancer: Higher prevalence in advanced forms of the disease. *J Clin Endocrinol Metable* 2013;98(9):E1562–6.
11. Peifer M, Hertwig F, Roels F, *et al*. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* 2015;526(7575):700–4.
12. Langford LA, Piatyszek MA, Xu R, *et al*. Telomerase activity in human brain tumours. *Lancet* 1995;346(8985):1267–8.
13. Hiyama E, Gollahon L, Kataoka T, *et al*. Telomerase activity in human breast tumors. *J Natl Cancer Inst* 1996; 88(2):116–22.
14. Albanell J, Lonardo F, Rusch V, *et al*. High telomerase activity in primary lung cancers: Association with increased cell proliferation rates and advanced pathologic stage. *J Natl Cancer Inst* 1997;89(21):1609–15.
15. Borah S, Xi L, Zaug AJ, *et al*. Cancer. TERT promoter mutations and telomerase reactivation in urothelial cancer. *Science* 2015;347(6225):1006–10.
16. Koneru B, Lopez G, Farooqi A, *et al*. Telomere maintenance mechanisms define clinical outcome in high-risk neuroblastoma. *Cancer Res* 2020;80(12):2663–75.
17. Henson JD, Cao Y, Huschtscha LI, *et al*. DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat Biotechnol* 2009;27(12):1181–5.
18. Murnane JP, Sabatier L, Marder BA, *et al*. Telomere dynamics in an immortal human cell line. *EMBO J* 1994;13(20):4953–62.
19. Bryan TM, Englezou A, Gupta J, *et al*. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO J* 1995;14(17):4240–8.
20. Yeager TR, Neumann AA, Englezou A, *et al*. Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res* 1999;59(17):4175–9.
21. Wu G, Lee WH, Chen PL. NBS1 and TRF1 colocalize at promyelocytic leukemia bodies during late S/G2 phases in immortalized telomerase-negative cells. Implication of NBS1 in alternative lengthening of telomeres. *J Biol Chem* 2000;275(39):30618–22.
22. Nabetani A, Yokoyama O, Ishikawa F. Localization of hRad9, hHus1, hRad1, and hRad17 and caffeine-sensitive DNA replication at the alternative lengthening of telomeres-associated promyelocytic leukemia body. *J Biol Chem* 2004;279(24):25849–57.
23. Heaphy CM, de Wilde RF, Jiao Y, *et al*. Altered telomeres in tumors with ATRX and DAXX mutations. *Science* 2011;333(6041):425.
24. Cesare AJ, Reddel RR. Alternative lengthening of telomeres: Models, mechanisms and implications. *Nat Rev Genet* 2010;11(5):319–30.
25. Sobinoff AP, Pickett HA. Mechanisms that drive telomere maintenance and recombination in human cancers. *Curr Opin Genet Dev* 2020;60:25–30.
26. Zhang JM, Zou L. Alternative lengthening of telomeres: From molecular mechanisms to therapeutic outlooks. *Cell Biosci* 2020;10:30.
27. Episkopou H, Draskovic I, Van Beneden A, *et al*. Alternative lengthening of telomeres is characterized by reduced compaction of telomeric chromatin. *Nucleic Acids Res* 2014;42(7):4391–405.
28. Clynes D, Jelinska C, Xella B, *et al*. Suppression of the alternative lengthening of telomere pathway by the chromatin remodelling factor ATRX. *Nat Commun* 2015;6:7538.
29. Goldberg AD, Banaszynski LA, Noh KM, *et al*. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 2010;140(5):678–91.
30. Lewis PW, Elsaesser SJ, Noh KM, *et al*. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci U S A* 2010;107(32):14075–80.
31. Cesare AJ, Kaul Z, Cohen SB, *et al*. Spontaneous occurrence of telomeric DNA damage response in the absence of chromosome fusions. *Nat Struct Mol Biol* 2009;16(12): 1244–51.
32. Conomos D, Stutz MD, Hills M, *et al*. Variant repeats are interspersed throughout the telomeres and recruit nuclear receptors in ALT cells. *J Cell Biol* 2012;199(6):893–906.
33. Dilley RL, Verma P, Cho NW, *et al*. Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature* 2016;539(7627):54–58.
34. Meeser A, Bartenhagen C, Werr L, *et al*. Reliable assessment of telomere maintenance mechanisms in neuroblastoma. *Cell Biosci* 2022;12(1):160.
35. Ackermann S, Cartolano M, Hero B, *et al*. A mechanistic classification of clinical phenotypes in neuroblastoma. *Science* 2018;362(6419):1165–70.
36. Chudasama P, Mughal SS, Sanders MA, *et al*. Integrative genomic and transcriptomic analysis of leiomyosarcoma. *Nat Commun* 2018;9(1):144.
37. Henson JD, Lau LM, Koch S, *et al*. The C-circle assay for alternative-lengthening-of-telomeres activity. *Methods* 2017;114:74–84.
38. Lau LM, Dagg RA, Henson JD, *et al*. Detection of alternative lengthening of telomeres by telomere quantitative PCR. *Nucleic Acids Res* 2013;41(2):e34.
39. Claude E, de Lhoneux G, Pierreux CE, *et al*. Detection of alternative lengthening of telomeres mechanism on tumor sections. *Mol Biomed* 2021;2(1):32.
40. Koneru B, Burrow TA, McCoy K, *et al*. Improved outcome for patients with alternative lengthening of telomeres (ALT) neuroblastoma randomized to tandem myeloablative therapy on COG ANBL0532. In *Advances in Neuroblastoma Research Meeting Abstract O6.3*. Amsterdam, Netherlands; 2023.
41. Koneru B, Farooqi A, Nguyen TH, *et al*. ALT neuroblastoma chemoresistance due to telomere dysfunction–induced ATM activation is reversible with ATM inhibitor AZD0156. *Sci Transl Med* 2021;13(607):eabd5750.

42. Macha SJ, Koneru B, Burrow TA, *et al*. Alternative lengthening of telomeres in cancer confers a vulnerability to reactivation of p53 function. *Cancer Res* 2022;82(18):3345–58.
43. Wang Y, Yang J, Wild AT, *et al*. G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat Commun* 2019;10(1):943.
44. Flynn RL, Cox KE, Jeitany M, *et al*. Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors. *Science* 2015;347(6219):273–7.
45. Frank L, Rademacher A, Mucke N, *et al*. ALT-FISH quantifies alternative lengthening of telomeres activity by imaging of single-stranded repeats. *Nucleic Acids Res* 2022;50(11):e61.
46. Chen YY, Dagg R, Zhang Y, *et al*. The C-circle biomarker is secreted by alternative-lengthening-of-telomeres positive cancer cells inside exosomes and provides a blood-based diagnostic for ALT activity. *Cancers (Basel)* 2021;13(21).
47. Lehman TE, Deal JL, Lee J, *et al*. HGG–48. The blood-based telomeric DNA C-circle assay can detect alternative lengthening of telomeres (ALT) in pediatric high-grade gliomas. *Neuro Oncol* 2024;26(Suppl 4):0.
48. Muyas F, Rodriguez MJG, Cascao R, *et al*. The ALT pathway generates telomere fusions that can be detected in the blood of cancer patients. *Nat Commun* 2024;15(1):82.

28

Click Chemistry-Mediated Enrichment of Tumor-Derived Extracellular Vesicles for RNA-Based Digital Scoring

Yazhen Zhu, Jacqueline Ziqian Yang, Anna Yaxuan Wei, Vivian Xufei Zuo, Judy Law, Sungyong You, Ju Dong Yang, Vatche G. Agopian, Edwin Posadas, Junseok Lee, and Hsian-Rong Tseng

28.1 Introduction

In the realm of liquid biopsy, which includes circulating tumor DNA, circulating tumor cells, and extracellular vesicles (EVs), EVs have garnered significant interest because of their distinctive characteristics.[1, 2] EVs, a heterogeneous collection of nanoparticles surrounded by phospholipid bilayer membranes, are known to be actively secreted by various cell types, particularly by tumor cells and those in the tumor microenvironment.[3, 4] These tumor-derived EVs are detectable in the peripheral blood from the early stages of the disease and persist throughout its progression. Typically, a substantial number of tumor EVs (between 10^8 and 10^{10} per mL) circulate in the peripheral blood.[5, 6] The half-life of these tumor EVs ranges from 1 to 5.5 hours, with 90% undergoing hepatic clearance.[7–9] The bilayer lipid membranes of tumor EVs carry surface markers that reflect of the parent tumor cells and shield delicate biomolecular cargos, such as nucleic acids and proteins.[10] The aforementioned characteristics have distinguished tumor EVs as a feasible alternative to tumor tissues for biomolecular analysis and systemic evaluation of solid tumors, thereby highlighting their potential for noninvasive cancer diagnostics and continuous treatment monitoring.[11, 12] Nonetheless, the extraction of tumor EVs remains a technical challenge due to their rare occurrence among the total EV population in blood samples.

One of the main obstacles hindering the clinical use of tumor EVs is the absence of reproducible and effective tumor EV enrichment techniques that ensures their intactness and purity at the end of enrichment procedures. Traditional methods such as ultracentrifugation,[13, 14] precipitation,[15] filtration,[16, 17] and size-based microfluidic enrichment[18–20] can separate total EVs in plasma or serum samples based on physical properties like size and density. Yet, these methods are incapable of separating tumor EVs from non-tumor EVs that are also present in the clinical samples. Recent studies have explored immunoaffinity-based capture techniques[13, 14, 21] for selectively isolating tumor EVs from different solid tumors. For example, EVs from pancreatic cancer can be isolated using beads coated with anti-GPC1 antibodies,[22] and EVs from glioblastoma can be enriched using an EGFRvIII antibody in a herringbone microfluidic device.[23] An antibody cocktail targeting surface protein biomarkers (i.e., EGFR, EpCAM, CD24, and GPA33) was employed with an integrated magneto-electrochemical device[24] to quantify tumor EVs in colorectal cancer (CRC) patients. However, these methods often involve antibody coating by processes such as physisorption,[25] amide coupling,[26] and biotin-streptavidin binding,[27] which may affect the efficiency and introduce potential biological interferences.[28–31] Moreover, the standard immunoaffinity-based methods for EV capture typically show suboptimal tumor EV enrichment performance due to the small number of antigens on each tumor EV, resulting in insufficient interactions between the tumor EV antigens and the capture antibodies.

To tackle the aforementioned challenges, our joint research team at the David Geffen School of Medicine at UCLA has led the way in developing click chemistry-mediated tumor EV enrichment technologies, specifically Click Chips[32, 33] and Click Beads.[34] These novel tumor EV enrichment platform technologies offer significant improvements in several areas: first, they improve the efficiency of labeling tumor EVs with click motif-grafted antibodies in small volumes in the liquid phase, significantly reducing the quantity of antibodies required and the time needed for labeling. Second, an antibody cocktail targeting multiple tumor surface markers can be effectively combined with click chemistry-mediated enrichment to address the challenges of tumor heterogeneity and the low density of antigens on tumor EV surfaces. Third, we utilize a click reaction,[35] a type of rapid bio-orthogonal organic reactions commonly used for biological labeling and bioconjugation, for EV enrichment.[36, 37] It is noteworthy that click chemistry was recently recognized with the 2022 Nobel Prize in Chemistry for its advances in diagnostics, pharmaceutical development, and medicine. In our approaches, we specifically used the inverse electron demand Diels–Alder reaction[38] between trans-cyclooctene (TCO) and tetrazine (Tz) motifs,[39] which effectively reduces biological disruptions typically seen in streptavidin-biotin interactions. Additionally, our recent research has shown that click chemistry-mediated enrichment of tumor EVs can be effectively combined with various downstream molecular analyses, such as mRNA profiling[32–34, 40, 41] and protein detection.[42]

Among the various biomolecular contents within tumor EVs, mRNA is a particularly promising target for implementation of cancer detection. Recent studies have demonstrated the feasibility of detecting oncogenic mutations using tumor EV-derived mRNA from pancreatic cancer[22] and glioblastoma.[23] This approach, using tumor EV-derived mRNA for

DOI: 10.1201/9781003449942-32

mutation analysis, has demonstrated improved sensitivity and better correlation with patient clinical outcomes.[43, 44] Additionally, enriched tumor EVs can undergo downstream profiling of tumor tissue lineage-specific mRNA markers,[32, 33] which are unique to individual tumor types such as hepatocellular carcinoma (HCC) and prostate cancer (PCa) but absent in blood cells and other tissues. This method could broaden the range of available biomarkers and confirm the tissue-of-origin of these tumor EVs. Notably, RNA-based EV profiling from tissue lineage can potentially be applied to almost all solid tumors without the need for individualized genotype assays based on primary tumor sequencing or dependency on acquired cancer-specific mutations, which could be unavailable in some solid tumors, including HCC and PCa. Overall, an ideal tumor EV-based mRNA assay would consist of two primary components: an efficient system for enriching tumor EVs and a highly sensitive mRNA profiling technology that can accurately measure mRNA levels.

To effectively analyze and measure the extremely low amounts of mRNA obtained from purified tumor EVs, advanced mRNA profiling techniques such as next-generation sequencing (NGS)[23] and reverse transcription quantitative PCR (RT-qPCR) are essential. Compared to mRNA profiling via NGS, which is relatively expensive and time-consuming, digital PCR machines offer a more efficient solution for measuring small quantities of mRNA against a high background. These machines utilize either droplet-based or nanoplate-based sample partitioning technologies. Although digital PCR is limited to detecting a small panel of mRNA targets, it operates at a lower cost and can deliver results within a few hours.

In this book chapter, we introduce a novel type of liquid biopsy assay known as EV digital scoring assays (Figure 28.1), designed for the absolute quantification of mRNA molecules in tumor EVs. These assays have the potential to serve as a diagnostic tool for measuring systemic tumor burden in peripheral blood, enabling noninvasive cancer detection and ongoing treatment monitoring. The workflow of the EV digital scoring assays involves two primary steps: Step 1: click chemistry-mediated enrichment of EVs using either Click Chips[32, 33] coupled with TCO-grafted antibodies targeting tumor-specific EV surface protein markers or EV Click Beads [34] coupled with the use of a TCO-grafted lipid compound (i.e., 1,2-distearoyl-sn-glycero-3-phosphoethanolamine-polyethylene glycol 1000-TCO, DSPE-PEG_{1000}-TCO) for those tumors' lack of tumor-specific EV surface protein markers. The DSPE-PEG_{1000}-TCO provides a universal approach to label EVs with TCO by incorporating its lipid motif onto EV membranes. Step 2: quantification of mRNA in the enriched tumor EVs using either RT-digital PCR[34] (QIAcuity, Qiagen, USA) or RT-droplet digital PCR[32, 33] (QX200, Bio-Rad Laboratories, Inc., USA). The resulting mRNA profiles, which include oncogenic mutations and tumor EV-specific mRNA markers, facilitate diagnostic applications in various solid tumors.

28.2 Materials

28.2.1 Chemicals and Reagents

All chemicals without indicated sources were purchased from MilliporeSigma and used as received. The cell line culture medium and serum-free culture media are purchased from Thermo Fisher Scientific, USA. Silica microbeads (10 mg, 6 × 108 beads, Bangs Laboratories, Inc., USA) were used for the preparation of EV Click Beads. For HCC EV purification, goat anti-human EpCAM (R&D Systems, Inc., USA), goat anti-human CD 147 (R&D Systems, Inc., USA), rabbit anti-human ASGPR1 (LifeSpan BioSciences, Inc., USA), and sheep anti-human GPC3 (R&D Systems, Inc., USA) antibodies were used. For the enrichment of PCa EVs, goat anti-human EpCAM (R&D Systems, Inc., USA) and sheep anti-human PSMA (R&D Systems, Inc., USA) antibodies were used.

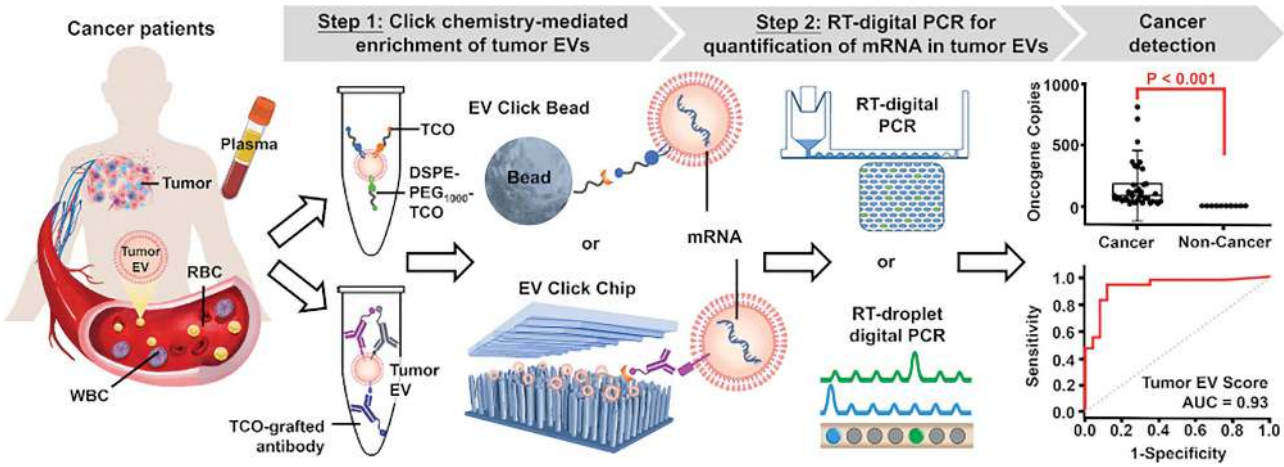


FIGURE 28.1 *General workflow of EV digital scoring assay.* Plasma samples from patients with different solid tumors were collected and subjected to EV digital scoring assay implemented through a two-step workflow—Step 1: click chemistry-mediated EV enrichment by either EV Click Chips or EV Click Beads, in conjunction with the use of TCO-grafted tumor-specific antibodies or TCO-grafted lipid motif (DSPE-PEG_{1000}-TCO), and Step 2: quantification of mRNA in the enriched tumor EVs using either RT-digital PCR (QIAcuity, Qiagen, USA) or RT-droplet digital PCR (QX200, Bio-Rad Laboratories, Inc., USA). The resulting mRNA signatures allow for diagnostic applications in individual solid tumors.

28.2.2 Cell Line Culture

Ewing sarcoma cell lines (i.e., A673 and ES 5838), pancreatic ductal adenocarcinoma (PDAC) cell lines (i.e., CFPAC-1 and AsPC-1), HCC cell lines (i.e., HepG2, Hep3B, and SNU387), and PCa cell lines (i.e., 22Rv1 and LNCaP) were obtained from the American Type Culture Collection (ATCC). C4–2B PCa cell line was kindly provided by Dr. Leland W. K. Chung. Ultrapure water (18 MΩ cm) was obtained using a Milli-Q purification system (Millipore Corp., MA) and used in preparation of all the solutions. All these cell lines were cultured in 18 EasYDish™ Dishes (Thermo Fisher Scientific, USA) under the conditions specified by ATCC using different culture mediums with 10% fetal bovine serum (FBS) and 100 U mL^{-1} penicillin-streptomycin in a humidified incubator with 5% CO_2. A673 and ES5838 Ewing sarcoma cells were cultured in Dulbecco's Modified Eagle Medium (DMEM). HepG2 and Hep3B HCC cell lines were cultured in Eagle's Minimum Essential Medium (EMEM). All other cell lines were cultured in Roswell Park Memorial Institute (RPMI) 1640 media. When cells reached 70%–80% confluency, the cell culture medium was changed to serum-free culture media (Thermo Fisher Scientific, USA) to starve the cells for 24–48 hours, which was finally collected for cell line-derived EV harvesting.

28.3 Methods

28.3.1 Preparation of Synthetic EV Samples as Model Systems

28.3.1.1 Harvesting Cell Line-Derived EVs from the Serum-Free Media

After collecting the cell line-derived EVs in the serum-free media, an initial centrifugation at 300 g (4 °C) for 10 min was applied to remove cells and cell debris. The supernatant was then transferred to new tubes and centrifuged at 2800 g (4 °C) for 10 min to further remove the remaining cellular debris. Finally, the supernatant was carefully transferred to Ultra-Clear Tubes (38.5 mL, Beckman Coulter, Inc., USA) for ultracentrifugation at 100,00 g (4 °C) for 70 min using an Optima L-100 XP Ultracentrifuge (Beckman Coulter, Inc., USA). The EV pellets at the bottom of the tubes were carefully collected and resuspended in200 µL phosphate buffered saline (PBS). The collected EVs were then treated with RNase[45, 46] at 37 °C for 30 min, followed by PBS wash, then the EV pellets were recollected by ultracentrifugation at 100,00 g (4 °C) for 70 min.

28.3.1.2 Preparation of Synthetic EV Samples

Synthetic EV plasma samples were prepared by spiking 10–50 µL of cell line-derived EVs into 90–200 µL of plasma from healthy donors (HDs). The absolute quantities of cell line-derived EVs in the synthetic EV samples were determined by RT-digital PCR (QIAcuity, Qiagen, USA) or RT-droplet digital PCR (QX200, Bio-Rad Laboratories, Inc., USA). These synthetic EV plasma samples were later used in the optimization of tumor EV enrichment by either EV Click Beads[34] or EV Click Chips.[32, 33]

28.3.2 Characterization of Cell Line-Derived EVs

The characterization of the cell line-derived EVs followed the guidelines outlined by the International Society for Extracellular Vesicles.[47] In short, these cell line-derived EVs were characterized using dynamic light scattering (DLS), scanning electron microscopy (SEM), transmission electron microscopy (TEM), and immunogold staining.

28.3.2.1 DLS

The EVs were placed into a disposable microcuvette and analyzed using a Zetasizer Nano instrument (Malvern Instruments Ltd., UK) at room temperature.

28.3.2.2 SEM

The EVs were deposited on substrates and fixed in 4% paraformaldehyde (PFA) for one hour, followed by sequential dehydration through 30%, 50%, 75%, 85%, 95%, and 100% ethanol solutions for 10 min each. After overnight lyophilization, the samples were sputter-coated with gold at room temperature. The EVs were visualized with a ZEISS Supra 40VP SEM at an accelerating voltage of 10 keV.

28.3.2.3 TEM

The EV samples were fixed in 4% PFA (in PBS) for 30 min. After depositing the fixed EV on 400-mesh carbon-coated copper grids, the grids were kept at room temperature for 10 min. Excess samples were blotted with filter paper and washed five times by water. Grids were dried for TEM imaging and the EVs were visualized by a T20 iCorr (FEI) high-resolution cryo-TEM at an accelerating voltage of 80 kV.

28.3.2.4 Immunogold Staining

EVs captured on EV Click Beads were fixed in 4% PFA. The samples were then incubated with anti-CD63 (Abcam, mouse, 1:100 dilution) for 30 min. After washing with DI water, these samples were treated with anti-mouse nanogold (12 nm, 1:50 dilution) for one hour. The gold-labeled samples were dropped onto carbon-coated copper grids and incubated for 10 min before being wiped off from the grids. After being rinsed five times using DI water, grids were then dried for TEM imaging.

28.3.3 Click Chemistry-Mediated Enrichment of Ewing Sarcoma or PDAC EVs by EV Click Beads

28.3.3.1 Preparation of EV Click Beads

Silica microbeads (10 mg, 6×10^8 beads, Bangs Laboratories, Inc., USA) were first treated with 2.0 N HNO_3 for 10 min to freshly introduce hydroxyl groups to the microbeads' surfaces. After drying, salinization of the microbeads was carried out in

600 μL of ethanol solution containing 4% of 3-aminopropyltriethoxysilane at room temperature for 45 min to give amino-functionalized silica microbeads. After reacting with tetrazine (Tz)-PEG-NHS ester (0.94 mg, 3.8×10^{-3} M) in DMSO/PBS (600 μL) for one hour, Tz-grafted EV Click Beads were obtained.

28.3.3.2 Preparation of DSPE-PEG$_{1000}$-TCO

TCO-PEG-NHS ester (5.0×10^{-3} M) was incubated with DSPE-PEG$_{1000}$-NH$_2$ (5.0×10^{-3} M) in dimethyl sulfoxide (DMSO)/PBS solution at room temperature for 30 min to give a solution of DSPE-PEG$_{1000}$-TCO conjugate, which was stored at −20 °C until use.

28.3.3.3 Enrichment of Ewing Sarcoma EVs or PDAC EVs by EV Click Beads

Ewing Sarcoma EVs or PDAC EVs in synthetic or clinical plasma samples (1 mL) are first labeled with TCO via the insertion of the lipid motif of DSPE-PEG$_{1000}$-TCO (50 nmol) into the membranes of the EVs in the plasma samples. Upon the addition of EV Click Beads, TCO-labeled EVs were immobilized onto EV Click Beads via click chemistry-mediated EV enrichment. The plasma samples were then subjected to rapid centrifugation at 300 g for 2 min, after which the Ewing Sarcoma or PDAC EVs on the EV Click Beads were collected at the bottom of an Eppendorf tube.

28.3.4 Click Chemistry-Mediated Enrichment of HCC EVs or PCa EVs by EV Click Chips

28.3.4.1 Fabrication of EV Click Chips

An EV Click Chip is composed of two functional components, including the Tz-grafted silicon nanowire substrates (SiNWS) and an overlaid polydimethylsiloxane (PDMS) chaotic mixer. A custom-designed chip holder[32] is employed to house a Tz-grafted SiNWS and a PDMS chaotic mixer to give the EV Click Chip.

28.3.4.1.1 The Fabrication of Tz-Grafted SiNWS

The fabrication was started from lithographical patterning of a p-type Si (100) wafer (Silicon Quest Int'l, USA), followed by the introduction of 10–15 μm long vertically aligned Si nanowires onto the patterned areas via Ag-nanoparticle templated wet etching process.[48] After drying under nitrogen flow, the resultant SiNWS were sealed in a vacuum desiccator for treatment with 3-mercaptopropyltrimethoxysilane vapor for 45 min to introduce thiol groups onto the SiNWS. The thiol-grafted SiNWS was then treated with orthopyridyl disulfide-PEG-Amine (OPSS-PEG-NH$_2$, 0.30 mg, 3.8 mM, Nanocs Inc., USA) in 200 μL of DMSO for two hours to introduce disulfide linkers with terminal amine groups. Subsequently, the amine-terminated SiNWS were incubated with Tz-sulfo-NHS ester (0.32 mg, 3.8 mM, Click Chemistry Tools, USA) in 200 μL of PBS (pH = 8.5) for one hour. The resulting Tz-grafted SiNWS were rinsed with DI water three times. After drying under nitrogen flow, the Tz-grafted SiNWS were stored at −20 °C for future use.

28.3.4.1.2 The Fabrication of PDMS Chaotic Mixers[48]

The mixers were fabricated from a photolithographically prepared master wafer prepared by inductively coupled plasma-reactive ion etching (ICP-RIE). To fabricate the master wafer, a protective layer of chromium and positive photoresist was first cast on top of a 4-inch silicon wafer; the silicon wafer was then exposed to UV light through a photomask of three rectangular microfluidic channels. Next, the exposed regions of the photoresist were dissolved away, and the exposed chromium was removed with an acid wash. ICP-RIE was then applied to etch the silicon to a depth of 115 μm. A similar procedure was then used to etch 40 μm deep herringbone ridges on top of the silicon wafer. Prior to making replicas via injection molding, the Si master was pretreated by exposure to trimethylchlorosilane vapor for 1 min. Well-mixed PDMS precursor (RTV615 A and B in a 10 to 1 ratio, GE Silicones) was injected to the mold and then incubated in an oven at 80 °C for 48 h to make a 5 mm thick slab. The resulting PDMS chaotic mixers were peeled off from the silicon master wafer/molds. During injection molding, two holes were also fabricated at the ends of the channel for insertion of tubing.

28.3.4.2 Preparation of TCO-antibody Conjugates

For efficient enrichment of HCC EVs or PCa EVs, a collection of HCC/PCa-specific antibodies were prepared by the following procedure. Each antibody was reacted with TCO-PEG4-NHS ester (0.5 mM, Click Chemistry Tools, USA) in PBS at room temperature for 30 min.

28.3.4.3 Enrichment of HCC EVs or PCa EVs by EV Click Chips

HCC EVs or PCa EVs in 1 mL of plasma samples are first labeled with TCO via the incubation with the respective cocktail of TCO-conjugated antibodies for 30 min at room temperature. After the assembly of EV Click Chips, the plasma/antibody mixtures were injected into EV Click Chips, in which HCC/PCa EVs in the plasma samples were captured through the click chemistry-mediated EV enrichment. After immobilizing HCC/PCa EVs onto the Tz-grafted SiNWS, 100 μL of 50 mM 1,4-dithiothreitol (DTT) solution in PBS was injected into the EV Click Chips at a flow rate of 1.0 mL/h. The cleavage of disulfide linkers resulted in the release of the HCC/PCa EVs, which were then collected in 1.5 mL RNase-free Eppendorf tubes for subsequent RNA extraction.

28.3.5 Extracting RNA from Enriched EVs and cDNA Synthesis

These enriched EVs, from Ewing Sarcoma, PDAC, HCC, or PCa, were lysed using 700 μL of QIAzol Lysis Reagent (Qiagen, USA). The miRNeasy Micro Kit (Qiagen, USA) was applied to extract and purify RNA from the isolated EVs according to the manufacturer's instructions.[32–34] The RNA concentration within the EVs was determined using a Qubit 3.0 Fluorometer (Thermo Fisher Scientific, USA) together with the Qubit RNA HS Assay (Thermo Fisher Scientific, USA). The extracted RNA from the enriched EVs was reversely transcribed to

cDNA using a Maxima H Minus Reverse Transcriptase Kit (Thermo Fisher Scientific, USA). The cDNA was prepared through synthesis at 55 °C for 30 min and subsequent denaturation at 85 °C for 5 min.

28.3.6 RT-Digital PCR

RT-digital PCR (QIAcuity, Qiagen, USA) was used for the quantification of EWS/FLI-1 rearrangements in Ewing sarcoma patients' plasma samples and KRAS mutations in PDAC patients' plasma samples. The preamplification of cDNA was conducted using 4.0 μL of cDNA as input to yield 10 μL of preamplified product for each gene, based on the manufacturer's guidelines provided with the Prelude PreAmp Master Mix Kit from Takara Bio. The reaction mixture, comprising 40 μL in total with 4 μL of preamplified product, was loaded into each well of a nanoplate (26 K, 24 wells). After this, the nanoplate was transferred to the QIAcuity instrument (Qiagen, USA) for the subsequent PCR process. The setting of the programmed Thermal Cycler was configured to 95 °C for 2 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 30 s. The readouts from digital PCR were analyzed using QIAcuity software by automatic measurement of positive and negative partitions.

28.3.7 RT-Droplet Digital PCR for Quantification of HCC and PCa-Specific mRNA Markers

For HCC EVs, the resultant cDNA was equally divided into five tubes for the quantification of the ten HCC-specific mRNA markers by RT-droplet digital PCR (QX200, Bio-Rad Laboratories, Inc., USA) with two fluorescence filters (FAM and VIC) in each tube. These ten HCC-specific mRNA markers include alpha-fetoprotein (*AFP*), glypican 3 (*GPC3*), albumin (*ALB*), apolipoprotein H (*APOH*), fatty acid binding protein 1 (*FABP1*), fibrinogen beta chain (*FGB*), fibrinogen gamma chain (*FGG*), alpha 2-HS glycoprotein (*AHSG*), retinol binding protein 4 (*RBP4*), and transferrin (*TF*). Similarly, for PCa EVs, the obtained cDNA was aliquoted into six tubes to detect the 11 PCa-specific mRNA markers by duplex digital PCR. These 11 PCa-specific mRNA markers include *ACP3*, *FOLH1*, *HOXB13*, *KLK2*, *KLK3*, *KLK4*, *MSMB*, *RLN1*, *SLC45A3*, *STEAP2*, and *TMPRSS2*. All the primers and probes were purchased from Thermo Fisher Scientific and verified by cell lines and cell line-derived EVs before testing each gene in the RT-droplet digital PCR. The copy numbers of gene transcripts in each assay were calculated using the QuantaSoft™ software based on the counts of positive droplets.

28.3.8 Enrollment of Cancer Patients and Control Cohorts

28.3.8.1 Ewing Sarcoma Patients

In total, 35 plasma samples from 28 patients with clinically confirmed Ewing sarcoma and 10 plasma samples from 10 HDs were analyzed. Among the 28 Ewing sarcoma patients, the EWS/FLI-1 rearrangements were compared between plasma samples and the respectively paired undecalcified formalin-fixed paraffin-embedded (FFPE) tissue samples of 12 Ewing sarcoma patients. All patients and HDs provided written informed consent to be enrolled in this study according to the UCLA Institutional Review Board (IRB) protocol (IRB #14-000197).

28.3.8.2 PDAC Patients

In total, 35 plasma samples from 19 patients with clinically confirmed PDAC and 10 plasma samples from 10 HDs were analyzed. Among the 19 pancreatic cancer patients, 13 have paired FFPE tissue samples. These paired FFPE tissue and plasma samples from 13 patients were compared for the three types of KRAS mutations. All patients and HDs provided written informed consent to be enrolled in this study according to the UCLA IRB protocol (IRB #14-000197).

28.3.8.3 HCC and Liver Cirrhosis Patients

All the participants (18+ years of age) in this study were enrolled through Ronald Reagan UCLA Medical Center and Cedars-Sinai Medical Center between October 2016 and October 2019. All patients and HDs provided written informed consent to be enrolled in this study according to the IRB protocol (IRB #14–000197, IRB#00000066). In total, we collected 62 plasma samples from two cohorts, including early-stage HCC cohort (n = 36): newly diagnosed, treatment-naive HCC patients (Barcelona Clinic Liver Cancer, BCLC stages: 0–A), and liver cirrhosis cohort (n = 26): patients with liver cirrhosis covering a diverse etiologies of hepatitis B virus (HBV), hepatitis C virus (HCV), alcoholic liver disease (ALD), and metabolic dysfunction–associated steatohepatitis (MASH). None of the enrolled patients was a part of any clinical trial, and patient allocation to each of the cohorts was defined by their disease states.

28.3.8.4 PCa Patients

All the participants in this study were enrolled in Cedars-Sinai Medical Center with IRB-approved protocols (IRB #Pro00042197, #Pro0033050, #Pro00051931) between January 2016 and December 2020. Samples were collected only after the participants provided written informed consent. We included two groups of PCa patients with a pathologically proven diagnosis of PCa for clinical validation of PCa EV digital scoring assay, which consisted of localized PCa (n = 20) and metastatic PCa (n = 20) patients. Patients with clinically localized PCa were defined as those who are on active surveillance or prior to radical prostatectomy (prostate intact) without any evidence of cancer metastasis either on next-generation imaging (PSMA PET/CT, ^{68}Ga or ^{18}F-based) or conventional radiographs (CT/MRI, ^{99}Tc-single-photon emission computed tomography [SPECT] bone scan). For patients classified by PSMA PET/CT, a metastatic lesion was defined by the presence of a PET-detectable lesion with an uptake greater than background and an SUV minimum of 10. This was consistent with previous studies utilizing PSMA PET as a screening modality for the detection of metastatic diseases. On the other hand, patients with metastatic PCa were defined as those who had metastatic PCa detected by conventional radiographs.

28.3.9 Blood Sample Processing

Peripheral venous blood samples were collected from cancer patients or HDs with written informed consent from each participant. Each blood sample was collected in a BD Vacutainer glass tube (BD Medical, Fisher Cat. #02–684–26) with acid citrate dextrose, which was then processed to obtain the final plasma by sequential centrifugation at 530 g for 10 min to remove blood cells, then 4600 g for 10 min, followed by 10,000 g for 10 min within four hours of collection. These plasma samples were then aliquoted and stored in −80 °C refrigerators. Prior to utilization, the stored plasma samples were thawed in a 37 °C water bath and then incubated with TCO-conjugated capturing reagents at room temperature for 30 min before being loaded onto the EV Click Chips/Beads for EV enrichment.

28.3.10 Statistical Analysis

The HCC EV Z score, which represents the likelihood estimate of the activation of ten HCC-specific mRNA markers, was computed from the RNA expression of these mRNA markers using a weighted Z-score method in R studio. After the mean centering of expression data across the samples, HCC EV Z scores were computed by the error-weighted mean of the expression values of the ten HCC-specific mRNA markers in each sample. Receiver operating characteristic (ROC) curve analysis was applied to evaluate the diagnostic performance for each parameter using MedCalc software. The expression percentiles of the 11 PCa-specific mRNA markers in primary PCa tissues (from The Cancer Genome Atlas, TCGA), PCa cell lines (from Cancer Cell Line Encyclopedia, CCLE), and immune cells (from Differentiation MAP dataset, DMAP) were compared using the Wilcoxon rank-sum test. The correlation of gene expression between PCa cell lines and PCa cell-derived EVs was computed as Pearson's correlation coefficient. Met Score, which represents the likelihood estimate of upregulation of 11 PCa-specific mRNA markers, was computed from the mRNA expression of these genes using a weighted Z-score method in R studio. After the median centering of expression data across all of the plasma samples, Met Scores were computed by the error-weighted mean of the expression values of the 11 PCa-specific mRNA markers in each sample. ROC curve analysis was used to evaluate the diagnostic performance of Met Score and serum prostate-specific antigen (PSA) concentration, which focused on the capability of each to distinguish localized PCa from metastatic PCa. Youden's index was used to identify the optimal cut-off of Met Score and PSA concentration. The area under ROC (AUROC) between Met Score and PSA concentration was compared using the paired DeLong's test. Pearson's correlation coefficients were computed to test the correlation between Met Score and continuous variables. A Mann–Whitney U test was used to compare the differences in Met Score between categorical variables. All statistical analyses were performed using R statistical software (version 4.1.0; R Foundation, Vienna, Austria), GraphPad Prism (version 9.2.0; GraphPad Software, Inc., CA, USA), and MATLAB (Mathworks, Natick, MA, USA) with two-sided tests and a significance level of 0.05.

28.4 Results

28.4.1 EV Digital Scoring Assay for Detecting Oncogenic Gene Alterations in PDAC and Ewing Sarcoma

We demonstrated the first EV digital scoring assay[34] (Figure 28.2) for noninvasive detection and quantification of oncogenic gene alterations, including KRAS mutations in PDAC and EWS/FLI-1 rearrangements in Ewing sarcoma. This EV digital scoring assay is carried out in a two-step workflow. Step 1: a lipid-based labeling approach was developed to label TCO onto EVs in the plasma samples using DSPE-PEG_{1000}-TCO. Tz-grafted microbeads (i.e., EV Click Beads) were employed to immobilize TCO-labeled EVs through click chemistry mediated enrichment where a bioorthogonal click reaction between TCO and Tz motifs led to effective EV enrichment. After rapid centrifugation at 300g for 2 min, the EVs captured onto the EV Click Beads were collected at the bottom of an Eppendorf tube. Step 2: the enriched EVs were lysed to release EV-derived mRNA, which was then subjected to RT-digital PCR for quantifying oncogenic gene alterations, including KRAS mutations in PDAC and EWS/FLI-1 rearrangements in Ewing sarcoma.

To test the clinical utility of this two-step workflow for quantifying KRAS mutations in PDAC patients, 35 plasma samples from 19 PDAC patients (some with multiple blood draws) and ten plasma samples from ten HDs were collected and analyzed. Each study required 1 mL of plasma per sample. Among the PDAC patients, 13 of them have paired FFPE tissue samples. We were able to compare the three KRAS mutation status (i.e., G12D, G12V, and G12R) in paired FFPE tissue and plasma samples from 13 patients. The results[34] showed that matched KRAS mutation subtypes were detected in EVs captured from plasma samples and FFPE tissues obtained at diagnosis in ten of the 13 PDAC patients. Among the remaining three PDAC patients with discordant KRAS mutation status, two patients had the KRAS mutations (one G12D, one G12R) detected in plasma samples but not detected in tissue samples, which might have resulted from the tissue heterogeneity. One patient had the KRAS G12D mutation detected in tissue sample but not in the plasma sample, presumably due to the copy number change of KRAS mutation over the period between tissue biopsy and blood draw. The boxplot (upper right, Figure 28.2A) showed that total copies of KRAS transcripts (mutated KRAS + wild-type KRAS) were significantly higher in the EVs captured from plasma samples of PDAC patients compared to the total copies of KRAS transcripts in EVs captured from plasma samples of HDs (P = 0.01). No mutated KRAS was detected in the plasma samples of HDs. We tracked the dynamic changes in a PDAC patient (PanC03) during treatment by correlating the total copies of KRAS transcripts and serum CA19–9 levels (U mL^{-1}) with CT images taken before and post-treatment. Serial plasma samples were collected from this patient 19 days prior to initiation of treatment and 84 and 210 days post-treatment. The corresponding serum CA19–9 levels and copies of KRAS transcripts for each timepoint were plotted in Figure 28.2B. At the final serial sample timepoint (210 days post-treatment), the patient's treatment response was

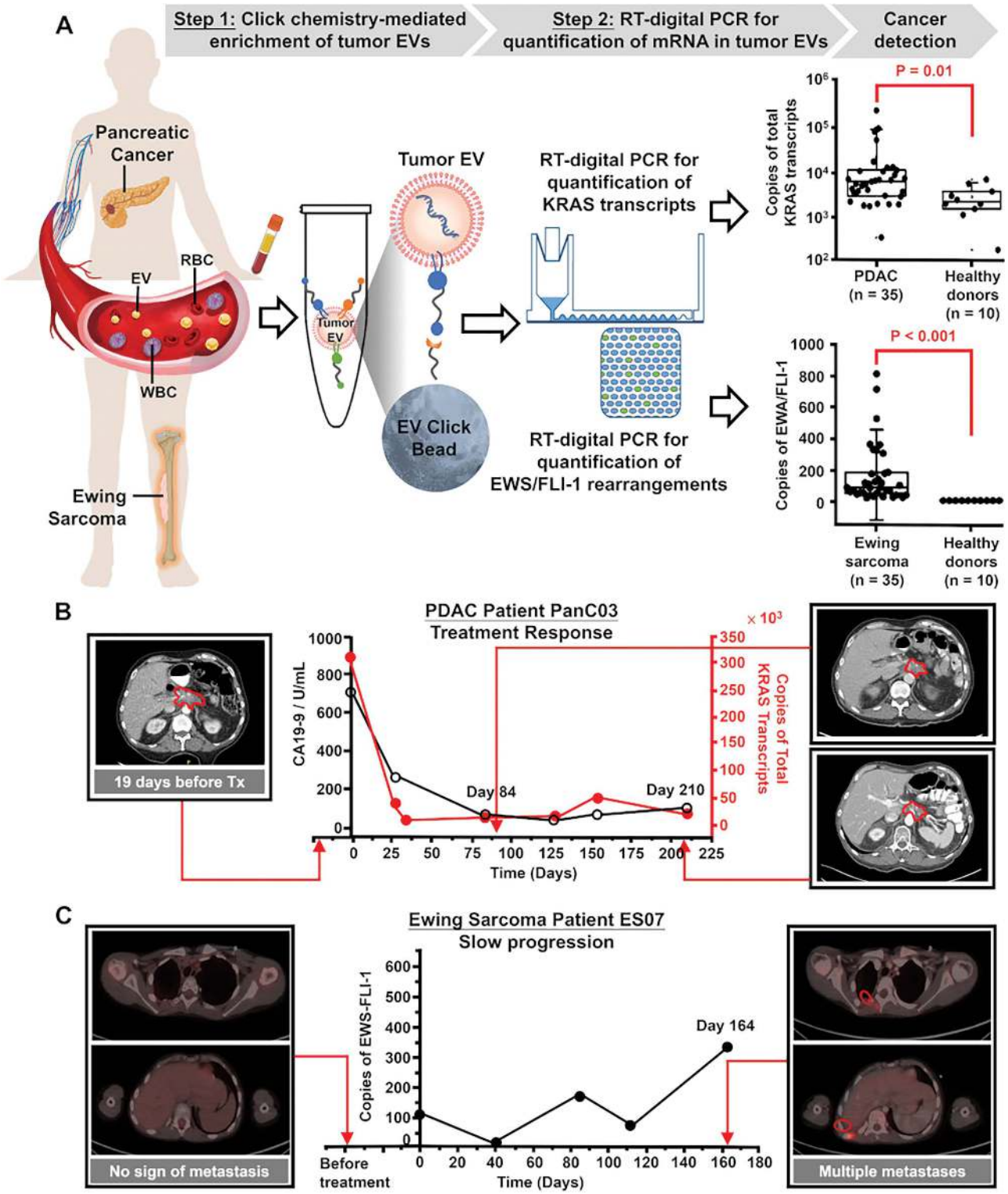


FIGURE 28.2 *Detection of oncogenic gene alterations by an EV digital scoring assays.* (A) Plasma samples from either pancreatic ductal adenocarcinoma (PDAC) or Ewing sarcoma patients were collected and subjected to a two-step workflow for the quantification of KRAS mutations and EWS/FLI-1 rearrangements, respectively. Step 1: after labeling EVs in the plasma samples with trans-cyclooctene (TCO) in the presence of DSPE-PEG_{1000} -TCO, the TCO-labeled EVs were immobilized onto EV Click Beads via click chemistry-mediated enrichment. Step 2: after extracting RNA from the enriched EVs, the EV-derived RNA was subjected to RT-digital PCR to quantify total KRAS transcripts (including KRAS mutations and wild type) in PDAC, and EWS/FLI-1 rearrangements in Ewing sarcoma. Boxplots showing the copy numbers of the oncogenic gene alterations can significantly differentiate PDAC patients from HDs, and Ewing sarcoma patients from HDs. (B) A plot of copies of total KRAS transcripts and CA19–9 levels (U mL^{-1}) versus times of the assay depicts the treatment response of a PDAC patient (PanC03). Computed tomography (CT) images taken 19 days before, 84 days, and 210 days after treatment. (C) A plot of copies of EWS/FLI-1 rearrangement versus times of the assay depicting the slow disease progression of a Ewing sarcoma patient (ES07). Positron emission tomography–computed tomography (PET/CT) images taken before and 164 days after treatment.

durable. The radiographic observations were consistent with the dynamic changes observed in both the total KRAS transcripts detected in EVs captured from serial plasma samples and the matched serum CA19–9 levels.

We then tested the clinical utility of this workflow for quantifying EWS/FLI-1 rearrangements in Ewing sarcoma. In total, 35 plasma samples from 28 patients with clinically confirmed Ewing sarcoma and ten plasma samples from ten HDs were analyzed. Among the 28 Ewing sarcoma patients, 12 patients have archived FFPE tissue samples. We therefore compared the most commonly observed EWS/FLI-1 rearrangement of paired tissue and plasma samples from 12 Ewing sarcoma patients, which showed that EWS/FLI-1 rearrangements were detected in the paired plasma and FFPE samples. The boxplot (lower right, Figure 28.2A) showed that EWS/FLI-1 rearrangements were detected in the EVs captured from all of the 35 Ewing sarcoma plasma samples but not detected in the EVs captured from HDs' plasma samples. Next, we aligned the copies of EWS/FLI-1 rearrangement with PET/CT images in order to monitor the dynamic changes in Ewing sarcoma in relation to treatment response. Serial plasma samples were collected from one Ewing sarcoma patient (ES-07) with confirmed EWS/FLI-1 rearrangement before treatment, and at 43, 85, 113, and 164 days post-treatment. The corresponding copies of EWS/FLI-1 rearrangement for each time point were depicted in Figure 28.2C. PET/CT images of the patient indicated the absence of tumor metastasis before treatment initiation. Subsequent to chemotherapy initiation, the patient exhibited a partial response after 43 days of treatment but relapsed at 164 days post-treatment with multiple metastases. These radiographic observations are consistent with the dynamic changes observed in the copies of EWS/FLI-1 rearrangement.

In this study, an EV digital scoring assay, which integrates the lipid-based EV labeling using DSPE-PEG 1000-TCO, click chemistry-mediated EV enrichment onto Click Beads, and absolute quantification of oncogenic gene alterations by RT-digital PCR, was developed to enable universal detection of EVs released by solid tumor of different origins, such as epithelial (e.g., PDAC) and mesenchymal (e.g., Ewing sarcoma) origin. Since lipid labeling bypasses the need for surface antigens, we adopted this two-step workflow[34] to quantitatively detect KRAS mutations in PDAC patients and EWS/FLI-1 rearrangements in Ewing sarcoma patients.

28.4.2 HCC EV Digital Scoring Assay for Detecting Early-Stage HCC from At-Risk Cirrhosis

We demonstrated HCC EV digital scoring assay[32] (Figure 28.3A) for noninvasive detection early-stage HCC from at-risk cirrhosis. HCC EV digital scoring assay is also carried out in a two-step workflow. Step 1: in conjunction with an antibody cocktail

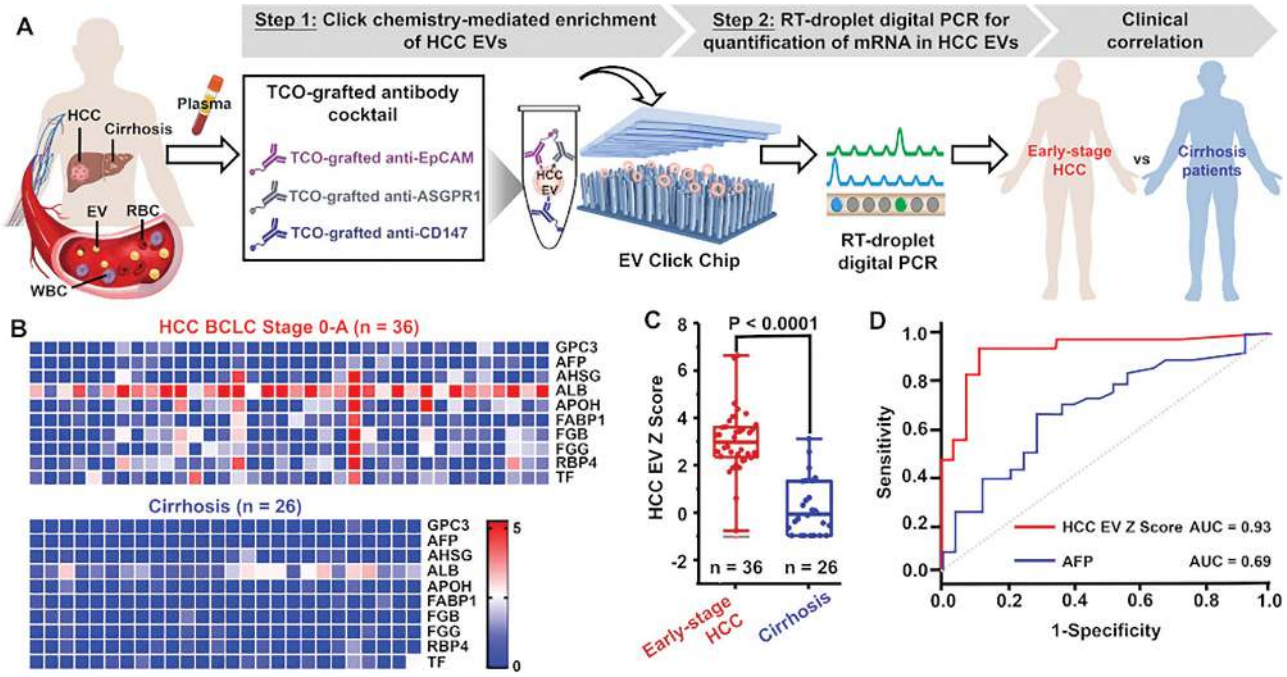


FIGURE 28.3 *HCC EV digital scoring assay for differentiating early-stage HCC from at-risk cirrhosis.* (A) A schematic illustration of a two-step workflow developed for conducting HCC EV digital scoring assay. Step 1: after labeling TCO onto HCC EVs in patient plasma samples with an antibody cocktail targeting three HCC-associated surface markers, including EpCAM, ASGPR1, and CD147, the TCO-labeled HCC EVs were immobilized onto the surfaces of EV Click Chips. By introducing 1,4-dithiothreitol (DTT) into EV Click Chip, disulfide cleavage-driven release of HCC EVs gave the enriched HCC EVs. Step 2: the enriched HCC EVs were subjected to downstream profiling of ten HCC-specific mRNA markers by RT-droplet digital PCR. The resultant HCC EV Z scores calculated from the ten HCC-specific mRNA signatures were used to distinguish HCC patients from at-risk cirrhotic patients. (B) Heatmaps displaying the expression of the ten HCC-specific mRNA markers in early-stage HCC and cirrhosis patients. (C) A boxplot showing the HCC EV Z scores detected from early-stage HCC ($n = 36$) and cirrhosis ($n = 26$) patients. (D) Comparison of the AUROC of HCC EV Z score (AUROC = 0.93, 95% CI, 0.86–1.00) and serum AFP level (AUROC = 0.69, 95% CI, 0.55–0.83) for differentiating early-stage HCC and at-risk liver cirrhosis.

targeting three HCC-associated surface markers,[49] including EpCAM, ASGPR1, and CD147, click chemistry-mediated HCC EV enrichment was developed to capture HCC EVs from plasma sample, followed by disulfide cleavage-driven[50] release of HCC EVs in the presence of a disulfide cleavage agent, DTT. It is important to note that the incorporation of densely packed silicon nanowires substrates dramatically increases the device surface area[48] interacting with HCC EVs. Furthermore, the overlaid PDMS chaotic mixer facilitates repeated physical contact[51] between SiNWS and the flow-through HCC EVs, further enhancing the performance of HCC EV capture. Step 2: mRNA payloads were extracted from the enriched HCC EVs for downstream molecular profiling of ten HCC-specific mRNA markers (i.e., *AFP*, *GPC3*, *ALB*, *APOH*, *FABP1*, *FGB*, *FGG*, *AHSG*, *RBP4*, and *TF*) using RT-droplet digital PCR. We computed HCC EV Z scores for each sample based on its ten HCC EV gene signatures using the weighted Z-score method,[52] which exhibited an outstanding performance to differentiate early-stage HCC from at-risk liver cirrhosis.

We collected 62 plasma samples from two cohorts, including early-stage HCC cohort (n = 36): newly diagnosed, treatment-naive HCC patients (BCLC stages: 0-A), and liver cirrhosis cohort (n = 26): patients with liver cirrhosis covering a diverse range of etiologies, including HBV, HCV, ALD, and MASH. A heatmap analysis revealed that early-stage HCC cohort had higher signals compared to the cirrhosis cohort (Figure 28.3B). The boxplot illustrated that the HCC EV Z score for the early-stage HCC cohort was significantly higher ($p < 0.0001$) than that for the cirrhosis cohort (Figure 28.3C). As patients with pre-existing chronic liver diseases are more prone to the development of HCC, a ROC curve analysis was performed to evaluate the capability of HCC EV Z score to distinguish early-stage HCC patients from cirrhosis patients. The AUROC was 0.93 (95% CI, 0.86–1.00; sensitivity = 94.4%, specificity = 88.5%, Figure 28.3D). This performance surpassed the current clinical standards for HCC detection, which is the use of serum AFP (AUROC = 0.69, 95% CI, 0.55–0.83; sensitivity = 66.7%, specificity = 72.0%). These results indicated that HCC EV Z score exhibited high sensitivity and specificity and could be a valuable tool for early HCC detection, which could increase the likelihood of providing curative-intent treatments that might eventually improve patient outcomes.

28.4.3 PCa EV Digital Scoring Assay for Differentiating Localized PCa from Metastatic PCa

We demonstrated PCa EV digital scoring assay[33] (Figure 28.4), which was capable of differentiating localized PCa from metastatic PCa. Similar to HCC EV digital scoring assay (Figure 28.3), PCa EV digital scoring assay is also carried out in a two-step workflow. Step 1: in conjunction with an antibody cocktail targeting two PCa-associated surface markers,[53–56] including EpCAM and PSMA, PCa EVs were enriched from plasma samples via click chemistry-mediated PCa EV capture, followed by disulfide cleavage-driven release of PCa EVs. Step 2: mRNA contents were extracted from the enriched PCa EVs for downstream molecular profiling of 11

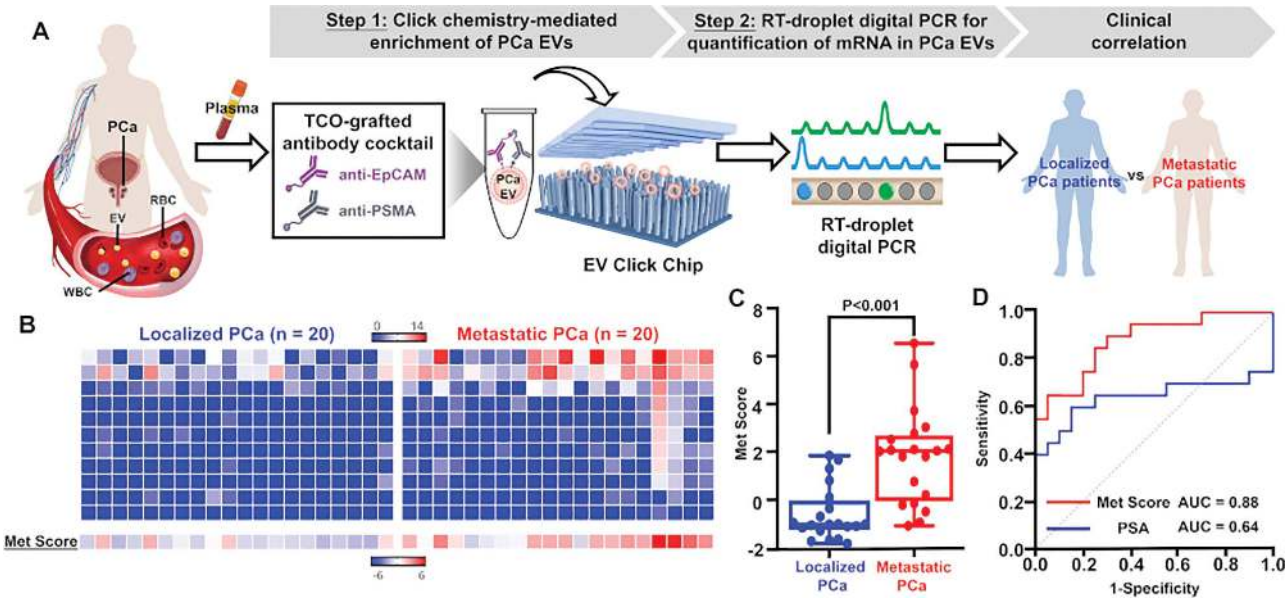


FIGURE 28.4 *PCa EV digital scoring assay for differentiating localized PCa from metastatic PCa.* (A) A schematic illustration of a two-step workflow developed for PCa EV digital scoring assay. Step 1: PCa EVs were enriched from plasma samples via click chemistry-mediated PCa EV capture, followed by disulfide cleavage-driven release of PCa EVs. Step 2: the enriched PCa EVs were subjected to downstream profiling of 11 PCa-specific gene markers by RT-droplet digital PCR. The resultant PCa-specific gene signatures were calculated by a weighted Z-score method to obtain a Met Score for individual patients. (B) Heatmaps displaying the expression of the 11 PCa-specific mRNA markers observed for localized PCa (n = 20) and metastatic PCa (n = 20) patients. (C) Boxplot showing the PCa EV Met Scores for localized and metastatic PCa patients. (D) AUROC for Met score (AUROC = 0.88, 95% CI, 0.78–0.98) and serum PSA concentration (AUROC = 0.64, 95% CI, 0.45–0.84) for differentiating localized PCa from metastatic PCa.

PCa-relevant genes (i.e., *ACP3*, *FOLH1*, *HOXB13*, *KLK2*, *KLK3*, *KLK4*, *MSMB*, *RLN1*, *SLC45A3*, *STEAP2*, and *TMPRSS2*) using RT-droplet digital PCR. The expression levels of the 11-gene panel were significantly higher in metastatic PCa cohort compared to the localized PCa cohort, as depicted in the heatmaps (Figure 28.4b). Based on the results obtained from RT-droplet digital PCR, a weighted Z-score method was employed to calculate a Met Score for individual patients in both cohorts. Summarizing the individual Met Scores, the boxplot (Figure 28.4c) showed that metastatic PCa patients had significantly higher Met Scores than localized PCa patients ($p < 0.0001$). To further assess the ability of the Met Score to differentiate between localized and metastatic PCa patients, a ROC curve analysis was performed, resulting in an AUROC of 0.88 (95% CI, 0.78–0.98, sensitivity = 85%, specificity = 75%). When compared to the AUROC of the current most widely used detection methodology for PCa (i.e., serum PSA concentration) (AUROC = 0.64, 95% CI, 0.45–0.84, sensitivity = 65%, specificity = 75%), the Met Score significantly outperformed the PSA concentration in its diagnostic power ($P = 0.03$). These findings suggested that the PCa EV digital scoring assay holds great promise as a supportive tool for early detection of disease progression in PCa patients, which could lead to improved patient care and treatment outcomes.

28.5 Discussions

Over the past five years, our joint research team from UCLA and Cedars Sinai Medical Center has successfully developed the three EV-based mRNA digital scoring assays, including EV digital scoring assay (Figure 28.2), HCC EV digital scoring assay (Figure 28.3), and PCa digital scoring assay (Figure 28.4). These assays operate on a general two-step process: an upstream click chemistry-mediated EV enrichment with either EV Click Chips or EV Click Beads, followed by digital counting of tumor-specific mRNA markers in the enriched tumor EVs, using either RT-digital PCR (QIAcuity, Qiagen) or RT-droplet digital PCR (QX200, Bio-Rad Laboratories, Inc.). Two types of TCO-labeling agents, DSPE-PEG1000-TCO and TCO-grafted tumor-specific antibodies, are used in conjunction with EV Click Chips or EV Click Beads. DSPE-PEG_{1000}-TCO is highly effective for labeling all EVs in plasma through a lipid membrane insertion mechanism, offering a universal solution for EV enrichment. Conversely, a cocktail of TCO-grafted antibodies that target various surface markers of solid tumors, such as HCC or PCa, allows for the selective enrichment of tumor EVs. This addresses the challenges of tumor heterogeneity and the typically low antigen density found on tumor EV surfaces. In short, click chemistry-mediated EV enrichment confers the first layer of sensitivity and specificity to these three EV-based mRNA digital scoring assays (Table 28.1).

Digital PCR technology leverages either droplet-based or nanoplate-based partitioning methods, which allow for absolute quantification of single nuclear acid molecules even amid high background noise. This technology is a cornerstone of our EV-based mRNA digital scoring assays, providing one of the most sensitive methods available for quantifying tumor-derived EVs and assessing systemic tumor burden. In the case of solid tumors marked by specific oncogenic alterations, such as KRAS mutations in pancreatic cancer and EWS-FLI1 rearrangements in Ewing sarcoma, we demonstrated the digital PCR-based counts of these oncogenic gene alterations can be a potential noninvasive diagnostic solution for (i) detecting solid tumor of either epithelial origin (e.g., pancreatic adenocarcinoma) or mesenchymal origin (e.g., Ewing sarcoma), and (ii) informing upfront treatment response or therapeutic effectiveness of salvage therapies, and relapse/recurrence monitoring for cancer patients in remission. For tumors like HCC and PCa, which lack specific oncogenic alterations, we demonstrated that digital PCR-based counting of HCC EV- and PCa EV-specific mRNA markers can be applied for noninvasive early detection of HCC from at-risk cirrhotic patients and distinguishing metastasis from localized PCa, respectively. These tumor EV-specific mRNA markers were bioinformatically selected from tumor tissue lineage-associated transcripts expressed in HCC and PCa cells but absent in blood cells and other tissues, enhancing the specificity of the diagnostics. Importantly, mRNA-based profiling of tumor EVs using this technology could be universally applied to virtually all solid tumors, independent of individual genetic alterations. Thus, digital PCR not only enhances the sensitivity and specificity of these assays but also expands the scope of actionable biomarkers in oncology, making it a transformative approach in cancer diagnostics.

TABLE 28.1

Technical Comparison Between EV Digital Scoring Assay and HCC/PCa EV Digital Scoring Assays

	EV Digital Scoring Assay	HCC/PCa EV Digital Scoring Assays
EV enrichment platforms	EV Click Beads - User friendly - Cost efficient	EV Click Chips - Large surface area resulting in higher EV enrichment performance
TCO-reagents for EV labeling	DSPE-PEG_{1000}-TCO - Universal solution for EV enrichment - Independent of EV surface protein markers	TCO-grafted antibodies - Specific for tumor EV enrichment - Compatible with multi-marker EV enrichment
RT-digital PCR platforms	RT-digital PCR - Absolute quantification of single mRNA molecules - No need for droplet transfer	RT-droplet digital PCR - Absolute quantification of single mRNA molecules - Require droplet transfer

28.6 Future Plan and Limitations

The successful development of the three EV-based mRNA digital scoring assays has proven to be swiftly applicable, facilitating the creation of a sensitive and biologically significant EV-based assay suitable for a wide array of noninvasive diagnostic purposes within oncology. These applications encompass early cancer detection, disease staging, and treatment monitoring. Moreover, the utilization of click chemistry-mediated EV enrichment, utilizing either EV Click Chips or EV Click Beads, holds promise in facilitating the isolation of tumor-specific EVs from diverse solid tumors through the targeting of corresponding cancer markers. This development lays the groundwork for the implementation of tumor-derived EV-based digital scoring assays, thereby extending the diagnostic reach within the field of oncology.

The proof-of-concept demonstration of the EV-based mRNA digital scoring assays shows promise but also has limitations. These assays were tested in relatively small preclinical patient cohorts and lacked sufficient longitudinal follow-up studies. For further development toward practical applications in cancer detection and treatment monitoring, more extensive validation, larger testing cohorts, and clearly defined longitudinal follow-up in clinical studies are essential. Our joint research team is actively working towards these objectives.

Acknowledgments

We acknowledge all the patients and healthy donors who participated in this study. We thank Lucy Ruoxi Shi from Harvard-Westlake School for her design and creation of the cartoon illustration for this study. This work is supported by National Institutes of Health (R01CA277530, R01CA255727, R01CA253651, R01CA253651–04S1, R01CA246304, U01EB-026421, K08CA259534, and R44CA288163).

REFERENCES

1. van Niel G, D'Angelo G, Raposo G. Shedding light on the cell biology of extracellular vesicles. *Nat Rev Mol Cell Biol*. 2018;19(4):213–28. Epub 20180118. doi: 10.1038/nrm.2017.125. PubMed PMID: 29339798.
2. Gao Z, Pang B, Li J, Gao N, Fan T, Li Y. Emerging role of exosomes in liquid biopsy for monitoring prostate cancer invasion and metastasis. *Front Cell Dev Biol*. 2021;9:679527. Epub 20210522. doi: 10.3389/fcell.2021.679527. PubMed PMID: 34017837; PMCID: PMC8129505.
3. Raposo G, Stoorvogel W. Extracellular vesicles: Exosomes, microvesicles, and friends. *J Cell Biol*. 2013;200(4):373–83. doi: 10.1083/jcb.201211138. PubMed PMID: 23420871; PMCID: 3575529.
4. Revenfeld AL, Baek R, Nielsen MH, Stensballe A, Varming K, Jorgensen M. Diagnostic and prognostic potential of extracellular vesicles in peripheral blood. *Clin Ther*. 2014;36(6):830–46. doi: 10.1016/j.clinthera.2014.05.008. PubMed PMID: 24952934.
5. Johnsen KB, Gudbergsson JM, Andresen TL, Simonsen JB. What is the blood concentration of extracellular vesicles? Implications for the use of extracellular vesicles as blood-borne biomarkers of cancer. *Biochim Biophys Acta Rev Cancer*. 2019;1871(1):109–16. Epub 20181207. doi: 10.1016/j.bbcan.2018.11.006. PubMed PMID: 30528756.
6. Tian Y, Gong M, Hu Y, Liu H, Zhang W, Zhang M, Hu X, Aubert D, Zhu S, Wu L, Yan X. Quality and efficiency assessment of six extracellular vesicle isolation methods by nano-flow cytometry. *J Extracell Vesicles*. 2020;9(1):1697028. Epub 20191129. doi: 10.1080/20013078.2019.1697028. PubMed PMID: 31839906; PMCID: PMC6896440.
7. Ferguson S, Weissleder R. Modeling EV kinetics for use in early cancer detection. *Adv Biosyst*. 2020;4(12):e1900305. Epub 20200512. doi: 10.1002/adbi.201900305. PubMed PMID: 32394646; PMCID: PMC7658022.
8. Kang M, Jordan V, Blenkiron C, Chamley LW. Biodistribution of extracellular vesicles following administration into animals: A systematic review. *J Extracell Vesicles*. 2021;10(8):e12085. Epub 20210624. doi: 10.1002/jev2.12085. PubMed PMID: 34194679; PMCID: PMC8224174.
9. Auber M, Svenningsen P. An estimate of extracellular vesicle secretion rates of human blood cells. *J Ex Bio*. 2022;1(6):e46. doi: https://doi.org/10.1002/jex2.46.
10. Jabalee J, Towle R, Garnis C. The role of extracellular vesicles in cancer: Cargo, function, and therapeutic implications. *Cells*. 2018;7(8). doi: 10.3390/cells7080093. PubMed PMID: 30071693; PMCID: 6115997.
11. King HW, Michael MZ, Gleadle JM. Hypoxic enhancement of exosome release by breast cancer cells. *BMC Cancer*. 2012;12:421. doi: 10.1186/1471-2407-12-421. PubMed PMID: 22998595; PMCID: 3488584.
12. Taylor DD, Gercel-Taylor C. The origin, function, and diagnostic potential of RNA within extracellular vesicles present in human biological fluids. *Front Genet*. 2013;4:142. doi: 10.3389/fgene.2013.00142. PubMed PMID: 23908664; PMCID: 3726994.
13. Tauro BJ, Greening DW, Mathias RA, Ji H, Mathivanan S, Scott AM, Simpson RJ. Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. *Methods*. 2012;56(2):293–304. doi: 10.1016/j.ymeth.2012.01.002. PubMed PMID: 22285593.
14. Kalra H, Adda CG, Liem M, Ang CS, Mechler A, Simpson RJ, Hulett MD, Mathivanan S. Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. *Proteomics*. 2013;13(22):3354–64. doi: 10.1002/pmic.201300282. PubMed PMID: 24115447.
15. Rekker K, Saare M, Roost AM, Kubo AL, Zarovni N, Chiesi A, Salumets A, Peters M. Comparison of serum exosome isolation methods for microRNA profiling. *Clin Biochem*. 2014;47(1–2):135–8. doi: 10.1016/j.clinbiochem.2013.10.020. PubMed PMID: 24183884.
16. Heinemann ML, Ilmer M, Silva LP, Hawke DH, Recio A, Vorontsova MA, Alt E, Vykoukal J. Benchtop isolation and characterization of functional exosomes by sequential filtration. *J Chromatogr A*. 2014;1371:125–35. doi: 10.1016/j.chroma.2014.10.026. PubMed PMID: WOS:000345810200014.

17. Heinemann ML, Vykoukal J. Sequential filtration: A gentle method for the isolation of functional extracellular vesicles. In: Kuo WP, Jia S, editors. *Extracellular Vesicles: Methods and Protocols*. Totowa: Humana Press Inc; 2017. p. 33–41.
18. Kanwar SS, Dunlay CJ, Simeone DM, Nagrath S. Microfluidic device (ExoChip) for on-chip isolation, quantification and characterization of circulating exosomes. *Lab Chip*. 2014;14(11):1891–900. doi: 10.1039/c4lc00136b. PubMed PMID: 24722878; PMCID: 4134440.
19. Zhao Z, Yang Y, Zeng Y, He M. A Microfluidic ExoSearch chip for multiplexed exosome detection towards blood-based ovarian cancer diagnosis. *Lab Chip*. 2016;16(3):489–96. doi: 10.1039/c5lc01117e. PubMed PMID: 26645590; PMCID: 4729647.
20. Wunsch BH, Smith JT, Gifford SM, Wang C, Brink M, Bruce RL, Austin RH, Stolovitzky G, Astier Y. Nanoscale lateral displacement arrays for the separation of exosomes and colloids down to 20 nm. *Nat Nanotechnol*. 2016;11(11):936–40. doi: 10.1038/nnano.2016.134. PubMed PMID: 27479757.
21. Yang F, Liao X, Tian Y, Li G. Exosome separation using microfluidic systems: Size-based, immunoaffinity-based and dynamic methodologies. *Biotechnol J*. 2017;12(4). doi: 10.1002/biot.201600699. PubMed PMID: 28166394.
22. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, LeBleu VS, Mittendorf EA, Weitz J, Rahbari N, Reissfelder C, Pilarsky C, Fraga MF, Piwnica-Worms D, Kalluri R. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature*. 2015;523(7559):177–82. doi: 10.1038/nature14581. PubMed PMID: 26106858; PMCID: 4825698.
23. Reategui E, van der Vos KE, Lai CP, Zeinali M, Atai NA, Aldikacti B, Floyd FP, Jr., A HK, Thapar V, Hochberg FH, Sequist LV, Nahed BV, Carter BS, Toner M, Balaj L, Ting DT, Breakefield XO, Stott SL. Engineered nanointerfaces for microfluidic isolation and molecular profiling of tumor-specific extracellular vesicles. *Nat Commun*. 2018;9(1):175. doi: 10.1038/s41467-017-02261-1. PubMed PMID: 29330365; PMCID: 5766611.
24. Park J, Park JS, Huang CH, Jo A, Cook K, Wang R, Lin HY, Van Deun J, Li H, Min J, Wang L, Yoon G, Carter BS, Balaj L, Choi GS, Castro CM, Weissleder R, Lee H. An integrated magneto-electrochemical device for the rapid profiling of tumour extracellular vesicles from blood plasma. *Nat Biomed Eng*. 2021;5(7):678–89. Epub 20210630. doi: 10.1038/s41551-021-00752-7. PubMed PMID: 34183802; PMCID: PMC8437135.
25. Kim K, Son T, Hong JS, Kwak TJ, Jeong MH, Weissleder R, Im H. Physisorption of affinity ligands facilitates extracellular vesicle detection with low non-specific binding to plasmonic gold substrates. *ACS Appl Mater Interfaces*. 2022. Epub 20220602. doi: 10.1021/acsami.2c07317. PubMed PMID: 35653580; PMCID: PMC9715849.
26. Zocco D, Zarovni N. Extraction and analysis of extracellular vesicle-associated miRNAs following antibody-based extracellular vesicle capture from plasma samples. *Methods Mol Biol*. 2017;1660:269–85. doi: 10.1007/978-1-4939-7253-1_22. PubMed PMID: 28828664.
27. Lane RE, Korbie D, Trau M, Hill MM. Purification protocols for extracellular vesicles. *Methods Mol Biol*. 2017;1660:111–30. doi: 10.1007/978-1-4939-7253-1_10. PubMed PMID: 28828652.
28. Oliveira JP, Prado AR, Keijok WJ, Antunes PWP, Yapuchura ER, Guimaraes MCC. Impact of conjugation strategies for targeting of antibodies in gold nanoparticles for ultrasensitive detection of 17beta-estradiol. *Sci Rep*. 2019;9(1):13859. Epub 20190925. doi: 10.1038/s41598-019-50424-5. PubMed PMID: 31554912; PMCID: PMC6761283.
29. Jung Y, Jeong JY, Chung BH. Recent advances in immobilization methods of antibodies on solid supports. *Analyst*. 2008;133(6):697–701. Epub 20080417. doi: 10.1039/b800014j. PubMed PMID: 18493668.
30. Luong JHT, Vashist SK. Chemistry of biotin-streptavidin and the growing concern of an emerging biotin interference in clinical immunoassays. *ACS Omega*. 2020;5(1):10–8. Epub 20191218. doi: 10.1021/acsomega.9b03013. PubMed PMID: 31956746; PMCID: PMC6963918.
31. Luong JHT, Male KB, Glennon JD. Biotin interference in immunoassays based on biotin-strept(avidin) chemistry: An emerging threat. *Biotechnol Adv*. 2019;37(5):634–41. Epub 20190311. doi: 10.1016/j.biotechadv.2019.03.007. PubMed PMID: 30872068.
32. Sun N, Lee YT, Zhang RY, Kao R, Teng PC, Yang Y, Yang P, Wang JJ, Smalley M, Chen PJ, Kim M, Chou SJ, Bao L, Wang J, Zhang X, Qi D, Palomique J, Nissen N, Han SB, Sadeghi S, Finn RS, Saab S, Busuttil RW, Markovic D, Elashoff D, Yu HH, Li H, Heaney AP, Posadas E, You S, Yang JD, Pei R, Agopian VG, Tseng HR, Zhu Y. Purification of HCC-specific extracellular vesicles on nanosubstrates for early HCC detection by digital scoring. *Nat Commun*. 2020;11(1):4489. Epub 20200909. doi: 10.1038/s41467-020-18311-0. PubMed PMID: 32895384; PMCID: PMC7477161.
33. Wang JJ, Sun N, Lee YT, Kim M, Vagner T, Rohena-Rivera K, Wang Z, Chen Z, Zhang RY, Lee J, Zhang C, Tang H, Widjaja J, Zhang TX, Qi D, Teng PC, Jan YJ, Hou KC, Hamann C, Sandler HM, Daskivich TJ, Luthringer DJ, Bhowmick NA, Pei R, You S, Vizio DD, Tseng HR, Chen JF, Zhu Y, Posadas EM. Prostate cancer extracellular vesicle digital scoring assay—a rapid noninvasive approach for quantification of disease-relevant mRNAs. *Nano Today*. 2023;48:101746. Epub 20230103. doi: 10.1016/j.nantod.2022.101746. PubMed PMID: 36711067; PMCID: PMC9879227.
34. Sun N, Tran BV, Peng Z, Wang J, Zhang C, Yang P, Zhang TX, Widjaja J, Zhang RY, Xia W, Keir A, She JW, Yu HH, Shyue JJ, Zhu H, Agopian VG, Pei R, Tomlinson JS, Toretsky JA, Jonas SJ, Federman N, Lu S, Tseng HR, Zhu Y. Coupling lipid labeling and click chemistry enables isolation of extracellular vesicles for noninvasive detection of oncogenic gene alterations. *Adv Sci (Weinh)*. 2022;9(14):e2105853. Epub 20220430. doi: 10.1002/advs.202105853. PubMed PMID: 35486030; PMCID: PMC9108594.
35. Kolb HC, Finn MG, Sharpless KB. Click chemistry: Diverse chemical function from a few good reactions. *Angew Chem Int Ed*. 2001;40(11):2004–21. doi: 10.1002/1521-3773(20010601)40:11<2004::aid-anie2004> 3.3.co;2-x. PubMed PMID: 11433435.
36. Blackman ML, Royzen M, Fox JM. Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J Am Chem Soc*. 2008;130(41):13518–9. Epub 20080918. doi: 10.1021/ja8053805. PubMed PMID: 18798613; PMCID: PMC2653060.

37. Karver MR, Weissleder R, Hilderbrand SA. Synthesis and evaluation of a series of 1,2,4,5-tetrazines for bioorthogonal conjugation. *Bioconjug Chem*. 2011;22(11):2263–70. Epub 20110929. doi: 10.1021/bc200295y. PubMed PMID: 21950520; PMCID: PMC3254257.
38. Blackman ML, Royzen M, Fox JM. Tetrazine ligation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J Am Chem Soc*. 2008;130(41):13518–9. doi: 10.1021/ja8053805.
39. Karver MR, Weissleder R, Hilderbrand SA. Synthesis and evaluation of a series of 1,2,4,5-tetrazines for bioorthogonal conjugation. *Bioconjug Chem*. 2011;22(11):2263–70. doi: 10.1021/bc200295y.
40. Dong J, Zhang RY, Sun N, Smalley M, Wu Z, Zhou A, Chou SJ, Jan YJ, Yang P, Bao L, Qi D, Tang X, Tseng P, Hua Y, Xu D, Kao R, Meng M, Zheng X, Liu Y, Vagner T, Chai X, Zhou D, Li M, Chiou SH, Zheng G, Di Vizio D, Agopian VG, Posadas E, Jonas SJ, Ju SP, Weiss PS, Zhao M, Tseng HR, Zhu Y. Bio-inspired nanoVilli chips for enhanced capture of tumor-derived extracellular vesicles: Toward non-invasive detection of gene alterations in non-small cell lung cancer. *ACS Appl Mater Interfaces*. 2019;11(15):13973–83. Epub 20190402. doi: 10.1021/acsami.9b01406. PubMed PMID: 30892008; PMCID: PMC6545291.
41. Dong J, Zhang RY, Sun N, Hu J, Smalley MD, Zhou A, Yue H, Rothermich W, Chen M, Chen J, Ye J, Teng PC, Qi D, Toretsky JA, Tomlinson JS, Li M, Weiss PS, Jonas SJ, Federman N, Wu L, Zhao M, Tseng HR, Zhu Y. Coupling nanostructured microchips with covalent chemistry enables purification of sarcoma-derived extracellular vesicles for downstream functional studies. *Adv Funct Mater*. 2020;30(49). Epub 20210706. doi: 10.1002/adfm.202003237. PubMed PMID: 34220409; PMCID: PMC8248519.
42. Sun N, Zhang C, Lee YT, Tran BV, Wang J, Kim H, Lee J, Zhang RY, Wang JJ, Hu J, Zhang Z, Alsudaney MS, Hou KC, Tang H, Zhang TX, Liang IY, Zhou Z, Chen M, Yeh AH, Li W, Zhou XJ, Chang HR, Han SB, Sadeghi S, Finn RS, Saab S, Busuttil RW, Noureddin M, Ayoub WS, Kuo A, Sundaram V, Al-Ghaieb B, Palomique J, Kosari K, Kim IK, Todo T, Nissen NN, Tomasi ML, You S, Posadas EM, Wu JX, Wadehra M, Sim MS, Li Y, Wang HL, French SW, Lu SC, Wu L, Pei R, Liang L, Yang JD, Agopian VG, Tseng HR, Zhu Y. HCC EV ECG score: An extracellular vesicle-based protein assay for detection of early-stage hepatocellular carcinoma. *Hepatology*. 2022:774–88. Epub 20220801. doi: 10.1002/hep.32692. PubMed PMID: 35908246.
43. Mohrmann L, Huang HJ, Hong DS, Tsimberidou AM, Fu S, Piha-Paul SA, Subbiah V, Karp DD, Naing A, Krug A, Enderle D, Priewasser T, Noerholm M, Eitan E, Coticchia C, Stoll G, Jordan LM, Eng C, Kopetz ES, Skog J, Meric-Bernstam F, Janku F. Liquid biopsies using plasma exosomal nucleic acids and plasma cell-free DNA compared with clinical outcomes of patients with advanced cancers. *Clin Cancer Res*. 2018;24(1):181–8. doi: 10.1158/1078-0432.CCR-17-2007. PubMed PMID: 29051321; PMCID: 5754253.
44. Castellanos-Rizaldos E, Grimm DG, Tadigotla V, Hurley J, Healy J, Neal PL, Sher M, Venkatesan R, Karlovich C, Raponi M, Krug A, Noerholm M, Tannous J, Tannous BA, Raez LE, Skog JK. Exosome-based detection of EGFR T790M in plasma from non-small cell lung cancer patients. *Clin Cancer Res*. 2018;24(12):2944–50. doi: 10.1158/1078–0432.CCR-17-3369. PubMed PMID: 29535126.
45. Huang X, Yuan T, Tschannen M, Sun Z, Jacob H, Du M, Liang M, Dittmar RL, Liu Y, Liang M, Kohli M, Thibodeau SN, Boardman L, Wang L. Characterization of human plasma-derived exosomal RNAs by deep sequencing. *BMC Genom*. 2013;14:319. Epub 20130510. doi: 10.1186/1471-2164-14-319. PubMed PMID: 23663360; PMCID: PMC3653748.
46. Wan Y, Cheng G, Liu X, Hao SJ, Nisic M, Zhu CD, Xia YQ, Li WQ, Wang ZG, Zhang WL, Rice SJ, Sebastian A, Albert I, Belani CP, Zheng SY. Rapid magnetic isolation of extracellular vesicles via lipid-based nanoprobes. *Nat Biomed Eng*. 2017;1. Epub 20170410. doi: 10.1038/s41551-017-0058. PubMed PMID: 28966872; PMCID: PMC5618714.
47. Thery C, Witwer KW, Aikawa E, Alcaraz MJ, Anderson JD, Andriantsitohaina R, Antoniou A, Arab T, Archer F, Atkin-Smith GK, Ayre DC, Bach JM, Bachurski D, Baharvand H, Balaj L, Baldacchino S, Bauer NN, Baxter AA, Bebawy M, Beckham C, Bedina Zavec A, Benmoussa A, Berardi AC, Bergese P, Bielska E, Blenkiron C, Bobis-Wozowicz S, Boilard E, Boireau W, Bongiovanni A, Borras FE, Bosch S, Boulanger CM, Breakefield X, Breglio AM, Brennan MA, Brigstock DR, Brisson A, Broekman ML, Bromberg JF, Bryl-Gorecka P, Buch S, Buck AH, Burger D, Busatto S, Buschmann D, Bussolati B, Buzas EI, Byrd JB, Camussi G, Carter DR, Caruso S, Chamley LW, Chang YT, Chen C, Chen S, Cheng L, Chin AR, Clayton A, Clerici SP, Cocks A, Cocucci E, Coffey RJ, Cordeiro-da-Silva A, Couch Y, Coumans FA, Coyle B, Crescitelli R, Criado MF, D'Souza-Schorey C, Das S, Datta Chaudhuri A, de Candia P, De Santana EF, De Wever O, Del Portillo HA, Demaret T, Deville S, Devitt A, Dhondt B, Di Vizio D, Dieterich LC, Dolo V, Dominguez Rubio AP, Dominici M, Dourado MR, Driedonks TA, Duarte FV, Duncan HM, Eichenberger RM, Ekstrom K, El Andaloussi S, Elie-Caille C, Erdbrugger U, Falcon-Perez JM, Fatima F, Fish JE, Flores-Bellver M, Forsonits A, Frelet-Barrand A, Fricke F, Fuhrmann G, Gabrielsson S, Gamez-Valero A, Gardiner C, Gartner K, Gaudin R, Gho YS, Giebel B, Gilbert C, Gimona M, Giusti I, Goberdhan DC, Gorgens A, Gorski SM, Greening DW, Gross JC, Gualerzi A, Gupta GN, Gustafson D, Handberg A, Haraszti RA, Harrison P, Hegyesi H, Hendrix A, Hill AF, Hochberg FH, Hoffmann KF, Holder B, Holthofer H, Hosseinkhani B, Hu G, Huang Y, Huber V, Hunt S, Ibrahim AG, Ikezu T, Inal JM, Isin M, Ivanova A, Jackson HK, Jacobsen S, Jay SM, Jayachandran M, Jenster G, Jiang L, Johnson SM, Jones JC, Jong A, Jovanovic-Talisman T, Jung S, Kalluri R, Kano SI, Kaur S, Kawamura Y, Keller ET, Khamari D, Khomyakova E, Khvorova A, Kierulf P, Kim KP, Kislinger T, Klingeborn M, Klinke DJ, 2nd, Kornek M, Kosanovic MM, Kovacs AF, Kramer-Albers EM, Krasemann S, Krause M, Kurochkin IV, Kusuma GD, Kuypers S, Laitinen S, Langevin SM, Languino LR, Lannigan J, Lasser C, Laurent LC, Lavieu G, Lazaro-Ibanez E, Le Lay S, Lee MS, Lee YXF, Lemos DS, Lenassi M, Leszczynska A, Li IT, Liao K, Libregts SF, Ligeti E, Lim R, Lim SK, Line A, Linnemannstons K, Llorente A, Lombard CA, Lorenowicz MJ, Lorincz AM, Lotvall J, Lovett J, Lowry MC, Loyer X, Lu Q, Lukomska B, Lunavat TR, Maas SL, Malhi H, Marcilla A, Mariani J, Mariscal J, Martens-Uzunova ES,

Martin-Jaular L, Martinez MC, Martins VR, Mathieu M, Mathivanan S, Maugeri M, McGinnis LK, McVey MJ, Meckes DG, Jr., Meehan KL, Mertens I, Minciacchi VR, Moller A, Moller Jorgensen M, Morales-Kastresana A, Morhayim J, Mullier F, Muraca M, Musante L, Mussack V, Muth DC, Myburgh KH, Najrana T, Nawaz M, Nazarenko I, Nejsum P, Neri C, Neri T, Nieuwland R, Nimrichter L, Nolan JP, Nolte-'t Hoen EN, Noren Hooten N, O'Driscoll L, O'Grady T, O'Loghlen A, Ochiya T, Olivier M, Ortiz A, Ortiz LA, Osteikoetxea X, Ostergaard O, Ostrowski M, Park J, Pegtel DM, Peinado H, Perut F, Pfaffl MW, Phinney DG, Pieters BC, Pink RC, Pisetsky DS, Pogge von Strandmann E, Polakovicova I, Poon IK, Powell BH, Prada I, Pulliam L, Quesenberry P, Radeghieri A, Raffai RL, Raimondo S, Rak J, Ramirez MI, Raposo G, Rayyan MS, Regev-Rudzki N, Ricklefs FL, Robbins PD, Roberts DD, Rodrigues SC, Rohde E, Rome S, Rouschop KM, Rughetti A, Russell AE, Saa P, Sahoo S, Salas-Huenuleo E, Sanchez C, Saugstad JA, Saul MJ, Schiffelers RM, Schneider R, Schoyen TH, Scott A, Shahaj E, Sharma S, Shatnyeva O, Shekari F, Shelke GV, Shetty AK, Shiba K, Siljander PR, Silva AM, Skowronek A, Snyder OL, 2nd, Soares RP, Sodar BW, Soekmadji C, Sotillo J, Stahl PD, Stoorvogel W, Stott SL, Strasser EF, Swift S, Tahara H, Tewari M, Timms K, Tiwari S, Tixeira R, Tkach M, Toh WS, Tomasini R, Torrecilhas AC, Tosar JP, Toxavidis V, Urbanelli L, Vader P, van Balkom BW, van der Grein SG, Van Deun J, van Herwijnen MJ, Van Keuren-Jensen K, van Niel G, van Royen ME, van Wijnen AJ, Vasconcelos MH, Vechetti IJ, Jr., Veit TD, Vella LJ, Velot E, Verweij FJ, Vestad B, Vinas JL, Visnovitz T, Vukman KV, Wahlgren J, Watson DC, Wauben MH, Weaver A, Webber JP, Weber V, Wehman AM, Weiss DJ, Welsh JA, Wendt S, Wheelock AM, Wiener Z, Witte L, Wolfram J, Xagorari A, Xander P, Xu J, Yan X, Yanez-Mo M, Yin H, Yuana Y, Zappulli V, Zarubova J, Zekas V, Zhang JY, Zhao Z, Zheng L, Zheutlin AR, Zickler AM, Zimmermann P, Zivkovic AM, Zocco D, Zuba-Surma EK. Minimal information for studies of extracellular vesicles 2018 (MISEV2018): A position statement of the international society for extracellular vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles*. 2018;7(1):1535750. Epub 20190115. doi: 10.1080/20013078.2018.1535750. PubMed PMID: 30637094; PMCID: PMC6322352.

48. Wang S, Wang H, Jiao J, Chen KJ, Owens GE, Kamei K, Sun J, Sherman DJ, Behrenbruch CP, Wu H, Tseng HR. Three-dimensional nanostructured substrates toward efficient capture of circulating tumor cells. *Angew Chem Int Ed Engl*. 2009;48(47):8970–3. doi: 10.1002/anie.200901668. PubMed PMID: 19847834; PMCID: PMC2878179.
49. Court CM, Hou S, Winograd P, Segel NH, Li QW, Zhu Y, Sadeghi S, Finn RS, Ganapathy E, Song M, French SW, Naini BV, Sho S, Kaldas FM, Busuttil RW, Tomlinson JS, Tseng HR, Agopian VG. A novel multimarker assay for the phenotypic profiling of circulating tumor cells in hepatocellular carcinoma. *Liver Transpl*. 2018;24(7):946–60. Epub 20180407. doi: 10.1002/lt.25062. PubMed PMID: 29624843; PMCID: PMC6097911.
50. Witt D. Recent developments in disulfide bond formation. *Synthesis*. 2008;2008(16):2491–509. Epub 20080724. doi: 10.1055/s-2008-1067188.
51. Wang S, Liu K, Liu J, Yu ZT, Xu X, Zhao L, Lee T, Lee EK, Reiss J, Lee YK, Chung LW, Huang J, Rettig M, Seligson D, Duraiswamy KN, Shen CK, Tseng HR. Highly efficient capture of circulating tumor cells by using nanostructured silicon substrates with integrated chaotic micromixers. *Angew Chem Int Ed Engl*. 2011;50(13):3084–8. Epub 20110304. doi: 10.1002/anie.201005853. PubMed PMID: 21374764; PMCID: PMC3085082.
52. Levine DM, Haynor DR, Castle JC, Stepaniants SB, Pellegrini M, Mao M, Johnson JM. Pathway and gene-set activation measurement from mRNA expression data: The tissue distribution of human pathways. *Genome Biol*. 2006;7(10):R93. Epub 20061017. doi: 10.1186/gb-2006-7-10-r93. PubMed PMID: 17044931; PMCID: PMC1794557.
53. Went PT, Lugli A, Meier S, Bundi M, Mirlacher M, Sauter G, Dirnhofer S. Frequent EpCam protein expression in human carcinomas. *Hum Pathol*. 2004;35(1):122–8. doi: 10.1016/j.humpath.2003.08.026. PubMed PMID: 14745734.
54. Spizzo G, Fong D, Wurm M, Ensinger C, Obrist P, Hofer C, Mazzoleni G, Gastl G, Went P. EpCAM expression in primary tumour tissues and metastases: An immunohistochemical analysis. *J Clin Pathol*. 2011;64(5):415–20. Epub 20110317. doi: 10.1136/jcp.2011.090274. PubMed PMID: 21415054; PMCID: PMC3088404.
55. Ross JS, Sheehan CE, Fisher HA, Kaufman RP, Jr., Kaur P, Gray K, Webb I, Gray GS, Mosher R, Kallakury BV. Correlation of primary tumor prostate-specific membrane antigen expression with disease recurrence in prostate cancer. *Clin Cancer Res*. 2003;9(17):6357–62. PubMed PMID: 14695135.
56. Bostwick DG, Pacelli A, Blute M, Roche P, Murphy GP. Prostate specific membrane antigen expression in prostatic intraepithelial neoplasia and adenocarcinoma: A study of 184 cases. *Cancer*. 1998;82(11):2256–61. doi: 10.1002/(sici)1097-0142(19980601)82:11<2256::aid-cncr22>3.0.co;2-s. PubMed PMID: 9610707.

29

Identifying Methods to Deliver Mutant p53-Based Diagnostics and Therapeutics to Target the Previously Undruggable

Gu Xiao, Outi Keinänen, Florencia Madorsky Rowdo, Samantha Delaney, Katherine W. Harmon, Brian M. Zeglis, and Jill Bargonetti

29.1 Introduction

Cancer is a leading cause of premature death worldwide [1]. The development of cancer results from cellular mutations in genes that disrupt the normal function of proteins that regulate cell survival and proliferation, with missense mutations in the p53 gene being a key cancer driver mutation [2–4]. In normal cells wild-type p53 protein functions as a tumor suppressor transcription factor. However, in a majority of cancers, including triple negative breast cancer (TNBC), missense mutant p53 (mtp53) has lost DNA binding and transactivation activity while becoming a stable protein that is considered undruggable [5–7]. Wild-type p53 (wtp53) protein levels are kept low via constitutive proteasomal degradation via the E3 ubiquitin ligase MDM2, but tumor cell missense mtp53 is stable and highly identifiable [8, 9]. The stable mtp53 not only impairs antitumor activity, but also acquires oncogenic properties termed "gain of function" (GOF), that include promoting genome instability and metastasis [7, 10–13]. Stabilized mtp53 is found in tumors of cancer patients and can be used as a biomarker [14]. One targeting approach for mtp53 has focused on reactivating the protein to gain back normal wtp53 transcriptional activity [15, 16]. Other methods have tried targeting mt53 for degradation through the use Hsp90 inhibitors [17] or blocking activated cholesterol biosynthesis pathways using statins [18]. The nuclear localization of mtp53, and the lack of understood oncogenic activity of mtp53 protein, has caused a lack of interest in targeting mtp53 as an oncogene [8]. We focus on the nuclear localization of mtp53 and the oncogenic activity of the protein as avenues to deliver theranostics.

This chapter focuses on the high stability of mtp53 as a method for diagnosing and targeting cancers with missense p53 mutations and uses TNBC cells to demonstrate the proof of principle. The mtp53 protein focus centers on the fact that the molecule contains an intact tetramerization domain (TD). The p53 protein consists of five functional domains: a transactivation domain (TAD residues 1–61), a proline-rich domain (Pro-rich residues 63–97), an often-mutated sequence-specific DNA binding domain (DBD residues 98–293), an oligomerization domain that confers the tetrameric structure necessary for wtp53 function (TD, residues 326–356), and a C-terminal domain (CTD, residues 363–393) that interacts with DNA in a sequence-nonspecific manner [19]. The p53 TD consists of a β-strand (Glu326-Arg333), a tight turn (Gly334), and an α-helix (Arg335-Gly356) [19, 20]. The most common type of *TP53* mutations are missense, with a high occurrence of six DBD p53 "hotspot" missense mutations (R175, G245, R248, R249, R273, and R282) [7]. Wild-type p53 binds to DNA site-specifically as a tetramer regulated by the p53 C-terminal region, and this region contains the TD that is present in large quantities in cancers with stable mtp53 [21, 22]. Stable mtp53 can form tetramers, but the tetramer formation is not required for the GOF associated chromatin interactions [23]. Moreover, mtp53 R273H, R280K, and L194F proteins are tightly associated with chromatin and exist as highly stable nuclear proteins [24, 25]. Therefore, we tested if delivery of labeled-p53 TD peptides, which could enter cancer cells and interact with the stable mtp53, would enable TNBC cell detection in both a tissue culture and a xenograft mouse model. Peptide-based cancer treatment can have high specificity, membrane permeability, and low toxicity [26], and we have shown that the p53 TD imaging agent for cancers expressing stable mtp53 has promise [27]. In fact, some potential uses are as a fluorescent molecular p53 TD probe to provide opportunities for diagnosis, imaging guided surgery, or as backbones for the delivery of therapy choices directly to tumor cell nuclei.

We designed a nucleus-penetrating, fluorophore-labeled p53-tetramerization-domain derived peptide, called Cy5p53-Tet, for the detection of TNBC with stable chromatin-associated mtp53 [27]. Cy5p53Tet is a peptide with a Cy5 fluorophore and an HIV-1 Tat sequence, Arg-Lys-Lys-Arg-Arg-Gln-Arg-Arg, that enables nuclear penetration [28]. In addition to nuclear penetration, HIV-1 Tat can also deliver a p53 TD interaction [29]. Therefore, the Cy5p53Tet design has both a p53 TD for interaction with endogenous stable mtp53 and HIV-1 Tat sequences for nuclear localization and increased p53 targeting.

Patient-derived cancer cell models are representative of the original tumor morphology, polarization, and genetic heterogeneity of parental tumors [30]. Patient-derived tumor organoids (PDTOs) are cells obtained from a patient's tumor that are grown in 3D cellular models in culture *ex vivo*. PTDOs recapitulate the original tissue architecture, mutational signatures, and tumor-clonal heterogeneity [31, 32]. In light of this, we tested Cy5p53Tet in both TNBC cell culture and PDTOs. We observed Cy5p53Tet penetrates cells and binds to high levels

DOI: 10.1201/9781003449942-33

of mtp53 for excellent detection [27, 33]. We also demonstrated that Cy5p53Tet enters into mtp53-expressing TNBC cells *in vivo* in mouse xenografts models [27, 33]. To achieve the goal of delivery of a wide variety of mtp53-based diagnostic and therapeutic payloads, we optimized Cy5p53tet by utilizing stapled peptides SPCy5p53Tet to increase peptide stability, and we describe this new design herein. We decreased the conformational flexibility of linear peptides and thus increased the Cy5p53Tet stability and reduced the peptide proteolysis. The novel stapled peptide targeting the mtp53 TD interface, SPCy5p53Tet, provides a new concept in theranostic cancer targeting design.

29.2 Materials

29.2.1 Peptides

Cy5p53Tet was purchased from JPT peptide (Germany) at a purity > 95%. The peptide Cy5p53Tet35 is 35 amino acids long with an N-terminal Cy5 fluorophore conjugation: H-CysCy5-RKKRRQRRGEYFTLQIRGRERFEMFRELNEALELK-OH. The peptide Cy5p53Tet37 is 37 amino acids long with an N-terminal Cy5 fluorophore conjugation: H-CysCy5-GRKKRRQRRRGEYFTLQIRGRERFEMFRELNEALELK-OH. The random scrambled peptide Cy5Scramble is 27 amino acids long designed by JPT peptide with an N-terminal Cy5 fluorophore conjugation: H-Cys(Cy5)-LGEFELQKRQKRERLNGRLRERRAEFR-OH.

Stapled peptides were purchased from CPC Scientific (San Jose CA) at a purity > 95%. SPCy5p53TetA has an N-terminal Cy5 fluorophore conjugation: betaA-RERFE-S5-FRE-S5-NEALELKD (stapled on S5 and S5; S5 stands for S-2–(4′-pentenyl) alanine) SpCy5p53TetB has an N-terminal Cy5 fluorophore conjugation: beta A-GRKKRRQRRR-betaA-RERFE-S5-FRE-S5-NEALELKD (stapled on S5 and S5; S5 stands for S-2–(4′-pentenyl) alanine).

29.2.2 Antibodies

The following antibodies were used: anti-p53 antibody (Proteintech Cat# 10442–1-AP), anti-MDM2 antibody (R&D System Cat# AF1244), anti-estrogen receptor α (Cell Signaling Technology CAT# #13258), and anti-β-actin antibody (Sigma-Aldrich CAT# A2066).

29.3 Methods

29.3.1 Peptide Synthesis

Cy5p53Tet was performed using state-of-the-art solid phase peptide synthesis on an automated peptide synthesizer (ABI 433A) using Fmoc-chemistry. A double coupling protocol was used applying a preloaded Lys-PHB resin, Fmoc-amino acids, HATU/Oxyma as coupling reagents, and piperidine for Fmoc protecting group removal. The peptide resin cleavage was performed using trifluoroacetic acid with 3% EDT and 3% water for 3 h. The crude peptide was purified using reverse phase chromatography by means of a linear acetonitrile water gradient and a C18 HPLC column. After freeze drying, Cy5 labeling peptide was performed by dissolving the peptide in buffer and acetonitrile. Cy5-maleimide was then added and the mixture was agitated for one hour. The final peptide was purified using the aforementioned method and freeze dried.

29.3.2 Stapled Peptides

Rink Amide MBHA resin was swelled in DMF. The Fmoc-amino acid and coupling reagent were added to the resin. The reaction was carried out under a nitrogen atmosphere at room temperature. After the reaction was complete, the suspension was filtered, and the resin was washed three times with DMF. 20% piperidine in DMF was added into the reaction vessel to remove the Fmoc group. The suspension was kept at room temperature for 0.5 h while a stream of nitrogen was bubbled through it. The suspension was filtered, and the resin was washed five times with DMF. After coupling all the amino acid to obtain full protected peptidyl resin, a full protected peptidyl-Resin was stapled by Grubbs first on S5 and S5. The fully protected peptidyl resin was cleaved by TFA solution for 3 h, filtered and precipitated by cool ether, centrifuged, and washed with ether three times. The product was dried under vacuum to obtain crude peptide. The crude product was dissolved in 30% acetonitrile in H_2O, and loaded onto the C18 column (1 inch, Luna C18, 5 μm). The separation was carried out with a gradient from 25% to 85% Mobile Phase B (0.1% TFA in [80% ACN + 20% H_2O]) over 60 minutes yielding 85% pure AG fractions. The AG fractions were loaded onto the C18 column (1 inch, Luna C18, 5 μm) and eluted with a gradient from 36% to 66% Mobile Phase B (0.1% TFA in [80% ACN + 20% H_2O]) over 60 minutes yielding 96.7% pure final fractions which were pooled and lyophilized.

29.3.3 Cell Culture

Human breast cancer cell lines MCF7 and MDA-MB-468 were purchased from American Type Culture Collection (ATCC). Cells were authenticated and were tested for mycoplasma free. MCF7 and MDA-MB-468 cells were grown in DMEM (Invitrogen) and supplemented with 10% FBS (Gemini) at 5% CO_2 in a 37 °C humidified incubator. All cells were supplemented with 50 U/mL penicillin, 50 μg/mL streptomycin (Mediatech), and 5 μg/mL plasmocin (InvivoGen) [27].

29.3.4 Cy5p53Tet Cellular Uptake by Live Cell Imaging

One day before imaging, 2×10^5 cells were seeded per well in a 12-well glass bottom plate (MatTek). Cells were incubated with 100 or 500 nM Cy5p53Tet at 37 °C for indicated time. Cy5p53Tet media was removed and the cells were washed three times with phosphate buffered saline (PBS) at room temperature and co-stained with 1 μg/mL Hoechst 33342 (Thermo Fisher) in PBS for 5 min. *Z*-stack images of stained cells were taken by confocal microscopy using a Nikon A1 confocal microscope with 60´ objective [27].

29.3.5 Protein Extraction

Cells were scraped off from plate and harvested at 1100 rpm for 5 min at 4 °C on Sorvall benchtop centrifuge. After washing three times with ice-cold PBS, cells were resuspended in RIPA buffer (0.1% SDS, 1% IGEPAL NP-40, 0.5% Deoxycholate, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM Tris-Cl pH 8.0, 1mM PMSF, 8.5 µg/ml Aprotinin, 2 µg/ml Leupeptin and phosphatase inhibitor cocktail). The cell suspension was incubated on ice for 30 min, vortexing every 5 min. Additional sonication of lysate (3x on ice 30 sec on/off at 98% amplitude) was done after the incubation. Samples were centrifuged at 13,000 rpm for 30 min at 4 °C and supernatant was collected. Protein extraction from frozen tumor was extracted by homogenization of tissue with RIPA buffer. RIPA buffer was added at a ratio of 0.4 ml per 0.1 g tissue (1:4 w/v). Incubation was carried out on ice for 10 min and the sample was vortexed every two minutes. Additional sonication of the lysate (3x on ice for 30 sec on/off at 98% amplitude) was carried out after the incubation. Samples were centrifuged at 13,00 rpm for 10 min at 4 °C. Cell extracts were run on 10% SDS-PAGE followed by electrotransfer onto polyvinylidene fluoride membrane (GE). The membrane was blocked with 5% non-fat milk (Bio-Rad) in either 1X PBS−0.1% Tween-20 or 1X TBS−0.1% Tween-20 followed by incubation with primary antibody overnight at 4 °C. The membrane was washed with either 1X PBS−0.1% Tween-20 or 1X TBS−0.1% Tween-20 and incubated with secondary antibody for 1 h at room temperature. The signal was detected by chemiluminescence with Super Signal Kit (Pierce) and autoradiography with Hyblot CL films (Denville Scientific) or Typhoon FLA7000 laser scanner (GE Healthcare) [34].

29.3.6 Organoid Establishment and Culture

Breast cancer PDTOs were generated by International Center for the Study of Breast Cancer Subtypes (ICSBCS) at Weill Cornell Medicine [33]. PDTOs were developed as previously described including some modifications. Fresh tissue samples were placed in CO_2 independent medium (Gibco) with GlutaMAX (1 × Invitrogen), 100 U/ml Antibiotic-Antimycotic (Thermo Fisher Scientific), and Primocin 100 µg/ml (InvivoGen). Tissue samples were washed in culture media two times before being placed in a 10 cm Petri dish for mechanical dissection. The dissected tissue was then enzymatically digested with collagenase IV media (DMEM [Gibco],100 U/mL penicillin, 100 µg/mL streptomycin [Gibco], 250 U/mL collagenase IV [LifeTechnologies], 100 µg/mL Primocin, and 10 µmol/L Rock inhibitor Y-27632 [Selleck Chemical Inc.]) in a volume of at least 20 times the tissue volume and incubated on a shaker at 200 rpm at 37 °C. Incubation time of the specimen was dependent on the amount of collected tissue and ranged from 20 to 60 min, until the majority of cell clusters were in suspension. After tissue digestion, Advanced DMEM/F12 media (Invitrogen) containing 1x GlutaMAX, 100 U/ml penicillin-streptomycin, and HEPES (10 mM, Gibco) (+++ media) was added to the suspension and the mixture was centrifuged at 300 rcf for 3 min. The pellet was then washed with +++ media and resuspended in optimized breast organoid-specific culture media. The final resuspended pellet was combined with Matrigel (Corning) in a 1:2 vol media: Matrigel, with five 80 µl droplets pipetted onto each well of a six-well suspension culture plate (Greiner Bio-One). The plate was placed into a cell culture incubator at 37 °C and 5% CO_2 for 30 min to polymerize the droplets before 3 ml of organoid-specific culture media was added to each well. The culture was maintained with fresh media changed twice a week. Dense cultures with organoids ranging in size from 200 to 500 µm were passaged every 10–30 days depending on the organoid line. During passaging, the organoid droplets were mixed with TrypLE Express (Gibco) and placed in a water bath at 37 °C for a maximum of 7 min. The resulting cell clusters and single cells were washed and replated [33].

29.3.7 Peptide Organoids Cellular Uptake by Live Cell Imaging

Organoids were extracted from Matrigel using Cell Recovery Solution (Corning) following manufacturer's protocol. Organoids were incubated with 500 nM Cy5p53Tet or Cy5Scramble at 37 °C for 4 h. After incubation, the cells were washed three times with PBS at room temperature and co-stained with NucBlue live ready probes reagent (Hoechst 33342) (Invitrogen). Cells were plated in 96-well CellCarrier Ultra imaging plates (PerkinElmer) and imaging was performed using the Operetta High Content Imaging System (PerkinElmer) [33].

29.3.8 IVIS Spectrum Imaging of Cy5p53Tet in Mice Bearing Bilateral MCF7/ MDA-MB-468 Xenografts

Female athymic nude mice (6–10 weeks old, 01B74-Athymic NCr-nu/nu) were obtained from Charles River Laboratories. Animals were supplemented with 17β-estradiol with dose of 0.72 mg/pellet (60-day release) into the neck seven days before MCF7 cells were subcutaneously implanted. 5×10^6 cells/mouse MCF7 cells were suspended in 100 µL of 1:1 media/matrigel basement membrane matrix (Corning) and injected subcutaneously on the left shoulder of each mouse ($n \geq 3$/group). After four weeks, 5×10^6 cells/mouse MDA-MB-468 cells were subcutaneously implanted in the right shoulder of mouse in a 100 µL of 1:1 media/matrigel basement membrane matrix. Imaging experiments were performed when the tumors reached a volume of ~50–250 mm^3 (after approximately three weeks). Injected into the tail vein of each mouse was 10 nmol Cy5p53Tet. Prior to *in vivo* imaging, the mice were anesthetized with 1.5%–2.0% isoflurane (Baxter Healthcare). Images were collected using an IVIS Spectrum (Perkin Elmer) 12 min, 30 min and 3 h following the administration of Cy5p53Tet. Epifluorescence exposure time on each side was identical, with multiple exposures ranging from 0.2 s to 2 s. Fluorescence imaging was obtained with excitation and emission wavelengths of 640 and 680 nm, respectively. Animals were sacrificed 40 min, 80 min, or 3 h after the injection of Cy5p53Tet, and epifluorescence images of the excised MCF7 and MDA-MB-468 xenografts were obtained using same condition as previously. Semiquantitative analysis of the Cy5p53Tet signal was conducted by measuring the average radiant efficiency [p/s/cm2/sr]/[µW/cm2] in regions of interest [27].

29.3.9 Ethics

All animal experiments were done in accordance with protocols approved by the Institutional Animal Care and Use Committees (IACUC) of Hunter College, Weill Cornell Medical College, and Memorial Sloan Kettering Cancer Center and followed the National Institutes of Health guidelines for animal welfare.

29.4 Results

29.4.1 Penetration of Cy5p53Tet into TNBC Cells Expressing High Levels of Mutant p53

Triple negative breast cancer lacks detectable estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) gene amplification, which makes TNBC very difficult to treat with conventional therapies. Eighty percent of the TNBC patients carry p53 mutations [35]. Our novel development of fluorescent conjugated peptides Cy5p53Tet uses the p53 TD as a targeting molecule and scaffold for imaging probes [27]. Cy5p53Tet is a peptide with a nucleus-penetrating HIV-TAT amino acid sequence along with a p53 TD sequence (Figure 29.1A and [27]). Peptides offer a multitude of cancer cell targeting advantages including rapid distribution, reduced or absent immunogenicity, ease and scalability of synthesis, and affordable of labeling [36]. We examined the uptake of Cy5p53Tet in TNBC cell culture lines to test proof of principle. Live cell imaging indicated that Cy5p53Tet strongly enters into MDA-MB-468 cells that express stable missense mtp53 R273H, but poorly enters ER+ MCF7 breast cancer cells that express wtp53. Significant higher intensity of Cy5p53Tet signal corresponds to mtp53-high cancer detection (Figure 29.1B and [27]). Merged fluorescent images of the cell staining with Cy5p53Tet and Hoechst demonstrates the significant Cy5p53Tet nuclear localization (Figure 29.1B and

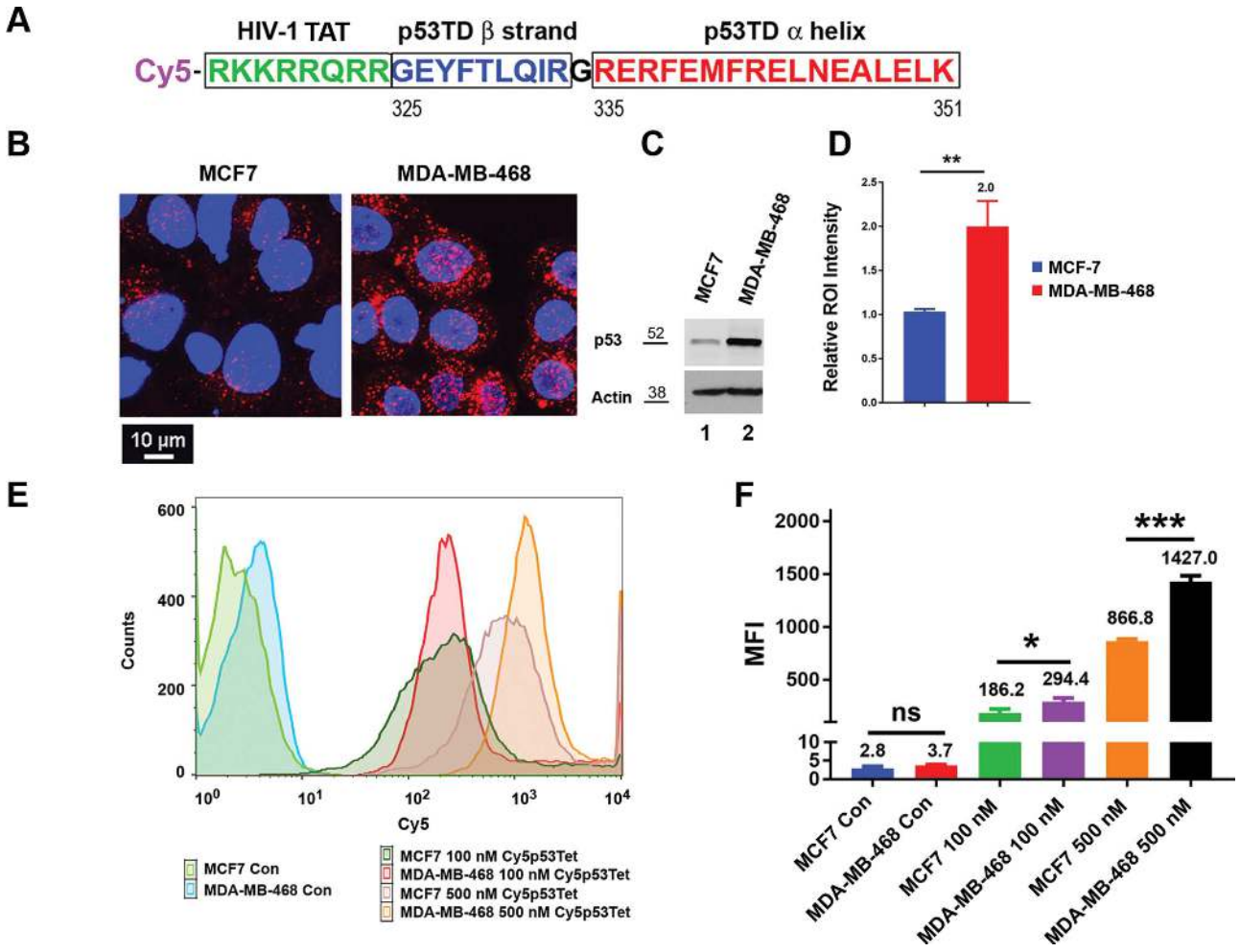


FIGURE 29.1 *Cy5p53Tet penetrates MDA-MB-468 TNBC cells expressing mtp53 R273H.* (A) Schematic of the structure of Cy5p53Tet. (B) Live cell imaging staining of MCF7 and MDA-MB-468 cells after 2 h of incubation with 500 nM Cy5p53Tet (red). Hoechst staining (blue) was used to stain the nuclei. Two independent experiments with biological replicates were performed. (C) p53 protein levels in MCF7 and MDA-MB-468 cells determined by western blot analysis before carrying out live cell imaging. (D) Quantification of Cy5p53Tet uptake in MCF7 and MDA-MB-468 cells via Nikon Element analysis. At least 200 cells per sample were measured by fluorescence microscopy. (E) Flow cytometry of MCF7 and MDA-MB-468 cells after incubation with 100 or 500 nM Cy5p53Tet for 2 h at 37 °C. FlowJo software was used to analyze the cytometric data. (F) Geometric MFI from the FACS experiments in (E). (Figure adapted with permission from [27]. Copyright: The American Chemical Society 2021.)

[27]). Western blot analysis shows the corresponding high versus low levels of p53 in MDA-MB-468 TNBC cells compared to ER+ MCF7 cells and correlates with the peptide-mediated detection (Figure 29.1C and D and [27]). Furthermore, the cellular uptake of Cy5p53Tet in MDA-MB-468 and MCF7 was analyzed using flow cytometry. The flow cytometry of MCF7 and MDA-MB-468 treated with vehicle or Cy5p53Tet at 100 or 500 nM for 2 h demonstrates the proof that mtp53-expressing cells uptake the Cy5p53Tet peptide (Figure 29.1E). The geometric mean fluorescence intensity (MFI) values indicate that uptake of Cy5p53Tet is significantly elevated in cells with mtp53 compared to cells with wtp53, further suggesting that a p53 TD peptide can be developed as a cancer screening tool (Figure 29.1F).

29.4.2 Penetration of Cy5p53Tet into TNBC Patient-Derived Tumor Organoids (PDTOs) Expressing High Levels of Mutant p53

Patient-derived tumor organoids, which are directly derived from patient cancer specimens, have been utilized in translational oncology, particularly in anti-cancer drug screening [37]. We tested if the diagnostic Cy5p53Tet peptide would penetrate into PDTOs. PDTOs with mtp53 Arg213stop p53 expression (ICSBCS007) and the TNBC organoid with missense mtp53 Arg248Trp (ICSBCS002) were compared for uptake of Cy5p53Tet compared to a random scrambled Cy5-labeled peptide (Cy5Scramble) in live cell imaging. The PDTOs with 500 nM Cy5p53Tet, compared to the scrambled Cy5Scramble peptide, showed higher uptake (Figure 29.2A and [33]). Western blot analysis provides the evidence that high levels of mtp53 in the missense mtp53 Arg248Trp (ICSBCS002) PDTO compared to mtp53 Arg213stop (ICSBCS007) correlate with increased imaging peptide uptake (Figure 29.2B and [33]). The intensity of the Cy5p53Tet signal and the mtp53 protein levels are positively correlated. A higher uptake of Cy5p53Tet in missense mtp53 Arg248Trp ICSBCS002 compared to mtp53 Arg213stop ICSBCS007 PDTO is supporting evidence of the proof of principle for using p53 TD peptide to detect cancers with high missense mtp53. However, it also shows the difficulty in detecting cancer cells with low levels of mtp53 LOF. Importantly, both ICSBCS002 and ICSBCS007 organoids have ten-fold higher uptake of Cy5p53Tet than of the nonspecific Cy5Scramble peptides, which indicates the diagnostic peptide specificity (Figure 29.2A and B and [33]). The accumulation of Cy5p53Tet versus the Cy5Scramble in PDTOs examined by extracting proteins from ICSBCS002 and ICSBCS007 after incubation with Cy5p53Tet or Cy5Scramble and then western blot is another means to detect peptide uptake level. Higher Cy5p53Tet is detected in ICSBCS002 compared to ICSBCS007 and only low levels of Cy5Scramble are detected (Figure 29.2C and [33]). We identified that Cy5p53Tet uptake between ICSBCS002 and ICSBCS007 PDTOs is correlated to the level of p53 expression. As such, Cy5p53Tet provides an image-based nondestructive approach to detect PDTOs based on mtp53 expression [33].

29.4.3 Cy5p53Tet Accumulates in TNBC mtp53 Xenograft Tumors

To investigate whether Cy5p53Tet could be used as a platform for the *in vivo* imaging of tumors with mtp53, we performed *in vivo* near-infrared fluorescence imaging (NIRF) experiments in mice bearing bilateral ER+ MCF7 and TNBC MDA-MB-468 xenografts that express wtp53 and mtp53R273H (Figure 29.3 and [27]). High levels of ER and low levels of wtp53 were found in MCF7 cells, while MDA-MB-468 cells produce high levels of mtp53 and undetectable ER (Figure 29.3A and

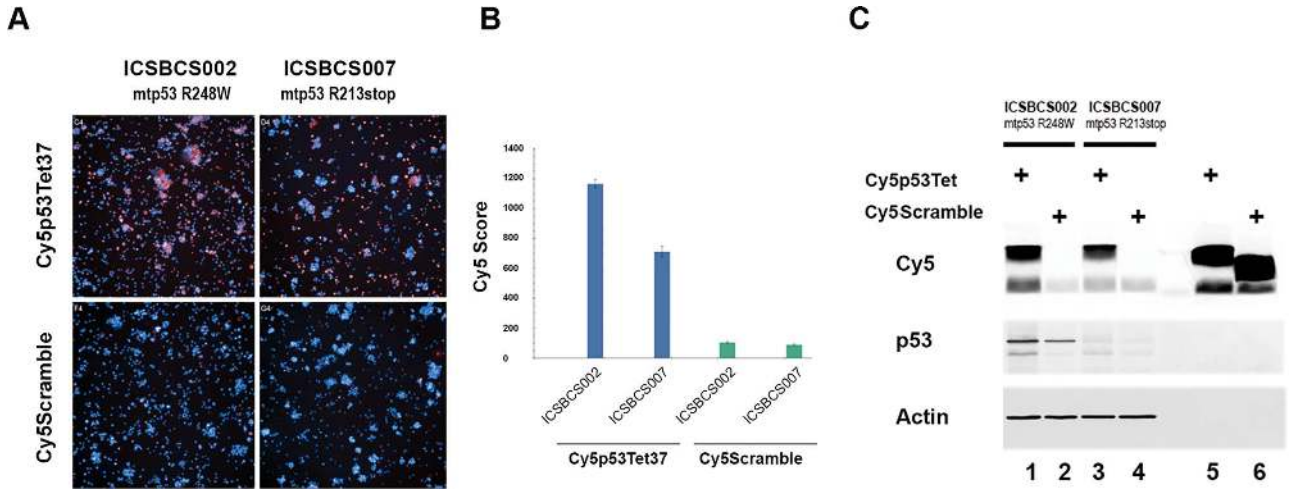


FIGURE 29.2 *Specific uptake of Cy5p53Tet in high levels of mtp53 breast cancer organoids.* (A) Representative images from high-content imaging analysis of mtp53 PDTO ICSBCS002 (mtp53 R248W) and ICSBCS007 (mtp53 R213 stop) treated with either 500 nM Cy5p53Tet or Cy5scramble peptides for 4 h. Hoechst staining was used to stain the nuclei (Cy5: red; Hoechst: blue). (B) Quantification of Cy5p53Tet and Cy5scramble uptake in ICSBCS002 and ICSBCS007 PDTOs via Operetta High Content Imaging System. (C) Whole cell extracts were prepared from breast cancer PDTO ICSBCS002 and ICSBCS007 treated with either 500 nM Cy5p53Tet or Cy5scramble peptides for 4h. 1 μg of cell extracts was run on 15% polyacrylamide gel and uptake of Cy5p53Tet or Cy5Scramble was examined by scanning the gel using Typhoon FLA 7000 biomolecular imager. Protein levels of p53 and actin were determined by western blot analysis. (Figure adapted with permission from [33]. Copyright: Elsevier, 2024.)

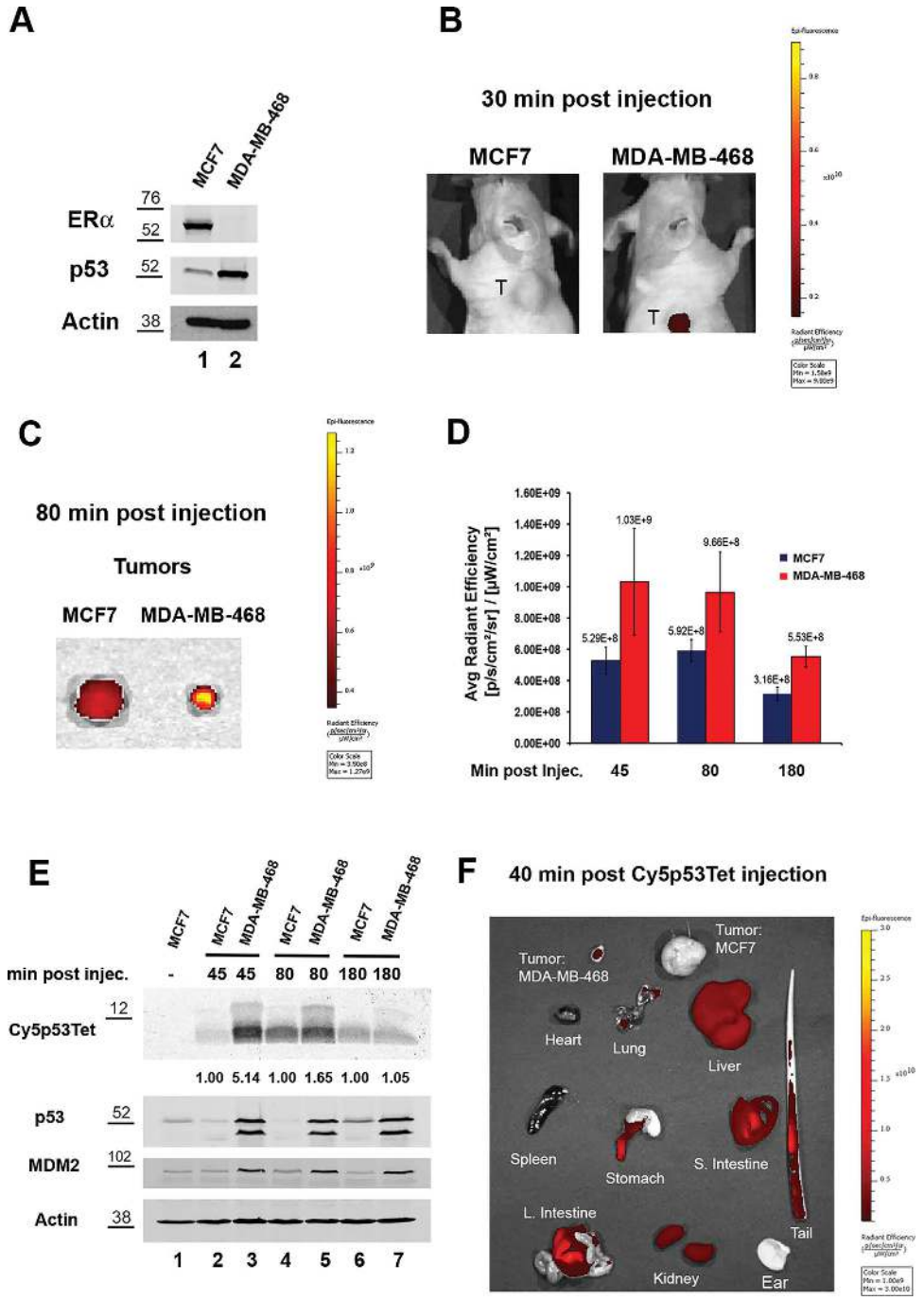


FIGURE 29.3 *Cy5p53Tet accumulates in mtp53-expressing xenografts.* (A) ER and p53 protein levels in MCF7 and MDA-MB-468 cells were determined by western blot analysis before implantation. (B) *In vivo* NIRF image acquired 30 min after the injection of Cy5p53Tet to a mouse bearing bilateral MCF7 and MDA-MB-468 xenografts (T = tumor). (C) *Ex vivo* NIRF imaging of MCF7 and MDA-MB-468 tumors excised 80 min after the administration of Cy5p53Tet. (D) Quantification of epifluorescence from tumors resected at 40, 80, and 180 min after the injection of Cy5p53Tet. (E) SDS-PAGE analysis of extracts collected from the tumors in the *in vivo* imaging experiment. Protein (25 μg) from each sample was run on a 12% polyacrylamide gel, and the fluorescence signal from Cy5p53Tet was interrogated and quantified. The expression levels of p53 and MDM2 in the same tumor samples were determined by western blot analysis. F) High uptake of Cy5p53Tet in mtp53, but not wtp53, expressing tumor visualized by ex vivo NIRF imaging and organ uptake. (Figure adapted with permission from [27]. Copyright: The American Chemical Society 2021.)

[27]). Fluorescence signal is clearly observed 12 minutes after the intravenous administration of Cy5p53Tet, in the MDA-MB-468 tumors, whereas no signal is present in the MCF7 tumors (Figure 29.3B and [27]). At 40, 80, and 180 min post injection, the mice were sacrificed, and the xenografts were resected and imaged *ex vivo* (Figure 29.3C and [27]). At each time point, the MDA-MB-468 mtp53-expressing tumors exhibit higher radiant efficiency than the MCF7 xenografts (Figure 29.3D and [27]). The ratios of the Cy5 signal in the MDA-MB-468 xenografts to that in the MCF7 tumors at 40,

80, and 180 min varied over time (Figure 29.3D and [27]). The degree to which Cy5p53Tet penetrated the tumor tissue was measured by extracting proteins from the tumors excised at 40, 80, and 180 min post injection (Figure 29.3E and [27]). The intensity of Cy5p53Tet in each sample was obtained by scanning on the Cy5 channel by a Typhoon FLA 7000 laser scanner and quantifying the signal intensity. In this analysis, the ratios of the Cy5 signal derived from the MDA-MB-468 cells to that in the MCF7 cells at 40, 80, and 180 min again varied over time. The expression levels of p53 and MDM2 following the administration of Cy5p53Tet were probed by western blot. The levels of both oncogenic proteins mtp53 and MDM2 are higher in the TNBC MDA-MB-468 cells (Figure 29.3E and [27]). Interestingly, after the administration of Cy5p53Tet in MCF7 cells, wtp53 and MDM2 levels demonstrated the classic oscillation pattern (Figure 29.3E, lanes 2, 4, 6 and [27]). Oscillations in p53 and MDM2 protein levels are known to keep wtp53 levels low. In conclusion, the elevated uptake of Cy5p53Tet in the mtp53-expressing tumors of the mice suggests that Cy5p53Tet preferentially targets tumors expressing mtp53. The permeation properties of Cy5p53Tet suggest that the peptide could hold potential for the detection of TNBC and could be a delivery vehicle for cancer treatments. The whole-body images and the images of the resected tissues confirm the specific uptake of the peptide in the MDA-MB-468 tumors compared to MCF7 tumors but also reveal substantially high levels of the signal in several healthy tissues, including the liver, large intestine, and small intestines (Figure 29.3F and [27]).

29.4.4 Stapled Peptides SPCy5p53Tet Accumulate at mtp53 TNBC Xenograft Tumor Sites

In light of these initial results, we embarked upon the development of a second-generation mtp53-targeting imaging agent aiming for enhanced stability and improved pharmacokinetic profiles. Along these lines, two strategies that have been used to improve the *in vivo* performance of peptides are the use of "stapled" peptides with enhanced stability [38] and the modification of the charge and size of the peptide to shift clearance from the hepatobiliary system to the renal system [39]. With respect to the former, stapled peptides are locked into a stabilized α-helical structure by hydrocarbon staple to stabilize their bioactive structure and enhance protease resistance and cellular penetrance [38]. We designed two p53TD stapled peptides based on p53TD α-helices bearing *N*-terminal Cy5 fluorophore and either were constructed without, or with, the HIV-1 TAT nuclear penetration sequence. The peptides SPCy5p53TetA and SPCy5p53TetB were stapled in the S5 position with S-2–(4′-pentenyl) alanine (Figure 29.4A). This alanine was added to the stapled peptides because N- and C-terminal alanine residues (slightly apolar) exhibit enhanced helicity [40]. The two stapled p53TD peptides were formulated with the p53TD α-helices amino acids that have been noted as required for p53 dimer and tetramer formation [19]. To investigate whether SPCy5p53Tet could be an improved imaging agent for tumors containing mtp53, *in vivo* and *ex vivo* NIRF imaging experiments were performed in mice bearing MDA-MB-468 xenografts that express mtp53R273H (Figure 29.4B). The biodistribution of SPCy5p53TetA and SPCy5p53TetB was evaluated via IVIS 1 and 2 h post-injection. For both peptides, the signal in the MDA-MB-468 tumors was significantly higher than that in the surrounding normal tissues. Interestingly, SPCy5p53TetA without the nuclear localization HIV-TAT sequence exhibited greater uptake in tumor tissue than SPCy5p53TetB (Figure 29.4B). At 1 and 2 h after injection, cohorts of mice were sacrificed and dissected. The uptake of SPCY5p53TetA and SPCy5p53TetB in the tumor and organs of interest was then assessed via *ex vivo* NIRF imaging (Figure 29.4C, data shown at 1 h post-injection of SPCy5p53TetA). These data indicate that SPCy5p53TetA was eliminated by the kidneys and exhibited significantly lower uptake in the liver than the uptake in the tumor (Figure 29.4C). In conclusion, the high tumor uptake and rapid renal clearance of SPCy5p53Tet suggest that the peptide has potential for the delivery of diagnostic or therapeutic payloads to tumors with mtp53.

29.5 Discussion

We demonstrated Cy5p53Tet peptide uptake in multiple models of TNBC, including 3D breast cancer PTDOs and mouse xenograft models. Importantly, we improved the *in vivo* performance of Cy5p53Tet by pivoting to stapled peptides. The goal of this change was to increase stability while obtaining sufficient cell permeability to target intracellular/nuclear mtp53 while avoiding off-target effects. The second-generation peptides SPCy5p53TetA and SPCy5TetB contain a stapled p53TD α-helix with all the amino acids that are required for dimer and tetramer formation. The difference between SPCy5p53TetA and SPCy5p53TetB is the presence (B) or absence (A) of an HIV-1 TAT sequence (i.e., eight cationic Arg and Lys residues have been shown to enable cell penetration) [41]. It is important to note that in the *in vivo* xenograft model of TNBC, SPCy5p53TetA produced higher uptake in tumor tissue compared to SPCy5p53TetB, showing that cellular uptake can be achieved even in the absence of adding an HIV-TAT sequence to peptides.

The US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) actively encourage the use of biomarker tests in the development and use of prescription drugs [42]. The SPCy5p53Tet could be developed as a TNBC-specific diagnostic to identify patients with high mtp53 expression which can indicate use for synthetic lethal interactions. We have identified that mtp53 may enhance sensitivity to inhibiting the enzyme poly-ADP ribose polymerase (PARP). When we profiled datasets from TCGA we found the p53 and PARP1 protein levels are significantly elevated in TNBC, suggesting that stable mtp53 and PARP co-associate as dual biomarkers [34]. Tissue microarrays (TMAs) show a positive correlation between stable p53 and high PARP1 [33] suggesting that high mtp53 could identify patients likely to respond to PARP inhibitors. Indeed, our experiments with breast cancer PDTOs show that in breast cancers with mtp53, but not wtp53, the combination of the PARPi talazoparib and the DNA damaging agent temozolomide induce synergistic cytotoxicity [25, 33].

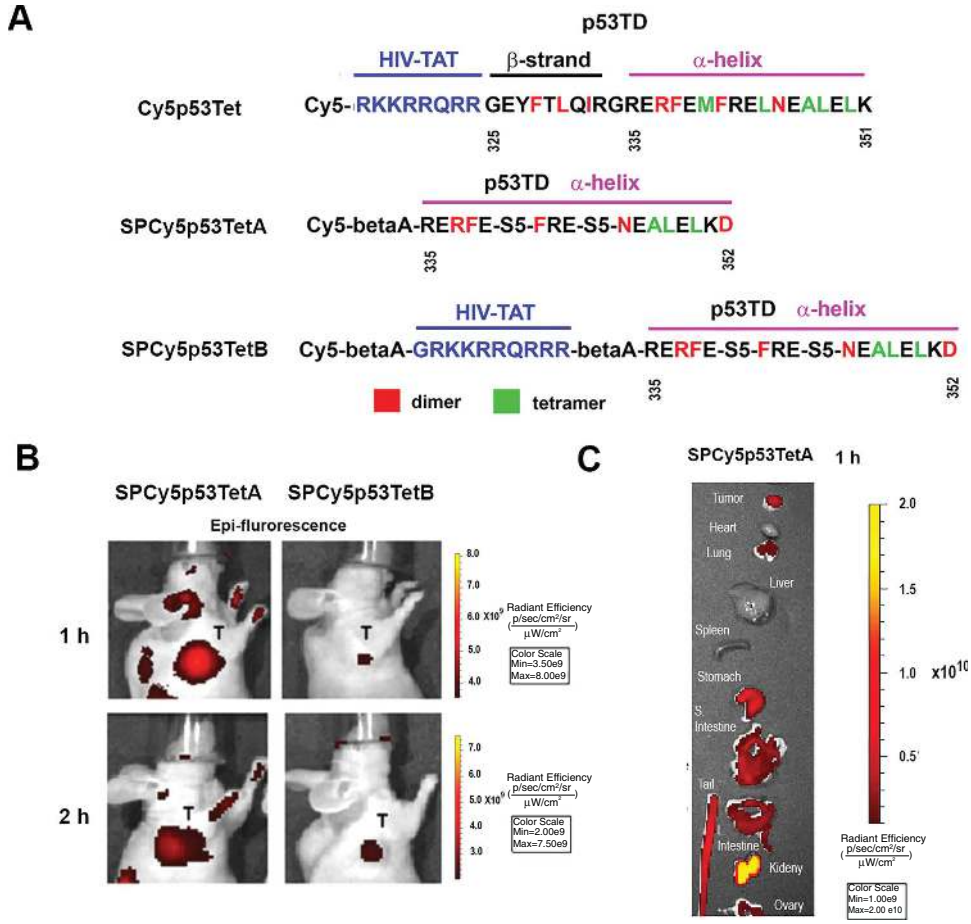


FIGURE 29.4 *Stapled peptides SPCy5p53TetA/B accumulate in mtp53-expressing MDA-MB-468 xenografts.* (A) Schematic of the structures of SPCy5p53TetA and SPCy5p53TetB compared to Cy5p53Tet. The sequence of SPCy5p53TetA and SPCy5p53TetB are shown and describe the fluorophore Cy5 conjugated to the p53TD α-helix (335–352) without SPCy5p53TetA or with SPCy5p53TetB, an N-terminal HIV-1 TAT (48–57). The stapled locations are on S5 and S5. The residues involved in dimer and tetramer formation are labeled in red (dimer) and green (tetramer) [19]. (B) Left: *In vivo* NIRF images obtained 1 and 2 h after the administration of SPCy5p53TetA or SPCy5p53TetB to mice bearing MDA-MB-468 xenografts. (C) *Ex vivo* NIRF imaging of MDA-MB-468 tumors and selected organs excised 1 h after the administration of SPCy5p53TetA.

Finally, over the last decade, a great deal of progress has been made on the preclinical development and clinical deployment of peptide-based radiopharmaceuticals for the diagnosis and treatment of cancer. The FDA recently approved both [^{68}Ga]Ga-DOTA-TOC in 2019 and [^{177}Lu]Lu-DOTA-TATE for the PET imaging and radiopharmaceutical therapy, respectively, of patients with somatostatin receptor (SSTR)-positive neuroendocrine tumors [43]. In the near future, we plan to explore the synthesis, and *in vivo* validation, of radiolabeled variants of SPp53Tet in murine models of TNBC. If the preclinical evaluation of these prove successful, they may indeed have a place in the clinic for the diagnostic imaging or radiopharmaceutical therapy of patients with mtp53-expressing malignancies.

29.6 Future Trends

SPCy5p53Tet showed promising results as an *in vivo* molecular imaging agent for the detection of xenograft mtp53-expressing tumor cells. Moving forward, we will continue our *in vivo* validation of SPCy5p53Tet peptides using multiple TNBC cell lines, patient-derived tumor organoids, and patient derived xenografts (PDX) models. Future studies will also include the synthesis of a chelator-bearing variant of SPCy5p53Tet for the creation of a radiopeptide for the PET imaging of mtp53-expressing tumors. Such an imaging agent could ultimately be used as a theranostic tool for the identification of patients with mtp53-expressing cancers who may respond to combination therapies based on PARP inhibitors and DNA damaging agents.

REFERENCES

1. Bray, F., et al., The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer*, 2021. 127(16): p. 3029–3030.
2. Hahn, W.C., et al., An expanded universe of cancer targets. *Cell*, 2021. 184(5): p. 1142–1155.
3. Marjanovic, N.D., R.A. Weinberg, and C.L. Chaffer, Cell plasticity and heterogeneity in cancer. *Clin Chem*, 2013. 59(1): p. 168–179.

4. Zhao, M., et al., Mutant p53 gains oncogenic functions through a chromosomal instability-induced cytosolic DNA response. *Nat Commun*, 2024. 15(1): p. 180.
5. Olivier, M., et al., The IARC TP53 database: New online mutation analysis and recommendations to users. *Hum Mutat*, 2002. 19(6): p. 607–14.
6. Kandoth, C., et al., Mutational landscape and significance across 12 major cancer types. *Nature*, 2013. 502(7471): p. 333–339.
7. Freed-Pastor, W.A. and C. Prives, Mutant p53: One name, many proteins. *Genes Dev*, 2012. 26(12): p. 1268–1286.
8. Hassin, O. and M. Oren, Drugging p53 in cancer: One protein, many targets. *Nat Rev Drug Discov*, 2023. 22(2): p. 127–144.
9. Levine, A.J., p53: 800 million years of evolution and 40 years of discovery. *Nat Rev Cancer*, 2020. 20(8): p. 471–480.
10. Bargonetti, J. and C. Prives, Gain-of-function mutant p53: History and speculation. *J Mol Cell Biol*, 2019. 11(7): p. 605–609.
11. Blandino, G. and S. Di Agostino, New therapeutic strategies to treat human cancers expressing mutant p53 proteins. *J Exp Clin Cancer Res*, 2018. 37(1): p. 30.
12. Marvalim, C., A. Datta, and S.C. Lee, Role of p53 in breast cancer progression: An insight into p53 targeted therapy. *Theranostics*, 2023. 13(4): p. 1421–1442.
13. Zhu, G., et al., Mutant p53 in cancer progression and targeted therapies. *Front Oncol*, 2020. 10: p. 595187.
14. Zhang, C., et al., Gain-of-function mutant p53 in cancer progression and therapy. *J Mol Cell Biol*, 2020. 12(9): p. 674–687.
15. Bykov, V.J.N., et al., Targeting mutant p53 for efficient cancer therapy. *Nat Rev Cancer*, 2018. 18(2): p. 89–102.
16. Zhang, Q., et al., APR-246 reactivates mutant p53 by targeting cysteines 124 and 277. *Cell Death Dis*, 2018. 9(5): p. 439.
17. Roh, J.L., et al., The Hsp90 inhibitor 17-(allylamino)-17-demethoxygeldanamycin increases cisplatin antitumor activity by inducing p53-mediated apoptosis in head and neck cancer. *Cell Death Dis*, 2013. 4(12): p. e956.
18. Chou, C.W., et al., Therapeutic effects of statins against lung adenocarcinoma via p53 mutant-mediated apoptosis. *Sci Rep*, 2019. 9(1): p. 20403.
19. Gencel-Augusto, J. and G. Lozano, p53 tetramerization: At the center of the dominant-negative effect of mutant p53. *Genes Dev*, 2020. 34(17–18): p. 1128–1146.
20. Tanaka, T., M. Watanabe, and K. Yamashita, Potential therapeutic targets of TP53 gene in the context of its classically canonical functions and its latest non-canonical functions in human cancer. *Oncotarget*, 2018. 9(22): p. 16234–16247.
21. Friedman, P.N., et al., The p53 protein is an unusually shaped tetramer that binds directly to DNA. *Proc Natl Acad Sci U S A*, 1993. 90(8): p. 3319–23.
22. Laptenko, O., et al., The p53 C terminus controls site-specific DNA binding and promotes structural changes within the central DNA binding domain. *Mol Cell*, 2015. 57(6): p. 1034–1046.
23. Annor, G.K., et al., Oligomerization of mutant p53 R273H is not required for gain-of-function chromatin associated activities. *Front Cell Dev Biol*, 2021. 9: p. 772315.
24. Polotskaia, A., et al., Proteome-wide analysis of mutant p53 targets in breast cancer identifies new levels of gain-of-function that influence PARP, PCNA, and MCM4. *Proc Natl Acad Sci U S A*, 2015. 112(11): p. E1220–9.
25. Qiu, W.G., et al., Identification, validation, and targeting of the mutant p53-PARP-MCM chromatin axis in triple negative breast cancer. *NPJ Breast Cancer*, 2017. 3.
26. Samec, T., et al., Peptide-based delivery of therapeutics in cancer treatment. *Mater Today Bio*, 2022. 14: p. 100248.
27. Xiao, G., et al., Targeting triple negative breast cancer with a nucleus-directed p53 tetramerization domain peptide. *Mol Pharm*, 2021. 18(1): p. 338–346.
28. Brooks, H., B. Lebleu, and E. Vivès, Tat peptide-mediated cellular delivery: Back to basics. *Adv Drug Deliv Rev*, 2005. 57(4): p. 559–577.
29. Gabizon, R., et al., Using peptides to study the interaction between the p53 tetramerization domain and HIV-1 Tat. *Biopolymers*, 2008. 90(2): p. 105–116.
30. Manduca, N., et al., 3D cancer models: One step closer to in vitro human studies. *Front Immunol*, 2023. 14: p. 1175503.
31. Podaza, E., et al., Next generation patient derived tumor organoids. *Transl Res*, 2022. 250: p. 84–97.
32. Rossi, G., A. Manfrin, and M.P. Lutolf, Progress and potential in organoid research. *Nat Rev Genet*, 2018. 19(11): p. 671–687.
33. Madorsky Rowdo, F.P., et al., Patient-derived tumor organoids with p53 mutations, and not wild-type p53, are sensitive to synergistic combination PARP inhibitor treatment. *Cancer Lett*, 2024. 584: p. 216608.
34. Xiao, G., et al., Gain-of-function mutant p53 R273H interacts with replicating DNA and PARP1 in breast cancer. *Cancer Res*, 2020. 80(3): p. 394–405.
35. Bareche, Y., et al., Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. *Ann Oncol*, 2018. 29(4): p. 895–902.
36. Staderini, M., et al., Peptides for optical medical imaging and steps towards therapy. *Bioorg Med Chem*, 2018. 26(10): p. 2816–2826.
37. Vivarelli, S., et al., Patient-derived tumor organoids for drug repositioning in cancer care: A promising approach in the era of tailored treatment. *Cancers (Basel)*, 2020. 12(12).
38. Walensky, L.D. and G.H. Bird, Hydrocarbon-stapled peptides: Principles, practice, and progress. *J Med Chem*, 2014. 57(15): p. 6275–6288.
39. Yang, S., et al., Clearance pathways of near-infrared-II contrast agents. *Theranostics*, 2022. 12(18): p. 7853–7883.
40. Chandramohan, A., et al., Design-rules for stapled peptides with in vivo activity and their application to Mdm2/X antagonists. *Nat Commun*, 2024. 15(1): p. 489.
41. Su, Y., et al., Membrane-bound dynamic structure of an arginine-rich cell-penetrating peptide, the protein transduction domain of HIV TAT, from solid-state NMR. *Biochemistry*, 2010. 49(29): p. 6009–6020.
42. D'Avó Luís, A.B. and M.K. Seo, Has the development of cancer biomarkers to guide treatment improved health outcomes? *Eur J Health Econ*, 2021. 22(5): p. 789–810.
43. Hennrich, U. and M. Benešová, [(68)Ga]Ga-DOTA-TOC: The first FDA-approved (68)Ga-radiopharmaceutical for PET imaging. *Pharmaceuticals (Basel)*, 2020. 13(3).

30

DNA Methylation Biomarkers for Preoperative Diagnostic of Thyroid Nodules

Byung-Wook Kim, Yuman Fong, John H. Yim, and Maria A. Hahn

30.1 Introduction

30.1.1 Preoperative Thyroid Cancer Diagnostics

Thyroid nodules are a common clinical problem. With the use of ultrasound, the prevalence of thyroid nodules is in the 19%–68% range (1, 2). The clinical importance of thyroid nodules rests on the need to diagnose thyroid cancer, which occurs in 7%–15% of these nodules (3).

According to the Surveillance, Epidemiology, and End Results Program (SEER), thyroid cancer is the 12th most common cancer type (4). In 2023, approximately 43,720 new cases of thyroid cancer were reported, representing 2.2% of all cancer cases in the US. The disease occurs approximately three times more frequently in women than men (4, 5). Thyroid cancer is the eighth highest cause of cancer death in women and the 14th highest cause of cancer death in men (5). The introduction of early detection of thyroid cancer by using ultrasound and fine-needle aspirations made a significant impact on thyroid cancer epidemiology (6, 7). According to SEER-8, from 1975 until 2019, the incidence of thyroid cancer has risen almost three times from 5.01 to 14.11 per 100,000 residents per year in the US, while thyroid cancer-related death remained at 0.5 per 100,000 residents per year (4).

Differentiated thyroid cancers, including papillary thyroid cancer (PTC) (83.6%), follicular cancer (10.8%), and medullary cancer (2.2%), comprise over 96.6% of newly identified thyroid cancers (8). PTC, follicular, and Hurthle cell cancers arise from thyroid hormone-producing follicular cells, while medullary thyroid cancers originate from parafollicular C cells. Thyroid cancer is associated with the highest survival rate among all cancers, with a five-year survival rate of 99.1% for women and 96.7% for men (4). This rate would be even higher if some differentiated thyroid cancers would not dedifferentiate into poorly differentiated thyroid cancer (PDTC) and further into undifferentiated thyroid cancer, anaplastic cancer, and become more aggressive. Anaplastic thyroid cancer is the most aggressive type of thyroid cancer, is detected in 1% of thyroid cancer patients in the US, and is responsible for 19.9% of thyroid cancer mortality (8).

Thyroid cancer is usually represented as a solid nodule on ultrasound. Ultrasound features of thyroid nodules are used to determine the risk of malignancy of the nodules by using the Thyroid Imaging Reporting and Data System (TI-RADS™) (9). To diagnose suspicious thyroid nodules according to ultrasound, fine-needle aspiration (FNA) of thyroid nodules is performed, and obtained thyroid nodule fragments are evaluated by a cytopathologist. The Bethesda System for Reporting Thyroid Cytopathology divides thyroid FNAs into six categories based on the cytopathological features: (i) non-diagnostic (ND), (ii) benign, (iii) atypia of undetermined significance (AUS), (iv) follicular neoplasm (FN), (v) suspicious for malignancy (SFM), and (vii) malignant. Each category is associated with a risk of malignancy (ROM) (10). The benign and malignant Bethesda categories are associated with high accuracy of diagnosis, 4% and 97% ROM, respectively. However, AUS, FN, and SFM categories have ROM ranging from 22% to 74% and are considered indeterminate. This indeterminate cytopathology is observed in up to 40% of thyroid FNAs (11) and represents a clinical challenge associated with overdiagnosis and overtreatment.

To diagnose indeterminate FNA, diagnostic lobectomy, removal of half of the thyroid, is performed. It has been estimated that for indeterminate nodule diagnosis, almost 75,500 surgeries were performed in the United States in 2010, of which approximately 50,000 were benign (12). It is important to note that these unnecessary thyroidectomies result in a wide range of avoidable complications. For instance, when thyroid tissue is removed, the remaining tissue is frequently unable to produce sufficient thyroid hormone, thereby necessitating lifelong thyroid hormone replacement (13, 14). Inadvertent removal and/or damage to the parathyroid glands can result in calcium imbalance. Furthermore, structural damage can also occur, including (1) damage to the recurrent laryngeal nerve (vocal cord paralysis) and (2) airway compromise (bilateral vocal cord paralysis), requiring a tracheostomy.

Considering that over 600,000 patients undergo FNA for preoperative thyroid nodule diagnostics in the United States yearly (15) and up to 40% of FNAs have indeterminate cytopathology (11), there is a significant number of patients requiring risky and very frequently unnecessary diagnostic thyroidectomies.

Molecular testing was introduced to reduce the number of diagnostic thyroidectomies on indeterminate nodules. The commercially available tests are based on genetic and genomic alterations and expression of mRNA, mitochondrial RNA (mtRNA), and microRNA (miRNA). However, as of today, due to the low PPV of these tests, many clinics keep performing diagnostic thyroidectomies without molecular testing.

DOI: 10.1201/9781003449942-34

30.1.2 Commercially Available Diagnostics for Indeterminate Thyroid FNA

There are five commercially available molecular tests for indeterminate thyroid FNAs: Afirma® Genomic Sequencing Classifier (GSC) (Veracyte Inc., CA, USA), ThyroSeq V3 (CBLPath, Inc., NY, USA), ThyGeNEXT/ThyraMIRv2 (Interpace Biosciences®, Inc., NJ, USA), Mir-THYpe (ONKOS Diagnostics Molecular LTDA, Brazil), and ThyroidPrint (GeneproDX, Chile) (Table 30.1).

One of the most widely used assays in the US is GSC, which is based on whole transcriptome sequencing with expression analysis of over 3600 genes, including 13 mitochondrial genes (Table 30.1). The RNA data are also used to evaluate loss-of-heterozygosity (LOH), gene fusions, and mutations (16, 17). In addition, BRAFV600E mutation is evaluated by Competitive Allele-Specific TaqMan PCR. According to the developers, the test has a sensitivity of 91% (95% confidence interval [CI], 79%–98%), a specificity of 68% (95% CI, 60%–76%), a negative predictive value (NPV) of 96% (95% CI, 90%–99%) and a positive predictive value (PPV) of 47% (95% CI, 36%–58%) (16). Independent validations reported PPV for GSC between 29% and 85.3% (15, 18–23).

Another very popular diagnostic test in the US is ThyroSeq V3 (Table 30.1). The test is also based on next-generation sequencing (NGS), such as GSC. However, ThyroSeq V3 is focused on the DNA and RNA analysis of 112 genes. The assay screens for 12,135 single-nucleotide variants (SNV) and insertions/deletions for over 120 gene fusions and for copy number abbreviation (CNV) in ten genomic regions and evaluates expression levels of 90 genes (24). According to the validation by the developers, the test demonstrated a sensitivity of 94% (95% CI, 86%–98%), specificity of 82% (95% CI, 75%–87%), NPV of 97% (95% CI, 93%–99%), and PPV of 66% (95% CI, 56%–75%) (25). For independent validations, ThyroSeq V3 PPV reported between 39% and 85% (15, 26–29).

Another test in the US is ThyGeNEXT and ThyraMIRv2, which represent a combination of two platforms (Table 30.1). The first platform is the NGS-based mutation classifier ThyGeNEXT that evaluates genetic changes in ten genes, 38 fusion transcripts, and the expression of four genes (30). The second platform, ThyraMIRv2, is an algorithmic microRNA risk classifier based on the analysis of the activity of ten miRNAs by using RT-PCR (31). According to the validation by Interpace Biosciences®, Inc., the ThyGeNEXT and ThyraMIRv2 performance is associated with sensitivity of 89% (95% CI, 82%–94%), specificity of 82% (95% CI, 77%–86%), NPV of 95% (95% CI, 92%–97%), and PPV of 66% (95% CI, 60%–72%). An independent validation of the approach reported a PPV of 53% (95% CI, 35%–70%) (32).

The Brazilian test Mir-THYpe offers thyroid nodule diagnosis based on quantitative RT-PCR of 11 miRNA (33, 34) (Table 30.1). The test's developers reported that it is associated with a sensitivity of 89.3% (95% CI, 82%–94.3%), specificity of 81.6% (95% CI, 76.7%–86%), NPV of 95% (95% CI, 91.7%–97%), and PPV of 66.2% (95% CI, 60.3%–71.7%).

TABLE 30.1

Commercially Available Molecular Diagnostics for Thyroid Nodules

	What Is Analyzed	Reported Accuracy	Accuracy per Independent Validation
GSC	mRNA and mtRNA expression, LOH, gene fusions, and mutations	Sensitivity: 91% (95% CI, 79%–98%) Specificity: 68% (95% CI, 60%–76%) NPV: 96% (95% CI, 90%–99%) **PPV: 47% (95% CI, 36%–58%)** (ref. 16)	Sensitivity: 91%–100% Specificity: 43%–94.3% NPV: 90%–100% **PPV: 29%–85%** (ref. 15, ref. 18–23).
ThyroSeq V3	CNV, DNA mutations, and RNA expression	Sensitivity: 94% (95% CI, 86%–98%) Specificity: 82% (95% CI, 75%–87%) NPV: 97% (95% CI, 93%–99%) **PPV: 66% (95% CI, 56%–75%)** (ref. 25)	Sensitivity: 90%–100% Specificity: 74%–90% NPV: 84%–100% **PPV: 39%–82%** (ref. 15, ref. 26–29)
ThyGeNEXT and ThyraMIRv2	ThyGeNEXT: DNA mutations, mRNA expression, gene fusions ThyraMIRv2: miRNA expression	Sensitivity: 98% (95% CI, 90%–99%) Specificity: 86% (95% CI, 79%–92%) NPV: 99% (95% CI, 94%–100%) **PPV: 75% (95% CI, 66%–83%)** (ref. 31)	Sensitivity: 50% (95% CI, 26%–74%) Specificity: 65% (95% CI, 43%–84%) NPV: 63% (95% CI, 49%–74%) **PPV: 53% (95% CI, 35%–70%)** (ref. 32)
Mir-THYpe	miRNA expression	Sensitivity: 89% (95% CI, 82%–94%) Specificity: 82% (95% CI, 77%–86%) NPV: 95% (95% CI, 92%–97%) **PPV: 66% (95% CI, 60%–72%)** (ref. 34)	N/A
ThyroidPrint	mRNA expression	Sensitivity: 91% (95% CI, 82%–96%) Specificity: 88% (95% CI, 82%–92%) NPV: 95% (95% CI, 91%–98%) **PPV: 78% (95% CI, 70%–84%)** (ref. 37)	N/A

Another test for indeterminate thyroid nodules, ThyroidPrint, is performed in Chile (Table 30.1). The test uses multiplexed quantitative PCR to evaluate the activity of ten genes (35–37). The performance of the test is associated with sensitivity of 91% (95% CI, 82%–96%), specificity of 88% (95% CI, 82%–92%), NPV of 95% (95% CI, 91%–98%), and PPV of 78 (95% CI, 70%–84%) (37).

While all these tests are associated with outstanding sensitivity and NPV and rarely miss a malignancy, reduced specificity and PPV of these tests result in unnecessary thyroidectomies. The cause of this overdiagnosis is a similarity of benign and malignant thyroid nodules. There is only one cancer-specific feature, BRAFV600E mutation, associated with 99.3% sensitivity (38). However, BRAFV600E is only detected in 45% of papillary thyroid cancer and is never observed in follicular thyroid cancers. Other reported changes between benign and malignant nodules like RET/PTC and PAX8-PPARγ1 rearrangement and RAS mutations can be classified as cancer-associated rather than cancer-specific (38–42).

Our team used DNA methylation to discriminate benign from malignant thyroid nodules to overcome this challenge. DNA methylation is strongly associated with cancer development and cancer-associated changes in gene activity (43). Our analysis revealed that benign nodules, malignant nodules, and adjacent thyroid tissue are associated with unique DNA methylation patterns that can potentially be used for thyroid nodule diagnostics (44). Based on DNA methylation specific to benign and malignant nodules, we developed a new diagnostic assay, the Diagnostic DNA Methylation Signature (DDMS) approach that is described in this chapter.

30.2 Materials

30.2.1 Chemicals

1. Water (AM9937, Ambion nuclease-free water, Thermo Fisher Scientific, NH, USA).
2. 96% Ethanol.
3. 80% Ethanol.
4. dsDNA Quantitation, High Sensitivity assay (Q32851, Thermo Fisher Scientific, NH, USA).
5. MspI 100 units/µl (R0106M, New England Biolabs, MA, USA; supplied with 10× NEB2 Buffer).
6. Deoxynucleoside triphosphates (11969064001, Roche).
7. Klenow Fragment (3′→5′ exo-) (M0212L, New England Biolabs, MA, USA).
8. Phenol/Chloroform/Isoamyl Alcohol (25:24:1 Mixture, pH 6.7/8.0, Liq.), Fisher BioReagents™ (BP1752I-100, Thermo Fisher Scientific, NH, USA).
9. Chloroform (BP1145–1, Thermo Fisher Scientific, NH, USA).
10. 20 mg/ml Glycogen from mussels (10901393001, Roche).
11. 3M Sodium Acetate pH 5.5 (AM9740, Thermo Fisher Scientific, NH, USA).
12. Methylated linker (Integrated DNA Technologies, IA, USA).
13. T4 ligase (M0202M, New England Biolabs, MA, USA; supplied 10× Reaction Buffer).
14. EZ DNA Methylation-Gold Kit (D5005, CA, USA).
15. AMPure XP beads (A63881, Beckman Coulter, CA, USA).
16. 10x PfuTurbo Cx HotStart DNA Polymerase Buffer (600410, Agilent Technologies, CA, USA; supplied 10× Reaction Buffer).
17. Primer F: 5′CAAGCAGAAGACGGCATACG, desalted (Integrated DNA Technologies, IA, USA).
18. Primer R: 5′AATGATACGGCGACCACCGA, desalted (Integrated DNA Technologies, IA, USA).
19. TE buffer: 10 mM Tris-HCl, pH 8.0, 1 mM EDTA.
20. Novex® TBE gel, 6% (EC6265BOX, Thermo Fisher Scientific, NH, USA).
21. 5x Novex® Hi-Density TBE sample buffer (LC6678, Thermo Fisher Scientific, NH, USA).
22. SYBR® Safe DNA gel stain (S33102, Thermo Fisher Scientific, NH, USA).
23. EB Buffer (19086, Qaigen Inc., CA, USA).
24. Amicon Ultra−0.5, Ultracel-30 membrane, 30 kDa (UFC503024, MilliporeSigma, MA, USA).

30.2.2 Samples

For DDMS, we optimized the sample processing with a minimum starting material of 5 nanograms (ng) of genomic DNA. This starting amount is within the range for the minimum amount of the material for commercially available tests, ThyroSeq V3 (2.5 ng of DNA) and GSC (5 ng of RNA) (24, 45). The starting DNA materials for the current protocol are 5 to 20 ng of DNA. The DNA amount is quantified by the dsDNA Quantitation, High Sensitivity assay on Qubit.

30.3 Methods

30.3.1 Library Preparation and NGS Sequencing

For DDMS, genomic library preparation consists of (1) MspI digestion of genomic DNA, (2) end-repair and A-Tailing, (3) ligation with a methylated linker, (4) bisulfite conversion of DNA fragments carrying methylated linker, (5) library amplification, and (6) removal of primer dimers by gel purification. The first step of the genomic library preparation, MspI digestion, is responsible for enriching CG-rich genomic fraction since MspI recognizes and cuts CCGG sites, which are frequently detected in CG-rich areas of the human and mouse genome.

30.3.1.1 MspI Digestion

1. Set up a 30 µl digestion reaction in 0.2ml of PCR tubes as follows:

	Amount
10x NEB2 buffer	3 µl
MspI, 100 000 U/ml	1 µl
5–20 ng DNA in water	26 µl

2. Mix the sample. Incubate overnight at 37 °C in a PCR-cycler machine.

30.3.1.2 End Repair and A-Tailing

1. Set up an end-repair reaction as follows:

	Amount
DNA from 2.1	30 µl
dNTPs (10 mM dATP, 1mM dCTP, 1mM dGTP	1 µl
Klenow Fragment (3′→5′ exo-)	1 µl

2. Incubate for 20 min at 30 °C. Incubate 20 min at 37 °C. Hold at 4 °C.
3. Perform phenol/chloroform purification followed by ethanol precipitation as follows:
   a. Transfer 32 µl of DNA mixture into 1.5 ml of an Eppendorf centrifuge tube.
   b. Add 368 µl of water. Add 400 µl phenol/chloroform/isoamyl alcohol (mix for 3 min).
   c. Centrifuge for 3 min at 18,000 × g at room temperature. Transfer supernatant (upper phase) to a new tube.
   d. Add 400 µl chloroform to the supernatant (upper phase). Mix for 3 min.
   e. Centrifuge for 3 min at 18,000 × g at room temperature. Transfer supernatant (upper phase) into a new tube.
   f. Add 1 µl of 20 mg/ml glycogen and 40 µl of 3 M sodium acetate pH 5.5 to the supernatant. Vortex for 10 sec.
   g. Add 1.1 ml of 96% ethanol. Mix sample. Incubate 30 min on dry ice.
   h. Centrifuge the DNA at 18,000 × g for 20 min at room temperature. Remove supernatant.
   i. Add 700 µl of 80% ethanol and incubate the pellet for 1 min. Remove the supernatant.
   j. Add 700 µl of 80% ethanol and incubate the pellet for 1 min. Remove the supernatant.
   k. Remove residual ethanol. Dry DNA pellet for 7 min at room temperature.
   l. Dissolve DNA in 7.5 µl of water.

30.3.1.3 Linker Ligation

1. Set up a ligation reaction as follows:

	Amount
DNA from 2.2	7.5 µl
10× T4 ligase reaction buffer	1 µl
0.75 µM linker	0.5 µl
T4 ligase, 2000 units/µl	1 µl

2. Incubate at 16 °C for 24–48 hours.
3. Add 20 µl of water into the ligation mixture.
4. Add 60 µl of AMPure XP beads. Mix well.
5. Incubate for 15 min at room temperature.
6. By using a magnet, separate beads from the supernatant and remove the supernatant.
7. Wash beads by using a magnetic rack:
   a. Add 150 µl of 80% ethanol. Incubate for 1 min at room temperature.
   b. Remove ethanol.
8. Repeat the washing step #7 in 80% of ethanol.
9. Dry beads for 7 min at room temperature.
10. Add 20 µl water.
11. Incubate for 2 min at room temperature.
12. Separate the beads from the supernatant using a magnet and transfer the supernatant into a new tube.

30.3.1.4 Bisulfite Conversion

For the CT conversion by using the EZ DNA Methylation-Gold Kit, we modified the manufacturer's protocol. These changes improve DNA recovery from the membrane. Particularly, we doubled the amount of M-Binding Buffer from 600 µl to 1200 µl.

1. To DNA from 2.3, add 130 µl of freshly prepared CT- mix.
2. Incubate:

98 °C	10 min
64 °C	2.5 hours
4 °C	Hold (maximum for 24 hours)

3. Mix 150 µl of CT-treated DNA with 1200 µl of M-binding buffer.
4. Apply 675 µl of the mixture into the Zymo-Spin IC column and centrifuge it at 10,000 × g for 30 seconds. Discard the flow-through.
5. Apply the remaining 675 µl of DNA mixture into the same Zymo-Spin IC column and centrifuge at 10,000 × g for 30 seconds. Discard the flow-through.
6. Apply 100 µl of M-Wash buffer to the column and spin 30 seconds at 10,000 × g. Discard the flow-through.
7. Apply 200 µl of M-Desulphonation Buffer into the column and incubate for 15 minutes at room temperature. After incubation, spin the column at 10,000 × g for 30 seconds. Discard the flow-through.
8. Wash the column with 200 µl of M-wash buffer by centrifugation at 10,000 × g for 30 seconds, twice. Each time, discard the flow-through.
9. Transfer the column to a tube. Add 20 µl of M-elution buffer. Incubate for 5 minutes at room temperature. Centrifuge it at 10,000 × g for 30 seconds.

30.3.1.5 Genomic Library Amplification

1. Set up a PCR reaction as follows:

	Amount
DNA from 2.4	20 µl
10x PfuTurbo Cx Hotstart DNA polymerase reaction buffer	5 µl
10 µM primers R + F	5 µl
2.5 mM dNTPs	5 µl
PfuTurbo Cx Hotstart DNA Pol. 2.5u/µl	1 µl
Water	14 µl

2. Mix well. Amplify samples by using the following conditions:

95 °C	5 min	
98 °C	30 sec	
98 °C	30 sec	repeat 15 times
65 °C	30 sec	
72 °C	30 sec	
72 °C	5 min	
4 °C	hold	

3. To the PCR mixture, add 90 µl of AMPure XP beads. Mix well. Incubate for 15 min at room temperature.
4. By using a magnet, separate beads from the supernatant and remove the supernatant.
5. Wash beads using a magnetic rack. Add 150 µl of 80% ethanol and incubate 1 min at room temperature. Remove ethanol.
6. Repeat step 5.
7. Dry beads for 7 min.
8. Add 20 µl water. Resuspend bead and incubate for 2 min at room temperature.
9. Separate the beads and supernatant by using a magnet.
10. Transfer the supernatant into a new tube.

30.3.1.6 Gel Purification

1. To the amplicon after step 2.5, add 2.5 µl Novex® Hi-Density TBE sample buffer (5×) to the samples. Mix well.
2. Load samples into the wells of a Novex® TBE Gel, 6%.
3. Run gel for 40 min at 200 V.
4. Dye the gel with SYBR® Safe DNA gel stain according to the manufacturer's instructions.
5. Under UV light, cut DNA fragments ranging from 150 to 500 bp (insert size range should be from 30 to 380 bp).
6. Add 500 µl of EB buffer. Smash gel in the buffer. Shake samples overnight at room temperature.
7. Concentrate the eluted DNA with an Amicon Ultra−0.5, Ultracel-30 membrane according to the manufacturer's instructions.
11. Quantify the genomic library with quantitative PCR.

30.3.1.7 NGS Sequencing

1. Perform Illumina paired-end sequencing for at least 50 million reads and perform base calling.
2. Using Bismark, align bisulfite-treated reads to HG19 and determine the methylation levels of individual cytosines (46).

30.3.2 Data Analysis

1. Remove information for cytosines that are associated with the sequencing coverage of less than ten reads.
2. Extract the DNA methylation information for 373 cytosines representing DDMS biomarkers.
3. For each location associated with the benign score, evaluate whether DNA methylation in the sample is associated with the benign-specific pattern (Table 30.1). If the DNA methylation value of the sample is associated with the benign-specific pattern, then the benign score of the sample increases by "1." If the DNA methylation value is not associated with the benign-specific pattern, the benign score remains the same. Perform the analysis for all locations associated with the benign score.
4. Evaluate the malignant score, which starts at "zero" at the beginning of the analysis. For each location associated with the malignant score, evaluate whether DNA methylation in the sample is associated with the malignant-specific pattern (Table 30.1). If the DNA methylation value of the sample is associated with the malignant-specific pattern, then the malignant score of the sample increases by "1." The malignant score remains the same if the DNA methylation value is not associated with the malignant-specific pattern. Perform the analysis for all locations associated with the malignant score.
5. Identify undiagnosable samples. Undiagnosable samples have both benign and malignant scores less than 5 (malignant score < 5 and benign score < 5).
6. Evaluate cancer risk scores: [cancer risk score = (malignant score + 1)/(benign score + 1)].
7. Classify the nodule. If its cancer score is one or more, it is malignant. If it is less than one, it is benign.

30.4 Results

Recently, our team developed a new thyroid nodule diagnostic approach (44) using DNA methylation differences between benign and malignant nodules. We identified 373 cytosines whose DNA methylation status can separate benign from malignant nodules (Figure 30.1).

These 373 genomic locations comprise two sets of cytosines. The first set, the benign signature, is represented by 137 cytosines, which undergo DNA methylation changes in benign thyroid nodules but not malignant ones (Table 30.1). At the same time, the second set, the malignant signature, is associated with 236 cytosines, which undergo DNA methylation

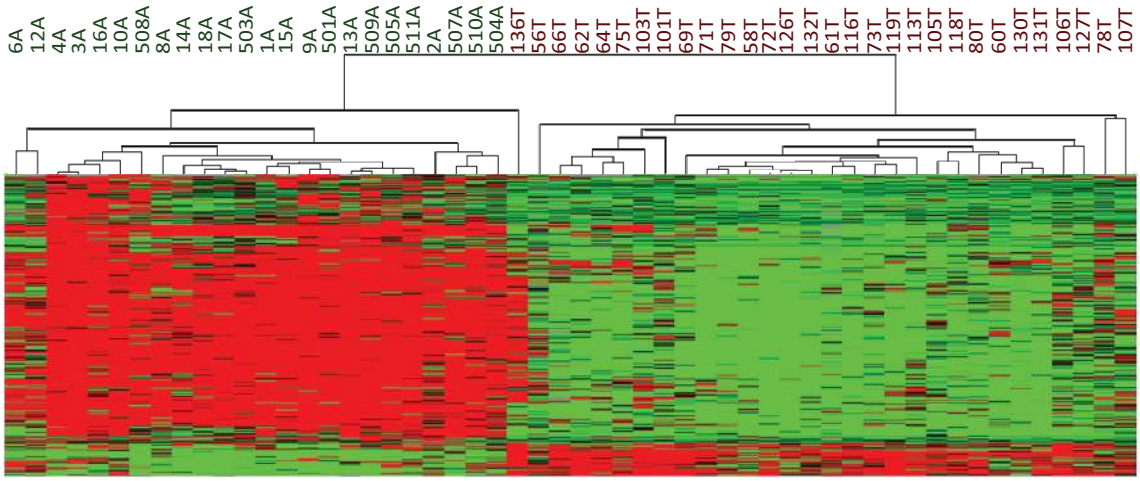


FIGURE 30.1 DNA methylation pattern of 373 cytosines separates benign thyroid nodules from malignant thyroid cancer. Each row represents a single cytosine. Each column indicates surgically resected thyroid nodule. Red, green, and black colors indicate high, low, and medium levels of DNA methylation, respectively. Abbreviation "A" is for benign thyroid nodule and "T" is for thyroid cancer.

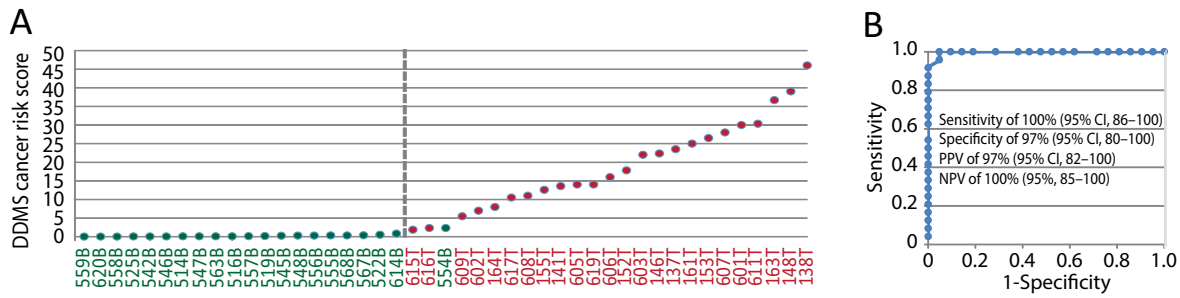


FIGURE 30.2 Performance of DDMS in the validation cohort. A. Cancer risk scores for the thyroid nodules in the validation cohort. The abbreviation "B" marks benign nodule, and "T" is for malignant thyroid nodules. A cancer risk score of 1.0 is indicated. B. Receiver operating characteristics (ROC) curve for DDMS in the validation cohort.

alterations in malignant nodules but not benign nodules (Table 30.1).

DDMS diagnostics contains three steps. The first step in the DDMS classification of thyroid nodules is determining the benign score. To determine the benign score, we take cytosine locations from the benign signature and evaluate how many of these cytosines have a DNA methylation pattern associated with benign nodules (see 30.3.2 Data Analysis). The benign score is the number of cytosines within the benign signature with a DNA methylation pattern specific to the benign nodule.

The second step is to estimate the malignant score. This score is calculated by taking 237 genomic locations associated with the malignant signature and determining how many of those have DNA methylation patterns associated with thyroid malignancies (see 30.3.2 Data Analysis).

The final step of the DDMS classification is calculating a cancer risk score. The last is defined as the ratio of the cancer score and the benign score [cancer risk score = (malignant score + 1)/(benign score + 1)]. According to the DDMS algorithm, nodules with a cancer risk score above or equal to 1.0 are classified as malignant, while thyroid nodules with a cancer score below 1.0 are classified as benign.

It is important to note that some thyroid specimens are not characterized by any of the tissue-specific changes or are associated with very few changes (benign score less than 5 and malignant score less than 5). According to the DDMS algorithm, this type of thyroid nodules is classified as undiagnosable. The reason for this definition is that (1) DDMS is based on the DNA methylation signatures specific to thyroid nodules, and (2) undiagnosable specimens are missing these signatures.

According to our validation, undiagnosable samples are frequently associated with contamination of the nodule cells with other cell types, including cells from the adjacent normal tissue and infiltration of lymphocytes (Hashimoto's thyroiditis) (44). In fact, according to our analysis, most adjacent thyroid tissue samples are undiagnosable.

We validated the DDMS assay using a retrospective cohort containing 37 benign and 41 malignant thyroid nodules, which were surgically resected. Of 78 specimens, 13 were undiagnosable, out of which eight were associated with Hashimoto's thyroiditis. According to DDMS, 30 out of 31 diagnosable benign nodules had a cancer risk score of less than 1.0 and were classified as benign (Figure 30.2A). At the same, according to DDMS, all 34 diagnosable malignant nodules had a

cancer risk score of 1.0 or more and were classified as malignant. Thus, DDMS is associated with an estimated sensitivity of 100% (95% CI, 86%–100%), specificity of 97% (95% CI, 80%–100%), PPV of 97% (95% CI, 82%–100%), and NPV of 100% (95% CI, 85%–100%) (Figure 30.2B). These data suggest that DDMS and DNA methylation biomarkers have a great potential for thyroid nodule diagnosis, which would lead to eradicating diagnostic thyroidectomies.

30.5 Discussion

Preoperative diagnosis is very challenging for thyroid nodules with indeterminate cytopathology. In fact, diagnosis of these nodules can be difficult even after these nodules are surgically resected. The most recent World Health Organization (WHO) classification introduced a definition of "tumors of uncertain malignant potential" for surgically resected nodules with borderline features between benign and malignant tumors (47). There is an urgent need for molecular biomarkers to differentiate thyroid nodules with indeterminate cytology. In fact, this diagnostic is needed not only for preoperative diagnostics of indeterminate nodules but also for postoperative diagnostics of tumors of uncertain malignant potential.

Current molecular diagnostic tests for thyroid nodules are associated with high NPV and low PPV values and work as rule-out tests (Table 30.1). The low PPV leads to overdiagnosis and results in unnecessary thyroidectomies for incorrectly diagnosed benign nodules. Also, the low PPV makes the current molecular tests less popular among physicians. Therefore, instead of molecular testing, these physicians perform a diagnostics thyroidectomy, removing half or the entire thyroid. After surgery, 85% of these cases are found to be benign (48). The new DDMS assay performs excellently in the cohort with surgically resected nodules. The PPV of DDMS is very promising compared to the currently available commercial tests. However, DDMS needs to be tested in the real world by using a prospective collection of thyroid nodules, which is currently undergoing. If DDMS keeps its performance in the prospective FNA collection, it can compete with current market leaders for thyroid nodule diagnostics.

Our test is associated with 16.7% of undiagnosable samples, whereas 10.2% of undiagnosable samples are associated with thyroiditis. To reduce the number of undiagnosable FNAs, we are considering combining DDMS testing with BRAFV600E mutation screening, which is currently incorporated in GSC and ThyroSeq V3 (16, 17, 24).

While other commercial tests are based on the search for cancer-specific alterations, DDMS identifies features specific to malignant and benign nodules. Both benign and malignant DDMS signatures represent positive controls for the substantial presence of thyroid nodule fragments in the FNA specimen. It is important to note that according to our DNA methylation data, there are differences between thyroid nodules and adjacent thyroid tissue (44). While other commercial tests like ThyroSeq V3 and GSC confirm the presence of thyroid epithelial cells by analyzing the expressions of genes expressed in adjacent thyroid and thyroid nodules (49, 50), DDMS is based on the features specific to thyroid nodules but not adjacent thyroid tissue. This fact results in the reduction of the false negative results associated with adjacent thyroid tissue, which can mistakenly be collected during thyroid FNA. It is important to note that thyroid FNAs are used to diagnose nodules with sizes as small as 5 mm. Since the procedure is performed under ultrasound guidance, there is a chance that the normal thyroid adjacent to the nodule will be collected. According to DDMS, specimens that do not have cancer-specific changes are not necessarily considered benign. These specimens will be considered undiagnosable if the nodules do not demonstrate features specific to benign nodules. Our data demonstrated that 62% of adjacent normal thyroid tissue were classified as non-diagnostic by DDMS (44). Thus, the introduction of the benign DDMS signature reduces false negative results.

In this regard, it is important to note that DDMS is based on the DNA methylation differences between benign and malignant thyroid nodules and not on the difference between malignant nodules and adjacent thyroid tissue. Our analysis indicated that benign and adjacent thyroid tissue have very different DNA methylation patterns, which makes the use of adjacent thyroid tissue unsuitable for developing preoperative thyroid nodule testing.

DDMS requires a very small starting material, at least 5 ng DNA, like the most popular preoperative thyroid cancer diagnostics, ThyroSeq V3 and GSC. However, unlike GSC testing, which is based on RNA and requires special handling, DDMS is based on stable epigenetic mark, DNA methylation. This fact provides an advantage for DDMS over RNA-based assays. Due to the stability of DNA methylation, DDMS can be performed on FFPE blocks and applied to analyze surgical specimens with uncertain malignant potential.

In summary, DDMS is a new, highly accurate test for diagnosing indeterminate thyroid nodules based on DNA methylation differences between benign and malignant nodules. While demonstrating outstanding performance in surgical specimens, the test must be validated in a prospective clinical study with actual FNA samples to prove its outstanding performance.

30.6 Future Trends

Epigenetics determines a cell's identity. In cancer, epigenetic alterations, including DNA methylation, reflect cancer-specific transcriptional alterations. From the perspective of cancer diagnostics, DNA methylation analysis holds great promise. In contrast to cancer-specific gene expression analysis, DNA methylation does not require special sample handling due to its stability. While genetic alterations are crucial in cancer development and are easily detected, not every cancer sample is characterized by cancer-specific genetic alterations. At the same time, some mutations detected in cancer, like RAS mutations, are frequently found in benign tumors; these facts lead to the limitations of cancer diagnostics based on genetics and expressional alterations and make analysis of DNA methylation an ideal approach for cancer diagnostics.

While our technology, DDMS, provides very high accuracy in diagnosing thyroid nodules, it does not classify thyroid cancers by subtype, such as papillary thyroid cancer,

follicular cancer, and anaplastic thyroid cancer. DDMS classification also does not reflect the metastatic potential of the tumor. The next step for thyroid cancer diagnostic development by using DNA methylation will be providing not only an accurate classification of thyroid nodules, but also classification of subtypes of thyroid cancer and determination of metastatic risk. These improvement strategies can possibly improve thyroid cancer diagnostics, reduce unnecessary diagnostic thyroidectomies, and provide a better treatment for thyroid cancer patients.

REFERENCES

1. Tan GH, Gharib H. Thyroid incidentalomas: Management approaches to nonpalpable nodules discovered incidentally on thyroid imaging. *Ann Intern Med*. 1997 Feb 1;126(3):226–31. doi: 10.7326/0003-4819-126-3-199702010-00009.
2. Guth S, Theune U, Aberle J, Galach A, Bamberger CM. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur J Clin Invest*. 2009 Aug;39(8):699–706. doi: 10.1111/j.1365-2362.2009.02162.x.
3. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM. Revised American thyroid association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. American thyroid association (ATA) guidelines taskforce on thyroid nodules and differentiated thyroid cancer. *Thyroid*. 2009 Nov;19(11):1167–214.
4. https://seer.cancer.gov/statfacts/html/thyro.html.
5. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. *CA Cancer J Clin*. 2024 Jan–Feb;74(1):12–49. doi: 10.3322/caac.21820. Epub 2024 Jan 17.
6. Davies L, Welch HG. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg*. 2014 Apr;140(4):317–22. doi: 10.1001/jamaoto.2014.1.
7. Boucai L, Zafereo M, Cabanillas ME. Thyroid cancer: A review. *JAMA*. 2024 Feb 6;331(5):425–35. doi: 10.1001/jama.2023.26348.
8. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA*. 2017 Apr 4;317(13):1338–48. doi: 10.1001/jama.2017.2719.
9. Kwak JY, Han KH, Yoon JH, Moon HJ, Son EJ, Park SH, Jung HK, Choi JS, Kim BM, Kim EK. Thyroid imaging reporting and data system for US features of nodules: A step in establishing better stratification of cancer risk. *Radiology*. 2011 Sep;260(3):892–9. doi: 10.1148/radiol.11110206. Epub 2011 Jul 19.
10. Ali SZ, Baloch ZW, Cochand-Priollet B, Schmitt FC, Vielh P, VanderLaan PA. The 2023 bethesda system for reporting thyroid cytopathology. *Thyroid*. 2023 Sep;33(9):1039–44. doi: 10.1089/thy.2023.0141. Epub 2023 Jul 8.
11. Prades JM, Querat C, Dumollard JM, Richard C, Gavid M, Timoshenko AA, Peoc'h M, Martin C. Thyroid nodule surgery: Predictive diagnostic value of fine-needle aspiration cytology and frozen section. *Eur Ann Otorhinolaryngol Head Neck Dis*. 2013 Sep;130(4):195–9. doi: 10.1016/j.anorl.2012.12.005. Epub 2013 May 20.
12. Li H, Robinson KA, Anton B, Saldanha IJ, Ladenson PW. Cost-effectiveness of a novel molecular test for cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metable*. 2011 Nov; 96(11): E1719–26. doi: 10.1210/jc.2011-0459. Epub 2011 Aug 24.
13. Chan AC, Lang BH, Wong KP. The pros and cons of routine central compartment neck dissection for clinically nodal negative (cN0) papillary thyroid cancer. *Gland Surg*. 2013 Nov;2(4):186–95. doi: 10.3978/j.issn.2227-684X.2013.10.10.
14. Hambleton C, Kandil E. Appropriate and accurate diagnosis of thyroid nodules: A review of thyroid fine-needle aspiration. *Int J Clin Exp Med*. 2013 Jun 26;6(6):413–22. Print 2013.
15. Livhits MJ, Zhu CY, Kuo EJ, Nguyen DT, Kim J, Tseng CH, Leung AM, Rao J, Levin M, Douek ML, Beckett KR, Cheung DS, Gofnung YA, Smooke-Praw S, Yeh MW. Effectiveness of molecular testing techniques for diagnosis of indeterminate thyroid nodules: A randomized clinical trial. *JAMA Oncol*. 2021 Jan 1;7(1):70–7. doi: 10.1001/jamaoncol.2020.5935.
16. Patel KN, Angell TE, Babiarz J, Barth NM, Blevins T, Duh QY, Ghossein RA, Harrell RM, Huang J, Kennedy GC, Kim SY, Kloos RT, LiVolsi VA, Randolph GW, Sadow PM, Shanik MH, Sosa JA, Traweek ST, Walsh PS, Whitney D, Yeh MW, Ladenson PW. Performance of a genomic sequencing classifier for the preoperative diagnosis of cytologically indeterminate thyroid nodules. *JAMA Surg*. 2018 Sep 1;153(9):817–24. doi: 10.1001/jamasurg.2018.1153. PMID: 29799911, PMCID: PMC6583881.
17. Hao Y, Duh QY, Kloos RT, Babiarz J, Harrell RM, Traweek ST, Kim SY, Fedorowicz G, Walsh PS, Sadow PM, Huang J, Kennedy GC. Identification of Hürthle cell cancers: Solving a clinical challenge with genomic sequencing and a trio of machine learning algorithms. *BMC Syst Biol*. 2019 Apr 5;13(Suppl 2):27. doi: 10.1186/s12918-019-0693-z. PMID: 30952205, PMCID: PMC6450053.
18. Angell TE, Heller HT, Cibas ES, Barletta JA, Kim MI, Krane JF, Marqusee E. Independent comparison of the afirma genomic sequencing classifier and gene expression classifier for cytologically indeterminate thyroid nodules. *Thyroid*. 2019 May;29(5):650–56. doi: 10.1089/thy.2018.0726. Epub 2019 Mar 22.
19. Endo M, Nabhan F, Porter K, Roll K, Shirley LA, Azaryan I, Tonkovich D, Perlick J, Ryan LE, Khawaja R, Meng S, Phay JE, Ringel MD, Sipos JA. Afirma gene sequencing classifier compared with gene expression classifier in indeterminate thyroid nodules. *Thyroid*. 2019 Aug;29(8):1115–24. doi: 10.1089/thy.2018.0733. Epub 2019 Jul 17.
20. San Martin VT, Lawrence L, Bena J, Madhun NZ, Berber E, Elsheikh TM, Nasr CE. Real-world comparison of afirma GEC and GSC for the assessment of cytologically indeterminate thyroid nodules. *J Clin Endocrinol Metable*. 2020 Mar 1;105(3):dgz099. doi: 10.1210/clinem/dgz099.
21. Vuong HG, Nguyen TPX, Hassell LA, Jung CK. Diagnostic performances of the afirma gene sequencing classifier in comparison with the gene expression classifier: A meta-analysis. *Cancer Cytopathol*. 2021 Mar;129(3):182–189. doi: 10.1002/cncy.22332. Epub 2020 Jul 29.
22. Zhang L, Smola B, Lew M, Pang J, Cantley R, Pantanowitz L, Heider A, Jing X. Performance of afirma genomic sequencing classifier vs gene expression classifier in

Bethesda category III thyroid nodules: An institutional experience. *Diagn Cytopathol*. 2021 Aug;49(8):921–7. doi: 10.1002/dc.24765. Epub 2021 May 22.

23. Polavarapu P, Fingeret A, Yuil-Valdes A, Olson D, Patel A, Shivaswamy V, Matthias TD, Goldner W. Comparison of afirma GEC and GSC to nodules without molecular testing in cytologically indeterminate thyroid nodules. *J Endocr Soc*. 2021 Oct 7;5(11):bvab148. doi: 10.1210/jendso/bvab148. eCollection 2021 Nov 1.
24. Nikiforova MN, Mercurio S, Wald AI, Barbi de Moura M, Callenberg K, Santana-Santos L, Gooding WE, Yip L, Ferris RL, Nikiforov YE. Analytical performance of the ThyroSeq v3 genomic classifier for cancer diagnosis in thyroid nodules. *Cancer*. 2018 Apr 15;124(8):1682–90. doi: 10.1002/cncr.31245. Epub 2018 Jan 18.
25. Steward DL, Carty SE, Sippel RS, Yang SP, Sosa JA, Sipos JA, Figge JJ, Mandel S, Haugen BR, Burman KD, Baloch ZW, Lloyd RV, Seethala RR, Gooding WE, Chiosea SI, Gomes-Lima C, Ferris RL, Folek JM, Khawaja RA, Kundra P, Loh KS, Marshall CB, Mayson S, McCoy KL, Nga ME, Ngiam KY, Nikiforova MN, Poehls JL, Ringel MD, Yang H, Yip L, Nikiforov YE. Performance of multigene genomic classifier in thyroid nodules with indeterminate cytology: A prospective blinded multicenter study. *JAMA Oncol*. 2019 Feb 1;5(2):204–12. doi: 10.1001/jamaoncol.2018.4616.
26. Desai D, Lepe M, Baloch ZW, Mandel SJ. ThyroS2eq v3 for Bethesda III and IV: An institutional experience. *Cancer Cytopathol*. 2021 Feb;129(2):164–70. doi: 10.1002/cncy.22362. Epub 2020 Oct 8.
27. Yang SP, Nga ME, Bundele MM, Chiosea SI, Tan SH, Lum JHY, Parameswaran R, Lim MY, Li H, Cheah WK, Sek KS, Tan ATH, Loh TKS, Ngiam KY, Tan WB, Huang X, Ho TWT, Lim KH, Lim CM, Singaporewalla RM, Rao AD, Rao NCL, Chua DYK, Chin DC, Wald AI, LiVolsi VA, Nikiforov YE, Tai ES. Performance of a multigene genomic classifier and clinical parameters in predicting malignancy in a Southeast Asian cohort of patients with cytologically indeterminate thyroid nodules. *Cancer Cytopathol*. 2024 Feb 6. doi: 10.1002/cncy.22796.
28. Desai D, Lepe M, Baloch ZW, Mandel SJ. ThyroSeq v3 for Bethesda III and IV: An institutional experience. *Cancer Cytopathol*. 2021 Feb;129(2):164–70. doi: 10.1002/cncy.22362. Epub 2020 Oct 8.
29. Jug R, Foo WC, Jones C, Ahmadi S, Jiang XS. High-risk and intermediate-high-risk results from the ThyroSeq v2 and v3 thyroid genomic classifier are associated with neoplasia: Independent performance assessment at an academic institution. *Cancer Cytopathol*. 2020 Aug;128(8):563–69. doi: 10.1002/cncy.22283. Epub 2020 Apr 2.
30. Ablordeppey KK, Timmaraju VA, Song-Yang JW, Yaqoob S, Narick C, Mireskandari A, Finkelstein SD, Kumar G. Development and analytical validation of an expanded mutation detection panel for next-generation sequencing of thyroid nodule aspirates. *J Mol Diagn*. 2020 Mar;22(3):355–67. doi: 10.1016/j.jmoldx.2019.11.003. Epub 2019 Dec 19.
31. Finkelstein SD, Sistrunk JW, Malchoff C, Thompson DV, Kumar G, Timmaraju VA, Repko B, Mireskandari A, Evoy-Goodman LA, Massoll NA, Lupo MA. A retrospective evaluation of the diagnostic performance of an interdependent pairwise MicroRNA expression analysis with a mutation panel in indeterminate thyroid nodules. *Thyroid*. 2022 Nov;32(11):1362–71. doi: 10.1089/thy.2022.0124. Epub 2022 Sep 21.
32. Kandil E, Metz TA, Issa PP, Aboueisha M, Omar M, Attia AS, Chabot B, Hussein M, Moroz K, Shama M, Toraih E. Diagnostic performance of afirma and interpace diagnostics genetic testing in indeterminate thyroid nodules: A single center study. *Cancers (Basel)*. 2023 Mar 31;15(7):2098. doi: 10.3390/cancers15072098.
33. Santos MTD, Buzolin AL, Gama RR, Silva ECAD, Dufloth RM, Figueiredo DLA, Carvalho AL. Molecular classification of thyroid nodules with indeterminate cytology: Development and validation of a highly sensitive and specific new miRNA-based classifier test using fine-needle aspiration smear slides. *Thyroid*. 2018 Dec;28(12):1618–26. doi: 10.1089/thy.2018.0254. Epub 2018 Nov 22.
34. Santos MT, Rodrigues BM, Shizukuda S, Oliveira AF, Oliveira M, Figueiredo DLA, Melo GM, Silva RA, Fainstein C, Dos Reis GF, Corbo R, Ramos HE, Camacho CP, Vaisman F, Vaisman M. Clinical decision support analysis of a microRNA-based thyroid molecular classifier: A real-world, prospective and multicentre validation study. *EBioMedicine*. 2022 Aug;82:104137. doi: 10.1016/j.ebiom.2022.104137. Epub 2022 Jul 1.
35. González HE, Martínez JR, Vargas-Salas S, Solar A, Veliz L, Cruz F, Arias T, Loyola S, Horvath E, Tala H, Traipe E, Meneses M, Marín L, Wohllk N, Diaz RE, Véliz J, Pineda P, Arroyo P, Mena N, Bracamonte M, Miranda G, Bruce E, Urra S. A 10-gene classifier for indeterminate thyroid nodules: Development and multicenter accuracy study. *Thyroid*. 2017 Aug;27(8):1058–67. doi: 10.1089/thy.2017.0067. Epub 2017 Jul 11.
36. Zafereo M, McIver B, Vargas-Salas S, Domínguez JM, Steward DL, Holsinger FC, Kandil E, Williams M, Cruz F, Loyola S, Solar A, Roa JC, León A, Droppelman N, Lobos M, Arias T, Kong CS, Busaidy N, Grubbs EG, Graham P, Stewart J, Tang A, Wang J, Orloff L, Henríquez M, Lagos M, Osorio M, Schachter D, Franco C, Medina F, Wohllk N, Diaz RE, Veliz J, Horvath E, Tala H, Pineda P, Arroyo P, Vasquez F, Traipe E, Marín L, Miranda G, Bruce E, Bracamonte M, Mena N, González HE. A thyroid genetic classifier correctly predicts benign nodules with indeterminate cytology: Two independent, multicenter, prospective validation trials. *Thyroid*. 2020 May;30(5):704–12. doi: 10.1089/thy.2019.0490. Epub 2020 Feb 11.
37. Olmos R, Domínguez JM, Vargas-Salas S, Mosso L, Fardella CE, González G, Baudrand R, Guarda F, Valenzuela F, Arteaga E, Forenzano P, Nilo F, Lustig N, Martínez A, López JM, Cruz F, Loyola S, Leon A, Droppelmann N, Montero P, Domínguez F, Camus M, Solar A, Zoroquiain P, Roa JC, Muñoz E, Bruce E, Gajardo R, Miranda G, Riquelme F, Mena N, González HE. ThyroidPrint: Clinical utility for indeterminate thyroid cytology. *Endocr Relat Cancer*. 2023 Oct 6;30(11):e220409. doi: 10.1530/ERC-22-0409. Print 2023 Nov 1.
38. Nikiforov YE, Nikiforova MN. Molecular genetics and diagnosis of thyroid cancer. *Nat Rev Endocrinol*. 2011 Aug 30;7(10):569–80. doi: 10.1038/nrendo.2011.142.
39. Marotta V, Guerra A, Sapio MR, Vitale M. RET/PTC rearrangement in benign and malignant thyroid diseases: A clinical standpoint. *Eur J Endocrinol*. 2011 Oct;165(4):499–507. doi: 10.1530/EJE-11-0499. Epub 2011 Jul 12.

40. Bardet S, Goardon N, Lequesne J, Vaur D, Ciappuccini R, Leconte A, Monpeyssen H, Saguet-Rysanek V, Clarisse B, Lasne-Cardon A, Ménégaux F, Leenhardt L, Buffet C. Diagnostic and prognostic value of a 7-panel mutation testing in thyroid nodules with indeterminate cytology: the SWEETMAC study. *Endocrine*. 2021 Feb;71(2):407–17. doi: 10.1007/s12020-020-02411-4. Epub 2020 Jul 7.
41. Marques AR, Espadinha C, Catarino AL, Moniz S, Pereira T, Sobrinho LG, Leite V. Expression of PAX8-PPAR gamma 1 rearrangements in both follicular thyroid carcinomas and adenomas. *J Clin Endocrinol Metable*. 2002 Aug;87(8):3947–52. doi: 10.1210/jcem.87.8.8756.
42. Cheung L, Messina M, Gill A, Clarkson A, Learoyd D, Delbridge L, Wentworth J, Philips J, Clifton-Bligh R, Robinson BG. Detection of the PAX8-PPAR gamma fusion oncogene in both follicular thyroid carcinomas and adenomas. *J Clin Endocrinol Metable*. 2003 Jan;88(1):354–7. doi: 10.1210/jc.2002-021020.
43. Hahn MA, Li AX, Wu X, Yang R, Drew DA, Rosenberg DW, Pfeifer GP. Loss of the polycomb mark from bivalent promoters leads to activation of cancer-promoting genes in colorectal tumors. *Cancer Res*. 2014 Jul 1;74(13):3617–29. doi: 10.1158/0008-5472.CAN-13-3147. Epub 2014 May 1.
44. Yim JH, Choi AH, Li AX, Qin H, Chang S, Tong ST, Chu P, Kim BW, Schmolze D, Lew R, Ibrahim Y, Poroyko VA, Salvatierra S, Baker A, Wang J, Wu X, Pfeifer GP, Fong Y, Hahn MA. Identification of tissue-specific DNA methylation signatures for thyroid nodule diagnostics. *Clin Cancer Res*. 2019 Jan 15;25(2):544–51. doi: 10.1158/1078-0432. CCR-18-0841. Epub 2018 Aug 9.
45. Hao Y, Choi Y, Babiarz JE, Kloos RT, Kennedy GC, Huang J, Walsh PS. Analytical verification performance of afirma genomic sequencing classifier in the diagnosis of cytologically indeterminate thyroid nodules. *Front Endocrinol (Lausanne)*. 2019 Jul 4;10:438. doi: 10.3389/fendo.2019.00438. eCollection 2019.
46. Krueger F, Andrews SR. Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011 Jun 1;27(11):1571–2. doi: 10.1093/bioinformatics/btr167. Epub 2011 Apr 14.
47. Lloyd RV, Osamura RY, Kloppel G, Rosai J (ed). *Tumours of the thyroid gland. WHO classification of tumours of endocrine organs* (4th). IARC Publications, Lyon, France, 2017: 65–143.
48. Sherman SI. Thyroid carcinoma. *Lancet*. 2003 Feb 8; 361(9356):501–11. doi: 10.1016/s0140-6736(03)12488-9. PMID: 12583960.
49. Nikiforov YE, Ohori NP, Hodak SP, Carty SE, LeBeau SO, Ferris RL, et al. Impact of mutational testing on the diagnosis and management of patients with cytologically indeterminate thyroid nodules: A prospective analysis of 1056 FNA samples. *J Clin Endocrinol Metable*. 2011;96:3390–7.
50. Chudova D, Wilde JI, Wang ET, Wang H, Rabbee N, Egidio CM, Reynolds J, Tom E, Pagan M, Rigl CT, Friedman L, Wang CC, Lanman RB, Zeiger M, Kebebew E, Rosai J, Fellegara G, LiVolsi VA, Kennedy GC. Molecular classification of thyroid nodules using high-dimensionality genomic data. *J Clin Endocrinol Metable*. 2010 Dec;95(12):5296–304. doi: 10.1210/jc.2010-1087. Epub 2010 Sep 8. PMID: 20826580.

31

Utilizing RADAR for the Integration of Genomic Data and Immunotherapy Targets in Precision Oncology for Multiple Myeloma

Alessandro Laganà, Sanjeev Sariya, David T. Melnekoff, Meghana Ram, and Samir Parekh

31.1 Introduction

Precision oncology is a subspecialty of precision medicine, a cutting-edge approach focused on delivering highly individualized diagnostic, prognostic, and therapeutic solutions for cancer patients [1–3]. This innovative paradigm tailors strategies to the unique molecular and genetic profile of each patient's tumor rather than the broader histology of the cancer. Central to precision oncology are actionable alterations—specific genetic or molecular aberrations that can be precisely targeted by therapeutic agents or serve as biomarkers to forecast a tumor's responsiveness to particular treatments.

The advent of precise, cost-effective high-throughput sequencing technologies over the past decade has been a catalyst for breakthroughs in precision oncology, which is now transitioning into standard of care, marking a new era in clinical oncology [4]. The pillars of precision oncology are the state-of-the-art technologies that generate extensive biological data and the sophisticated computational frameworks that interpret this data. Sequencing of DNA and RNA identifies the aberrations within the cancer cell genome and transcriptome, informing the classification of cancer subtypes, the gauging of disease risk, and the customization of tailored treatment regimens [5].

Despite the gaps in our comprehension of cancer's complexities and the current reality that only a small segment of patients reaps the full benefits of such targeted approaches, the translation of research discoveries to clinical interventions is accelerating. This has been crucial in the swift progression of novel treatment modalities and their tailored clinical deployment.

Multiple myeloma (MM) is a cancer of the plasma cells found in bone marrow that produce antibodies [6]. Despite progress in understanding MM's genetic landscape and treatment advancements, the majority of the patients face a relapsing clinical course and run out of FDA-approved options, which underscores the potential benefits of precision medicine approaches [7]. In a recent pilot study at our center, we demonstrated that next-generation sequencing (NGS) can guide the treatment of patients with relapsed/refractory MM [8]. Using a new software platform, DAPHNI, we combined whole-exome sequencing (WES) with RNA-seq to pinpoint actionable changes like single nucleotide variants (SNV), copy number alterations (CNA), and disrupted gene expression and pathways. The study highlighted that while DNA changes can guide treatment, *RNA analysis often uncovers additional therapeutic options not detectable by DNA analysis alone*. Out of 21 patients treated according to our pipeline's recommendations, 16 (around 76%) showed a clinical response, defined as a 25% to 49% reduction in disease markers (Figure 31.1A). These patients were treated with various drugs, including trametinib, which targets MEK in the MAPK pathway when mutations in NRAS, KRAS, or BRAF are present, and venetoclax, prescribed for patients with elevated BCL2 expression.

Building upon our initial work, the DAPHNI platform has undergone significant advancements and has evolved into the research and data analysis resource for precision oncology (RADAR) system. (1) First, we created a faster and scalable pipeline employing Nextflow, a workflow tool that offers data-driven and reproducible computational pipelines, enabling us to handle complex analyses with ease and adaptability.

(2) Our collection of actionable alterations has expanded to include a broader range of databases beyond Clinical Interpretation of Variants in Cancer (CIViC) [9]. The analysis of SNV and CNA data now uncovers the complex subclonal structure of each tumor. Gene expression analysis has been refined for more precise pathway activity measurement. Additionally, we have implemented new indicators for disease progression risk and genomic stability. (3) We have also broadened our scope to include germline variant screening, focusing on recent results from our group identifying patients at higher risk for cancer predisposition. (4) We are now evaluating predictive markers for response to emerging immunotherapies like chimeric antigen receptor (CAR) T cells and bispecific antibodies, which have shown remarkable efficacy in treatment-resistant MM. Our platform offers insights into genetic and transcriptomic variations critical for tailoring immunotherapy sequences.

Based on our experience with integrative analysis of multi-omics data, we designed RADAR to incorporate additional layers of different data types, which can be both omics data and immune assays for enhancing patient stratification and treatment guidance. This includes integrating flow cytometry of the bone marrow tumor microenvironment (TME) and T cell killing assays.

DOI: 10.1201/9781003449942-35

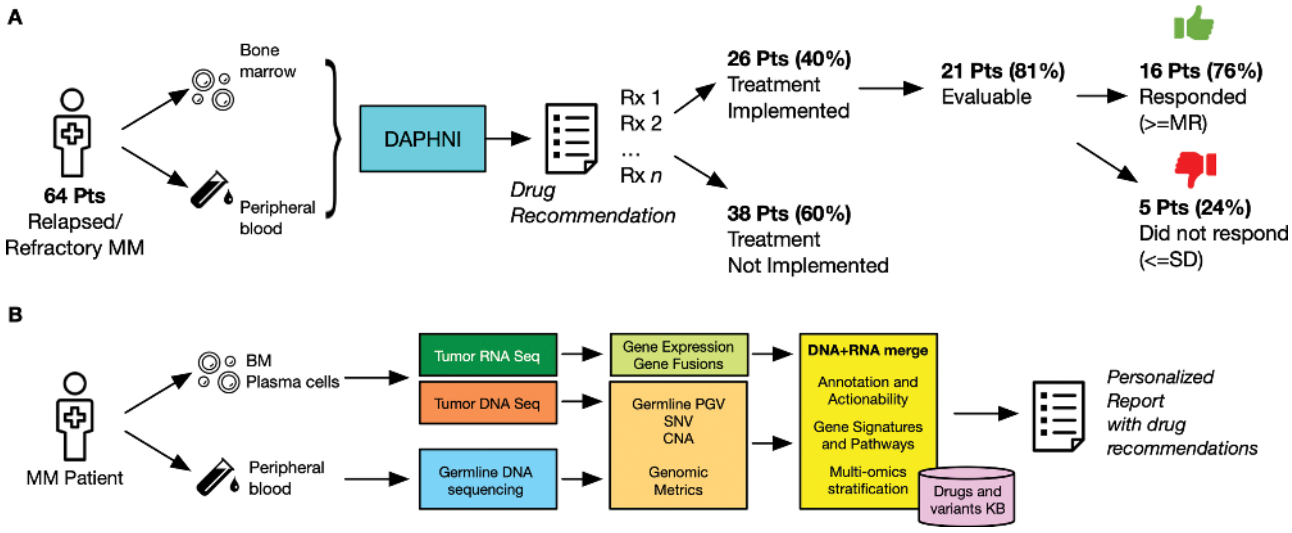


FIGURE 31.1 *Schematic of the DAPHNI trial and RADAR workflow.* (A) We recruited 64 patients (pts) with relapsed and/or refractory multiple myeloma (MM) treated at the Mount Sinai Hospital. RNAseq was performed on the majority, with WES and targeted DNA panels used for subsets of the group. Analysis through the DAPHNI pipeline generated treatment suggestions. Following these recommendations, a proportion of patients experienced clinical benefits, with the majority showing a response as per IMWG standards. Others maintained stable disease status or did not respond to the treatment regimen. (B) RADAR is a precision oncology platform specifically designed for MM. In this platform, both bone marrow (BM) tumor plasma cells and peripheral blood cells undergo whole-exome sequencing (WES) or whole genome sequencing (WGS). Additionally, the tumor is subjected to RNA-seq for a comprehensive transcriptomic analysis. The sequencing data is then analyzed to identify somatic single nucleotide variants (SNVs), copy number alterations (CNAs), and pathogenic germline variants (PGVs). Gene expression and gene fusions are detected from the RNA-seq data. The DNA and RNA data are annotated, and actionable variants, gene signatures, and pathways are investigated. This data is used to categorize the patient into one of 12 disease subtypes. The workflow results in personalized reports that highlight actionable variants, prognostic findings, and provide therapy recommendations.

This chapter outlines RADAR's workflow, from the processing of raw sequencing data to the generation of personalized patient reports, and presents case studies to illustrate its clinical and research applications.

31.2 Materials

RADAR is a platform designed to advance precision medicine for myeloma, through the analysis of NGS data from patients' bone marrow and blood samples. It features two comprehensive workflows for DNA and RNA analysis, respectively, both implemented using Nextflow [10]. Optimized for WES and RNA-seq data, RADAR can also accommodate whole-genome sequencing (WGS). The platform operates on Linux-based systems with support for Python and R, and integrates a mix of third-party and proprietary tools for data processing and analysis. Detailed descriptions of each workflow step are provided in the Methods section.

31.3 Methods

31.3.1 Quality Control

An initial step in the RADAR pipeline focuses on quality control (QC) and data filtering, crucial for ensuring the integrity of our sequencing reads. We use the tool fastp for tasks like base correction, trimming low-quality ends, and discarding reads with low Phred quality score [11]. Additionally, we leverage NGS-Checkmate to verify that tumor and normal samples match correctly, addressing potential mix-ups in sample handling [12]. This step is crucial for the accuracy of our analysis.

31.3.2 Pre-Processing of DNA Data

The DNA workflow begins with aligning high-quality reads to the reference genome using the Burrows-Wheeler Aligner (BWA), specifically utilizing the GRCh38 version of the Ensembl reference genome [13]. Subsequently, the Genome Analysis Toolkit (GATK) Best Practices for variant discovery are employed [14]. During pre-processing, functions such as "FixMateInformation" ensure proper formatting and synchronization of mate-pairing information, "MarkDuplicates" identifies and labels duplicate reads to prevent overestimation of read depth, and "BaseRecalibrator" adjusts base quality scores to minimize systematic errors in variant calling. Additionally, we use bamCoverage from deepTools to generate a coverage track in bigWig format, which aids in visualizing the distribution of sequencing reads across the genome. The "CollectMultipleMetrics" function is then utilized to produce comprehensive metrics on alignment, base distribution by cycle, insert size, quality by cycle, and quality distribution, providing a detailed overview of data quality essential for downstream analysis.

31.3.3 Germline Variant Discovery

We detect germline variants in pre-processed normal BAM files using the GATK tool "HaplotypeCaller," which calls

single nucleotide polymorphisms (SNPs) and indels simultaneously through local *de novo* assembly of haplotypes in active regions. Variants are then scored with "CNNScoreVariants," utilizing a convolutional neural network approach to assess their quality. Subsequently, "FilterVariantTranches" applies filters based on a machine learning model to remove artifact-prone variants. The high-confidence variants, outputted in VCF format, are annotated using Funcotator, which provides detailed insights into each variant's functional impact, linking genomic data with available biological information and potential clinical relevance.

31.3.4 Somatic Mutation Calling

For the identification of somatic single nucleotide variants, RADAR utilizes a consensus methodology harnessing the tools muTect2, Strelka2, and Lancet [15–17]. This dual-mode module can retain mutations identified by all tools for a comprehensive analysis or only those confirmed by at least two tools for enhanced specificity.

MuTect2 is executed with default settings, and the subsequent step involves the GATK function "FilterMutectCalls" to refine the call set based on multiple metrics for increased accuracy.

Lancet processes each chromosome individually, adhering to default guidelines. Post-processing includes using "UpdateVCFSequenceDictionary" to synchronize VCFs with the reference genome and "GatherVcfs" to merge the chromosome-specific files into a singular variant file.

Strelka2's precursor, Manta, identifies structural variant candidates, which Strelka2 further refines to pinpoint SNVs and indels with precision [18].

Subsequently, a custom script merges VCFs from the three tools, performing either the union or the intersection of variants. To address varying VAFs for identical mutations across tools, "bam-readcount" recalculates VAFs from BAM files for uniformity.

The consensus VCF undergoes filtering to retain only SNVs exceeding a VAF of 0.05 and with a minimum of ten tumor reads. Funcotator then annotates the filtered consensus VCF to provide functional context to each variant.

This systematic approach ensures that only the most relevant and verifiable somatic mutations are considered, generating high-quality data for downstream analysis and patient reports.

31.3.5 Identification of Copy Number Alterations

For copy number alterations calling in our RADAR platform, we employ an approach that begins with snp-pileup to generate a pileup of SNPs from sequencing data, ensuring high-quality bases and mapping. This step is crucial for accurately identifying regions with potential copy number changes. The tool FACETS then analyzes the pileup data to identify and quantify CNAs, taking into consideration tumor purity and ploidy, which are critical for distinguishing between true CNAs and artifacts due to sample heterogeneity [19]. The integration of these tools provides a comprehensive picture of CNAs, contributing to a more complete understanding of the genomic landscape of each tumor. Then, we produce a bed file representing CNAs, using a custom script. This bed file is then intersected with a reference genome annotation using bedtools to align our findings with known genomic features. Lastly, we apply another custom R script to annotate the intersected bed file with gene and transcript IDs, using additional genomic resources to provide context to the CNA, including chromosomal bands and sub-bands.

31.3.6 Calculation of Genomic Metrics

For the microsatellite instability (MSI) analysis, the tool msisensor is employed to assess MSI status across the genome [20]. This tool scans the normal and tumor samples for changes in microsatellite lengths, which are indicative of MSI, a form of genomic instability often found in cancer.

The mutational burden is quantified using a specific R script that processes the variant calling output VCF file to provide a measure of the tumor's mutational load reported as the number of non-silent somatic mutations per megabase sequenced (Muts/MB), along with a summary of the mutation types and their frequencies, providing a comprehensive mutational profile.

For the ScarScore component, we employ a method to evaluate the genomic scar left by the homologous recombination deficiency (HRD) in the cancer genome. This method, called scarHRD, analyzes the CNA data from the FACETS tool, to assess the presence and extent of these genomic scars, which indicate a tumor's history of DNA repair deficiencies [21].

31.3.7 Processing of RNA Data

The first step in the RNA workflow is QC, which is performed using fastp. Specifically, we filter out low-quality reads and trim unwanted sequences from the raw data. Similarly to the QC of DNA data, here we aim to correct base errors, cut low-quality tails, and trim polyX sequences.

Following QC, the next step is alignment, which is performed using STAR, a fast RNA-seq aligner that uses a strategy of maximal mappable prefix (MMP) to align reads to a reference genome [22, 23]. We use parameters optimized for aligning RNA-seq data, taking into consideration key aspects such as intron sizes, splice junctions, and chimeric alignments. To ensure accurate mapping of reads across intronic regions, we define specific maximum and minimum intron sizes. We also set a maximum gap limit for paired-end reads alignment. Special parameters for handling splice junctions and chimeric alignments are set to accurately map reads that span exon-exon junctions or originate from fusion transcripts. We include stringent parameters for handling mismatches and multi-mapping reads to ensure the quality and specificity of the alignment. We generate output files that include information about unmapped reads and the alignment in BAM format.

After alignment, the next step is to mark duplicates, which is performed using GATK's "MarkDuplicates" tool. This tool identifies duplicate reads, which are reads that originate from a single fragment of cDNA and are typically the result of PCR amplification.

31.3.8 Detection of Gene Fusions

Our RNA-seq workflow incorporates a step to identify gene fusions using the tool Arriba, which is specifically designed to detect fusion events in RNA-seq data with high sensitivity and specificity [24].

We provide Arriba with a comprehensive annotation file, which contains detailed information about the genes and their transcripts on the reference genome. This allows Arriba to map the reads accurately and identify potential fusions involving any known genes.

We also use a blacklist file in our analysis. This file contains a list of recurrent artifacts, false positives, and non-pathogenic but recurrent gene fusions, which helps to filter out potential false-positive fusion calls.

Arriba analyzes the aligned reads provided in the sorted BAM file, which is the output from the previous alignment step. It then generates two output files: one containing potential fusion events and another with discarded fusion events that didn't pass the set criteria.

To increase the sensitivity of the fusion detection, we enable a parameter that instructs Arriba to consider both split reads and discordant mates. This ensures that potential fusions are not missed, thereby enhancing the comprehensiveness of our fusion detection process.

31.3.9 Gene Expression Quantification

To quantify gene expression, we employ the featureCounts tool, part of the Subread package [25]. This tool is designed to count the number of reads that map to each gene in the genome, providing a measure of gene expression. The input for featureCounts is the BAM file that contains the aligned reads, which is the output from the previous step of marking duplicates, along with a comprehensive gene annotation file. Quantification is performed at the gene level, which means that all the reads mapping to different isoforms or transcripts of a single gene are collated together to give a single count for that gene. Through this process, we generate a comprehensive profile of gene expression for each sample. The output is a text file that contains the count of reads for each gene, which serves as a measure of gene expression.

31.3.10 Modeling of Intra-Tumor Heterogeneity

To model intra-tumor heterogeneity (ITH) and provide insights into the evolutionary history of a tumor, RADAR uses the PhyloWGS algorithm, which incorporates a multi-chain Markov Chain Monte Carlo (MCMC) approach [26].

This module starts with parsing CNV data from FACETS through a Python script that formats CNV information and associates it with cellularity measurements. Subsequent steps involve preparing the input data for PhyloWGS by consolidating CNV and somatic SNV data, ensuring that each data type is correctly formatted for analysis.

Next, concurrent MCMC chains are run, taking advantage of parallel computing to thoroughly sample the posterior distribution over possible evolutionary trees. The number of chains is typically set to match the available CPU cores to optimize resource utilization without shortening runtime. Each chain runs independently, enhancing the robustness of the results. The output from these chains is then consolidated, forming a comprehensive evolutionary history that details the genetic variations and subclonal structures within the tumor.

31.3.11 Merging DNA and RNA Primary Results

In the integration phase of our RADAR pipeline, we merge primary analysis results from DNA and RNA samples to streamline cohort-level analyses and facilitate the visualization of data. For the DNA samples, we employ an R script that combines individual metrics—such as mutational burden, MSI score, and ScarScore—from each sample into a cohesive CSV file. These files are annotated with sample IDs for easy identification within the cohort.

On the RNA side, gene counts from individual RNA-seq analyses are consolidated into a comprehensive table that displays the normalized counts across all samples. Normalization is conducted using the variance stabilizing transformation (VST) from the DESeq2 package, which stabilizes the variance across the range of mean values to make the data ready for statistical analysis [27]. To account for potential batch effects, which can introduce technical variability into the data, we apply the ComBat method, a widely used tool for correcting batch effects in genomic data [28]. Once the data are normalized and adjusted, Z scores are calculated to measure the standard deviation of each sample's gene expression levels from the cohort mean, providing a standardized measure of relative gene expression for downstream analyses such as plotting and identifying differential expression patterns.

31.3.12 Clinical Actionability Annotation

Clinically actionable alterations from SNV, CNA, gene fusion, and gene expression analyses are identified using the Variant Interpretation for Cancer Consortium (VICC) meta-knowledgebase [29]. This resource amalgamates data from six distinct sources: Cancer Genome Interpreter (CGI), Clinical Interpretation of Variants in Cancer (CIViC), Jackson Laboratory Clinical Knowledgebase (Jax-CKB), MolecularMatch, OncoKB, and Precision Medicine Knowledge Base (PMKB), which curate and interpret genomic variants and their implications in cancer [9, 30–34]. A dedicated Python script accesses the VICC MKB via its API to filter and annotate variants that have clinical relevance. The results are formatted into CSV tables, detailing each variant's associated drugs, clinical significance, cancer type, and evidence level.

These tables are further annotated to show the specific subclones harboring each actionable variant, as modeled by PhyloWGS. This enables the annotation of the tumor evolutionary tree in patient reports with clinically actionable variants.

31.3.13 Calculation of Gene Expression Signatures

RADAR includes the calculation of specific gene expression signatures to aid in prognosis and therapy recommendations.

One such signature is the GEP70 score, originally developed for microarray data, which has been adapted for RNA-seq through an R script according to the original publication's methodology [35, 36]. This score is used to assess risk and predict outcomes in MM. Additionally, RADAR calculates signatures to determine sensitivity to the XPO1 inhibitor selinexor and the BCL2 inhibitor venetoclax. The method for the Selinexor signature is detailed in our prior publication, while the approach for venetoclax, still under study, follows a similar protocol [37].

31.3.14 Calculation of Pathways Activity

RADAR incorporates the tool PROGENy to determine the activity of biological pathways based on gene expression data from patient samples [38]. PROGENy (Pathway RespOnsive GENes for activity inference from gene expression yields) predicts the activity of pathways in a manner independent of their upstream signaling, using a model trained on perturbation experiments where the pathway of interest was specifically activated. This model extrapolates the potential activation states of pathways from the expression patterns of a selected set of responsive genes. In RADAR, we utilize PROGENy to translate the complex gene expression data into actionable insights, enabling us to infer which signaling pathways are active in a patient's tumor and potentially guide targeted therapy decisions. This analysis helps to identify key drivers of tumorigenesis and potential therapeutic targets within the cellular signaling networks.

31.3.15 Multiomics Patient Stratification

RADAR utilizes our multi-omics network model of MM, known as MM-PSN, to categorize patients into one of 12 disease subtypes [39]. These subtypes, along with the machine learning classifier that leverages SNV, CNA, and gene expression data, are detailed in our recent publication. Each update to the patient cohort triggers a reclassification, ensuring that new samples are accurately assigned. More in-depth information on the classifier are provided in the supplementary material of the MM-PSN article.

31.3.16 Generation of Personalized Patient Reports

The final step in the RADAR pipeline is the generation of personalized patient reports, formatted as HTML or PDF using R Markdown and the kableExtra package. The results from the analyses described in the prior Methods sections and detailed in the Results section are compiled into these reports. These reports include a variety of tables and plots, utilizing packages such as ggplot2 for general visualizations, circlize for circus plots, and igraph, ggraph, and tidygraph for depicting the tumor clonal landscape. Survival curves of MM-PSN subtypes are created using survminer for Kaplan–Meier plots. Each component of the reports is designed to present complex information in an accessible and interpretable format, with specific sections detailed in the Results section.

31.4 Results

31.4.1 A Brief Overview of the RADAR Workflow

The RADAR workflow processes WES and RNA-seq data from matched tumor and normal samples, such as CD138+ plasma cells from bone marrow for MM, with normal samples taken from saliva or blood. Figure 31.1B presents the workflow, which is outlined briefly here and examined in greater detail in the subsequent sections. Detailed workflow steps are provided in the Methods section.

The process begins with rigorous quality control to ensure the accuracy and integrity of the data, which includes correcting base sequencing errors, trimming sequences with low-quality ends, and confirming the correct matching of tumor-normal sample pairs.

The DNA pipeline begins by aligning tumor and normal sequencing reads to the reference genome using the Burrows-Wheeler Aligner [13]. After alignment, GATK tools are used for refinement, fixing any data inconsistencies or duplications, thus preparing for accurate genetic variation identification and analysis [14]. Germline variants from normal samples are detected using a haplotype-based approach, followed by variant scoring and filtering for precision. These variants are then annotated with Funcotator, which provides detailed insights into their functional implications.

The tumor-normal aligned files undergo mutation analysis using muTect2, Strelka2, and Lancet for SNV calling [15–17]. This module can operate in union or consensus mode, retaining all mutations from the three tools for downstream analysis or just those identified by at least two tools. This dual-mode approach allows for the flexibility to either minimize false negatives or employ a more stringent method to reduce false positives. SNVs are filtered, keeping only those with a minimum of ten reads for the alternate allele and a VAF of no less than 5%.

For identifying somatic CNAs, the tool FACETS is employed, which also helps estimate tumor purity and ploidy [19]. These SNVs and CNAs undergo annotation, similar to germline variants. PhyloWGS then utilizes this information to reconstruct the clonal population structures within the tumor, offering a detailed picture of its evolutionary history [26].

The DNA analysis is rounded off with metrics such as mutational burden, reflecting the overall amount of genetic changes in the tumor, microsatellite instability, a condition of genetic predisposition to mutations that results from impaired DNA mismatch repair (MMR), and the genomic scar score, summarizing homologous recombination deficiency by incorporating loss of heterozygosity–homologous recombination deficiency (LOH-HRD), large-scale transitions (LST), and number of telomeric allelic imbalances (AI) [20, 21, 40]. All these genomic metrics generated at the single sample level are then pooled to create a comprehensive dataset at the cohort level and facilitate comparisons between patients.

The RNA analysis pipeline initiates with raw read alignment using STAR, followed by duplicate marking and gene fusion detection with Arriba [22–24]. Gene expression is quantified at the individual gene level by FeatureCounts for each sample, and these results are then integrated into an existing

aggregated gene expression table, continually augmenting it with new data from subsequent samples processed by the pipeline [25].

These data undergo normalization with variance stabilizing transformation from DESeq2, and batch effects are adjusted to ensure that the data is comparable across different samples and runs, effectively neutralizing technical variations that could confound true biological differences [27]. Z scores are also computed to flag dysregulated genes for in-depth analysis.

RNA findings such as gene fusions and dysregulated genes, along with DNA findings like SNVs and CNAs, are annotated using the VICC meta-knowledgebase to identify clinically actionable alterations [29]. RNA metrics like GEP70 and MM patient similarity network (MM-PSN) scores are computed for risk assessment and disease subtyping, along with the prediction of pathway activity [35, 36, 38, 39].

Ultimately, personalized reports integrating key findings are generated, enhanced with visuals, a breakdown of the analytic process in lay terms, and relevant literature references to support the analyses and outcomes.

The following sections detail the primary components of this workflow applied to MM, elaborating on their clinical impact and supported by relevant case studies. For coherence and clarity, we will discuss these components following the order in which they appear in a patient's report.

31.4.2 Genomic Metrics, Hallmark Aberrations, and Risk Stratification in Multiple Myeloma

Our personalized reports begin with useful genomic and transcriptomic metrics for disease characterization and risk stratification.

Risk stratification in MM is crucial for determining prognosis and tailoring treatment strategies. Recent guidelines have refined these systems to account for genetic aberrations and other risk factors. The Revised International Staging System (R-ISS) now includes lactate dehydrogenase (LDH) levels and specific cytogenetic abnormalities, such as t(4;14) and p53 deletion, which are considered high-risk factors [41]. These advancements enable more personalized approaches to therapy. For example, the mSMART system from Mayo Clinic offers treatment recommendations based on risk categorization, refining therapeutic strategies for MM [42]. This approach influences critical clinical decisions, like stem cell transplantation and maintenance therapy, and is instrumental in enhancing patient outcomes.

The Genomic Summary in the report outlines the unique genetic composition of the tumor. Each aspect of this summary reveals a different aspect of the disease, offering insights that can inform prognosis and guide treatment decisions.

The tumor mutation burden quantifies the number of mutations present in the DNA of the myeloma cells, reported as the number of non-silent somatic mutations per megabase sequenced (Figure 31.2A). A higher mutation burden can influence the disease's response to certain treatments and gives insight into the mutational landscape of the tumor.

The scarHRD score assesses the tumor's DNA repair capabilities. Faulty DNA repair mechanisms can lead to genetic alterations in the tumor, highlighting potential vulnerabilities that treatments might target. In MM, the scarHRD score has been proposed as a potential prognostic marker, with a low HRD score associated with a better prognosis [43] (Figure 31.2B).

Microsatellite instability is measured to gauge genome hypermutability, a propensity of the tumor to accumulate mutations due to DNA repair system deficiencies, which may suggest higher efficacy for specific treatments [40] (Figure 31.2C).

The reports also feature the MM-PSN classification, a network-based multi-omics model for categorizing MM, derived from our recent study on a large cohort of MM patients from the MMRF CoMMpass study [39]. This system groups patients based on SNVs, CNAs, chromosomal translocations, gene expression, and gene fusions, identifying specific pattern of genetic and molecular changes and offering a robust prognostic indicator. Survival curves for each identified group are plotted, with the curve relevant to the patient's specific group highlighted to offer insight into historical survival rates and characteristics unique to that subtype (Figure 31.2D).

Furthermore, the reports include an RNA-seq adaptation of the GEP70 risk score, originally developed using microarrays. This score evaluates gene expression patterns to predict the disease's aggressiveness and prospective trajectory [35, 36] (Figure 31.2E).

Another section of the report is focused to detailing MM hallmarks, which are the genetic features that are frequently observed in patients and present a clearer picture of the tumor's genetic landscape and can provide crucial insights for care and treatment planning. These features include critical chromosomal translocations and genomic aberrations that signify therapeutic and prognostic relevance, including hyperdiploidy, the amplification of chromosome 1q, and biallelic deletion of TP53. Complementing this section is a global CNA profile plot, which provides a visual snapshot of changes in the number of DNA copies within the tumor across different chromosomal locations. This plot effectively maps the gain and loss of genetic material, offering a comprehensive view of the tumor's genetic variances. This visualization not only aids in understanding the complexity of the disease but also highlights potential areas of genetic instability that may be targeted for personalized treatment strategies.

31.4.3 Actionable Alterations for Targeted Therapies and Our Two-Tiered System

At the heart of RADAR is the association between specific genetic and molecular changes with their clinical actionability. An aberration is considered actionable if it can be targeted by a specific therapy, acts as a biomarker, or serves as a substantial diagnostic or prognostic factor. Within our system, we utilize the VICC knowledge base, which compiles actionable alterations from various respected sources, including the Cancer Genome Interpreter, CIViC, Jax-CKB, MolecularMatch, OncoKB, and PMKB [9, 29–34]. These resources provide a wide spectrum of data supporting the clinical relevance of genetic variations.

A quintessential example of actionability is the presence of the BCR-ABL fusion gene in chronic myeloid leukemia (CML), which is targeted by tyrosine kinase inhibitors like

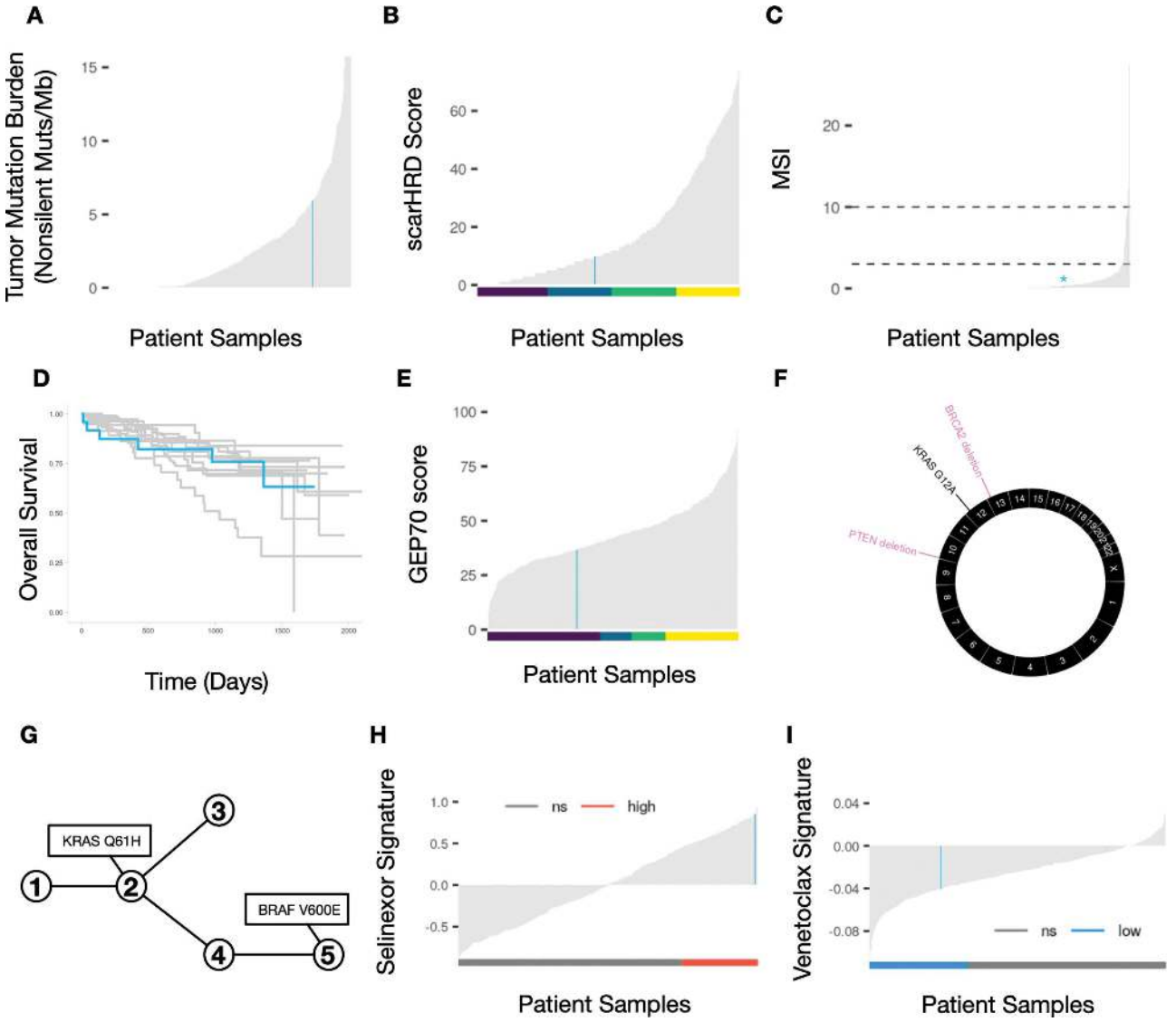


FIGURE 31.2 *Selected elements of RADAR's reports.* (A) Tumor mutation burden (TMB) quantifies the number of mutations present in the DNA of the myeloma cells. TMB is reported as the number of nonsilent somatic mutations per megabase of DNA. (B) The scarHRD score evaluates the tumor's capacity to repair its own DNA. A low scarHRD score is associated with a better prognosis. (C) Microsatellite instability assesses the presence of specific repetitive DNA sequences in the tumor that have undergone changes. A high MSI suggests a malfunction in the tumor's DNA repair system. (D) The plot displays survival curves for each of the 12 MM-PSN subgroups, which have been derived based on multi-omics network analysis. These curves indicate how patients with tumors in these groups have fared historically over time. The blue curve represents the group where this particular tumor belongs, which in the case illustrated is called 2d:tMMSET/1q+/15q+. Based on our analysis, patients with tumors in this group historically had an average overall survival of 32.99 months. (E) The GEP70 risk score analyzes the expression patterns of 70 specific genes in the tumor. It is useful for categorizing the aggressiveness of the disease and estimating its potential course, assisting in devising a suitable treatment plan. The tumor in the figure has a GEP70 risk score of 36.87, which falls into the low-risk category (dark purple). (F) Circos plot of actionable alterations. This plot offers a visual guide to the tumor's genome. The circle is a genome's map, with each section representing a chromosome. Distinctive labels and colors on this circle highlight particular DNA changes that might influence treatment choices. These changes can show areas where DNA is mutated, is missing, or has extra copies, genes that are more or less active than usual, and potential targets for treatments like immunotherapy. (G) Tumor phylogeny in RADAR's reports. An example of a tumor phylogenetic tree, illustrating the relationships between different subclones within the tumor. On this tree, RADAR highlights actionable genetic changes; that is, alterations that might be relevant for treatment decisions. (H–I) Gene expression plots in RADAR's reports. These plots illustrate gene expression for specific signatures or genes of interest. Each bar represents the tumor under investigation, compared to all other tumors analyzed (gray background). Color-coded regions at the bottom of each plot indicate zones where gene activity suggests potential benefit from specific treatments. The selinexor drug gene expression signature (H) combines the expression of WNT10A, ETV7, and DUSP1. Upregulation of these genes is associated with a better response to selinexor. Tumors in the high quartile (red area) are likely to be more sensitive to selinexor. The venetoclax drug gene expression signature (I) combines the expression of ATP1B3, MYL2, CNR1, FKBPL, LRRK2-DT, and LINC02541. Downregulation of this signature is associated with a better response to venetoclax. Tumors in the low quartile (blue area) are likely to be more sensitive to venetoclax. *Note*: Each blue bar in plots A, B, H, and I, and the star in plot C, represents the specific tumor under investigation compared to all the other tumors analyzed so far (gray background).

imatinib and dasatinib [44]. Another well-established actionable alteration is the BRAF V600E mutation in melanoma, effectively treated with BRAF inhibitors [45].

In MM, actionable genetic alterations include translocations and mutations that directly impact treatment decisions. For example, the translocation t(11;14) involving CCND1 can predict a response to BCL2 inhibitors; for example, venetoclax, and the presence of a deletion of chromosome 17p (del17p), is often associated with a poor prognosis and may influence the intensity of therapy [46, 47]. Mutations in the MAPK pathway, such as those in KRAS, NRAS, and BRAF, which are frequently identified in multiple myeloma, also have implications for targeted treatments [48]. Inhibitors like trametinib and binimetinib, which are designed to target MEK, a downstream component of the MAPK pathway, and pan-RTK inhibitors that target a range of receptor tyrosine kinases, are examples of how these mutations can be specifically addressed in treatment protocols.

In RADAR, we have curated a short list of Tier I top-choice actionable alterations, which include well-established therapies and emerging therapeutic strategies in MM. These include MEK inhibitors, such as trametinib, and pan-RTK inhibitors, such as regorafenib, which have shown efficacy against specific MAPK pathway mutations. Gene expression markers also guide the use of the XPO1 inhibitor selinexor and the BCL2 inhibitor venetoclax, tailoring therapy based on individual patient profiles. Additionally, for those with t(4;14) translocations leading to aberrations in FGFR3, FGFR inhibitors like erdafitinib and infigratinib are of particular interest.

Tier II actionable alterations are suggested as supplementary possibilities with less clinical evidence backing their potential effectiveness, drawn from a variety of cancers. These include agents like copanlisib for the PI3K pathway in cases of gene expression-verified pathway activation or when resistance to MAPK inhibitors emerges. PARP inhibitors, such as olaparib, are considered for those with a genetic profile suggesting HRD, possibly indicated by an elevated scar score reflecting genomic instability. However, the employment of Tier II drugs is off-label and based on emerging research rather than established guidelines. Their inclusion in RADAR provides physicians with potential treatment paths for those MM patients who have exhausted all standard options, offering a basis for innovative, investigational approaches in challenging cases.

Drug recommendations are provided in a two-part table for Tier I and II. This table offers a tailored list of potential drug options along with the corresponding genetic markers identified in the tumor's makeup and references to the supporting evidence. Furthermore, clinically actionable alterations are displayed in a circos plot, which offers a visual guide to the tumor's genome (Figure 31.2F).

31.4.4 Elucidating Intra-Tumoral Clonal Heterogeneity for Precision Prognostication

Intra-tumor heterogeneity refers to the presence of genetically diverse cancer cells within a single tumor, arising from mutations that begin in a single cell and spread [49, 50]. These mutations, while mostly neutral, can include "drivers" that promote unchecked growth. As a tumor evolves, its cells may follow a straight or branching path of development, leading to different subclonal populations that may cooperate to foster growth or compete, affecting the tumor's behavior and response to therapy [51]. Studies have linked this heterogeneity to varied clinical outcomes and resistance to treatment, underscoring its importance in understanding and managing cancer. For instance, studies in chronic lymphocytic leukemia and MM have shown that greater subclonal diversity can correlate with poorer clinical outcomes [52].

ITH contributes significantly to metastasis. For instance, studies on renal cell carcinoma show that genetic diversity and mutation patterns vary across different tumor regions, affecting prognosis [53]. This diversity complicates treatment since a single biopsy may not capture the tumor's full genetic landscape, highlighting the challenges in creating effective therapies.

In RADAR, we employ sophisticated computational methods like PhyloWGS to dissect the clonal landscape of tumors, parsing SNVs and CNAs [26]. This technique enables us to interpret somatic variations, grouping them into subclones that mirror the tumor's lineage and growth. This process reveals the genetic evolutionary branches within the cancer, offering insights into its progression and the development of subpopulations with distinct survival advantages or treatment resistances.

In patient reports, the "Tumor Phylogeny" plot depicts this evolutionary tree, illustrating the model that best fits the observed data and showcasing the relationships between the different subclones within the tumor, from founding clone to smaller subclones (Figure 31.2G). On this tree, we highlight actionable alterations that might be relevant for treatment decisions. These changes are not isolated to just one group but can be found across different branches, reflecting the shared heritage and diversity of the tumor cells. Along with the plot, we provide information on the estimated tumor purity, indicating how much of the sample we analyzed is made up of cancer cells as opposed to normal cells.

In RADAR, we are now integrating evolutionary modeling to track tumor evolution through multiple longitudinal samples from the same patient, similar to the methodology outlined in Section 31.4.8. This modeling provides an important tool for physicians to observe the trajectory of genetic aberrations over time, revealing the impact of treatments on the tumor's architecture. This insight is crucial for planning future therapy directions, as it may inform the choice of subsequent treatments based on the tumor's evolving genetic landscape.

31.4.5 Informing Therapy Selection with RNA Expression Signatures

In RADAR, we have incorporated novel biomarker signatures predicting response to specific drugs, translating the findings from our research into actionable clinical applications. Our recent work has uncovered a three-gene signature (WNT10A, DUSP1, ETV7) associated with progression-free survival (PFS) in patients treated with the XPO1 inhibitor selinexor, an approved therapy for MM [37, 54]. This signature, validated across independent trials at Mount Sinai, provides a potential tool for selinexor treatment selection. The clinical application

of this signature allows physicians to navigate therapeutic choices more effectively, offering the therapy to patients most likely to respond positively and sparing unlikely responders from selinexor's associated toxicities.

Our ongoing work is investigating venetoclax, a selective BCL2 inhibitor showing promise as a therapy for MM [46, 55]. While venetoclax has proven effective in MM patients with the t(11;14) translocation, this translocation alone doesn't definitively predict long-term benefit. In our ongoing study, we have identified a novel six-gene signature predictive of venetoclax response. This signature helps identify patients more likely to relapse on venetoclax, allowing clinicians to select alternate treatments with a higher success likelihood. Our reports incorporate both signatures as barplots, wherein each bar represents the tumor under examination (blue line) compared against all other tumors analyzed to date (gray background) (Figure 31.2H and I). These bars rank the expression of the biomarker signatures. Color-coded regions at each plot's bottom mark zones where signature activity suggests a potential benefit from the specific treatment.

The modular architecture of RADAR facilitates the easy incorporation of any signature and gene expression profile of interest, including those related to immunotherapy targets. These are described in greater detail in the following section.

31.4.6 Identifying Candidates for Immune-Based Interventions

In recent years, innovative therapeutics harnessing the immune system, such as CAR T cells and bispecific antibodies, have achieved remarkable responses in relapsed/refractory MM patients [56, 57]. CAR T cells, or chimeric antigen receptor T cells, are engineered T cells designed to express a receptor specific to a target antigen on cancer cells, enhancing the body's immune response to the cancer. On the other hand, bispecific antibodies are a class of artificial proteins that can bind to two different types of antigen simultaneously. Often, one part is designed to bind to a cancer cell, while the other part is designed to engage immune cells, effectively bringing them into close proximity and promoting a direct, targeted immune response against the cancer cells (Figure 31.3A).

In MM, there are two approved CAR T cell therapies, idecabtagene vicleucel (Ide-cel) and ciltacabtagene autoleucel (Cilta-cel), both targeting B cell maturation antigen or BCMA, a surface molecule expressed by myeloma cells encoded by the gene TNFRSF17 [58–60].

The first bispecific antibody approved by the FDA for the treatment of MM, teclistamab, also binds to BCMA on the surface of myeloma cells and to surface marker CD3 on immune system T cells, bringing the myeloma and T cells in closer proximity to each other, which activates the T cells and primes them to kill the myeloma cells [61]. In addition to teclistamab, the FDA has also granted breakthrough therapy designation to another bispecific antibody for relapsed or refractory MM targeting BCMA, called elranatamab [62]. Furthermore, talquetamab, which targets GPRC5D (G-protein coupled receptor class C group 5 member D), has joined this list with accelerated FDA approval for the treatment of relapsed or refractory MM [63].

These approvals represent significant advancements in the treatment of MM, offering new therapeutic options for patients who have not responded to other treatments. However, it's important to note that the use of bispecific antibodies among the multiple lines of treatment for MM is still being determined, and they may be used as a bridge to CAR T cell therapy or as an alternative to CAR T treatment that is more readily available.

CAR T and bispecific antibodies targeting different antigens are currently being investigated in clinical trials, including those targeting Fc receptor homolog 5 (FcRH), such as cevostamab, which is expressed on the surface of MM cells, making it another promising target for immune-based therapies [64].

Research to determine reliable markers for the identification of the optimal immune-based approach for therapy in MM, including sequencing of therapeutics targeting the same or different antigens, is still ongoing. In a recent work, we have shown that the sequential use of different T-cell redirection therapies can lead to deep and durable responses in RRMM patients following relapse after bispecific antibody therapy [65].

However, one obvious feature to determine the effectiveness of such treatments is the expression of targets on the surface of myeloma cells. In our previous work, we have shown that a particular subtype of MM patients exhibits frequent deletions in specific chromosome areas, including the locus of TNFRSF17 on chromosome 16 [39]. Such patients may be at higher risk of relapse from BCMA-directed therapies, as they only have one functional copy of the gene. Furthermore, recent studies have shown that tumors may escape from immune therapies by mutating or losing the targeted antigen. This has been observed particularly for bispecific antibodies targeting BCMA and GPRC5D [66, 67].

In an initial step towards assessing the efficacy of these therapies, we have added a unique section in RADAR's reports focusing on immunotherapy targets. This section presents the expression levels of key targets—BCMA, GPRC5D, and FCRH5—in the tumor being examined. These levels are then compared against those found in all other tumors analyzed to date, mirroring the process used for other biomarker signatures described previously (Figure 31.3B and D).

It's important to note that while RNA expression provides valuable insight, it's not a perfect marker for sensitivity to antigen-directed therapy. This is because RNA expression doesn't always equate to surface expression of the encoded proteins. Nevertheless, it gives an indication of whether the gene is actively being expressed.

In addition to this, our reports also present any SNVs or CNAs that affect the loci encoding these targeted antigens. Current panels do not routinely include these loci, yet their integration into clinical practice is poised to have a significant impact. This information is particularly crucial when evaluating patients who relapse following these therapies, as it can help clarify the mechanisms of tumor-associated relapse and guide the decision-making process for the next line of therapy.

As more evidence becomes available suggesting novel markers of response and resistance to immune-based therapies, RADAR's architecture will be ready for their incorporation.

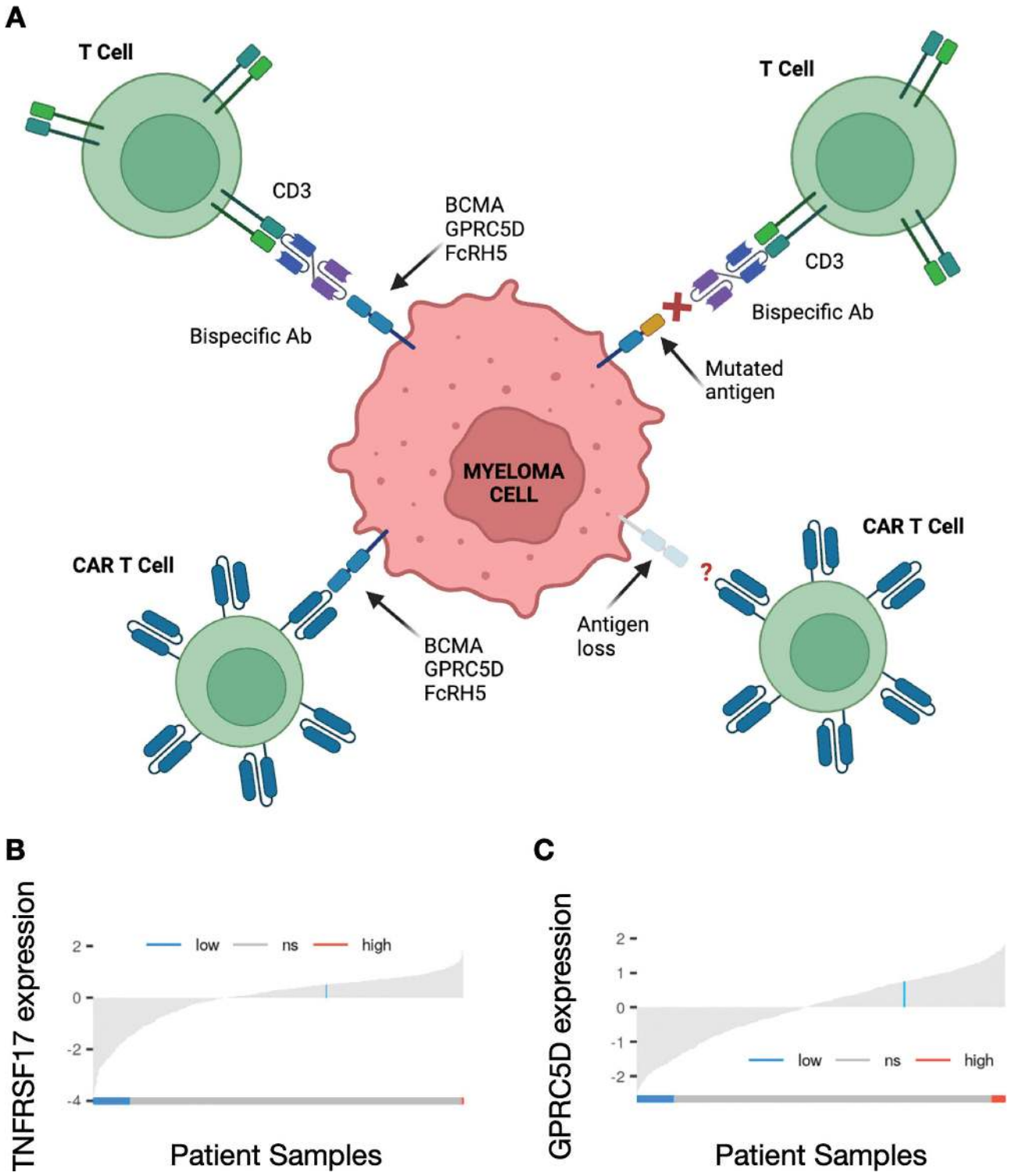


FIGURE 31.3 *CAR T cell and bispecific antibody therapies in MM.* (A) This figure illustrates the targeting mechanisms of CAR T cell and bispecific antibody therapies, as well as the escape mechanisms employed by cancer cells, such as antigen mutation and antigen loss. The targeting mechanisms involve the use of engineered T cells or bispecific antibodies to specifically recognize and kill cancer cells. It highlights three prominent targets on MM cells—BCMA, GPRC5D, and FcRH5. Also depicted is the CD3-mediated T cell redirection therapy by bispecific antibodies. The escape mechanisms highlight how cancer cells can adapt to evade these therapies, emphasizing the complexity of treating MM. (B) The plot shows the expression of the TNFRSF17 gene, which encodes for the immunotherapy target BCMA. (C) The plot shows the expression of the GPRC5D gene, whose protein product is an immunotherapy target. In both (B) and (C), the tumors fall in the gray area, which is not associated with a specific treatment outcome. Created in BioRender. Lagana, A. (2024) https:// BioRender.com/u49g225

We are additionally evaluating the expansion of our system to include additional assays that may broaden the scope of the immunological analysis, including investigation of the bone marrow tumor microenvironment by techniques such as spectral flow cytometry and the incorporation of T cell killing assays.

31.4.7 Incorporating Germline Screening into RADAR to Uncover Inherited Predisposition to Multiple Myeloma

The growing body of evidence supporting an inherited predisposition to MM has led to the incorporation of germline screening into RADAR. Epidemiological studies have shown that first-degree relatives of MM patients have a two- to four-fold increase in risk of developing MM or its precursor disease monoclonal gammopathy of unknown significance (MGUS) [68]. Genome-wide association studies (GWAS) have identified common SNPs as well as rare high-penetrance variants that collectively explain approximately 16% of the estimated heritability of MM [69].

In a recent study, we analyzed MMRF CoMMpass data and identified 895 newly diagnosed MM (NDMM) patients for whom WES of germline DNA was available [70]. We identified pathogenic/likely-pathogenic germline variants (PGV) in hereditary cancer genes (HCG) in 8.8% of these patients. Most PGV involved DNA damage repair (DDR) genes, with homologous recombination (HR) genes being the most commonly mutated.

Interestingly, MM patients with a family history of hematologic malignancy in a first- or second-degree relative were significantly more likely to carry PGV. Additionally, the likelihood of being diagnosed with MM before age 40 was significantly higher in PGV carriers. In terms of survival analysis, DDR-PGV carriers had a significant progression-free survival advantage over non-carriers.

These findings underscore the importance of germline screening in MM patients. The identification of PGV in HCG can provide valuable insights into the genetic predisposition of MM, especially in those with a family history of hematologic malignancy and in those diagnosed before age 40. Furthermore, the detection of PGV can have significant implications for treatment strategies and prognosis, as carriers of PGV in DDR genes have been observed to have a significant PFS advantage.

With the incorporation of germline screening into RADAR, we are now able to detect PGV in our patients. This will not only enhance our understanding of the genetic landscape of MM but also enable us to tailor treatment strategies based on the genetic profile of each patient.

31.4.8 Case Study: Longitudinal Assessment of Actionable Alterations in Post-CAR T Settings

A recent case study illustrates our platform's clinical impact, particularly following immunotherapy failure. We reported on a 61-year-old male with relapsed/refractory MM who, after progressing post-anti-BCMA CAR T therapy, presented with extramedullary disease [71] (Figure 31.4). The patient developed subcutaneous nodules and elevated levels of lambda free light chains. Utilizing RADAR, we identified a BRAF (V600E) dominant subclone in a sample collected prior to CAR T therapy, which persisted throughout treatment and was also detected in the cutaneous plasmacytoma at relapse. Based on *in vitro* experiments and on previous studies from our team, we implemented a triple MAPK inhibition strategy with a combination of a BRAF monomer inhibitor (dabrafenib), a MEK inhibitor (trametinib), and a BRAF dimer inhibitor (regorafenib), to inhibit the MAPK signaling pathway effectively [72]. This therapy was intended to overcome the limitations of BRAF monomer selective inhibitors with or without MEK inhibitors, which had been previously attempted on MM patients with equivocal outcomes on disease progression and survival due to feedback activation of the pathway via BRAF dimers induction. Preclinical studies showed that the addition of the BRAF dimer inhibitor regorafenib in the triple MAPK inhibition strategy was able to potently inhibit dimeric BRAF and relieve the induced negative feedback, overcoming this feedback recovery caused by the standard approach of inhibiting RAF monomers alone [72]. This resulted in more efficacious and tolerable treatment in preclinical BRAF (V600E) cancer models.

The patient initially responded well to the treatment, achieving a good partial response for 110 days, which allowed him to participate in subsequent clinical trials and eventually achieve a stringent complete response.

We further applied RADAR to the patient's post-MAPK inhibition samples to track clonal changes and evaluate pathway activity via gene expression analysis, coupled with further *in vitro* testing. Our findings suggested that the patient's plasma cell clones might have shifted their dependency to an alternative pathway, as indicated by increased PI3K signaling transcriptional output. In fact, treatment of the MM cells with the PI3K inhibitor copanlisib showed a significant reduction of a key growth signal in the MM cells, suggesting that these cells had become more reliant on the PI3K pathway for survival.

This study demonstrated RADAR's role in modeling data to guide treatment for MM patients after CAR T therapy, highlighting the platform's strength in tracking clonal evolution and assessing pathways for targeted therapy, showcasing the customized approach it offers in the complex landscape of MM treatment.

31.4.9 RADAR's Application to Subtype-Specific Therapy Recommendations

Here, we present another use of our platform to identify subtype-specific therapy recommendations in the context of our research study on a large cohort of newly diagnosed MM patients. In the MM-PSN project mentioned earlier in the context of prognostic indicators, we have performed network-based integration of multiple omics data from 655 MM patients [39]. Our model stratified patients into three main groups and 12 subgroups, each characterized by distinct genomic and molecular features as well as prognostic implications. Our analysis recapitulated previous MM biology and pathogenesis, such as hyperdiploidy and translocations involving the

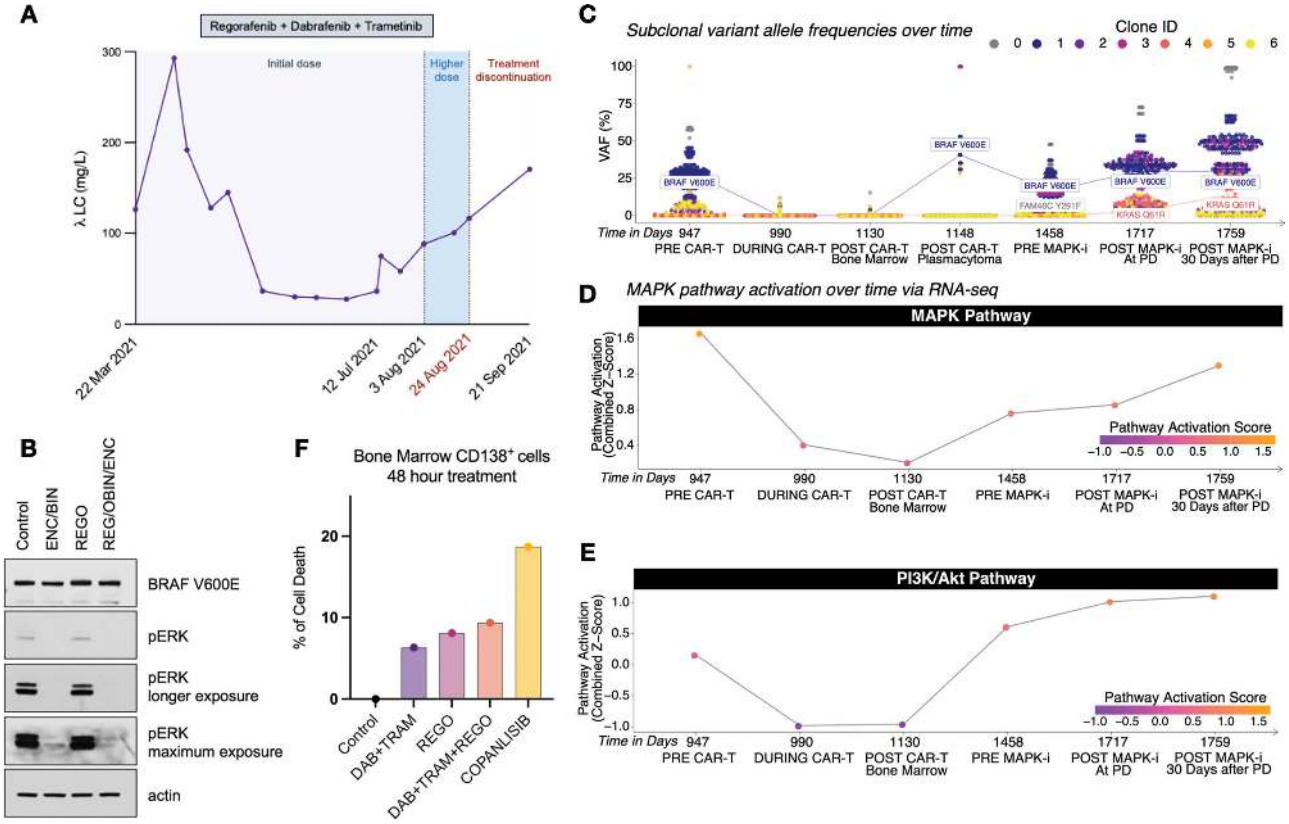


FIGURE 31.4 *Case study on clonality modeling in MM.* This figure details a patient's response to triple therapy regimen targeting the MAPK pathway (regorafenib, dabrafenib, and encorafenib). (A) Patient's lambda light chain (LC) levels from the start of treatment until discontinuation. (B) Western blot of the patient's CD138+ plasma cells treated for 48 h with encorafenib (ENC), binimetinib (BIN), regorafenib (REG), or a combination of all three. (C) Temporal evolution of MAPK alterations, illustrating the trajectory of the BRAF V600E mutation-bearing subclone in response to treatment and subsequent emergence of a KRAS Q61R mutation after triple MAPK inhibition. (D) Increasing MAPK pathway activation over time, represented as a combined Z score, in response to the triple MAPK inhibition therapy. (E) Elevated PI3K/Akt pathway activation throughout the treatment, also measured as a combined Z score. (F) Viability of CD138+ cells after 48 h of *in vitro* treatment with ENC, BIN, REG, a combination of these drugs, or copanlisib alone.

immunoglobulin locus as the main genetic driver events, and revealed novel subgroups defined by specific patterns of co-occurring alterations.

Our study applied RADAR's multiomics approach to MM-PSN patients, evaluating therapeutic implications through subgroup enrichment. We augmented RADAR with data from the Cancer Cell Line Encyclopedia (CCLE) and the Genomic of Drug Sensitivity in Cancer (GDSC) databases to pinpoint DNA and RNA alterations linked to potential treatments [73, 74].

Specifically, in Group 1 (Hyperdiploidy), drugs targeting the MET gene—amplified in certain subgroups—showed promise. Our approach also illuminated the utility of CDK4/6 inhibitors in subgroups with CCND1 gain and suggested PI3K inhibitors for those with PIK3CA alterations.

For Group 2 (Translocation of NSD2 or MAF), the analysis pointed to the efficacy of tazemetostat and selumetinib in subgroups with specific deletions affecting chromatin regulation and tumor suppressor genes. Additionally, FGFR inhibitors emerged as recommended options due to FGFR3 mutations.

In Group 3 (Translocation of CCND1), we highlighted a subset of patients who may benefit from BCL2 inhibition, aligning with the heightened expression of BCL2 and a deeper understanding of resistance mechanisms through MCL1 expression patterns. Our in-depth investigation on venetoclax sensitivity led to a promising gene expression signature for predicting sensitivity to venetoclax, as detailed in Section 31.4.5.

Lastly, we matched gene expression changes with potential CAR T cell and immunotherapy targets, including BCMA-targeted therapies, and explored new targets such as TACI and CD20. These insights hold significant potential for advancing personalized medicine and guiding clinical trial designs, as we anticipate diverse therapeutic responses across patient subgroups due to their distinct genomic signatures. RADAR's capabilities extend to pinpointing tailored treatment strategies for specific myeloma subtypes, underscoring its value in precision oncology.

31.5 Discussion

In this chapter we have described the robust capabilities of the RADAR platform in advancing precision medicine for myeloma. Through the integration of diverse data types,

RADAR has advanced the implementation of precision medicine for our MM patients, offering personalized therapeutic strategies anchored in their individual genetic profiles. This chapter has underscored the utility of genomic and transcriptomic metrics in stratifying risk, and the influence of intra-tumor heterogeneity in shaping disease progression and response to treatment, with a practical application illustrated through a case study of longitudinal assessment post-CAR T therapy for clinical decision-making. We have highlighted the important role of germline screening in identifying predispositions, as well as the potential of integrating RNA expression signatures as an additional layer for identifying therapeutic targets and understanding the disease's biology. Furthermore, we have showcased how novel treatments such as CAR T cells and bispecific antibodies are revolutionizing MM therapy, emphasizing the importance of target antigen expression on myeloma cells and the necessity for continuous research to understand and overcome mechanisms of resistance. This exploration into immunotherapies underscores RADAR's adaptability and readiness to incorporate new data and assays, reflecting its continuous evolution alongside the advancing landscape of MM treatment.

Emerging immunotherapies offer significant progress in managing MM, leveraging the power of the immune system. CAR T cells and bispecific antibodies are among these innovative treatments, with a focus on targeting key antigens such as BCMA and GPRC5D to fight the disease. The development and FDA approval of treatments such as Ide-cel, Cilta-cel, teclistamab, and talquetamab offer new hope, especially for patients with advanced disease. Ongoing research focuses on optimizing these therapies, including determining the most effective sequences and combinations, and understanding the mechanisms of resistance, such as antigen loss or mutation. Furthermore, recent studies on personalized neoantigen vaccines show potential to stimulate an immune response that may reduce the risk of disease relapse. Results from a phase I trial with pancreatic cancer patients, for example, underscore the promise of neoantigen vaccines tailored from individual tumor mutations, which, when combined with immunotherapy and chemotherapy, have shown potential in delaying disease recurrence [75]. Similarly, our research on MM highlights that neoantigen-specific T-cell responses can be induced, opening a pathway to personalized cancer vaccines even in cancers with moderate mutation rates [76].

Building on this, we aim to refine treatment selection and adapt to the evolving genetic landscape of MM, focusing particularly on individual patient immune responses and the dynamic changes within the tumor microenvironment. This adaptation will necessitate robust multimodal data integration, including single-cell RNA sequencing (scRNA-seq) and cell-free DNA (cfDNA) analysis, to provide a detailed view of tumor evolution and immune evasion mechanisms.

31.6 Future Trends

Looking ahead, the direct incorporation and modeling of longitudinal data into RADAR, as illustrated in the pilot post-CAR T case study, will deepen our understanding of how tumors evolve under treatment pressure, informing adjustments in therapeutic strategies in a patient-specific manner. The integration of treatment and response data in the model will allow us not only to reveal novel patterns of genomic and transcriptomic changes associated with disease progression and drug resistance, but also to better determine the next treatment in each individual patient's journey, enhancing the precision of follow-up interventions. This approach aims to align therapeutic regimens more closely with the evolving nature of MM, ultimately aiming to extend remission periods and improve overall patient survival.

The future of RADAR will see significant advancements through the incorporation of cutting-edge technologies. Expanding RADAR to include bone marrow TME analysis via spectral flow cytometry and scRNA-seq, alongside *in vitro* assays such as CAR T cell killing assays, represents a transformative step towards personalizing patient care. These techniques will allow us to gain an unprecedented understanding of the complex cellular interactions within the TME, which are crucial to MM progression and treatment resistance.

ScRNA-seq will be particularly valuable in dissecting intratumor heterogeneity and understanding the intricate interplay of various cell types in the TME. This level of detail will be essential in identifying new therapeutic targets and overcoming drug resistance mechanisms. Moreover, cfDNA will play an important role in monitoring minimal residual disease (MRD) and detecting early relapse. By providing real-time insights into tumor burden and genetic alterations, cfDNA analysis will further enhance the use of immunotherapies and other targeted treatments [77].

Integrating these technologies presents challenges, particularly due to the high-dimensional nature of the data and the potential noise that can obscure critical signals. This will require sophisticated computational tools for effective analysis and seamless integration with clinical information. Nevertheless, these efforts hold great promise in enabling a more adaptive approach to MM treatment, one that evolves with the disease and tailors therapeutic interventions to each patient's journey. This approach will enhance our understanding of ITH, aid in biomarker discovery, and lead to the development of more targeted and adaptive therapies for MM.

REFERENCES

1. Schwartzberg L, Kim ES, Liu D, et al. Precision oncology: Who, how, what, when, and when not? *Am Soc Clin Oncol Educ Book*. 2017;37:160–169.
2. Malone ER, Oliva M, Sabatini PJB, et al. Molecular profiling for precision cancer therapies. *Genome Med*. 2020;12:8.
3. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med*. 2015;372:793–795.
4. Mery B, Vallard A, Rowinski E, et al. High-throughput sequencing in clinical oncology: From past to present. *Swiss Med Wkly*. 2019;149:w20057.
5. Bode AM, Dong Z. Recent advances in precision oncology research. *NPJ Precis Oncol [Internet]*. 2018;2. http://dx.doi.org/10.1038/s41698-018-0055-0.
6. Kumar SK, Rajkumar V, Kyle RA, et al. Multiple myeloma. *Nat Rev Dis Primers*. 2017;3:17046.

7. Kint N, Vlayen S, Delforge M. The treatment of multiple myeloma in an era of precision medicine. *Expert Rev Precis Med Drug Dev*. 2019;4:153–162.
8. Laganà A, Beno I, Melnekoff D, et al. Precision medicine for relapsed multiple myeloma on the basis of an integrative multiomics approach. *JCO Precis Oncol [Internet]*. 2018;2018. http://dx.doi.org/10.1200/PO.18.00019.
9. Griffith M, Spies NC, Krysiak K, et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat Genet*. 2017;49:170–174.
10. Di Tommaso P, Chatzou M, Floden EW, et al. Nextflow enables reproducible computational workflows. *Nat Biotechnol*. 2017;35:316–319.
11. Chen S, Zhou Y, Chen Y, et al. fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34:i884–i890.
12. Lee S, Lee S, Ouellette S, et al. NGSCheckMate: Software for validating sample identity in next-generation sequencing studies within and across data types. *Nucleic Acids Res*. 2017;45:e103.
13. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*. 2009;25:1754–1760.
14. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr Protoc Bioinformatics*. 2013;43:11.10.1–11.10.33.
15. Benjamin D, Sato T, Cibulskis K, et al. Calling somatic SNVs and Indels with Mutect2. *bioRxiv [Internet]*. 2019. http://dx.doi.org/10.1101/861054.
16. Kim S, Scheffler K, Halpern AL, et al. Strelka2: Fast and accurate calling of germline and somatic variants. *Nat Methods*. 2018;15:591–594.
17. Narzisi G, Corvelo A, Arora K, et al. Genome-wide somatic variant calling using localized colored de Bruijn graphs. *Commun Biol*. 2018;1:20.
18. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: Rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2016;32:1220–1222.
19. Shen R, Seshan VE. FACETS: Allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res*. 2016;44:e131.
20. Niu B, Ye K, Zhang Q, et al. MSIsensor: Microsatellite instability detection using paired tumor-normal sequence data. *Bioinformatics*. 2014;30:1015–1016.
21. Sztupinszki Z, Diossy M, Krzystanek M, et al. Migrating the SNP array-based homologous recombination deficiency measures to next generation sequencing data of breast cancer. *NPJ Breast Cancer [Internet]*. 2018;4. http://dx.doi.org/10.1038/s41523-018-0066-6.
22. Dobin A, Davis CA, Schlesinger F, et al. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
23. Dobin A, Gingeras TR. Optimizing RNA-seq mapping with STAR. *Methods Mol Biol*. 2016;1415:245–262.
24. Uhrig S, Ellermann J, Walther T, et al. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res*. 2021;31:448–460.
25. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–930.
26. Deshwar AG, Vembu S, Yung CK, et al. PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol*. 2015;16:35.
27. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
28. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8:118–127.
29. Wagner AH, Walsh B, Mayfield G, et al. A harmonized meta-knowledgebase of clinical interpretations of somatic genomic variants in cancer. *Nat Genet*. 2020;52:448–457.
30. Tamborero D, Rubio-Perez C, Deu-Pons J, et al. Cancer genome interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med [Internet]*. 2018;10. http://dx.doi.org/10.1186/s13073-018-0531-8.
31. Duren R, Neeley S, Welsh JW, et al. MolecularMatch LAB integrates knowledgebases for collaborative clinical interpretation of variation in cancer. *Cancer Genet*. 2017;214:45.
32. Chakravarty D, Gao J, Phillips S, et al. OncoKB: A precision oncology knowledge base. *JCO Precis Oncol*. 2017:1–16.
33. Huang L, Fernandes H, Zia H, et al. The cancer precision medicine knowledge base for structured clinical-grade mutations and interpretations. *J Am Med Inform Assoc*. 2017;24:513–519.
34. Patterson SE, Liu R, Statz CM, et al. The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. *Hum Genomics*. 2016;10:4.
35. Shaughnessy JD Jr, Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood*. 2007;109:2276–2284.
36. van Laar R, Flinchum R, Brown N, et al. Translating a gene expression signature for multiple myeloma prognosis into a robust high-throughput assay for clinical use. *BMC Med Genomics*. 2014;7:25.
37. Restrepo P, Bhalla S, Ghodke-Puranik Y, et al. A three-gene signature predicts response to selinexor in multiple myeloma. *JCO Precis Oncol*. 2022;6:e2200147.
38. Schubert M, Klinger B, Klünemann M, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun*. 2018;9:20.
39. Bhalla S, Melnekoff DT, Aleman A, et al. Patient similarity network of newly diagnosed multiple myeloma identifies patient subgroups with distinct genetic features and clinical implications. *Sci Adv*. 2021;7:eabg9551.
40. Latham A, Srinivasan P, Kemel Y, et al. Microsatellite instability is associated with the presence of Lynch syndrome pan-cancer. *J Clin Oncol*. 2019;37:286–295.
41. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised international staging system for multiple myeloma: A report from international myeloma working group. *J Clin Oncol*. 2015;33:2863–2869.
42. Van Laar RK, Borrelo I, Jabalayan D, et al. MyPRSR molecular subtypes of multiple myeloma represent all high-risk FISH translocations included in the mSMART 2.0 and R-ISS guidelines. *Blood*. 2016;128:3264–3264.
43. Samur MK, Aktas Samur A, Fulciniti M, et al. Genome-wide somatic alterations in multiple myeloma reveal a superior outcome group. *J Clin Oncol*. 2020;38:3107–3118.

44. An X, Tiwari AK, Sun Y, et al. BCR-ABL tyrosine kinase inhibitors in the treatment of Philadelphia chromosome positive chronic myeloid leukemia: A review. *Leuk Res*. 2010;34:1255–1268.
45. Sponghini AP, Rondonotti D, Giavarra M, et al. Safety and efficacy of vemurafenib in BRAF V600E mutation-positive metastatic melanomas. *J Transl Med*. 2015;13:P7.
46. Ehsan H, Wahab A, Shah Z, et al. Role of venetoclax in the treatment of relapsed and refractory multiple myeloma. *J Hematol (Brossard)*. 2021;10:89–97.
47. Rajkumar SV. Multiple myeloma: 2022 update on diagnosis, risk stratification, and management. *Am J Hematol*. 2022;97:1086–1107.
48. Kortüm KM, Mai EK, Hanafiah NH, et al. Targeted sequencing of relapsed/refractory myeloma patients identifies an enrichment of mutations in MAPK and Cereblon pathways. *Clin Lymphoma Myeloma Leuk*. 2015;15:e83.
49. Laganà A. Computational approaches for the investigation of intra-tumor heterogeneity and clonal evolution from bulk sequencing data in precision oncology applications. *Adv Exp Med Biol*. 2022;1361:101–118.
50. McGranahan N, Swanton C. Biological and therapeutic impact of intratumor heterogeneity in cancer evolution. *Cancer Cell*. 2015;28:141.
51. Keats JJ, Chesi M, Egan JB, et al. Clonal competition with alternating dominance in multiple myeloma. *Blood*. 2012;120:1067–1076.
52. Landau DA, Carter SL, Stojanov P, et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152:714–726.
53. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366:883–892.
54. Chari A, Vogl DT, Gavriatopoulou M, et al. Oral selinexor-dexamethasone for triple-class refractory multiple myeloma. *N Engl J Med*. 2019;381:727–738.
55. Gupta VA, Barwick BG, Matulis SM, et al. Venetoclax sensitivity in multiple myeloma is associated with B-cell gene expression. *Blood*. 2021;137:3604–3615.
56. Sheykhhasan M, Ahmadieh-Yazdi A, Vicidomini R, et al. CAR T therapies in multiple myeloma: Unleashing the future. *Cancer Gene Ther [Internet]*. 2024. http://dx.doi.org/10.1038/s41417-024-00750-2.
57. Zhou X, Xiao X, Kortuem KM, et al. Bispecific antibodies in the treatment of multiple myeloma. *Hematol Oncol Clin North Am [Internet]*. 2024. http://dx.doi.org/10.1016/j.hoc.2023.12.003.
58. Munshi NC, Anderson LD Jr, Shah N, et al. Idecabtagene vicleucel in relapsed and refractory multiple myeloma. *N Engl J Med*. 2021;384:705–716.
59. Berdeja JG, Madduri D, Usmani SZ, et al. Ciltacabtagene autoleucel, a B-cell maturation antigen-directed chimeric antigen receptor T-cell therapy in patients with relapsed or refractory multiple myeloma (CARTITUDE-1): A phase 1b/2 open-label study. *Lancet*. 2021;398:314–324.
60. San-Miguel J, Dhakal B, Yong K, et al. Cilta-cel or standard care in lenalidomide-refractory multiple myeloma. *N Engl J Med*. 2023;389:335–347.
61. Moreau P, Garfall AL, van de Donk NWCJ, et al. Teclistamab in relapsed or refractory multiple myeloma. *N Engl J Med*. 2022;387:495–505.
62. Bahlis NJ, Costello CL, Raje NS, et al. Elranatamab in relapsed or refractory multiple myeloma: The MagnetisMM-1 phase 1 trial. *Nat Med*. 2023;29:2570–2576.
63. Rose S. MonumenTAL results for talquetamab in myeloma. *Cancer Discov*. 2023;13:250–251.
64. Trudel S, Cohen AD, Krishnan AY, et al. Cevostamab monotherapy continues to show clinically meaningful activity and manageable safety in patients with heavily pretreated relapsed/refractory multiple myeloma (RRMM): Updated results from an ongoing phase I study. *Blood*. 2021;138:157–157.
65. Mouhieddine TH, Van Oekelen O, Melnekoff DT, et al. Sequencing T-cell redirection therapies leads to deep and durable responses in patients with relapsed/refractory myeloma. *Blood Adv*. 2023;7:1056–1064.
66. Lee H, Ahn S, Maity R, et al. Mechanisms of antigen escape from BCMA- or GPRC5D-targeted immunotherapies in multiple myeloma. *Nat Med*. 2023;29:2295–2306.
67. Derrien J, Gastineau S, Frigout A, et al. Acquired resistance to a GPRC5D-directed T-cell engager in multiple myeloma is mediated by genetic or epigenetic target inactivation. *Nat Cancer*. 2023;4:1536–1543.
68. Landgren O, Kristinsson SY, Goldin LR, et al. Risk of plasma cell and lymphoproliferative disorders among 14621 first-degree relatives of 4458 patients with monoclonal gammopathy of undetermined significance in Sweden. *Blood*. 2009;114:791–795.
69. Went M, Sud A, Försti A, et al. Author correction: Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma. *Nat Commun*. 2019;10:213.
70. Thibaud S, Etra A, Subaran R, et al. Pathogenic germline variants in multiple myeloma. *Blood*. 2021;138:399–399.
71. Elnaggar M, Agte S, Restrepo P, et al. Triple MAPK inhibition salvaged a relapsed post-BCMA CAR-T cell therapy multiple myeloma patient with a BRAF V600E subclonal mutation. *J Hematol Oncol*. 2022;15:109.
72. Adamopoulos C, Ahmed TA, Tucker MR, et al. Exploiting allosteric properties of RAF and MEK inhibitors to target therapy-resistant tumors driven by oncogenic BRAF signaling. *Cancer Discov*. 2021;11:1716–1735.
73. Barretina J, Caponigro G, Stransky N, et al. The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483:603–607.
74. Yang W, Soares J, Greninger P, et al. Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2013;41:D955–D961.
75. Rojas LA, Sethna Z, Soares KC, et al. Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature*. 2023;618:144–150.
76. Perumal D, Imai N, Laganà A, et al. Mutation-derived neoantigen-specific T-cell responses in multiple myeloma. *Clin Cancer Res*. 2020;26:450–464.
77. Deshpande S, Tytarenko RG, Wang Y, et al. Monitoring treatment response and disease progression in myeloma with circulating cell-free DNA. *Eur J Haematol*. 2021;106:230–240.

32

Nanopore Whole Transcriptome Sequencing Offers the Potential for Accessible Classification of Pediatric Cancers

Thomas B. Alexander and Jeremy R. Wang

32.1 Introduction

32.1.1 Rationale for Novel Diagnostic Approaches

Over 80% of pediatric cancers occur in low- and middle-income countries (LMICs), where most health centers lack access to the key tools required for accurate diagnosis, disease classification, or appropriate treatment.[1–4] Consequently, while over 80% of children with cancer in high-income countries (HICs) like the United States will survive, less than 30% of children with cancer in LMICs will survive. This discrepancy represents perhaps one of the largest health-related disparities in global health. Deficiencies in several domains contribute to this survival disparity, including multilevel financial hardship, a trained and dedicated workforce, access to medications, and diagnostic testing. Improving access to accurate diagnostics is critical to ensure a solid foundation for accurate treatment and program monitoring in LMICs. Recognizing the diagnostic barriers and the growing disparity, the Global Initiative for Childhood Cancer has placed diagnostics high on their list of priority areas for innovation.[5]

A comprehensive clinical diagnosis of childhood cancer has traditionally relied on multiple costly and time-consuming techniques such as immunohistochemistry (IHC), flow cytometry, cytogenetics, fluorescence *in situ* hybridization (FISH), targeted PCR panels, and microarrays.[6–8] Most facilities in LMICs lack access to some or all of these traditional diagnostic tests.[1, 9–14]

Genomic subtypes of leukemia have been shown to have distinct patterns of gene expression as revealed by microarray or, more recently, whole transcriptome sequencing (WTS).[15–19] Gene expression profiling (GEP) is a powerful tool for discovery, and gene expression panels provide clinically actionable information in many subtypes of malignancies.[15, 17, 20–24] Comprehensive classification of acute leukemia using single short-read WTS is established in the research setting and moving toward some clinical settings, allowing determination of leukemia lineage, genomic subtypes, small mutations, and aneuploidy.[18, 25] However, the capital, maintenance, and operating costs of short-read WTS limit the global scalability.

32.1.2 Rationale for Nanopore Technology

Oxford Nanopore Technologies' (ONT, "nanopore") long-read sequencing platforms are a potential molecular diagnostic tool for diverse cancers. In contrast to most commercial sequencing approaches that utilize sequencing-by-synthesis, fluorescence, and optical data capture, ONT is an electric current–based approach. In ONT, strands of DNA or RNA are passed through small electrically charged pores. As nucleic acids (and their modifications) pass through the pores, they cause different perturbations to the electric current, which is translated into base calls. ONT can work on native RNA or DNA and on PCR-amplified cDNA or DNA. When sequencing native DNA or RNA, modifications such as methylation can also be detected from the same raw data.

The platform requires minimal capital investment and limited lab infrastructure and space, and it is amenable to single or multiplex sample runs with costs reduced by an order magnitude compared with next-generation sequencing-by-synthesis technology.[26, 27] Development of assays using ONT could lead to financially accessible, scalable, and cost-effective solutions across various settings, with near-term clinical implications. The cost of ONT's MinION sequencer is $1500 USD, including an external operating computer. This approach obviates the need for massive scale capital investments (hundreds of thousands USD startup costs per site), all of which are required using alternative higher cost short-read sequencing or molecular capture technologies (e.g., Illumina or NanoString). Given the financial accessibility of this technical approach, pediatric oncology researchers that we collaborate with in Malawi, Pakistan, India, Brazil, and Mexico have recently demonstrated the capacity to acquire nanopore sequencing and reagents, learn and perform library preparation protocols, generate high-quality sequencing data, and jointly analyze the results for classification and discovery using pediatric cancer specimens. This chapter summarizes the materials, methods, and results from specimens sequenced at the University of North Carolina.

32.2 Materials

32.2.1 Specimens

Samples representing the breadth of common pediatric leukemia, lymphoma, and solid tumors were acquired from tissue repositories from multiple pediatric cancer centers (Table 32.1). Patients were consented for secondary genetic research approved through their respective institutional review

DOI: 10.1201/9781003449942-36

TABLE 32.1

The Distribution of Specimens Sequenced by Tumor Type

Diagnosis	Abbreviation	# Specimens
B-cell Acute Lymphoblastic Leukemia	B-ALL	176
T-cell Acute Lymphoblastic Leukemia	T-ALL	35
Acute Myeloid Leukemia	AML	84
Rhabdomyosarcoma	RMS	38
Neuroblastoma	NB	33
Ewing Sarcoma	ES	38
Burkitt Lymphoma	BL	23
T-Lymphoblastic Lymphoma	TLL	3
Diffuse Large B Cell Lymphoma	DLBCL	34
Wilm's Tumor	WT	18
Anaplastic Large Cell Lymphoma	ALCL	13
Desmoplastic Small Round Cell Tumors	DSRCT	10
Classical Hodgkins Lymphoma	CHL	17
Nodular Lymphocyte-Predominant Hodgkin Lymphoma	NLPHL	14

boards. Specimen types were varied and included cryopreserved peripheral blood and bone marrow mononuclear cells (for leukemia), flash-frozen xenograft tissue (from some solid tumors), and formalin-fixed paraffin-embedded (FFPE) slides and scrolls, some stained (from most lymphomas and solid tumors). A total of 536 specimens were collected and analyzed using this approach; the distribution of cancer and specimen types is shown in Table 32.1.

32.2.2 RNA Extraction, Quality Control, and Library Preparation Materials

RNA was extracted from cryopreserved specimens using Zymo Quick DNA/RNA Miniprep kit or Zymo Quick DNA/RNA Magbead kit. FFPE specimens were extracted using the Zymo Quick DNA/RNA FFPE kit. RNA was quantified using Invitrogen Qubit RNA kit and quality was assessed by 1% agarose gel or Bioanalyzer (Agilent). Libraries were prepared for nanopore sequencing using preparation kits from Oxford Nanopore Technologies, including Direct RNA Sequencing Kit (SQK-RNA002), Direct cDNA Sequencing Kit (SQK-DCS109), and PCR-cDNA Sequencing Kits with or without barcoding (SQK-PCB109, SQK-PCB110, and SQK-PCS109). Additional reagents required for reverse transcription, PCR, and reaction cleanup include Maxima H- Reverse Transcriptase (Thermo Fisher Scientific), LongAmp PCR Master Mix (2X, New England Biosciences), and Ampure XP SPRI beads (Beckham Coulter).

32.2.3 Sequencing and Computational Hardware

Oxford Nanopore Technologies' (ONT) platforms used for sequencing included MinION Mk1B, MinION Mk1C, GridION, and PromethION P2 Solo using Flongle (FLG001), MinION (FLO-MIN109), or PromethION (FLO-PRO002) flow cells with kit 9 or kit 10 chemistry (*109, *110 kits). Basecalling was performed either on the sequencing device (MinION Mk1C and GridION) or on attached computer with GPU following ONT's recommended host computer specifications. Sequence analysis, including alignment, quantification, and machine learning-based classification, were performed on UNC's Research Computing cluster, a high-performance UNIX cluster running the SLURM scheduler. However, sample classification using a pretrained model requires modest computational resources broadly available in consumer laptop/desktop computers; post-basecalling informatics and classification can be performed in five to ten minutes on a single CPU thread using less than 4 GB of main memory (RAM).

32.3 Methods

32.3.1 Specimen Collection and Storage

A wide range of specimen collection, preservation, and storage conditions are compatible with the nanopore RNA sequencing protocol. We have validated recovery of high-quality intact RNA from cryopreserved mononuclear cells isolated from peripheral blood or bone marrow aspirate using standard density-gradient centrifugation (Ficoll). Alternatively, mononuclear cell lysate can be preserved in Zymo DNA/RNA Shield at –20 °C, obviating the need for cryopreservation. In either case, 1 million cells are recommended to yield sufficient RNA for sequencing.

Solid tumor biopsies for primary diagnosis are often preserved as formalin-fixed paraffin-embedded sections. Especially in low-resource settings, cryopreservation of solid tissue is rare. FFPE preparation processes vary widely, including variation in formalin fixative, fixation time, and drying. Similarly, FFPE specimens are stored variably at 4 °C or ambient temperature indefinitely. The cohort we discuss includes FFPE specimens stored for more than ten years, and the described protocol produces robust results independent of FFPE protocol variation and storage conditions. Recovery of RNA from FFPE sections or slides varies based on the thickness and surface area of the captured tissue, but typically two

(2) 5 μm sections or one (1) 10 μm section produces sufficient RNA for sequencing. Sections mounted on slides, including stained slides (e.g., common hematoxylin and eosin; H&E) can be removed into a microfuge tube for extraction by scraping with a sterile, nuclease-free razor blade (blade can be treated with RNase-Zap [Thermo Fisher Scientific] or similar to reduce the risk of RNase contamination).

32.3.2 RNA Quality Control

To proceed, extracted RNA should exceed 20 ng/μl, measured by fluorometry (e.g., Invitrogen Qubit). Most samples, especially fresh-frozen, typically exceed 100 ng/μl. RNA from fresh or cryopreserved samples, or samples stored in Zymo DNA/RNA Shield, is expected to be largely intact. Visualized on an agarose gel, it should show two sharp bands representing intact large- (23S) and small-subunit (18S) ribosomal RNA, corresponding to a high RNA integrity number (RIN) if analyzed by capillary electrophoresis (e.g., Agilent Bioanalyzer or Tapestation). RNA extracted from FFPE samples is typically severely degraded and rarely shows fragments > 500 nt by electrophoresis. However, we have observed no practical lower limit on the fragment size of degraded RNA as long as the RNA quantity is sufficient and quality control for fragment size by electrophoresis is typically unnecessary for FFPE-derived RNA.

32.3.3 Library Preparation and Sequencing

Samples were prepared for sequencing using ONT PCR-cDNA Barcoding Kit (SQK-PCB109 or SQK-PCB110) according to the manufacturer's recommended protocol, with the noted modifications. Briefly, 100 ng total RNA was reverse-transcribed using Maxima H-Reverse Transcriptase (Thermo Fisher Scientific) with initiation primer targeting the poly-A tail and a strand-switching primer to enable amplification of full-length poly-adenylated mRNA fragments. The resulting cDNA was amplified using barcoded PCR primers. FFPE samples were amplified for 20 cycles and fresh-frozen for 14 cycles to produce sufficient cDNA for sequencing while minimizing PCR bias. Amplified cDNA was purified using Ampure XP (SPRI) beads (Beckman Coulter). cDNA libraries were quantified using Qubit fluorometer (Invitrogen) dsDNA high sensitivity (HS) kit. Barcoded samples were mixed in equimolar volumes and a total of 50 fmol of the multiplexed library was used for rapid adapter ligation and loaded on the flow cell. Flow cells were loaded and sequencing initiated according to the manufacturer's protocol. Sequencing, optionally including real-time basecalling, was initiated and monitored using ONT's MinKNOW software.

32.3.4 Computational Analysis and Classification

Raw signal data was base-called, demultiplexed, and trimmed using Guppy in high-accuracy mode. Base-called reads were aligned to the Ensembl (v110) GRCh38 cDNA reference transcriptome using minimap2 (with -x map-ont).[28] Reads were assigned 1:1 to unique transcripts by choosing the best-mapping transcript target for each read (with the highest total matching residues) or a fractional count where multiple transcripts have equivalent alignment accuracy. Initial quality controls evaluated the total number of reads per sample, alignment rate (%), and calculated Shannon entropy[29] of the transcript abundances.

To account for sparsity from low-coverage sequencing, the initial gene expression table was filtered to exclude genes absent (zero reads) in greater than 1% of samples. We developed a novel composite machine learning framework consisting of binary pairwise and one-vs-all partial least-squares (PLS) regressions followed by multi-class support vector machine (SVM) classification. A set of binary partial least-squares regressions were performed on each pair (e.g., B-ALL and T-ALL) and each class against all others (e.g., AML and non-AML), for each of ($n \in [5,11]$)) output components. The PLS vectors form a feature matrix to train a support vector machine (SVM) classifier. The SVM outputs probabilities for each class (lineage, tumor type, or molecular subtype) for each tested sample.

32.3.5 Evaluation and Validation

For all cases, we compare the results from our machine learning approach to the standard of care diagnosis made at the institutions providing the specimens to our lab (UNC Hospitals, St. Jude Children's Research Hospital, or Children's Wisconsin), including morphology, immunohistochemistry, fluorescence *in situ* hybridization, and molecular diagnostic evidence such as targeted PCR or traditional high-throughput sequencing. We performed leave-one-out cross-validation, excluding patient-matched samples from training when testing technical and biological replicates.

32.4 Results

We developed a novel composite classification approach to predict tumor type/cell lineage and molecular subtypes directly from gene expression profiles. Training features include relative gene expression values derived from the generated nanopore sequencing data and augmented with 1261 transcriptomes across a subset of tumor types sequenced on the Illumina platform publicly available on the St. Jude Children's Research Hospital Cloud "https://www.stjude.cloud/). A composite machine learning model consisting of partial least-squares regression and support vector machine stages (PLS-SVM) was identified as the most robust classification approach for these low-coverage nanopore RNA sequencing-based expression profile data. PLS regression is well suited to learning binary classifications among these heterogeneous data sets as maximally informative features are trained to carry greater weight in the model.

32.4.1 Leukemia Lineage Classification

An initial step is in the diagnostic classification and treatment selection of acute leukemia determining whether the case is B-cell acute lymphoblastic leukemia (B-ALL), T-cell acute lymphoblastic leukemia (T-ALL), or acute myeloid leukemia (AML). We have performed nanopore WTS on 395 specimens

of acute leukemia, including 216 distinct cases and 179 technical and biological replicates. Our classification model identifies the correct lineage in 98.2% (388/395) of cases (Figure 32.1). We further stratified these by prediction probability from our model, resulting in 96.2% (380/395) with a confidence of > 0.8. This subset achieves an accuracy of 99.2% (377/380). All samples with confidence > 0.95 (359/395, 90.9%) are correctly classified.

32.4.2 B-ALL and AML Genomic Subtype Classification

B-ALL and AML prognosis and treatment is partially determined by genomic subtype. For core clinically relevant genomics subtypes of B-ALL and AML, 323 samples had known molecular subtypes from the primary diagnostic workup. Due to limited sample sizes for many uncommon genomic subtypes and broad expression-level similarity among some subsets, we classify B-ALL into six groups (*KMT2A*r, Ph/Ph-like, *TCF3::PBX1*, *ETV6::RUNX1*, Hyperdiploid/Near haploid, and Other). Similarly, for AML, we use *KMT2A*r, Core binding factor rearranged, and Other. In both cases, we group known molecular subtypes not falling into one of the remaining groups into "Other." We observed an overall accuracy of 89.8% (290/323), and 78.6% (254/323) are classifiable with prediction probability of > 0.8 with an accuracy of 96.1% (Figure 32.2).

32.4.3 Lymphoma Classification (FFPE Derived)

The classification of pediatric lymphoma can be challenging in locations without experienced pathologists and reliable immunohistochemistry. We organized a cohort of pediatric lymphoma specimens from several centers in the United States and collaborators in Malawi. All lymphoma cases were stored as FFPE and therefore all our training and cross-validation is from RNA derived from FFPE. With our discovery effort, we showed an overall accuracy of 89.7% (104/116), and 72.4% of cases (84/116) are classifiable with a prediction probability of > 0.6 with an accuracy of 97.6% (82/84) (Figure 32.3).

32.4.4 Solid Tumor Classification (FFPE Derived)

Like our lymphoma cohort, our pediatric solid tumor cohort was primarily made up of cases from FFPE derived RNA, with a few exceptions where we used frozen tissue from patient derived xenograft models obtained from the Childhood Solid Tumor Network at St. Jude. We achieve an overall accuracy of 94.9% (131/138), and 85.5% of cases (118/138) are classifiable

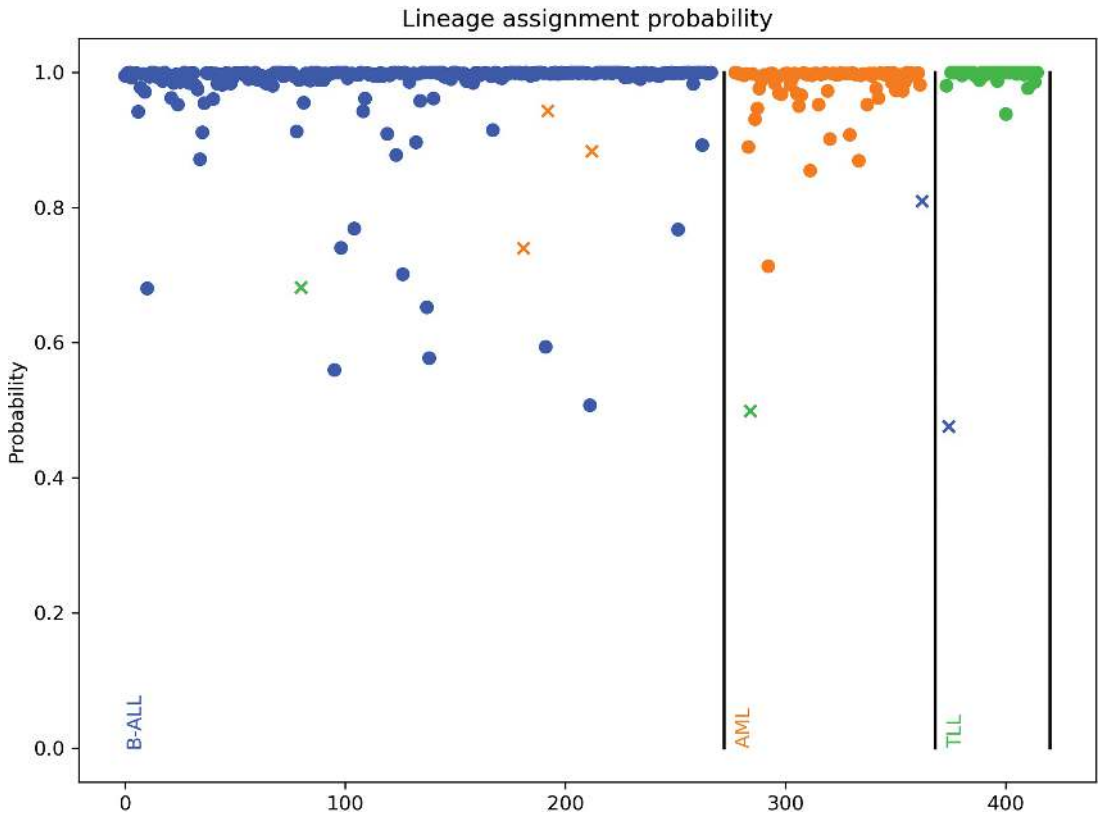


FIGURE 32.1 Pediatric acute leukemia samples stratified by prediction probability. Dots represent lineage classifications consistent with the standard-of-care diagnosis (typically lineage is determined by flow cytometry) and Xs represent inconsistent calls, where the color indicates the predicted lineage. Our model's (uncalibrated) prediction probabilities (on the Y-axis) clearly stratify samples by accuracy, achieving 100% accuracy above a cutoff of 0.91.

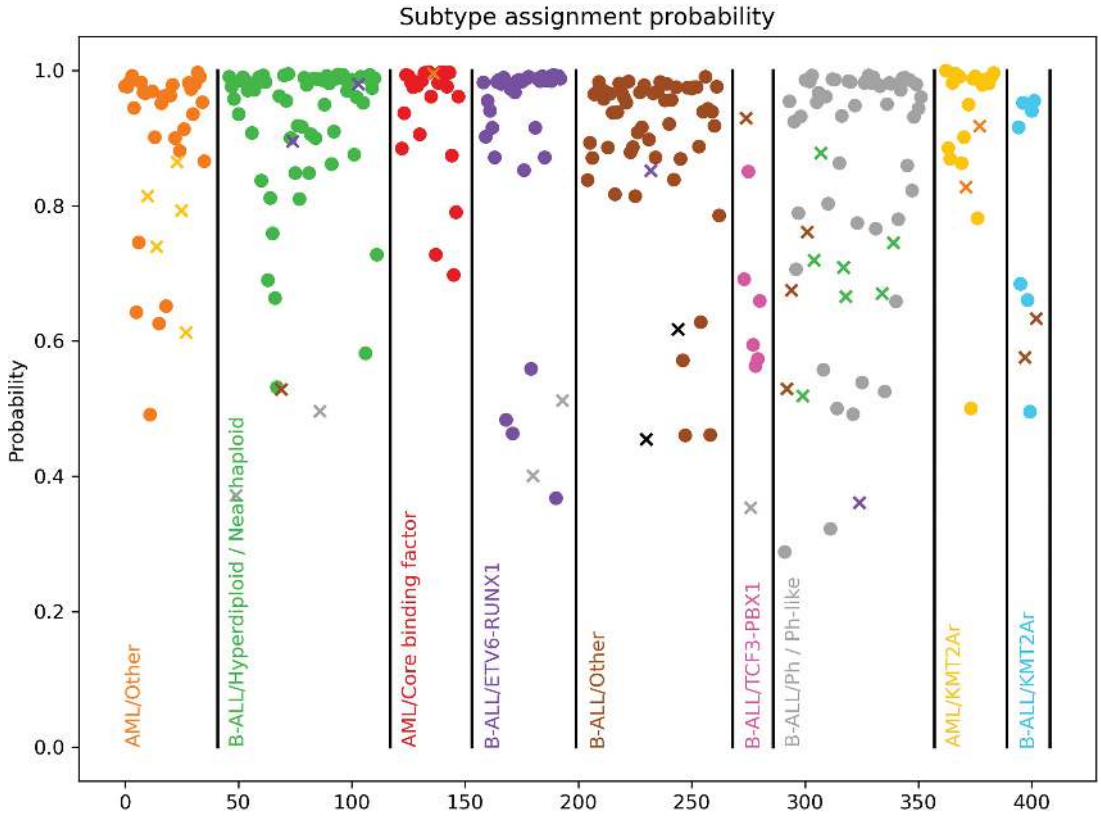


FIGURE 32.2 Molecular subtypes of pediatric acute leukemia stratified by prediction probability. Dots represent primary classifications consistent with the standard-of-care diagnosis and Xs represent inconsistent calls, where the color indicates the predicted tumor type. Our model's (uncalibrated) prediction probabilities (on the Y-axis) clearly stratify samples by accuracy.

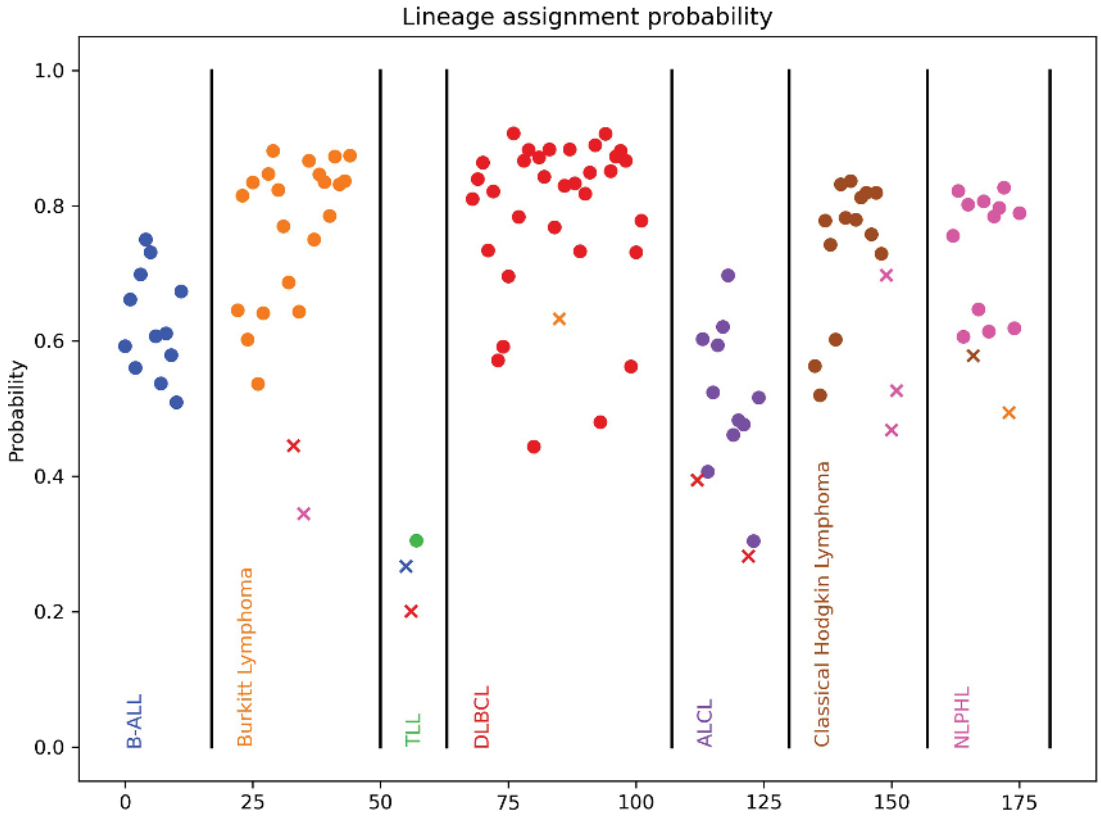


FIGURE 32.3 Pediatric lymphoma cases stratified by prediction probability. Dots represent primary classifications consistent with the standard-of-care diagnosis and Xs represent inconsistent calls, where the color indicates the predicted tumor type. Our model's (uncalibrated) prediction probabilities (on the Y-axis) clearly stratify samples by accuracy.

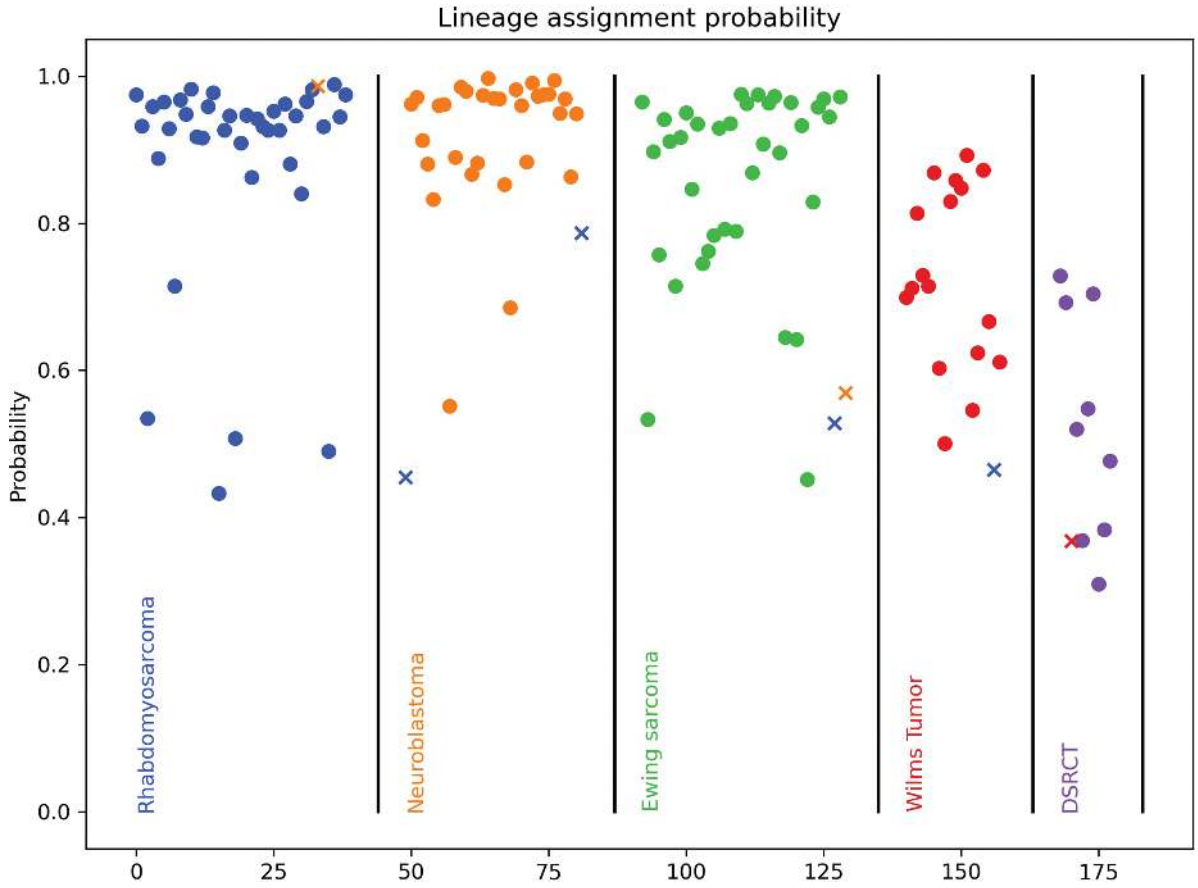


FIGURE 32.4 Pediatric solid tumor cases stratified by prediction probability. Dots represent primary classifications consistent with the standard-of-care diagnosis and Xs represent inconsistent calls, where the color indicates the predicted tumor type. Our model's (uncalibrated) prediction probabilities (on the Y-axis) clearly stratify samples by accuracy.

with a prediction probability of > 0.6 with an accuracy of 98.3% (116/118) (Figure 32.4).

32.5 Discussion

This chapter demonstrates the potential of low coverage nanopore transcriptome sequencing coupled with a novel machine learning approach to robustly characterize diverse pediatric cancers. By design, this approach is applicable to a wide variety of cancer types—this chapter shows the most common pediatric leukemias, lymphomas, and common solid tumors—with minimal modification, and robust to heterogeneity in specimen handling (incl. fresh-frozen, preserved lysate, and FFPE), library preparation, sequencing platform, and sequencing depth.

32.5.1 Informatics and Classification Model Parameters

Our novel machine learning approach robustly classifies tumor type and molecular subtype among these heterogeneous data. This model is sensitive to the feature (gene) filtering prior to model training. We exclude genes for which we observe no reads in greater than 1% of samples. This effectively excludes the vast majority of lowly or sporadically expressed genes, including those that may be characteristic of a particular tumor type. Low coverage sequencing results in sparse sampling of the transcriptome such that genes with low expression are often not observed at all. Without this pre-filtering, these sparsely observed genes can cause overfitting of the machine learning model and, in fact, we see lower accuracy using the same model when this gene filtering cutoff is relaxed. Partial least-squares regression effectively identifies a latent feature space that maximizes binary discrimination between classes despite heterogeneity across sequencing platforms, cDNA generation, and library preparation, including a combination Illumina and ONT sequence data.

32.5.2 Accuracy and Prediction Probability

This chapter presents classification results across 15 tumor types across leukemias, lymphomas, and solid tumors, and nine genomic subtypes of B-ALL and AML. The number of specimens for these distinct classes ranges from three to 176. As such, there is a clear relationship between the number of specimens available for model training and both the overall accuracy and the discriminatory power of the resulting prediction probabilities. The prediction probabilities reported are output from our composite model and do not represent a

"calibrated" probability of accurate classification. Our leukemia lineage classification has uniformly higher confidence and consistency with the clinical diagnosis than, for example, lymphoma cases, where there are far fewer specimens per tumor type. As the cohort of specimens analyzed using this approach expands, we expect the performance will continue to improve, especially among more tumor types with fewer samples. Similarly, we currently group B-ALL and AML molecular subtypes into relatively coarse bins that do not necessarily coincide with disease risk groups (e.g., hyperdiploid/near haploid) in part due to limited power of the current model to distinguish these groups based on a small number of cases. Expansion of the "Other" clusters into their precise (and biologically distinct) molecular diagnoses will be particularly important to better align our model with clinically relevant risk groups.

32.5.3 Advantages and Limitations

The cancer "diagnostic gap" between LMICs and HICs is a sobering, critical challenge to the improvement of outcomes for patients with cancer globally. Closing this gap is the primary motivation for development of the research methods presented in this chapter. ONT sequencing, utilizing the methods we and others have developed, provides potential for improving diagnostic classification across resource settings. The primary potential benefit of ONT WTS for leukemia, lymphoma, and pediatric solid tumor classification is the low coverage required and therefore low cost per sample, as well as the minimal laboratory setup and capital cost. As we are demonstrating with collaborators, this technology can be acquired without large capital costs, while staff with basic molecular lab foundation can be easily trained to perform the protocol. Moreover, the computational requirements to support base calling, transcript alignment, and classification are limited without the requirement of high-performance computing clusters. This approach may have application in locations with limited or inconsistent access to advanced diagnostic tests.

If the high accuracy of classifiable sample can be validated across the broad range of genomic subtypes of acute leukemia and in different geographic and resource settings, the addition of lineage histologic subtype and limited genomic subtype information has the potential to provide a large relative advantage to the childhood cancer community. The calculation of relative advantage includes comparison with locally available standards of care in the context of use.

Current ONT approaches for whole transcriptome sequencing have limited sensitivity for direct identification of fusions transcripts. The limitation is primarily related to the poly-A dependent library preparations. Direct identification of fusion transcripts requires sufficiently intact RNA, efficient reverse-transcription, and PCR to sequence from 5' tail through the fusion breakpoint. Often, the fusion can be directly identified, especially for cases with RNA quality and sequenced at higher depth. However, in cases with breakpoint many kilobases away from the 5' tail (e.g., *BCR::ABL1*), it is challenging to reliably identify the fusion. Further development is also required for WTS derived digital karyotype, which is required to improve accuracy of aneuploid driven classification of cancers, such as B-ALL. Despite these limitations, we and others are pushing boundaries of ONT to understand flexibility of specimen input, potential sensitivity, and specificity of classification for a variety of cancer subtypes and cost requirements determined by sequencing depth and computational requirements.

32.6 Future Trends

Over the last two decades, classification of cancer increasingly relies on molecular classification. Next-generation sequencing solutions such as WTS or WGS for B-ALL and AML, or methylation profiling for CNS malignancies, are the most established approaches. These will provide a solution in settings with sufficient resources and will provide a benchmark against which other technologies will be measured. However, the growth in cancer control and treatment programs globally demand development of molecular diagnostic solutions that are accessible, scalable, and sustainable. Technical and computation advancement on the ONT platform will provide some of these solutions.

WGS for leukemia classification on ONT could be another tool for genomic classification of acute leukemia. Potential advantages could be rapid turnaround, easier resolution of structural variants than short-read sequencing, and capturing additional information such as pharmacogenomics on a single assay. The capital and computational costs are higher than mutation panels or non-adaptive WTS or WGS approaches but remain considerably lower than alternative WGS or WTS solutions from short-read sequencing. With current technology, this is not a solution for the lowest resource settings, but with ongoing improvements and healthcare systems development broadly, this may become accessible in the intermediate term.

Methylation profiling from FFPE-based samples on Illumina platforms is revolutionizing diagnostic classification for central nervous system tumors.[30, 31] Methylation profile with ONT using FFPE-derived specimens has current technical limitations. Methylation calling on ONT is based upon direct sequencing of native nucleic acid (rather than a bisulfite or other conversion step). Therefore, the residual DNA damage and residual crosslinks on the native DNA introduced by clinical FFPE protocols limited the throughput on current nanopore flow cells. Despite this limitation, early publications show the potential of methylation profiling.[32] It is likely that technical advances in DNA preparation from FFPE and/or pores on the ONT flow cells will, in the future, offer a revolutionary opportunity to utilize the throughput, cost, and scalable benefits of ONT.

This chapter describes the potential of ONT for tumor classification across pediatric tumor types and specimen storage types. But early potential is a long way from hardened validation, clinical use, and health benefit. Current technical protocols, computational approaches, and training datasets will require iterative improvement and expansion. For this reason, we are actively collaborating with scientists and translational researchers globally to collectively push this effort. Researchers in Karachi, Pakistan; Lilongwe, Malawi; Barretos, Brazil; Kolkata, India; Mumbai, India; Kampala, Uganda; Yerevan, Armenia; and Accra, Ghana are actively contributing

to the discussion through local ONT sequencing, contribution of sequencing data, performance validation experiments, and development of new computational approaches. This type of global collaboration for the *collective development* of new diagnostic approaches, without shipping specimens internationally, is consistent with the goals of closing the diagnostic gap with resource-appropriate, scalable, accessible solutions.

REFERENCES

1. Bhakta N, Force LM, Allemani C, et al. Childhood cancer burden: A review of global estimates. *Lancet Oncol* 2019; 20(1): e42–53.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; 68(6): 394–424.
3. Bray F, Soerjomataram I. The changing global Burden of cancer: Transitions in human development and implications for cancer prevention and control. In: Gelband H, Jha P, Sankaranarayanan R, Horton S, eds. *Cancer: Disease Control Priorities* (Third Edition, Volume 3). Washington, DC: International Bank for Reconstruction and Development/The World Bank; 2015.
4. Ward ZJ, Yeh JM, Bhakta N, Frazier AL, Atun R. Estimating the total incidence of global childhood cancer: A simulation-based analysis. *Lancet Oncol* 2019; 20(4): 483–93.
5. WHO Global Initiative for Childhood Cancer: An Overview. https://www.who.int/docs/default-source/documents/health-topics/cancer/who-childhood-cancer-overview-booklet.pdf?sfvrsn=83cf4552_1&download=true.
6. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 2016; 127(20): 2391–405.
7. Iacobucci I, Mullighan CG. Genetic basis of acute lymphoblastic leukemia. *J Clin Oncol* 2017; 35(9): 975–83.
8. Narayanan D, Weinberg OK. How I investigate acute myeloid leukemia. *Int J Lab Hematol* 2020; 42(1): 3–15.
9. Makau-Barasa LK, Greene SB, Othieno-Abinya NA, Wheeler S, Skinner A, Bennett AV. Improving access to cancer testing and treatment in Kenya. *J Glob Oncol* 2018; 4: 1–8.
10. Bansal D, Davidson A, Supriyadi E, Njuguna F, Ribeiro RC, Kaspers GJL. SIOP PODC adapted risk stratification and treatment guidelines: Recommendations for acute myeloid leukemia in resource-limited settings. *Pediatr Blood Cancer* 2019: e28087.
11. Dayton V, Nguyen CK, Van TT, et al. Evaluation of opportunities to improve hematopathology diagnosis for Vietnam pathologists. *Am J Clin Pathol* 2017; 148(6): 529–37.
12. Fadoo Z, Nisar I, Yousuf F, et al. Clinical features and induction outcome of childhood acute lymphoblastic leukemia in a lower/middle income population: A multi-institutional report from Pakistan. *Pediatr Blood Cancer* 2015; 62(10): 1700–8.
13. Rubagumya F, Xu MJ, May L, et al. Outcomes of low-intensity treatment of acute lymphoblastic leukemia at Butaro cancer center of excellence in Rwanda. *J Glob Oncol* 2018; 4: 1–11.
14. Slone JS, Chunda-Liyoka C, Perez M, et al. Pediatric malignancies, treatment outcomes and abandonment of pediatric cancer treatment in Zambia. *PLoS One* 2014; 9(2): e89102.
15. Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: A genome-wide classification study. *Lancet Oncol* 2009; 10(2): 125–34.
16. Ferrando AA, Neuberg DS, Staunton J, et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell* 2002; 1(1): 75–87.
17. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. *Science* 1999; 286(5439): 531–7.
18. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet* 2019; 51(2): 296–307.
19. Yeoh EJ, Ross ME, Shurtleff SA, et al. Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell* 2002; 1(2): 133–43.
20. Harrell JC, Prat A, Parker JS, et al. Genomic analysis identifies unique signatures predictive of brain, lung, and liver relapse. *Breast Cancer Res Treat* 2012; 132(2): 523–35.
21. Moffitt RA, Marayati R, Flate EL, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* 2015; 47(10): 1168–78.
22. Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med* 2009; 360(5): 470–80.
23. Puvanesarajah S, Nyante SJ, Kuzmiak CM, et al. PAM50 and risk of recurrence scores for interval breast cancers. *Cancer Prev Res (Phila)* 2018; 11(6): 327–36.
24. Stratford JK, Bentrem DJ, Anderson JM, et al. A six-gene signature predicts survival of patients with localized pancreatic ductal adenocarcinoma. *PLoS Med* 2010; 7(7): e1000307.
25. Arindrarto W, Borras DM, de Groen RAL, et al. Comprehensive diagnostics of acute myeloid leukemia by whole transcriptome RNA sequencing. *Leukemia* 2021; 35: 47–61. https://doi.org/10.1038/s41375-020-0762-8
26. Magdy T, Kuo HH, Burridge PW. Precise and cost-effective nanopore sequencing for post-GWAS fine-mapping and causal variant identification. *iScience* 2020; 23(4): 100971.
27. Stockton JD, Nieto T, Wroe E, et al. Rapid, highly accurate and cost-effective open-source simultaneous complete HLA typing & phasing of Class I & II alleles using nanopore sequencing. *HLA* 2020; 96(2): 163–178. doi: 10.1111/tan.13926. PMID: 32419382.
28. Li H. Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics* 2018; 34(18): 3094–100.
29. Shannon CE, Weaver W. *The Mathematical Theory of Communication*. Urbana: University of Illinois Press; 1949.
30. Wu Z, Abdullaev Z, Pratt D, et al. Impact of the methylation classifier and ancillary methods on CNS tumor diagnostics. *Neuro Oncol* 2022; 24(4): 571–81.
31. Pickles JC, Fairchild AR, Stone TJ, et al. DNA methylation-based profiling for paediatric CNS tumour diagnosis and treatment: A population-based study. *Lancet Child Adolesc Health* 2020; 4(2): 121–30.
32. Afflerbach AK, Albers A, Appelt A, et al. Nanopore sequencing from formalin-fixed paraffin-embedded specimens for copy-number profiling and methylation-based CNS tumor classification. *Acta Neuropathol* 2024; 147(1): 74.

33

Quantitative Detection of Cancer Nucleic Acid Biomarkers with a Selective Solid-State Nanopore Assay

Adam R. Hall, Sara Abu Jalboush, Ian D. Wadsworth, Mohamed Amin Elaguech, and Komal Sethi

33.1 Introduction

As cancer initiates and progresses, impacted cells experience changes in the type, structure, and abundance of specific nucleic acids related to cell survival. Because these molecules can provide a window into disease state, their detection and quantification have extraordinary potential value to patient health. However, by their nature, some types of nucleic acids are difficult to probe, including chemical variants because of their subtle differences with canonical bases and microRNAs because of their small size. Additionally, for all such biomarkers in general, high sensitivity and specificity must be achieved even while reducing cost and measurement time. Consequently, there is a pressing need for new approaches for the quantitative analysis of nucleic acid cancer biomarkers.

A promising candidate technology for this purpose is the solid-state (SS-) nanopore [1]. SS-nanopore analysis functions using the concept of resistive pulse sensing (RPS) [2] wherein the passage of objects through a narrow aperture produces interruptions in pathway resistance (i.e., ionic current) that can be measured and interpreted to report on the nature of the object. This is the same mechanism upon which Coulter counter technology has been built successfully to quantify and characterize cells [3], but the scale of components is reduced by several orders of magnitude to yield one of a limited number of approaches capable of probing single biomolecules. An SS-nanopore (Figure 33.1a) consists of a single nanometer-scale opening formed in an insulating membrane (typically silicon nitride, though other materials can be used) tens of nanometers thick. This membrane, supported by a frame for easy handling, is mounted in a custom 3D-printed flow cell (Figure 33.1b) between two static reservoirs of electrolyte solution that are added manually by pipette. Electrodes are then placed into the reservoirs on either side of the SS-nanopore chip. In this arrangement the opening is the only connection point between the isolated reservoirs so that when a voltage difference is applied between the electrodes, an electric field is established that is highly convergent at the narrow pore. This has two major effects. First, a baseline ionic current is produced and recorded (see Methods), the magnitude of which is determined by both the composition of the electrolyte solution and the dimensions of the pore itself [4] as the only significant element of resistance. Second, the strong field gradient approaching the pore can transport charged molecules through the opening electrically and without forced fluid flow into the opposite chamber (Figure 33.1c). It is during this translocation that the threading molecule temporarily displaces ions contributing to the measured baseline current, resulting in a transient reduction that then returns to the baseline value after the passage is complete. These current interruptions are referred to as translocation "events" (Figure 33.1c, bottom) and their characteristics—like amplitude, duration, and rate—can be used to infer information on the molecules, including cross-sectional area [5], molecular weight [6], and concentration [7].

SS-nanopore detection has been applied successfully to a broad range of potential nucleic acid biomarkers [8–11] but nearly all such demonstrations have been carried out in controlled, synthetic specimens that do not approximate the complexity of clinical samples well. The major reason for this is the lack of intrinsic selectivity in the conventional measurement: because all translocating molecules can produce events, separating out the signals for any specific class of biomolecule from others in a mixture presents significant challenges. In response, we have developed [12] a selective SS-nanopore assay designed to produce events only when target molecules pass through the pore, enabling the specific evaluation of one biomarker among a background of interfering material. The assay employs two distinct molecular components: (1) a short, duplex nucleic acid containing in its structure a single biotin moiety and (2) a mutant variant of the small (~54 kDa), globular protein streptavidin featuring a single active biotin-binding pocket instead of the usual four [13] (monovalent streptavidin, or MS). When either of these components passes through the pore individually (Figure 33.1d, *i* and *ii*), their small size and rapid translocation speed makes the resulting transient current interruptions they produce impossible to resolve with conventional instrumentation. However, when the two molecules bind together via the extraordinarily strong biotin-streptavidin interaction, the resulting large, asymmetric nucleoprotein complex is hindered sterically by an SS-nanopore of appropriate size, passing more slowly than the constituents individually and yielding measurable events (Figure 33.1d, *iii*) that can be used to determine the amount of complex present quantitatively.

Of the experimental factors that impact the performance of this selective assay, pore diameter is the most critical. The smallest apparent nanopore diameter that is viable for the assay is ~7.5 nm, in good agreement with the minimum combined cross section of a duplex nucleic acid (~2.2 nm) and

DOI: 10.1201/9781003449942-37

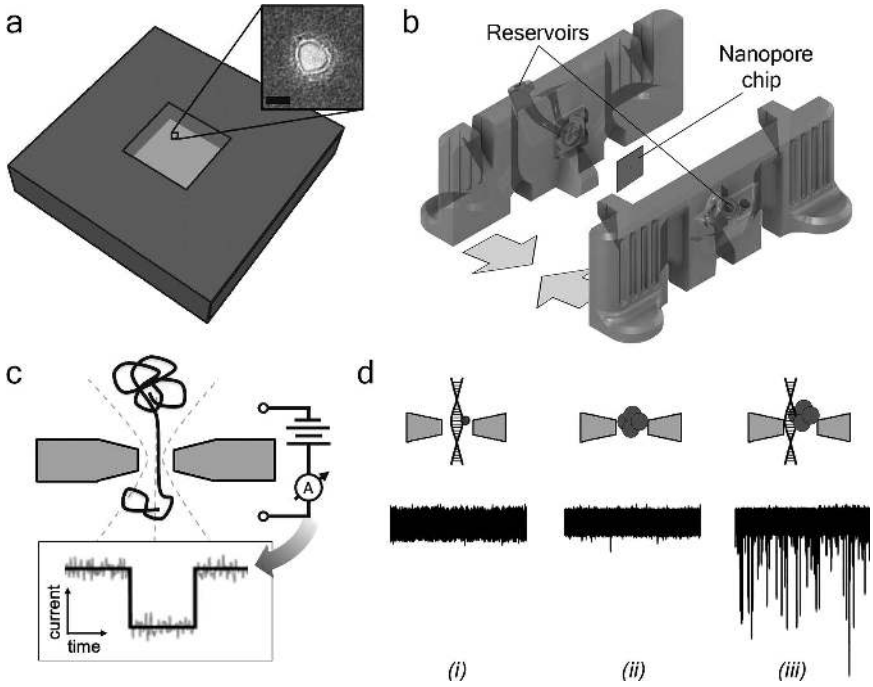


FIGURE 33.1 (a) Schematic representation of an SS-nanopore chip featuring a silicon nitride window. Inset: transmission electron micrograph of a typical SS-nanopore (scale bar = 10 nm). (b) Schematic of the custom 3D-printed flow cell used to house the SS-nanopore chip between the static electrolyte and sample reservoirs. (c) Schematic of conventional SS-nanopore detection showing electrical translocation of a molecule through a pore (top) and the resulting ionic current interruption ("event," bottom). (d) Schematic (above) and representative current traces (below) demonstrating the selective assay. Neither biotinylated nucleic acid fragments *(i)* nor MS *(ii)* produce significant events, in contrast to nucleoprotein complexes consisting of both elements *(iii)*.

the MS tetramer (~5.6 nm according to crystallography [14]). For any aperture smaller than this, there is no orientation of the complex that can pass through the pore without having to deform, making signal instability and eventual fouling of the device typical. As nanopore diameter is increased, steric hindrance is reduced and more nucleoprotein complexes are able to translocate rapidly without producing a detectable event. As a result, the measured event rate is observed to decrease exponentially with pore size [15]. However, under some conditions [16], the assay can deliver high selectivity (i.e., a nucleoprotein complex rate more than an order of magnitude greater than the combined rates of the constituents alone) with SS-nanopores up to 16 nm in diameter. A second important experimental variable affecting the assay is the length of the duplex nucleic acid construct. While we have demonstrated efficacy across a broad range of sizes [15] at least from 23 bp to 250 bp, there are two competing effects that must be considered. First, nucleoprotein complexes formed from larger nucleic acids are more resolvable, as evidenced by the monotonic increase observed in the measured event rate as a function of fragment size for a given concentration of complex [15]. For this reason, small nucleic acid constructs yield a reduced sensitivity. Second, while nucleoprotein complexes formed from large nucleic acid constructs produce a high event rate, the nucleic acid fragments themselves are also more easily resolved. As a result, selectivity is gradually reduced, counteracting assay effectiveness.

In this chapter, we will describe extensions to the selective SS-nanopore assay that have enabled its application to three specific types of cancer-relevant nucleic acid biomarkers: DNA base modifications and lesions, microRNAs, and conserved target sequence motifs within large nucleic acids.

33.2 Materials

33.2.1 Device Materials and Instrumentation

Solid-state nanopores were fabricated using an Orion PLUS helium ion microscope (Zeiss, Peabody, MA). Devices were cleaned with air plasma (Harrick Plasma, Ithaca, NY) prior to use. Nanopore measurements were performed using an Axopatch 200B patch-clamp amplifier (Molecular Devices LLC, San Jose, CA). Gels were imaged with an E-Gel Imager system (Invitrogen, Carlsbad, CA).

33.2.2 Biologicals

Synthetic DNA and RNA oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA) and resuspended in DNase- and RNase-free water (Invitrogen). MS was produced using either an established protocol [17] or one developed in-house [16]. M13mp18 single-strand DNA, T4 β-glucosyltransferase (βGT), *E. coli* uracil DNA glycosylase (UDG), human oxoguanine glycosylase (hOGG1), human alkyladenine DNA glycosylase (hAAG), EndoIV, bovine serum albumin, NEB2 buffer, and mung bean nuclease (MBN) were purchased from New England Biolabs, (Ipswich, MA).

Biotinylated dUTP, dGTP, dCTP, and dATP were acquired from Perkin Elmer (Waltham, MA). T4 polymerase (exo-) was purchased from Lucigen (Middleton, WI). Ten-eleven translocase (TET) was expressed in insect cells as described previously [18]. An APE1 mutant (D308A) with reduced 3′–5′ exonuclease activity [19] and a TDG mutant (N157D) lacking recognition for 5-carboxylcytosine (5caC) [20] were expressed in *E. coli* as reported previously [21]. HIV-1 BH10 non-infectious molecular clone (pBH10) [22] was obtained from the National Institutes of Health AIDS Reagent Program.

33.2.3 Chemicals

Agarose, TE buffer, HEPES buffer, $MgCl_2$, and dithiothreitol were purchased from Thermo Fisher Scientific (Waltham, MA). UDP-azide-glucose was purchased from Active Motif (Carlsbad, CA). Sulfo-dibenzocyclooctyne-biotin was acquired from Sigma-Aldrich (St. Louis, MO). GelRed nucleic acid stain was purchased from Phenix Research Products (Candler, NC). Phenol/chloroform/isoamyl alcohol (25:24:1) and chloroform were purchased from Acros Organics (Norris, NJ).

33.3 Methods

33.3.1 Device Fabrication

Each solid-state nanopore was fabricated in a 20 μm^2 silicon nitride (SiN) membrane (20 nm thick) supported by a 4 × 4 mm silicon frame (Norcada Inc., Edmonton, AB, Canada) using helium ion milling [23].

33.3.2 SS-Nanopore Measurements

Before measurement, a single nanopore was rinsed thoroughly with ethanol and DI water, dried, and exposed to a 30 W air plasma for 2 min on each side. The chip was then loaded into a custom 3D-printed flow cell and contacted on both sides with measurement buffer. The nanopore size was estimated from a current-voltage curve [8]. Then biomolecules were loaded by pipetting into one chamber (*cis*- side) and a positive bias was applied to the opposing chamber (*trans*- side) via Ag/AgCl electrodes placed in the reservoirs. Ionic currents were detected with Axopatch 200B patch-clamp amplifier (see Device Materials and Instrumentation), recorded, and later analyzed using custom software written in LabView (National Instruments, Austin, TX).

33.3.3 Biotin Labeling of Modified Bases in DNA

5-Hydroxymethylcytosine (5hmC) was biotin-labeled using a procedure developed previously [24] in which βGT attached UDP-azide-glucose and at the site and biotin was conjugated using copper-free click chemistry. Labeled DNA was purified using QIAquick purification kit (Qiagen, Germantown, MD) and redispersed in water. Other modified bases were labeled by excising the target base with a glycosylase (UDG for uracil, hOGG1 for oxoG, TDG for T:G mismatch, and hAAg for 1,N^6-Ethenoadenine) and then inserting the matching biotinylated base using a gap-filling polymerase (details described previously [25]). Selective labeling of cytosine demethylation bases was achieved through incorporation of TET to oxidize bases, βGT to block 5hmC, and a mutant TDG to select against 5caC [21].

33.3.4 Nucleic Acid Motif Extraction

Internal sequences of single-strand DNA (M13mp18) or RNA (pBH10) were annealed with a complementary biotinylated DNA oligo and then flanking single-strand regions were digested with MBN. The digestion reaction was quenched by phenol-chloroform extraction and cleaned up by nucleotide removal kit (DNA) or ethanol precipitation (RNA).

33.4 Results

33.4.1 Quantification of Single Base Modifications and Lesions in DNA

Chemical variants of DNA bases occur frequently in genomic DNA as both epigenetic modifications with crucial roles in biology [26] and lesions that must be repaired to ensure cellular function [27]. Alterations in the abundance of these modified bases have been associated with the development and progression of cancer as well as other diseases. However, their study by current analytical methods can be difficult due to artifacts, sample size requirements, high cost, and other challenges. The selective SS-nanopore assay can be used to assess base modifications quantitatively by implementing protocols that tag their locations with a biotin moiety, thereby enabling MS to bind only to fragments with the target base.

33.4.1.1 SS-Nanopore Quantification of 5-Hydroxymethylcytosine (5hmC)

The selective SS-nanopore assay was first used [28] to quantify global levels of 5hmC in murine brain genomic DNA (gDNA) by incorporating an established biotin labeling scheme [24]. Because of the assay length limitation (< 250 bp), gDNA was fragmented before labeling and subsequent MS-binding and analysis by SS-nanopore. Modified fragment concentration could then be determined from a concentration-dependent calibration curve measured with a synthetic monobiotinylated duplex DNA of the same length or interpolated from a series of measurements on such constructs of different length (Figure 33.2a). To transform this concentration into a figure of merit, the ratio of 5hmC:C, R, could be determined using the expression:

$$R = \frac{\left(c' \big/ c_o\right)}{2Yla}$$

where c' is the measured partial concentration of labeled 5hmC fragments, c_o is the total DNA concentration introduced, Y is the biotin labeling yield, l is the mean fragment

length, and α is the fractional cytosine content of the DNA. This general expression can be applied to any modified base assuming it is sufficiently rare such that fragments will have either 1 or 0 target bases. This expression yielded R values that matched LC-MS/MS analysis of the same samples quantitatively (Figure 33.2b), validating the quantitative capacity of our SS-nanopore assay for 5hmC. The same approach was subsequently used to investigate 5hmC in gDNA from human breast tissue and cell lines [15], finding that the reduction in the base reported in literature [29] is associated with invasive but not early-stage cancer cells.

33.4.1.2 SS-Nanopore Quantification of Other DNA Base Modifications

To target non-5hmC base modifications, a modular labeling approach was developed [25] in which a specific modified base is recognized and excised using a DNA glycosylase, an AP endonuclease is used to cleave the phosphodiester bond 5′ to the AP site and process the exposed 3′ phosphate, and a polymerase lacking 3′–5′ exonuclease activity is used to fill the gap with the cognate biotin-dNTP (Figure 33.2c). Labeling and SS-nanopore analysis was validated initially with four distinct glycosylases (UDG for uracil, hOGG1 for oxoG, TDG for T:G mismatch, and hAAg for 1,N^6-Ethenoadenine) and later adaptations were developed to distinguish DNA bases associated with the cytosine demethylation pathway [21], including 5-methylcytosine, 5hmC, 5-formylcytosine, and 5caC.

33.4.2 Quantification of Target MicroRNA Sequences

MicroRNAs are short (20–25 nt), single-strand, non-coding RNAs that serve diverse roles in regulating cellular protein expression and have emerged as potent biomarkers for cancer that serve diagnostic, staging, profiling, and prognostic purposes [30, 31]. The current gold standard for quantification of microRNAs from biological fluids is reverse transcription-quantitative polymerase chain reaction (RT-qPCR) [32]. However, despite its essential diagnostic importance, the approach has limitations that stem from the amplification process, including long processing times, potential biases, and complexity [33–35]. This has created the need for a versatile, amplification-free microRNA quantification approach.

The selective SS-nanopore assay could be applied to this challenge via the observation that nucleoprotein complexes featuring single-strand nucleic acid fragments do not produce significant events whereas those featuring duplex nucleic acids of the same length do [36]. This is a result of differences in stericity with the single-strand molecule being relatively compact in comparison and interacting less with the pore walls during translocation. Leveraging this difference, a target microRNA within a mixture could be assayed by interfacing it with a complementary biotinylated single-strand DNA probe along with MS to form a heteroduplex nucleoprotein complex; in the absence of target sequences, no detectable complexes would be formed (Figure 33.3a).

MicroRNA detection by this approach was first validated with a synthetic RNA oligonucleotide (miR155, associated with lung cancer) in clean measurement buffer wherein sensitivity down to 10 nM concentrations was demonstrated [36]. After altering experimental conditions to improve the sensitivity and selectivity of the assay [16], analyses of a second synthetic microRNA (cel-miR-54) were conducted, this time spiked into human plasma to provide a complex background. These measurements achieved detection down to 1 nM (Figure 33.3b), demonstrating the capacity of the assay to detect a sequence-specific microRNA biomarker in biofluid at clinically relevant concentrations.

33.4.3 Quantification of Sequence Motifs

While annealed microRNAs are naturally within a range of sizes amenable to direct analysis with the selective SS-nanopore assay, identification of target motifs within much larger constructs could also hold tremendous value for cancer diagnostics. Applications for this capability include mRNA analysis to report on oncogenes or tumor suppressor genes [37] and the detection of cancer-associated viral infections [38]. To address this need, a strategy to extract sequence

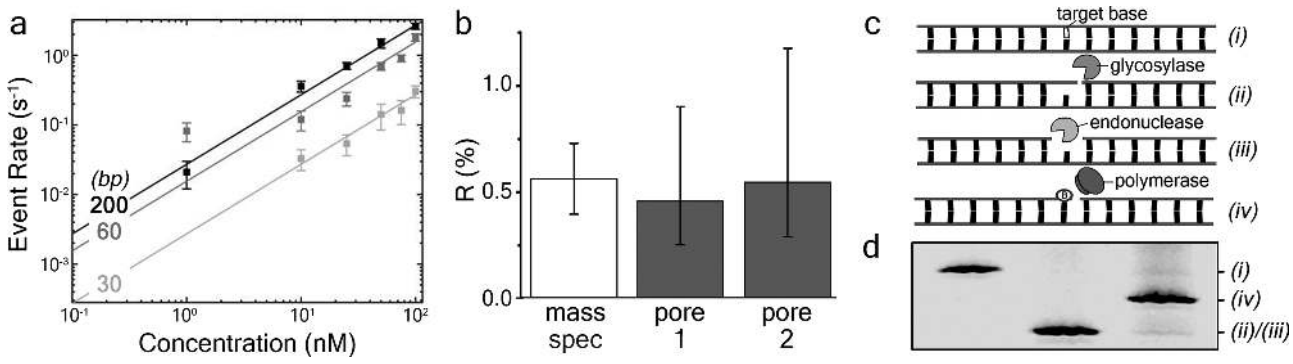


FIGURE 33.2 (a) Example length-dependent calibration curves (event rate vs. concentration) for synthetic DNA nucleoprotein complexes. (b) 5hmC/C ratio (R) for murine brain gDNA measured by conventional LC-MS/MS (white) and by two separate SS-nanopore analyses (gray). (c) Schematic method for modular biotin labeling of DNA base modifications: a target modification in a DNA *(i)* is excised by a specific glycosylase *(ii)* and prepped by an endonuclease *(iii)* before the missing base is replaced with a biotinylated nucleotide by a gap-filling polymerase *(iv)*. (d) Example denaturing gel analysis (following the modified strand) of an oligonucleotide with a single modified base labeled with the steps shown in (c).

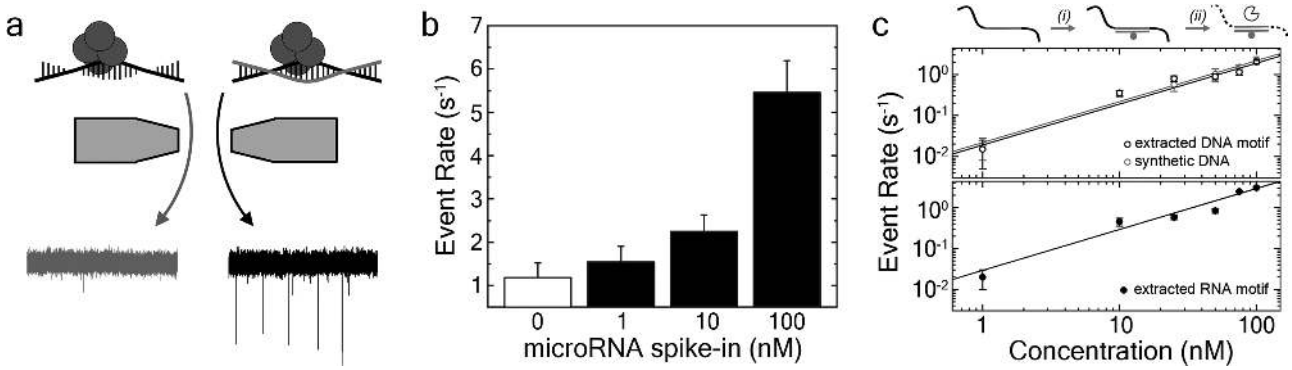


FIGURE 33.3 (a) MicroRNA detection scheme showing that single-strand nucleoprotein complexes do not produce events (left) whereas heteroduplex complexes do (right). (b) Example detection of microRNA (synthetic cel-miR-54) spiked into human plasma. Control (no spike-in, white) is the rate associated with off-target material in the plasma. (c) Top: scheme for motif extraction in which a target sequence in a large single-strand nucleic acid is annealed with a synthetic biotinylated probe *(i)* and flanking regions are digested *(ii)* to leave a duplex fragment. Bottom: event rate vs. concentration for fragments extracted from a long DNA (upper) and a long RNA (lower) template. Control (synthetic DNA construct) shown in gray in upper plot match extracted trends. Lines are linear fits to the data.

motifs from large single-strand nucleic acids has been developed [39] (Figure 33.3c, top). Here, a synthetic biotinylated DNA probe is designed with complementarity to a target internal sequence. Following annealing of the probe to the motif, flanking regions are digested with MBN, an endonuclease that catalyzes preferential digestion of single-strand DNA and RNA. This yields a biotinylated duplex amenable to selective SS-nanopore analysis with a size determined by the probe length. Multiple probes, each designed for a different motif in a single target nucleic acid, can also be used simultaneously [39], providing an "amplified" signal without physical amplification.

To validate the approach, a target 60 nt motif within a single-strand DNA genome (bacteriophage M13mp18; 7,249 nt) was initially probed. Gel analysis confirmed the efficient extraction of the (biotinylated) fragment with the nuclease approach and subsequent concentration-dependent SS-nanopore analysis of the construct after MS conjugation gave results that were indistinguishable from those achieved with a synthetic 60 bp duplex (Figure 33.3c, upper). Finally, to extend the approach to RNA targets, the process was repeated for a separate 60 nt motif in single-strand HIV-1B RNA (~10,000 nt). After minor protocol modifications to account for nuclease digestion efficiency of RNA vs. DNA, analysis of the extracted 60 bp biotinylated heteroduplex yielded a concentration-dependent event rate that matched well with DNA measurements (Figure 33.3c, lower).

33.5 Discussion

While SS-nanopores allow analysis at the molecular level, their indiscriminate detection creates a dilemma in the need to identify which events are associated with which molecules. This can necessitate complex or computationally demanding methods and create uncertainty when signal differences are subtle. The selective SS-nanopore assay enables target molecules among a background to be assessed by the platform with specificity. The fundamental approach relies on differences in the dynamics of two molecular components—a nucleic acid fragment and a globular protein—when they are translocated separately compared to when they are translocated as a stable complex. The result is a nearly binary detection mechanism that produces events only when both constituents are present.

One limiting factor for the assay is pore stability. SS-nanopores can be fabricated with exquisite precision but can grow in diameter over time under voltage due to electromigration and other stresses. Because of the critical relationship between assay performance and pore diameter [16], this can result in growing quantification error. Possible solutions to this challenge may include the incorporation of coatings [40] to limit pore deterioration or the application of alternative materials that are inherently more resistant to wear [41]. Alternatively, a quantitative understanding of the diameter dependence can in principle be used to account for changes across long measurement times.

Despite this limitation, the discriminatory power of the assay has been employed in several ways to target specific nucleic acid biomarkers associated with cancer and other diseases. First, epigenetic base modifications/lesions can be assessed by implementing specific biotin labeling techniques, enabling MS binding to and subsequent detection of DNA fragments that contain them. This has been used to quantify modifications like 5hmC in genomic DNA, providing an alternative to expensive mass spectrometry analyses, and specific targeting of up to seven independent base modifications has been demonstrated. While quantitative and simple to perform, a limitation of the approach is its inability to define precisely where in the sequence the modifications occur. Second, microRNAs can be assessed by leveraging the additional observation that the assay can differentiate nucleoprotein complexes composed of single- and double-strand nucleic acids with the latter producing events exclusively. By providing a synthetic biotinylated DNA probe with complementarity to a target microRNA sequence, duplex formation is limited by the concentration of that microRNA, allowing its quantity to be determined. The approach is rapid and simple compared to conventional amplification-based techniques but has the limitation of targeting

only a single sequence at a time. Finally, conserved sequence motifs within large DNAs or RNAs can be assayed by implementing a nuclease-based extraction technique, allowing quantitative assessment of mRNAs, viruses, and more. A limitation of this approach is that it is currently validated only for single-strand nucleic acids.

33.6 Future Trends

Each embodiment of the assay is poised for further development. For example, future iterations of the base modification/lesion assay may be improved by implementing targeted amplification or capture protocols, allowing the assay to be applied to the detection of cancer-related mutations at targeted positions in genomic DNA. Future developments of microRNA quantification will be focused on multiplexing and could include expansion to parallel SS-nanopores or the use of fluorescent tags to couple analysis with optical detection. This would enable the assay to be applied to the quantitative detection of panels of established microRNA cancer biomarkers [42, 43]. Finally, further development of sequence motif detection should also enable double-strand genomes to be probed, expanding capabilities to assess genetic mutations and other features relevant to cancer emergence and progression.

REFERENCES

1. Dekker C 2007 Solid-state nanopores *Nat. Nanotechnol.* 2 209–15
2. DeBlois R W and Bean C P 1970 Counting and sizing of submicron particles by the resistive pulse technique *Rev. Sci. Instrum.* 41 909–16
3. Graham M D 2003 The coulter principle: Foundation of an industry *J. Assoc. Lab. Autom.* 8 72–81
4. Smeets R M M, Keyser U F, Krapf D, Wu M Y, Dekker N H and Dekker C 2006 Salt dependence of ion transport and DNA translocation through solid-state nanopores *Nano Lett.* 6 89–95
5. Smeets R M M, Kowalczyk S W, Hall A R, Dekker N H and Dekker C 2009 Translocation of RecA-coated double-stranded DNA through solid-state nanopores *Nano Lett.* 9 3089–95
6. Rivas F, Zahid O K, Reesink H L, Peal B T, Nixon A J, DeAngelis P L, Skardal A, Rahbar E and Hall A R 2018 Label-free analysis of physiological hyaluronan size distribution with a solid-state nanopore sensor *Nat. Commun.* 9 1037
7. Charron M, Briggs K, King S, Waugh M and Tabard-Cossa V 2019 Precise DNA concentration measurements with nanopores by controlled counting *Anal. Chem.* 91 12228–37
8. Wanunu M, Dadosh T, Ray V, Jin J, McReynolds L and Drndic M 2010 Rapid electronic detection of probe-specific microRNAs using thin nanopore sensors *Nat. Nanotechnol.* 5 807–14
9. Wanunu M, Cohen-Karni D, Johnson R R, Fields L, Benner J, Peterman N, Zheng Y, Klein M L and Drndic M 2011 Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules *J. Am. Chem. Soc.* 133 486–92
10. Shim J, Humphreys G I, Venkatesan B M, Munz J M, Zou X, Sathe C, Schulten K, Kosari F, Nardulli A M, Vasmatzis G and Bashir R 2013 Detection and quantification of methylation in DNA using solid-state nanopores *Sci. Rep.* 3 1389
11. Rozevsky Y, Gilboa T, van Kooten X F, Kobelt D, Huttner D, Stein U and Meller A 2020 Quantification of mRNA expression using single-molecule nanopore sensing *ACS Nano* 14 13964–74
12. Carlsen A T, Zahid O K, Ruzicka J A, Taylor E W and Hall A R 2014 Selective detection and quantification of modified DNA with solid-state nanopores *Nano Lett.* 14 5488–92
13. Howarth M, Chinnapen D J-F, Gerrow K, Dorrestein P C, Grandy M R, Kelleher N L, El-Husseini A and Ting A Y 2006 A monovalent streptavidin with a single femtomolar biotin binding site *Nat. Methods.* 3 267–73
14. Hendrickson W A, Pähler A, Smith J L, Satow Y, Merritt E A and Phizackerley R P 1989 Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation. *Proc. Natl. Acad. Sci. U. S. A.* 86 2190–4
15. Zahid O K, Rivas F, Wang F, Sethi K, Reiss K, Bearden S and Hall A R 2021 Solid-state nanopore analysis of human genomic DNA shows unaltered global 5-hydroxymethylcytosine content associated with early-stage breast cancer *Nanomedicine Nanotechnol. Biol. Med.* 35 102407
16. Abu Jalboush S, Wadsworth I D, Sethi K, Rogers L C, Hollis T and Hall A R 2024 Improving the performance of selective solid-state nanopore sensing using a polyhistidine-tagged monovalent streptavidin *ACS Sens.* 9 1602–10
17. Fairhead M, Krndija D, Lowe E D and Howarth M 2014 Plug-and-play pairing via defined divalent streptavidins *J. Mol. Biol.* 426 199–214
18. Liu M Y, Torabifard H, Crawford D J, DeNizio J E, Cao X-J, Garcia B A, Cisneros G A and Kohli R M 2017 Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine *Nat. Chem. Biol.* 13 181–7
19. Masuda Y, Bennett R A O and Demple B 1998 Rapid dissociation of human apurinic endonuclease (Ape1) from incised DNA induced by magnesium *J. Biol. Chem.* 273 30360–5
20. Maiti A and Drohat A C 2011 Thymine DNA glycosylase can rapidly excise 5-Formylcytosine and 5-Carboxylcytosine *J. Biol. Chem.* 286 35334–8
21. Wang F, Zahid O K, Ghanty U, Kohli R M and Hall A R 2020 Modular affinity-labeling of the cytosine demethylation base elements in DNA *Sci. Rep.* 10 20253
22. Hahn B H, Shaw G M, Arya S K, Popovic M, Gallo R C and Wong-Staal F 1984 Molecular cloning and characterization of the HTLV-III virus associated with AIDS *Nature.* 312 166–9
23. Yang J, Ferranti D C, Stern L A, Sanford C A, Huang J, Ren Z, Qin L-C and Hall A R 2011 Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection *Nanotechnology.* 22 285310
24. Song C X, Szulwach K E, Fu Y, Dai Q, Yi C Q, Li X K, Li Y J, Chen C H, Zhang W, Jian X, Wang J, Zhang L, Looney T J, Zhang B C, Godley L A, Hicks L M, Lahn B T, Jin P and He C A 2011 Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine *Nat. Biotechnol.* 29 68–72

25. Wang F, Zahid O K, Swain B E, Parsonage D, Hollis T, Harvey S, Perrino F W, Kohli R M, Taylor E W and Hall A R 2017 Solid-state nanopore analysis of diverse DNA base modifications using a modular enzymatic labeling process *Nano Lett*. 17 7110–6
26. Liyanage V R B, Jarmasz J S, Murugeshan N, Del Bigio M R, Rastegar M and Davie J R 2014 DNA modifications: Function and applications in normal and disease states *Biology*. 3 670–723
27. Ames B N, Shigenaga M K and Gold L S 1993 DNA lesions, inducible DNA repair, and cell division: Three key factors in mutagenesis and carcinogenesis *Environ. Health Perspect*. 101 Suppl 5 35–44
28. Zahid O K, Zhao B S, He C and Hall A R 2016 Quantifying mammalian genomic DNA hydroxymethylcytosine content using solid-state nanopores *Sci. Rep*. 6 29565
29. Jin S-G, Jiang Y, Qiu R, Rauch T A, Wang Y, Schackert G, Krex D, Lu Q and Pfeifer G P 2011 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations *Cancer Res*. 71 7360–5
30. Hayes J, Peruzzi P P and Lawler S 2014 MicroRNAs in cancer: Biomarkers, functions and therapy *Trends Mol. Med*. 20 460–9
31. Quirico L and Orso F 2020 The power of microRNAs as diagnostic and prognostic biomarkers in liquid biopsies *Cancer Drug Resist*. 3 117–39
32. Thomou T, Mori M A, Dreyfuss J M, Konishi M, Sakaguchi M, Wolfrum C, Rao T N, Winnay J N, Garcia-Martin R, Grinspoon S K, Gorden P and Kahn C R 2017 Adipose-derived circulating miRNAs regulate gene expression in other tissues *Nature* 542 450–5
33. Kanagawa T 2003 Bias and artifacts in multitemplate polymerase chain reactions (PCR) *J. Biosci. Bioeng*. 96 317–23
34. Chen C, Tan R, Wong L, Fekete R and Halsey J 2011 Quantitation of microRNAs by real-time RT-qPCR *Methods Mol. Biol. Clifton NJ*. 687 113–34
35. Le M N and Nguyen T A 2023 Innovative microRNA quantification by qPCR *Mol. Ther. Nucleic Acids* 31 628–30
36. Zahid O K, Wang F, Ruzicka J A, Taylor E W and Hall A R 2016 Sequence-specific recognition of MicroRNAs and other short nucleic acids with solid-state nanopores *Nano Lett*. 16 2033–9
37. Kontomanolis E N, Koutras A, Syllaios A, Schizas D, Mastoraki A, Garmpis N, Diakosavvas M, Angelou K, Tsatsaris G, Pagkalos A, Ntounis T and Fasoulakis Z 2020 Role of oncogenes and tumor-suppressor genes in carcinogenesis: A review *ANTICANCER Res*. 40 6009–15
38. McLaughlin-Drubin M E and Munger K 2008 Viruses associated with human cancer *Biochim. Biophys. ACTA-Mol. BASIS Dis*. 1782 127–50
39. Sethi K, Dailey G P, Zahid O K, Taylor E W, Ruzicka J A and Hall A R 2021 Direct detection of conserved viral sequences and other nucleic acid motifs with solid-state nanopores *ACS Nano*. 15 8474–83
40. Yusko E C, Johnson J M, Majd S, Prangkio P, Rollings R C, Li J, Yang J and Mayer M 2011 Controlling protein translocation through nanopores with bio-inspired fluid walls *Nat. Nanotechnol*. 6 253–60
41. Larkin J, Henley R, Bell D C, Cohen-Karni T, Rosenstein J K and Wanunu M 2013 Slow DNA transport through nanopores in hafnium oxide membranes *ACS Nano*. 7 10121–8
42. Puerta-Gil P, García-Baquero R, Jia A Y, Ocaña S, Alvarez-Múgica M, Alvarez-Ossorio J L, Cordon-Cardo C, Cava F and Sánchez-Carbayo M 2012 miR-143, miR-222, and miR-452 are useful as tumor stratification and noninvasive diagnostic biomarkers for bladder cancer *Am. J. Pathol*. 180 1808–15
43. Mouliere F, El Messaoudi S, Pang D, Dritschilo A and Thierry A R 2014 Multi-marker analysis of circulating cell-free DNA toward personalized medicine for colorectal cancer *Mol. Oncol*. 8 927–41

34

Isolation and Characterization of Cell-Free RNA from Liquid Biopsy Taken from Cancer Patients

Real Sumayya Abdul Sattar, Luisa Quezada Ziesse, and Subhajyoti De

34.1 Introduction

Cancer is a major cause of mortality among adults in the USA. A major bottleneck in improving cancer care is the difficulty in timely detection of malignancies and tracking response to anti-tumor treatment [1]. The imaging-based tests routinely used for detection and monitoring of solid tumors are typically organ-selective and lack molecular details. Conventional tissue biopsies are the gold standard methods for histopathological assessment, but they are laborious, invasive, and limited to accessibility of the site of malignancy; furthermore, tissue biopsies may not be effective in capturing intra-tumor heterogeneity, tumor evolution during treatment, and metastasis [2]. Recently, noninvasive cancer diagnosis based on liquid biopsy is emerging as a potential option [3]. Tumor and other tissues shed intact cells, DNA, RNA, proteins, and metabolites in blood and other body fluids, which can be examined to harness information about the tumor type, drug targets, and tumor burden [4–7]. Liquid biopsies allow repeated sampling and real-time monitoring of tumor progression in response to treatment [4, 8, 9]. Recent studies have shown that liquid biopsies from not just peripheral blood, but also cerebrospinal fluid, urine, sputum, and saliva can be used for cancer diagnosis and monitoring [10].

Most of the currently approved liquid biopsy-based tests focus on profiling circulating tumor cells, tumor-derived DNA, or protein biomarkers from serum or plasma for cancer detection, while the full potential of cell-free RNA (cfRNA) for clinical management of cancer patients remains to be systematically assessed. Cell-free RNA is abundant in blood [11, 12] and primarily comes from transcribed gene regions, and thus has information about transcriptome and bioprocesses in the cell types of origin. While circulating microRNAs in blood have been examined as cancer biomarkers for decades [13, 14], recent studies highlight the potential of cell-free mRNAs as emerging biomarkers for cancer detection, predict the origin of tumor, and determine tumor subtype [11, 15, 16]. However, long cell-free RNAs such as mRNAs and long non-coding RNAs in circulation are prone to degradation, and careful handling is needed to isolate and characterize them [4, 5]. Furthermore, next-generation sequencing and data analysis of cfRNA are non-trivial. Here we present an integrated experimental and computational workflow to isolate and characterize cell-free RNA from serum taken from cancer patients, as summarized in Figure 34.1.

34.2 Materials

34.2.1 Sample Collection

The blood samples were obtained from deidentified samples under an Institutional Review Board (IRB) approved protocol and was collected in EDTA tubes as well as CPT tubes.

34.2.2 Isolation and Quantification Kit

The cfRNA isolation kit was purchased from Qiagen (55114). RNase-free DNase Set was purchased from Qiagen (79254). Qubit RNA High Sensitivity Kit was purchased from Qiagen (Q32852), and Agilent RNA 6000 Pico Kit was purchased from Agilent (5607–1514).

34.3 Methods

34.3.1 Study Design and Sample Collection

The blood samples were obtained from deidentified samples under an Institutional Review Board (IRB) approved protocol. 4–8 mL of whole blood was collected in EDTA tubes as well as CPT tubes and were centrifuged at 3000 rpm at 4 °C for 10 min. The separated plasma was transferred to a new tube and centrifuged at 12,000 rpm for 15 min at 4 °C to remove cellular contaminations. Hemolysis was checked by visual inspection of samples against white background, and utmost care was taken to exclude all the samples with hemolysis. The plasma was again transferred to a fresh tube and RNase inhibitors were added and stored at −80 °C until further process. The samples were processed as early as possible to avoid the degradation of cfRNA over time. Multiple cycles of freeze thaw were avoided for all the samples.

34.3.2 Cell-Free RNA Isolation and Quantification

The total cfRNA was isolated using the circulating miRNA protocol from QIAamp Circulating Nucleic Acid Kit (Qiagen, 55114) as manufacturer's recommendation with some optimization. An in-column digestion with DNase was given to the extracted material using RNase-free DNase Set (Qiagen, 79254). The cfRNA was eluted with AV buffer and stored at −80 °C until further analysis. The cfRNA was quantified by

DOI: 10.1201/9781003449942-38

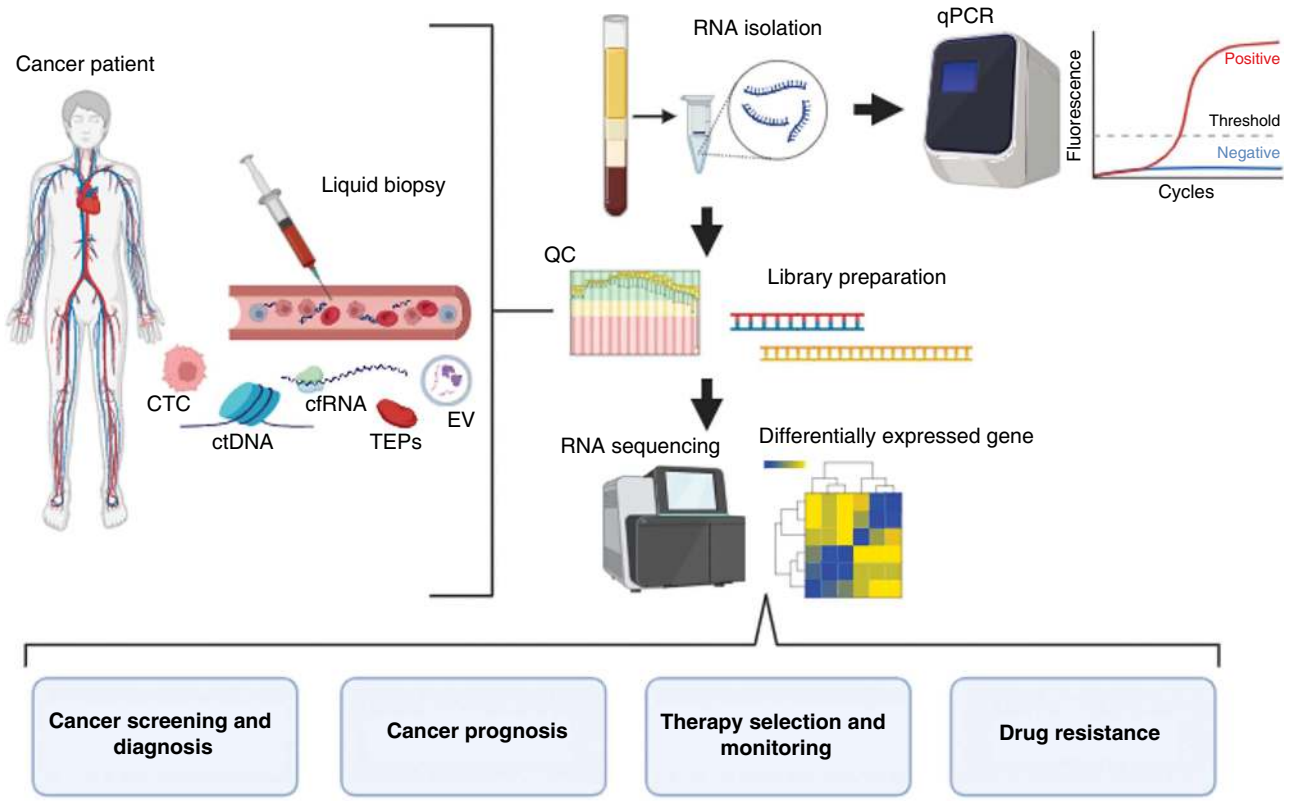


FIGURE 34.1 An overview of the overall workflow for the cell-free RNA profiling and analysis from cancer patients. The panel below indicates possible applications in the context of cancer diagnosis and monitoring treatment. CTC: circulating tumor cell, ctDNA: circulating tumor DNA, cfRNA: cell-free RNA, TEPs: tumor educated platelets, EV: extracellular vesicle, QC: quality control. (Created using BioRender.)

Qubit and bioanalyzer using Qubit RNA High Sensitivity kit (Qiagen, Q32852) and Agilent RNA 6000 Pico kit (Agilent, 5607–1514) respectively. The cfRNA were kept at ambient temperature throughout the quantification processes.

34.3.3 Target Region Selection and qRT-PCR

The primers against selected gene regions of interest were designed using NCBI primer BLAST and amplified using quantitative polymerase chain reaction (qPCR). The target region or transcript to be amplified by qPCR was selected on the following basis:

1. The primers from exon spanning junctions were designed using NCBI primer BLAST to avoid DNA contamination. The size of the amplicon was between 70 and 150 bp, the primer length was 18–25 nucleotide, and GC content of 40%–60%.
2. To avoid the false reporting of absence of the signal from the amplicon, the exons with reported exon skipping were excluded.
3. The signals for commonly used housekeeping genes were first checked using qPCR. The housekeeping genes are constitutively expressed with stable expression under normal physiological, diseased, or experimental conditions. As majority of the total RNA belongs to rRNA (> 80%–90%), we first checked for 18S rRNA amplification using specific primers. This was then followed by ACTB, GAPDH, B2M, and TBP, and also SMAD3, which is a cancer-associated gene.

The qPCR was carried out with Agilent PCR to check for detectable expressions of housekeeping genes and cancer specific genes from the extracted cfRNA using QuantiNova SYBR Green RT-PCR Kit (Qiagen, 208152). The primers for qPCR were designed using NCBI primer BLAST and purchased from IDT. One-step real-time RT-PCR was performed accordingly: 10 min of reverse transcription at 50 °C, followed by 2 min of PCR initial heat activation at 95 °C followed by 40 cycles of denaturation for 5 s at 95 °C and a combined annealing and extension for 10 s at 60 °C.

34.3.4 Cell-Free RNA Sequencing and Transcriptomic Analysis

CfRNA was sequenced using the Ovation® SoLo RNA-seq library kit with rRNA depletion, following the exosome lncRNA-seq workflow (Novogene), which enabled unbiased detection of most annotated coding and noncoding RNAs. The total RNA was subjected to DNase treatment followed by reverse transcription. The first strand cDNA was fragmented and then the second strand of cDNA was synthesized. After end-repairing, A-tailing, and adapter ligation and purification,

the first round of amplification was performed to remove ribosomal RNA by gene-specific InDA-C primers. Finally, the library was ready for sequencing after size selection, the second round of PCR amplification and purification. The library was checked with Qubit and qPCR for quantification and bioanalyzer for size distribution detection. Quantified libraries were pooled and sequenced on Illumina platforms, according to effective library concentration and data amount required. The QC of library was checked with Qubit and qPCR for quantification and bioanalyzer for size distribution detection. Quantified libraries were then pooled and sequenced on Illumina platforms, according to effective library concentration and data amount required. All samples prepared for sequencing were pathogen-free and had concentration of 5 ng, volume of 10 mL, fragments distribution of 25–4000 nt, FU > 10 with no peaks > 4000 nt.

Paired end reads from the fastq files were QCed and then mapped to the human reference transcriptome (hg38) using Salmon [17] with default setting and gene-by-gene TPM values were calculated. Alongside, we followed the STAR [18] RNA sequencing pipeline to map the reads to the transcribed regions of the human genome (GRCh38), and then use Integrative Genomic Viewer (IGV) [19, 20] to visualize this alignment, identify splicing patterns, and annotate oncogenic genomic variants in these regions.

In addition to our collected in-house samples, we downloaded additional cfRNA sequencing data for 180 cfRNA samples across five different cancer types from a publicly available study (GSE174302, Chen S, et al., 2022). This out-house cohort included samples from esophageal cancer (n = 29), lung cancer (n = 34), liver cancer (n = 26), stomach cancer (n = 37), and colorectal cancer (n = 54). Sequencing data for those samples was processed as noted previously.

A principal component analysis (PCA) [21–23] was performed using prcomp package combining in-house and out-house samples to assess the sources of biological and technical variations. A single-sample gene set enrichment analysis (ssGSEA) using a published approach [24, 25] was performed using the hallmark gene sets from MSigDB Collections to assess molecular signatures of cancer-related and other cellular processes in the cell-free transcriptome samples. All analyses were performed in R.

34.4 Results

34.4.1 Cell-Free RNA Isolation and Quantification from Cancer Patients

The primary goal of the study was to establish a standardized experimental workflow to obtain high-quality cell-free RNA from liquid biopsy samples from cancer patients in our laboratory settings and detect cancer-associated gene and pathway-level molecular signatures.

We obtained 4–8 ml of plasma from ten deidentified cancer patients under an IRB-approved protocol to establish the workflow for isolation and characterization of cfRNA. The cohort included four different cancer types, including six patients with kidney cancer and one each with liver, stomach, and penile tumor. Seven out of the ten samples had a detectable yield when quantified using bioanalyzer. Tumor stage and treatment status for the patients were not available. The isolated cfRNA had yield ranging from 28.6 pg/µl to 9.35 ng/µl per 4 to 8 ml of plasma/serum when analyzed by Qubit and varied across different patients. Both the CPT tubes and EDTA tubes gave comparably good yield of cfRNA. The summary of the sample, cancer types, yield, and QC has been summarized in Figure 34.2A.

Since cfRNA present in liquid biopsy samples is expected to be partially degraded, we assessed RNA integrity number (RIN) for our samples (Figure 34.2B–E). The extent of RNA integrity varied between the patient in our cohort (RIN: 1.3–2.9, Figure 34.2B–E). The cfRNA fragment size distribution ranged 25–4000 nt, FU > 10, and no peaks > 4000 nt. Peaks for 18S and 28S are prominently present indicating overall RNA integrity. All ten cfRNA samples were free from genomic contamination. Seven out of ten samples passed the initial QC for the aforementioned criteria, but upon sequencing all samples provided useful data regarding activities of cancer-associated genes and pathways, as discussed later.

34.4.2 Identification of Housekeeping and Cancer-Associated Gene Signals Using qPCR

Next, we assessed whether individual gene transcripts have sufficient RNA integrity and detectable abundance using qPCR. As a "rule of thumb," qPCR cycle threshold (Ct) values < 29 indicate high expression of the targeted transcript, while values of 30–37 and 38–40 indicate moderate and low expression of target transcripts, respectively [26].

Since a majority of the total RNA belongs to rRNA (> 80%–90%) and mRNA transcripts of housekeeping genes such as ACTB, GAPDH, B2M, and TBP are expressed by all tissue types and also present at relatively high abundance in both diseased and normal tissues, we checked for their expression by qPCR. The primer sequences for the target genes are summarized in Figure 34.3A. As expected, 18S rRNA had the lowest Ct value, indicating highest abundance, followed by housekeeping genes such as ACTB, GAPDH, B2M, and TBP, as expected. CfRNA from two patients had Ct values of 13.65 and 19.85 for 18S rRNA region. We assessed ACTB expression in four patients and observed a Ct value of 19.98, 26.29, 30.98, and 31.95 respectively. GAPDH, B2M, and TBP had Ct values of 27.99, 25.08, and 36.91 respectively (Figure 34.3A).

These observations suggested that the abundance and integrity of cfRNA in the blood in our cohort was reasonable for detection of gene-level signals. However, these initial observations also highlighted the limitation of the qPCR-based approach, that it is low throughput, is laborious, and may not be suitable for detection of low-abundance transcripts.

34.4.3 Cell-Free RNA Sequencing and Transcriptomic Analysis

Next we sequenced total cfRNA using the exosome lncRNA-seq protocol on the Illumina platform, which enabled detection of coding and noncoding RNAs from low-input samples such as ours. All samples passed library preparation and fastq

A

Sample codes	Cancer type	CfRNA conc (ng/µl)	RIN	Nucleotide	FU	cfRNA QC
cfR_1	Liver tumor	6.32	2.9	25-4000	>15	Pass
cfR_2	kidney tumor	2.3	2.3	25-4000	>10	Pass
cfR_3	Kidney tumor	9.35	2.3	25-4000	>10	Pass
cfR_4	Pyloric Antrum CA (Stomach)	5.93	1.3	25-4000	>25-200	Pass
cfR_5	Renal mass	4.26	1.5	25-4000	>10	Pass
cfR_6	Kidney tumor	7.55	2.3	25-4000	>10	Pass
cfR_7	Renal mass	Too low	Fail	-	-	Fail
cfR_8	Penile Mass	Too low	Fail	-	-	Fail
cfR_9	Kidney Mass	4.42	2.5	25-750	>10->200	Pass
cfR_10	Unidentified	Too low	Fail	-	-	Fail

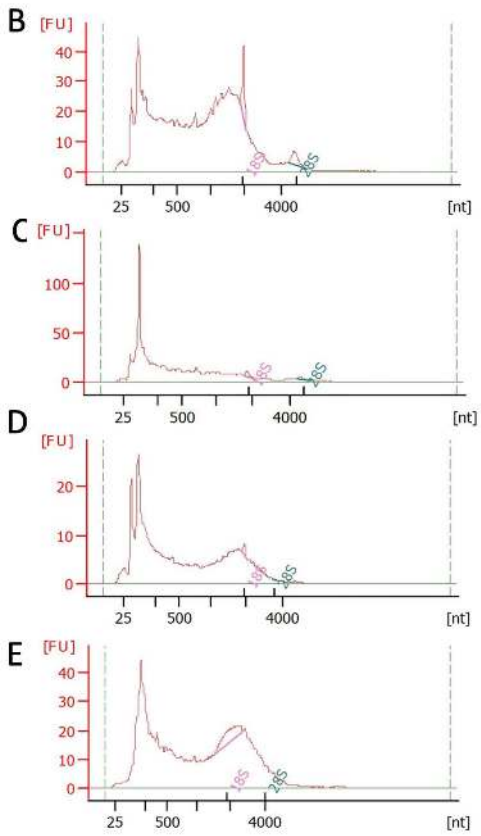


FIGURE 34.2 (A) Cancer type, cfRNA yield, and QC of the cfRNA samples in the in-house cohort. Nucleotide: cfRNA fragment size (nt), FU: Fluorescence Units, cfRNA QC: Cell-free RNA quality check. (B–E) Assessment of cell-free RNA size distribution and integrity using bioanalyzer after DNase digestion. Plots shown for four representative samples. Peaks for 18S and 28S are prominently present, indicating overall RNA integrity.

QC was acceptable, and transcript-level abundance was estimated using a published approach (see Methods).

We detected expression of 19,211 known protein coding genes in the samples, suggesting that cfRNA has representation from nearly the entire transcriptome. Furthermore, gene expression estimates (TPM) of the housekeeping genes such as ACTB, GAPDH, B2M, and TBP were consistent with the qPCR results (Figure 34.3B and C). Consistency of the results from two independent technologies, and their agreement with expected patterns of abundance of these genes, suggest that both cfRNA sequencing and targeted qPCR can provide a reasonable measure of transcript-level abundance in liquid biopsy samples, with the former providing a transcriptome-wide coverage.

We next sought to benchmark our cfRNA transcriptomic characteristics against that of other published studies and assess any concerns about batch effects. We obtained cfRNA sequencing data for 180 samples from five different cancer types from GSE174302 [27], and analyzed using the aforementioned pipeline. Our out-house multi-cancer cohort included 29 samples from esophageal cancer, 34 from lung cancer, 26 from liver cancer, 37 from stomach cancer, and 54 from colorectal cancer. The number of detectable transcripts and gene-level expression estimates were comparable, with high expression reported for the housekeeping genes.

A PCA jointly analyzing transcriptomes of the in-house and out-house samples indicated that there were no detectable batch effects between the in-house and out-house samples, and rather the samples segregated based on the cancer types (Figure 34.3D and E).

Furthermore, the cfRNA reads had adequate coverage of the exons and exon-exon junctions, at least for the highly expressed genes, sufficient to identify the transcript isoform structures. ACTB locus is shown as an example, where all exons had detectable cfRNA representation and alignment showed a primary transcript isoform, as expected (Figure 34.3F and G). This result indicated that our cell-free RNA extraction and processing protocol meets the standard practice, and that gene expression patterns detectable from the cell-free RNA could inform about the cell-free transcriptome, including gene expression signatures attributed to cancer types.

34.4.4 Identification of cfRNA Biomarkers across Different Cancer Types

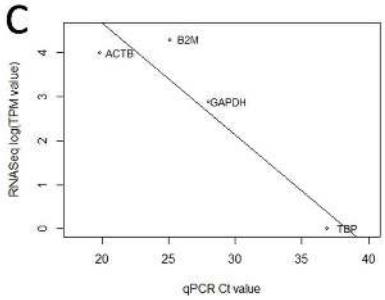
Consistent detection of housekeeping genes and no major batch effects in the meta-analysis motivated us to look for cancer-associated biomarkers. First, as an example, we considered examining SMAD3, which is a well-known cancer gene overexpressed in multiple cancer types, and had well-characterized

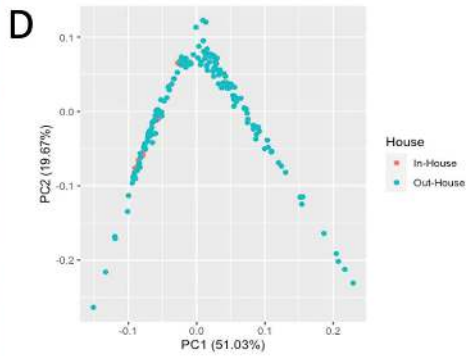
A

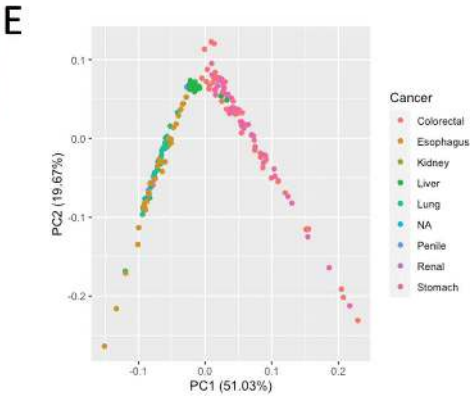
Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Exon	Product size (bp)	Ct values
TBP	ACA ACA GCC TGC CAC CTT A	GAA TAG GCT GTG GGG TCA GT	2 and 3	120	36.91
ACTB	GAG CAC AGA GCC TCG CCT TT	TCA TCA TCC ATG GTG AGC TGG	1 and 2	70	19.98, 26.29, 30.98, 31.95
18S rRNA	CTCAACACGGGAA ACCTCAC	CGCTCCACCAACTA AGAACG	Ribosomal subunit	110	13.65, 19.85
GAPDH	GGTCACCAGGGCT GCTTTTA	TTCCCGTTCTCAGCC TTGAC	2 and 3	147	27.99
B2M	TGT CTT TCA GCA AGG ACT GGT	ATG CTG CTT ACA TGT CTC GAT	2 and 3	147	25.08

B

Gene	Ct Values cfR_3	TPM cfR_3
TBP	36.91	0
ACTB	19.98	9672.90
GAPDH	27.99	744.78
B2M	25.03	19212.09







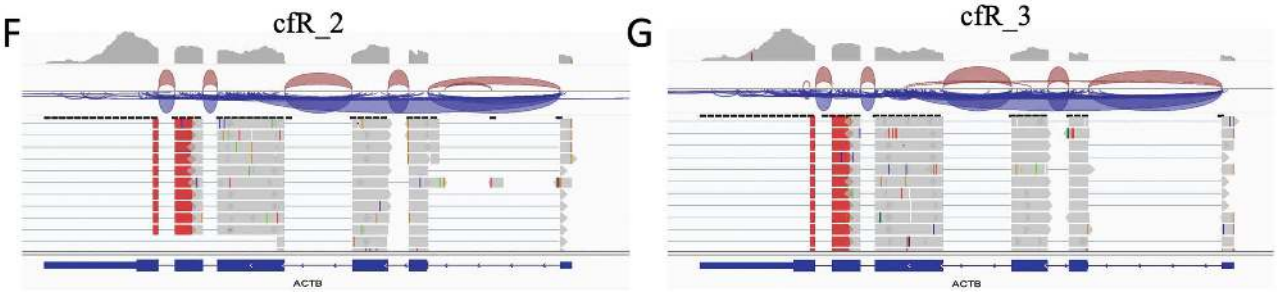


FIGURE 34.3 (A) Selected housekeeping genes, their qPCR target regions, and observed Ct values. (B) Table of Ct and TPM values for sample cfR_3. (C) Scatter plot showing inverse correspondence between Ct and TPM values in the cfRNA samples. (D–E) PCA plot showing patterns of gene expression variation among the cfRNA samples stratified by (D) cohort and (E) cancer types. (F–G) Transcript isoform structures of ACTB gene in the cfR_2 and cfR_3 samples.

in-house primers. We assessed SMAD3 expression in cfRNA in three patients and observed Ct values of 22.62, 23.56, and 36.92 respectively, indicating its moderately low abundance in those samples. qPCR was able to detect SMAD3 transcripts in the cfRNA samples which had a very low concentration of 0.22 ng in Qubit and failed the bioanalyzer QC. Based on cfRNAseq data we could detect SMAD3 expression in 10/10 samples in our cfRNA cohort and 143/180 of the out-house cohort.

We then performed ssGSEA for our in-house samples using the cancer hallmark gene sets from MSigDB Collections. This analysis identified molecular signatures of cancer-related and other cellular processes in the cell-free transcriptome in this cohort (Figure 34.4A). We found that the enrichment patterns for the hallmark pathways formed two major clusters—one cluster was prominently upregulated throughout all the samples in our cohort. This included the TGF beta signaling, apoptosis, and androgen response pathway among others. Those represented mostly tumor suppressor functions, DNA damage response activities, and inflammation. Some pathways such as KRAS signaling, epithelial mesenchymal transition, and others were substantially underrepresented in this cohort. When we examined clustering patterns among the cancer samples, there is no prominent grouping observed—which is not surprising since the samples were mostly from the genital-urinary system.

We then examined the alignment of sequencing reads at the base-pair resolution to detect somatic mutations, including oncogenic mutations in known cancer genes. As an example, we showcase KRAS gene region in Figure 34.4B. KRAS expression in cfRNA was relatively low, and the known

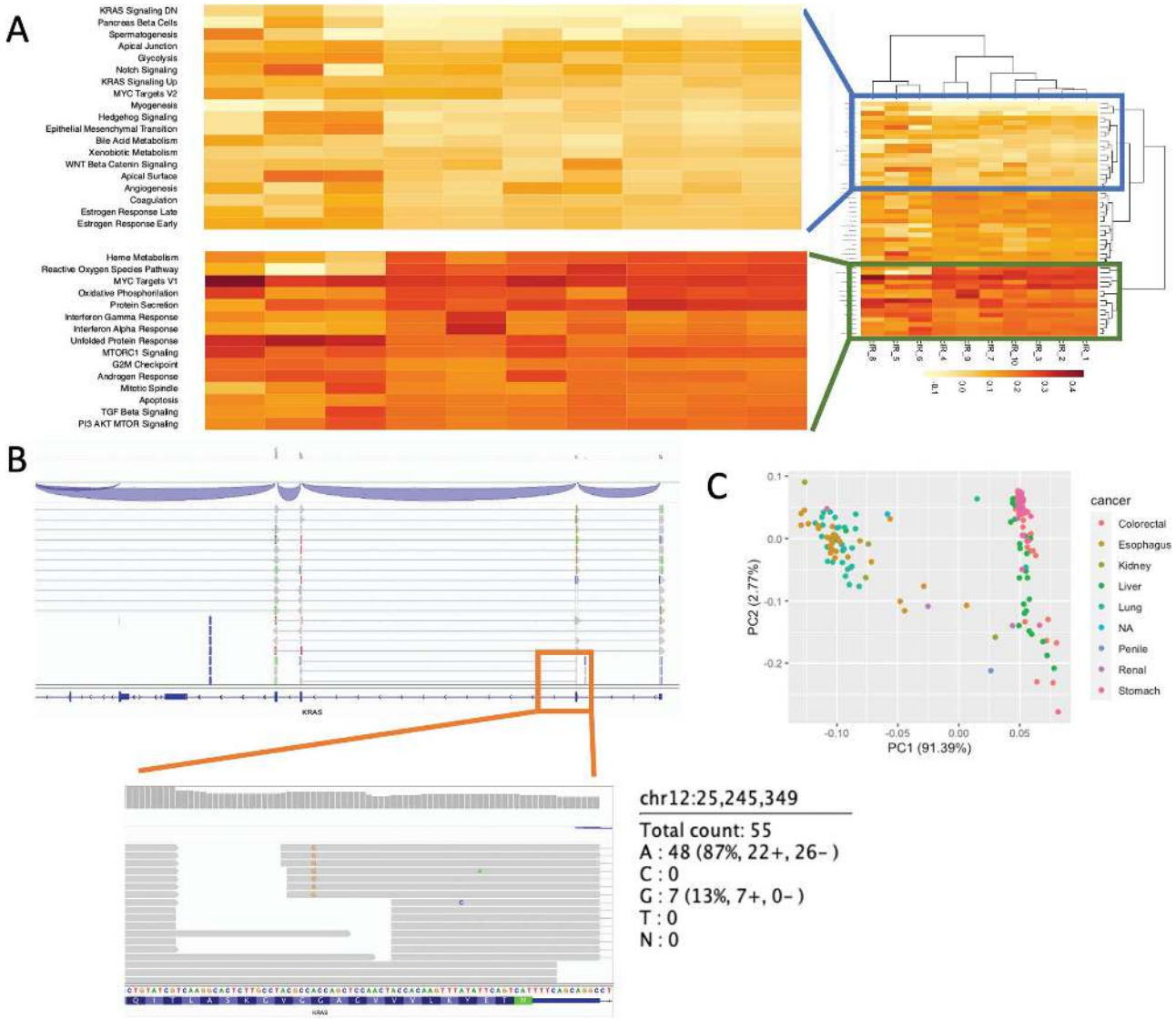


FIGURE 34.4 (A) Downregulated (top) and upregulated (bottom) hallmark pathways in the ssGSEA of our in-house samples. (B) IGV plot of KRAS gene of a colorectal cancer sample obtained from an out-house cohort shows somatic mutation at Codon 12. (C) PCA plot showing patterns of variation in hallmark ssGSEA signatures among the cfRNA samples stratified by cancer types.

primary transcript isoform structure was detected. While our in-house samples did not have any KRAS mutation, we highlight KRAS G12 somatic mutation (7/55 reads) in a colorectal cancer sample in the other cohort. The estimated allele frequency for the mutant allele in cfRNA was high, indicating its utility as a biomarker.

A PCA of in-house and out-house samples based on the ssGSEA scores stratified the samples based on cancer types into two groups (Figure 34.4c)—one dominated by gastrointestinal tumors and the other by lung and esophageal tumors. These results are similar to that observed at the gene-level in Figure 34.3, but signify emergence of greater differences between the two groups when activities of the cancer hallmark pathways are taken into consideration. Taken together, these observations suggest that cfRNA has both mutational and pathway-level signatures associated with specific cancer types.

34.5 Discussion

In this study we have established an optimized workflow for extraction, quantitation, and characterization of cfRNA obtained from blood of cancer patients in our laboratory settings. Our analyses indicate that isolated cfRNA had reasonable RNA integrity and purity for downstream processing and analyses, and that the batch effect is low. The qPCR and sequencing analyses provided estimates of transcript-level abundance for 18S rRNA and housekeeping genes, which are consistent with the expected results.

Unlike qPCR, which is targeted, laborious, and has high cost-per-basepair, cfRNA sequencing allowed high throughput detection of most high and moderately expressed transcripts from >19K known genes in the samples, with

good benchmarking and low cost-per-base pair estimates. Furthermore, genomic and pathway-level analyses of cfRNA sequencing enabled detection of transcript isoform structures and potential mutations therein, as well as pathway and overall transcriptome-level signatures that stratified the samples based on cancer types and/or associated organ systems. These results suggested that genomic analysis of cfRNA has potential to not only identify malignancies, but also provide information about cancer types and targetable mutations therein, similar to that reported elsewhere [12, 14, 28].

34.6 Future Trends

While the focus of this work has been to set up a workflow for isolation and initial characterization of cell-free RNA from liquid biopsy samples, future work will assess the utility of the biomarkers relevant for classifying tumor subtypes, evaluating treatment response, and other aspects. cfRNA sequencing has some limitations in not being informative for detecting chromosomal copy number alterations and non-coding mutations, which could be learnt from genomic analysis of cell-free DNA and could be used in clinical trials [6, 7, 10, 29].Nevertheless, both cfDNA and cfRNA harbor many emerging molecular signatures about cellular processes in tumor and non-malignant tissues and systemic responses [30]. While tumor-derived cfDNA is currently used for cancer diagnosis [7], joint analysis of cell-free DNA and RNA fractions have potentials to offer complementary information about genetic and non-genetic attributes of tumors and their ecosystems, which may be relevant for clinical management of cancer patients.

Acknowledgments and Funding

The authors acknowledge financial support from New Jersey Commission for Cancer Research, New Jersey Alliance for Clinical and Translational Sciences and R35GM149224 (SD). The authors thank the other members of the laboratories of SD, and also Rutgers Cancer Institute for helpful discussions.

REFERENCES

1. D. Crosby *et al*., "Early detection of cancer," *Science*, vol. 375, no. 6586, p. eaay9040, Mar. 2022, doi: 10.1126/science.aay9040.
2. L. W. Elmore *et al*., "Blueprint for cancer research: Critical gaps and opportunities," *CA. Cancer J. Clin*., vol. 71, no. 2, pp. 107–139, Mar. 2021, doi: 10.3322/caac.21652.
3. M. Ilié and P. Hofman, "Pros: Can tissue biopsy be replaced by liquid biopsy?" *Transl. Lung Cancer Res*., vol. 5, no. 4, pp. 420–423, Aug. 2016, doi: 10.21037/tlcr.2016.08.06.
4. L. Cabús, J. Lagarde, J. Curado, E. Lizano, and J. Pérez-Boza, "Current challenges and best practices for cell-free long RNA biomarker discovery," *Biomark. Res*., vol. 10, no. 1, p. 62, Aug. 2022, doi: 10.1186/s40364-022-00409-w.
5. S. Connal *et al*., "Liquid biopsies: The future of cancer early detection," *J. Transl. Med*., vol. 21, no. 1, p. 118, Feb. 2023, doi: 10.1186/s12967-023-03960-8.
6. K. Pantel and C. Alix-Panabières, "Liquid biopsy and minimal residual disease—latest advances and implications for cure," *Nat. Rev. Clin. Oncol*., vol. 16, no. 7, pp. 409–424, Jul. 2019, doi: 10.1038/s41571-019-0187-3.
7. J. C. M. Wan *et al*., "Liquid biopsies come of age: Towards implementation of circulating tumour DNA," *Nat. Rev. Cancer*, vol. 17, no. 4, pp. 223–238, Apr. 2017, doi: 10.1038/nrc.2017.7.
8. M. V. Martinez-Dominguez *et al*., "Current technologies for RNA-directed liquid diagnostics," *Cancers*, vol. 13, no. 20, p. 5060, Oct. 2021, doi: 10.3390/cancers13205060.
9. A. M. Miller *et al*., "Tracking tumour evolution in glioma through liquid biopsies of cerebrospinal fluid," *Nature*, vol. 565, no. 7741, pp. 654–658, Jan. 2019, doi: 10.1038/s41586-019-0882-3.
10. A. Tivey, M. Church, D. Rothwell, C. Dive, and N. Cook, "Circulating tumour DNA—looking beyond the blood," *Nat. Rev. Clin. Oncol*., vol. 19, no. 9, pp. 600–612, Sep. 2022, doi: 10.1038/s41571-022-00660-y.
11. M. H. Larson *et al*., "A comprehensive characterization of the cell-free transcriptome reveals tissue- and subtype-specific biomarkers for cancer detection," *Nat. Commun*., vol. 12, no. 1, p. 2357, Apr. 2021, doi: 10.1038/s41467-021-22444-1.
12. B. Roskams-Hieter *et al*., "Plasma cell-free RNA profiling distinguishes cancers from pre-malignant conditions in solid and hematologic malignancies," *NPJ Precis. Oncol*., vol. 6, no. 1, p. 28, Apr. 2022, doi: 10.1038/s41698-022-00270-y.
13. X. Chen *et al*., "Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases," *Cell Res*., vol. 18, no. 10, pp. 997–1006, Oct. 2008, doi: 10.1038/cr.2008.282.
14. J. Matsuzaki *et al*., "Prediction of tissue-of-origin of early stage cancers using serum miRNomes," *JNCI Cancer Spectr*., vol. 7, no. 1, p. pkac080, Jan. 2023, doi: 10.1093/jncics/pkac080.
15. L. J. Albrecht *et al*., "Circulating cell-free messenger RNA enables non-invasive pan-tumour monitoring of melanoma therapy independent of the mutational genotype," *Clin. Transl. Med*., vol. 12, no. 11, p. e1090, Nov. 2022, doi: 10.1002/ctm2.1090.
16. W. T. Kim *et al*., "UBE2C cell-free RNA in urine can discriminate between bladder cancer and hematuria," *Oncotarget*, vol. 7, no. 36, pp. 58193–58202, Sep. 2016, doi: 10.18632/oncotarget.11277.
17. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, and C. Kingsford, "Salmon provides fast and bias-aware quantification of transcript expression," *Nat. Methods*, vol. 14, no. 4, pp. 417–419, Apr. 2017, doi: 10.1038/nmeth.4197.
18. A. Dobin *et al*., "STAR: Ultrafast universal RNA-seq aligner," *Bioinformatics*, vol. 29, no. 1, pp. 15–21, Jan. 2013, doi: 10.1093/bioinformatics/bts635.
19. H. Li *et al*., "The sequence alignment/map format and SAMtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, Aug. 2009, doi: 10.1093/bioinformatics/btp352.
20. J. T. Robinson *et al*., "Integrative genomics viewer," *Nat. Biotechnol*., vol. 29, no. 1, pp. 24–26, Jan. 2011, doi: 10.1038/nbt.1754.
21. M. Horikoshi and Y. Tang, "ggfortify: Data visualization tools for statistical analysis results," 2018, [Online], 1, 28, http://CRAN.R-project.org/package=ggfortify.Rpackage version0.4.

22. Y. Tang, M. Horikoshi, and W. Li, "ggfortify: Unified interface to visualize statistical results of popular R packages," *R J.*, vol. 8, no. 2, p. 474, 2016, doi: 10.32614/RJ-2016-060.
23. H. Wickham, *ggplot2. in Use R!* Cham: Springer International Publishing, 2016, doi: 10.1007/978-3-319-24277-4.
24. H. Bengtsson, "matrixStats: Functions that apply to rows and columns of matrices (and to vectors)," 2017, [Online], https://github.com/HenrikBengtsson/matrixStats.
25. Z. Gu, L. Gu, R. Eils, M. Schlesner, and B. Brors, "*circlize* implements and enhances circular visualization in R," *Bioinformatics*, vol. 30, no. 19, pp. 2811–2812, Oct. 2014, doi: 10.1093/bioinformatics/btu393.
26. C. A. Heid, J. Stevens, K. J. Livak, and P. M. Williams, "Real time quantitative PCR," *Genome Res.*, vol. 6, no. 10, pp. 986–994, Oct. 1996, doi: 10.1101/gr.6.10.986.
27. S. Chen *et al.*, "Cancer type classification using plasma cell-free RNAs derived from human and microbes," *eLife*, vol. 11, p. e75181, Jul. 2022, doi: 10.7554/eLife.75181.
28. C. Ning *et al.*, "A comprehensive evaluation of full-spectrum cell-free RNAs highlights cell-free RNA fragments for early-stage hepatocellular carcinoma detection," *EBioMedicine*, vol. 93, p. 104645, Jul. 2023, doi: 10.1016/j.ebiom.2023.104645.
29. J. Li *et al.*, "Cell-free DNA copy number variations in plasma from colorectal cancer patients," *Mol. Oncol.*, vol. 11, no. 8, pp. 1099–1111, Aug. 2017, doi: 10.1002/1878-0261.12077.
30. S. De, "Signatures beyond oncogenic mutations in cell-free DNA sequencing for non-invasive, early detection of cancer," *Front. Genet.*, vol. 12, p. 759832, 2021, doi: 10.3389/fgene.2021.759832.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Section 4

Biomarker–Molecular Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

35

The Use of Dielectrophoresis to Recover Cancer-Derived Nanoparticles Straight from Undiluted Human Plasma for Cancer Detection Applications

Delaney Shea, Jason Ware, Michael I. Henderson, Ella Stimson, Mehrzad Sasanpour, Sarah Mitchell, Christian Ross, Shelby Nicholas, and Stuart D. Ibsen

35.1 Introduction

35.1.1 Tumor-Derived Nanoparticles as Cancer Biomarkers

Circulating cancer-derived nanoparticles are a rich source of biomarkers for cancer detection, prognosis, and monitoring due to their direct derivatization from solid tumors. These particles enter the circulation through several mechanisms including active secretion and passive release (Figure 35.1A). Active secretion includes the release of extracellular vesicles (EVs), such as exosomes, and is mediated by numerous cellular pathways including endosomal sorting complexes and ceramide-dependent pathways, which are currently under investigation [1]. EVs are excellent targets for blood-based cancer diagnostics as they serve as carriers for proteins, lipids, and nucleic acids. They are oversecreted by tumors relative to healthy tissue, possibly to modulate the tumor microenvironment to promote tumor growth, immunosuppression, and eventual metastasis. There is growing evidence that EVs support cellular crosstalk and migration in combination with cellular expression of matrix metalloproteinases to promote cell extravasation [2].

Passive release of nanoparticles can begin to occur as soon as the tumor grows bigger than 4 mm in diameter as cells in the poorly vascularized hypoxic regions are unable to receive sufficient oxygen and nutrients for survival [3] and begin to undergo necrosis [4]. The resulting cell lysis passively releases intracellular contents such as cell-free DNA (cf-DNA) and organelles into the extracellular space and into circulation where they may be broken down into smaller nanoparticles from shear forces (Figure 35.1A) [5].

Another circulating particle type of interest originates from bacteria. Bacteria can also colonize tumors, affecting not only the aggressiveness of the cancer but also how it responds to therapy [6]. The necrotic and hypoxic tumor regions are shielded from the body's immune system providing a haven for bacteria. Bacteria also secrete nanoparticles called bacterial membrane vesicles (BMVs), commonly called outer membrane vesicles (OMVs) when referring specifically to vesicles derived from gram-negative bacteria [7]. These ~100 nm BMVs hold potential as a biomarker for bacterial identification, as they contain proteins and nucleic acids specific to their parent strain (Figure 35.1A). Additionally, each single bacterium will produce many BMVs, and the small size of the BMVs allows them to escape the tumor into the bloodstream, from which they can be isolated and characterized. The rapid identification of intra-tumoral bacteria would add valuable information to a biomarker panel.

The main challenge in using these cancer biomarker-rich nanoparticles for blood-based cancer diagnostics is that they are difficult to recover from plasma due to their low buoyant density and small size [9]. Current methods for isolation include commercial kits and techniques like size exclusion chromatography and ultracentrifugation, but often require 1 mL of plasma or more to obtain enough material for thorough biomarker analysis. These methods place strain on the particles compromising the integrity of the EV membrane. These techniques are also generally expensive and can take multiple days to complete, often requiring further downstream analysis for proteomic characterization after isolation. Consequently, there is a critical unmet need for a platform technology that can efficiently isolate and characterize tumor-derived nanoparticles from plasma for use in clinical diagnostics.

One promising method for nanoparticle recovery is an electrokinetic technique called dielectrophoresis (DEP) that pulls the nanoparticles from plasma using alternating current (AC) electric fields (Figure 35.1B). Our work utilizing this technique has shown the ability of DEP to recover nanoparticles from plasma quickly and efficiently, which is the characteristic needed for clinical translation [10].

35.1.2 Dielectrophoresis Theory and Applications

DEP uses a nonuniform electric field to impart forces on nanoparticles suspended in a fluid. The magnitude of this force is determined by several factors, including particle size and differences in dielectric properties between the particle and surrounding media. The properties of EV nanoparticles make them ideal candidates for DEP-based collection. This

DOI: 10.1201/9781003449942-40

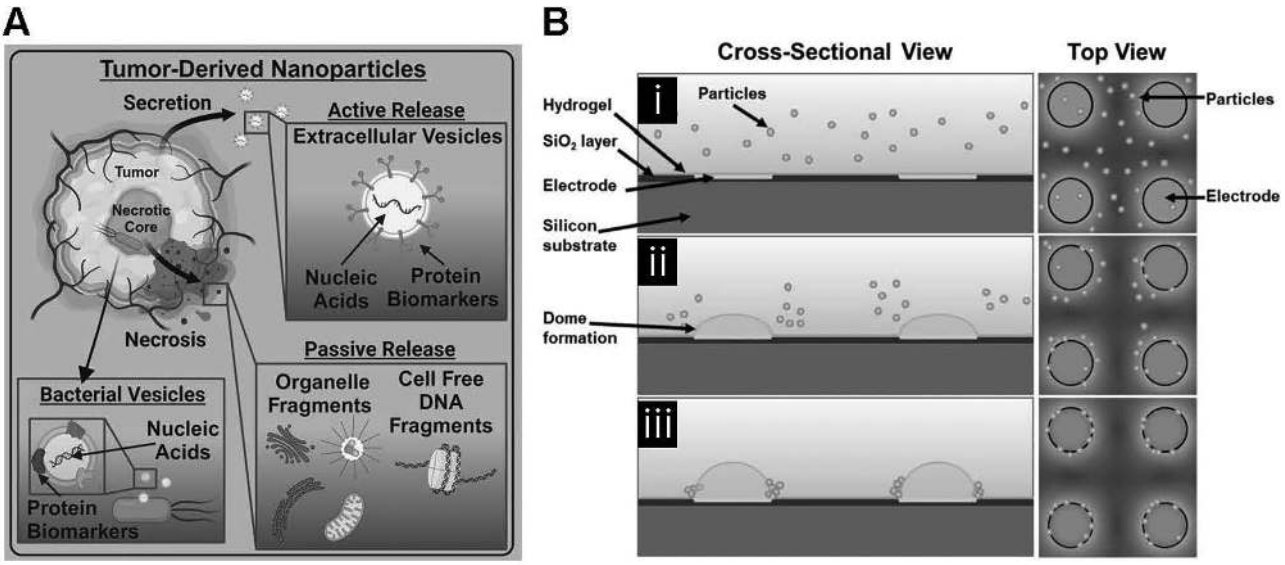


FIGURE 35.1 *Cancer-derived nanoparticle production and dielectrophoresis-based collection.* (A) Cancer-derived nanoparticles can enter circulation through both active secretion and passive release through cellular lysis events in the necrotic tumor region. Bacteria inhabit the necrotic tumor region and can release nanoparticles as well. (B) The dielectrophoresis-based collection uses an electrode array to create dielectrophoretic forces that recover desired nanoparticles straight from human plasma. (i) The sample is loaded into the microfluidic chamber. (ii) An AC electric signal is applied to the electrode array causing dome formation in the hydrogel layer which enhances the electric field gradient and drives particle collection. (iii) Particles are collected at the electrode edge. (Reproduced and adapted under the terms of the Creative Commons Attribution [8]. Copyright 2023, The Authors. Electrophoresis published by Wiley-VCH GmbH.)

includes exosomes which are lipid/protein vesicles ranging in diameter from 50 to 150 nm [11] that consist of cytoplasmic content, including nucleic acids and proteins, encapsulated by a phospholipid bilayer membrane which hosts transmembrane proteins from the cell. The lipid bilayer membrane is 2–3 nm thick and contains lipid hydrocarbon chains with a low dielectric constant. The hydrophilic polar lipid headgroups are more conductive and have a higher dielectric constant [12]. The dielectric properties of the lipid bilayer are a key factor impacting how the EVs respond to the nonuniform electric field used in DEP.

The dielectric properties of proteins can play a significant role in DEP isolation of EVs as well. Various transmembrane, lipid-anchored, peripheral membrane and tetraspanin proteins are highly enriched in exosomes [13], and their dielectric constant is a complex function of different factors including their structure, the local pH, temperature, and the dielectric permittivity of the medium [14]. Studies show that the dielectric constant inside the proteins is relatively low, due to the hydrophobic residues inside the protein, whereas the protein's surface has a higher dielectric constant because of the charged residues [15]. All these factors contribute to the resulting force that will be imparted on an EV particle in a particular electric field configuration and driving frequency.

35.1.3 The Dielectrophoresis Force

EVs are single compartment vesicles which can be approximated by a homogenous spherical dielectric particle. The DEP force experienced by such a particle in a spatially nonuniform electric field is given by the following equation:

$$F_{DEP} = \varepsilon_0 \varepsilon_m R^3 \mathbb{R}\left[CM^*\right] \nabla \left|E\right|^2 \tag{35.1}$$

where F_{DEP} is the DEP force, ε_0 is the electric permittivity of free space, ε_m is the electric permittivity of the media, R is the radius of the particle, $\mathbb{R}\left[CM^*\right]$ represents the real part of the complex Clausius–Mossotti factor, and $\nabla \left|E\right|^2$ is the electric field factor, which is equivalent to the gradient of the electric field squared. Positive DEP is a term used to describe particle motion towards a region of high electric field factor while negative DEP is used to describe movement of polarized particles away from regions of high electric field factor. For DEP cause motion of an EV particle, the DEP force must be greater than the sum of all disruptive factors such as Brownian motion, electroosmotic flow, and thermally driven convection.

DEP can collect particles within a limited size range dictated by the particle's radius and dielectric properties. The first particle variable dictating the DEP force is the radius where the force is proportional to the cube of the radius (Equation 35.1). Therefore, the electric field factor must increase as the particle radius decreases to generate a DEP force strong enough to overcome the disruptive force and isolate a particle. Electric field strengths required to manipulate extracellular vesicles (50–1000 nm) can be readily generated in microfluidic systems. The contribution of disruptive forces, such as the drag force, that reduce particle collection can be significant when the particle size is below 50 nm [16]. The collection of single molecular

proteins (~7 nm) requires higher electric field strength and often requires high voltages applied across pointed microelectrodes to maximize the electric field factor. This presents a challenge from electrolysis and bubble formation when collecting from a high-conductance medium like plasma.

The particle's dielectric properties are the second particle variable dictating DEP collection and are approximated using the Clausius–Mossotti factor. This frequency-dependent term describes the ability of a dielectric particle to be polarized in response to an electric field and how this compares to the polarizability of the surrounding media. This factor influences the magnitude of the DEP force as well as determines the direction of the DEP force on a particle dictating whether a particle will undergo positive or negative DEP when a certain frequency of AC electric field is applied. This allows particles with different dielectric properties to be separated from one another by localizing to high or low regions of the electric field. However, the Clausius–Mossotti factor becomes less predictive for submicron particles, and new theories are emerging to predict the dielectrophoretic force of these small particles [17].

The particle radius influences the electrical conductivity of the particle. By decreasing the particle size, the electrical conductivity of the particle increases [18]. This means that particles of the same material composition can be localized to different regions of the electric field separating them based on their size. The applied frequency can be chosen to pull certain particles towards the electrodes or push them away. Particle size and the Clausius–Mossotti factor allow DEP to select subsets of particles with particular size ranges from the overall particle population based on the applied frequency of the electric field. This enables analysis of different subpopulations of circulating nanoparticles and enables separation of smaller diagnostically relevant particles, such as EVs, from larger particles such as red blood cells and other cellular debris.

35.1.4 Electrode Array Design to Achieve DEP Particle Collection

As mentioned earlier, the magnitude of the DEP force depends strongly on the applied nonuniform electric field, which is determined by the electrode geometry [16]. The DEP force can be increased by increasing the gradient of the electric field and by increasing the applied voltage. Increasing the voltage too high will result in undesirable electrolysis at the electrodes causing bubble formation. To perfrom DEP without applying voltages that cause excessive electrolysis, the electrode array needs to be designed to produce a high electric field gradient. To increase the gradient of the electric field, the electrode dimension can be reduced, requiring the use of modern microelectrode fabrication techniques. These include photolithography, which has enabled the production of a new class of dielectrophoresis devices capable of generating electric fields strong enough to manipulate particles by applying only a few volts. These devices have enabled dielectrophoresis manipulation of particles in biological samples including cells, lipid vesicles, and proteins [19].

Electrodes can be designed and oriented relative to one another in such a manner as to generate a strong electric field gradient around the perimeter of the electrodes, which allows for uniform particle collection across an array of hundreds of electrodes. The devices used in our research consist of an array of symmetrical microelectrodes shown in Figure 35.2A and B, which are arranged in an alternating "checkerboard" pattern in such a manner that each adjacent electrode is connected to the opposite pole. When an alternating current is applied between the two sets of microelectrodes, an electric field gradient is generated. The resulting electric field factor has been simulated using COMSOL and is depicted in Figure 35.2C. The dark region shown around the perimeter of the electrode represents the region of the device with the strongest electric field factor. This high field factor region is where positive dielectrophoresis collection is expected to occur. Alternatively, the dark region in the upper right, which is between each set of four electrodes in the array, represents the low electric field factor region where negative dielectrophoresis is expected to occur. Particles moving under the influence of DEP follow the line patterns depicted in Figure 35.2C.

To protect the microelectrodes from degradation and reduce electrochemical reactions and electrolysis, a protective hydrogel is coated onto the surface of the device. This hydrogel layer has also been shown to restrict the surface area of the electrode in contact with the bulk solution. As the alternating current is applied, the hydrogel forms a dome shape across the surface of the electrodes as shown in Figure 35.1B, leaving only a thin ring around the electrode edge in contact with the sample and thereby increasing the local electric field factor around the edge of the electrode.

35.2 Materials

The microfluidic devices used for DEP isolation were purchased from Biological Dynamics (San Diego, CA, USA). These devices are comprised of a microfluidic chamber with a microelectrode array located at the bottom surface of the chamber. The AC signal used to generate the DEP force was produced by a Keysight 33210A 10 MHz Function/Arbitrary Waveform Generator and a Newtons4th Ltd (LPA01) power amplifier. A Zeiss Axio Imager D2 Microscope and Zeiss Zen (blue edition) software was used to obtain brightfield and fluorescence images of the DEP collected nanoparticles on-chip. The DEP chip microfluidics setup included microbore tubing for the input and output ports, disposable 23-gauge Luer-lock blunt needles for input reservoirs, and 1 mL plastic syringes used with a KD Scientific Syringe pump to control the flow of reagents and samples across the channels.

Dulbecco's phosphate buffered saline 1X (no calcium, no magnesium, Gibco™) and UltraPure™ distilled water (DNAse- and RNAse-Free, Invitrogen™) were purchased from Thermo Fisher Scientific (Massachusetts, USA) and used to make 0.5x PBS wash buffer. 2% milk blocking buffer was prepared using spray-dried skim milk (BD™ DIFCO™) purchased from Midland Scientific Inc. (Nebraska, USA) diluted in 0.5X PBS wash buffer. All primary and secondary antibodies and DNA stains were commercially purchased and diluted in 2% milk blocking buffer prior to loading on the DEP chip to reduce nonspecific binding of the immunostains.

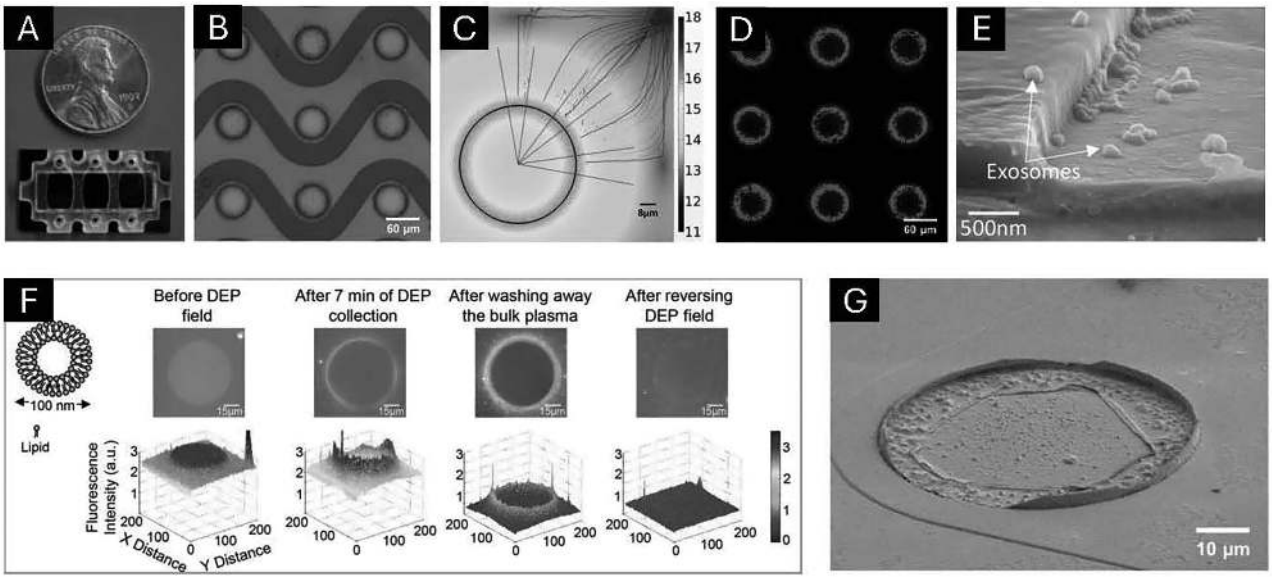


FIGURE 35.2 *Dielectrophoresis-based collection of nanoparticles from plasma.* (A) The microfluidic device contains three channels each with their own microelectrode array. (Reproduced and adapted under the terms of the Creative Commons Attribution [20]. Copyright 2022, The Authors. Electrophoresis published by Wiley-VCH GmbH.) (B) A brightfield microscopy image of the electrode array showing nine circular electrodes. The entire channel contains over 700 electrodes. (Reproduced and adapted under the terms of the Creative Commons Attribution 3.0 Unported License [21]. Copyright 2021, The Authors. Lab on a Chip published by The Royal Society of Chemistry.) (C) The electric field factor intensity around a single electrode shows that the highest strength is around the electrode edge. Particles travel along the trajectories shown by the dotted lines. The bar on the right side represents the base-10 exponents of the electric field gradient magnitude in V^2/m^3. (Reproduced and adapted under the terms of the Creative Commons Attribution [20]. Copyright 2022, The Authors. Electrophoresis published by Wiley-VCH GmbH.) (D) Fluorescence image of fluorescent polystyrene beads collected from high-conducting 0.5X PBS buffer showing collection around the electrode edge. (Reproduced and adapted under the terms of the Creative Commons Attribution 3.0 Unported License [21]. Copyright 2021, The Authors. Lab on a Chip published by The Royal Society of Chemistry.) (E) Scanning electron microscopy image of EVs collected from a buffer showing their original morphology preserved during the freeze-drying process for SEM imaging. (Adapted with permission from [11]. Copyright 2017 American Chemical Society. F. Sequence of events in the protocol for nanoparticle collection and purification from plasma. Reproduced and adapted under the terms of the Creative Commons Attribution [10]. Copyright 2015, Small published by Wiley-VCH GmbH.) (G) SEM image of an electrode after collection from plasma showing a ring of collected material around the electrode edge. (Reproduced and adapted under the terms of the Creative Commons Attribution 3.0 Unported License [21]. Copyright 2021, The Authors. Lab on a Chip published by The Royal Society of Chemistry.)

35.3 Methods

The microelectrode array chips for electric field gradient generation were 2 cm × 3 cm in size and contained over 700 platinum microelectrodes (60 µm diameter) (Figure 35.2A and B) [11]. The general protocol we have developed for collecting particles from plasma samples using DEP is as follows (Figure 35.2F). First, 0.5X PBS was added to the microfluidic chamber for 15 min to hydrate the hydrogel coating on the microelectrode array chip. Then, the chip was pretreated at a range of 6–10 V_{pp} at a frequency of 14 kHz using the function generator and power amplifier so that the hydrogel layer formed a closed dome structure over each electrode, which increased the collection of the particles [8]. The pretreatment process was conducted for 10 min. Then a 25 µL plasma sample was loaded into the microfluidic chamber at a rate of 7 µL per minute using the syringe pump. Once loaded, an AC signal of 10 V_{pp} at a frequency of 14 kHz was applied to the microelectrode array for 10 min to generate DEP forces on the biological nanoparticles in the sample. The positive DEP force draws the particles in the direction of increasing electric field gradient, which was greatest around the electrode edges (Figure 35.2C and D). With the DEP force still being applied, the bulk uncollected plasma was washed away with 0.5X PBS, thereby purifying the collected nanoparticles that remained around the electrodes. Scanning electron microscopy (SEM) images of the EV particles collected from buffer are shown in Figure 35.2E. Figure 35.2G shows an SEM image of all the material collected from plasma around the electrode edge.

The biomarkers present on the collected EVs were then quantified through fluorescence immunostaining (Figure 35.3). The quantification process started by loading a blocking buffer containing 2% fat-free milk in 0.5X PBS into the chamber to help reduce nonspecific binding of antibodies. The primary antibodies, which were diluted in blocking buffer, were then incubated with the sample at room temperature. The sample was washed with 0.5X PBS and then incubated with fluorescent secondary antibodies. Finally, the sample was washed with 0.5X PBS to remove unbound secondary antibodies before taking final fluorescence images.

To quantify our fluorescent images, we have developed a custom analysis program that automatically identifies a region of interest around each electrode as the area for quantification of our target nanoparticles while defining the area outside this region as the background [21]. The fluorescent background noise is subtracted from the collected fluorescence around the

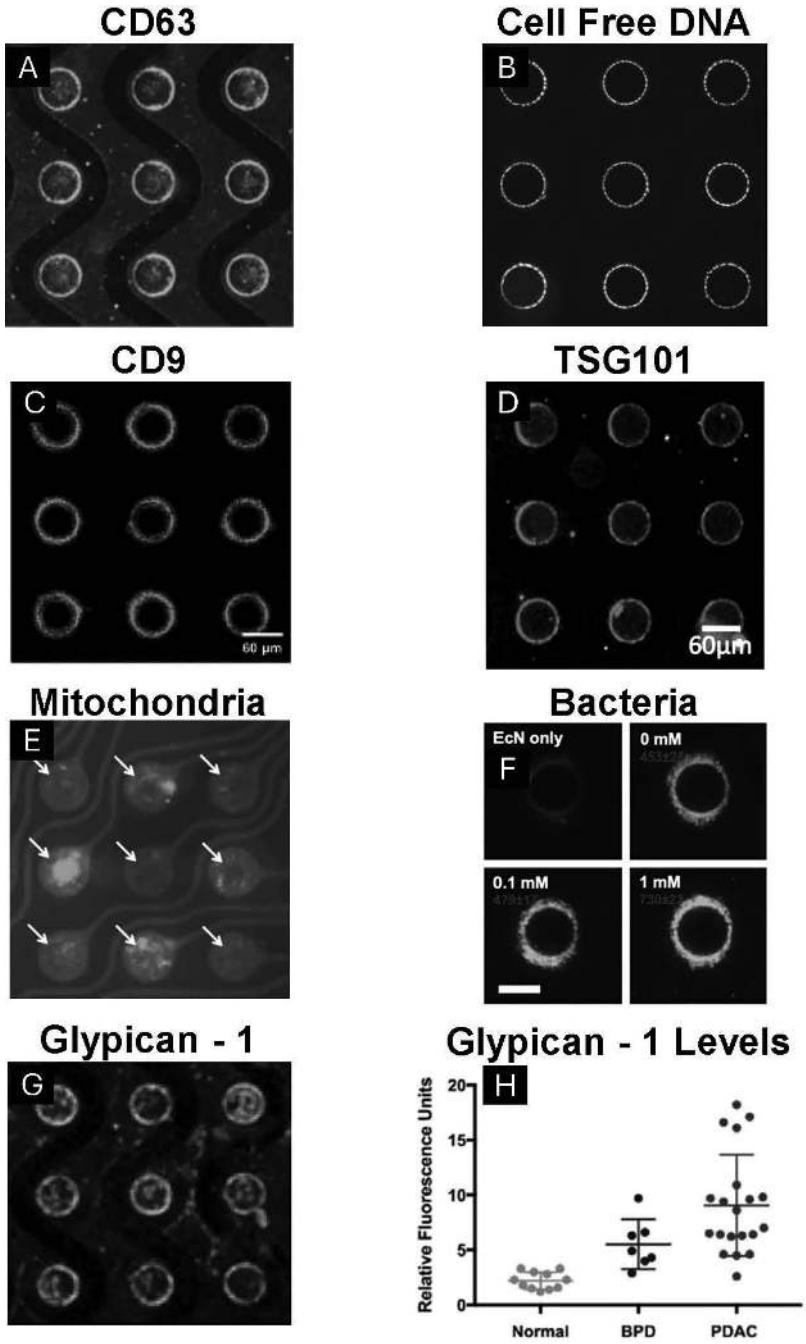


FIGURE 35.3 *The collection of different nanoparticle related biomarkers on the DEP chip.* (A) Collection of the exosome biomarker CD63 from breast cancer patient plasma. (Adapted with permission from [11]. Copyright 2017 American Chemical Society.) (B) Collection of cell-free DNA from serum. (Adapted with permission from [22]. Copyright 2013 Electrophoresis published by Wiley-VCH GmbH.) (C) Collection of the EV biomarker CD9. (Reproduced and adapted under the terms of the Creative Commons Attribution 3.0 Unported License [21]. Copyright 2021, The Authors. Lab on a Chip published by The Royal Society of Chemistry.) (D) Collection of the internal exosome biomarker TSG101 which required permeabilization of the exosome outer membrane using saponin treatment. (Adapted with permission from [11]. Copyright 2017 American Chemical Society.) (E) Collection of mitochondria using DEP. Here the collection occurred at the center of the electrodes shown by the arrows. (Adapted with permission from [22]. Copyright 2013 Electrophoresis published by Wiley-VCH GmbH.) (F) Collection of whole bacteria from synthetic saliva solution. The bacteria were genetically designed to produce a fluorescent green signal when exposed to lactate, an oral cancer related biomarker. (Reproduced and adapted under the terms of the Creative Commons Attribution [23]. Copyright 2022, Electrophoresis published by Wiley-VCH GmbH.) (G) Collection of cancer-related EVs carrying the biomarker Glypican-1 from pancreatic cancer patient plasma. (Adapted with permission from [24]. Copyright 2018 American Chemical Society.) (H) When applied to a cohort of patients, the EV-associated Glypican-1 levels showed a significantly higher level of expression in pancreatic cancer patients (PDAC) compared to healthy controls. Patients with benign pancreatic disease (BPD) did not show a significant difference with PDAC. (Adapted with permission from [24]. Copyright 2018 American Chemical Society.)

electrode allowing precise quantification of the collected signal. The program automatically identifies debris and artifacts in the image and removes those electrodes from the analysis.

Variations in collection efficiency can occur from sample to sample due to differences in plasma conductivity between different patients. The use of an internal standard, spiked into the plasma at a known concentration, can reduce variability between readings by enabling normalization of the biomarker signal between different samples [21]. Commercially available fluorescent polystyrene beads that are 100 nm in diameter have been shown to work for this application [21].

35.4 Results

35.4.1 Collection of Synthetic Nanoparticles from Plasma

We first demonstrated DEP collection using synthetic nanoparticles designed for drug delivery applications where we were able to control their surface and structural characteristics including size, density, and surface stiffness, which influenced their dielectric properties (Figure 35.2F).

Synthetic liposomes were one of the nanoparticle types that had the highest collection efficiency [23]. Both liposomes that were filled with buffer and liposomes that were filled with a gel material showed strong collection. These could be recovered from human plasma samples forming rings of collection around the electrode edges quantifiable by fluorescence imaging following washes to remove the bulk plasma (Figure 35.2F). This application of DEP to high conductance body fluids, such as plasma, makes this approach an excellent candidate for clinical diagnostics translation by reducing or eliminating sample preparation steps. The DEP force could be reversed to push the collected push particles off the electrode surface for further off-chip analysis [23].

The collection of the liposomes showed for the first time that lipid-based nanoparticles could be collected by DEP from human plasma [23]. These synthetic lipid-based nanoparticles lack the transmembrane proteins of EVs but have similar diameters and like EVs are largely comprised of lipids, indicating that EVs would also be collected by DEP from plasma.

35.4.2 DEP-Based Collection of EVs from High Conductance Buffer

The success that DEP had in collecting drug delivery liposome nanoparticles from plasma made it promising as a platform technology to enable rapid, label-free isolation of a range of lipid-based biologically derived nanoparticles from complex body fluids such as serum, plasma, and even whole blood. The ability to quickly recover EVs from plasma would be an important advancement for the field of liquid biopsy where blood-based biomarkers can be key to early disease detection and eliminate the need for invasive diagnostic tissue biopsies.

Our work has shown that this microarray chip device creates dielectrophoretic forces that successfully recover biologically derived nanoparticles containing a wide range of different biomarkers and characteristics from high conductance buffer and plasma (Figure 35.3). The presence of exosomes within the collected EV particles was confirmed by staining for the surface marker CD63, the internal marker TSG101, and observing a lack of staining for calnexin [11, 25].

The DEP collection process did not damage the EV membranes as observed through scanning electron microscopy imaging (Figure 35.2E) and by our observations that TSG101 on the inside of the exosomes could only be immunostained after permeabilization of the exosome outer membrane using saponin treatment (Figure 35.3D) [11]. The EVs are uniquely stabilized on the collecting electrode surfaces, allowing them to be permeabilized without them falling completely apart and diffusing away, which would make immunostaining much more difficult. A lack of a stabilizing surface makes labeling of internal protein content a challenge for techniques, such as flow cytometry, that maintain the EVs in free-floating form.

We also demonstrated that these DEP-collected EVs had RNA content that could be analyzed, showing tumor-related mRNA sequences for the EGFRvIII mutation known to be present in the glioblastoma cell line used as the source for the EVs [11]. mRNA analysis was conducted off-chip using RT-PCR procedures after recovering the collected EVs from the electrode surface using low frequency electrical pulses applied to the electrode array.

35.4.3 DEP-Based Collection of EVs and Cancer-Derived Nanoparticles from Undiluted Plasma

In a manner similar to how DEP was able to recover EVs from high-conducting buffer, we have demonstrated it is also capable of EV recovery from undiluted human plasma. This was first shown by spiking fluorescently labeled EVs collected from a glioblastoma cell line into human plasma and tracking their collection through the entire process [11].

Endogenous EVs were also collected by DEP from cancer patient plasma and were similarly concentrated around the electrode edge, which was particularly important because the locally high concentrations increased the signal-to-noise ratio. This also put the particles in known locations that our custom quantification software was able to use for precise biomarker detection and quantification [21]. The measurement of Glypican-1 content in the collected EVs demonstrated a significant difference in biomarker content between pancreatic cancer patients and healthy controls (Figure 35.3G and H) [24]. This shows the potential of DEP-based technology for clinical cancer diagnostic applications.

In addition to EV biomarkers, we have demonstrated that DEP also collects nanoparticles that carry circulating cell-free DNA fragments [26] and protein aggregates [27]. The cf-DNA isolated from blood-based samples provides important information for cancer detection applications due to the presence of certain cancer-related DNA mutations. In the case of chronic lymphocytic leukemia (CLL), DNA point mutations are traditionally detected by isolating peripheral blood mononuclear cells (PBMCs) from blood through centrifugation, then performing PCR amplification and sequencing directly on the cells. In a study of CLL patients, results from mutation information and analysis gained by PBMC isolation was

replicated using the cell-free DNA isolated by DEP from the same patient plasma. These results showed that cancer-derived DNA is indeed carried by nanoparticles in circulation [28]. Mutations for SF3B1, NOTCH1, and TP53 were identified in the 12 patients using traditional Qiagen kit based DNA isolation followed by Sanger sequencing. DEP recovered sufficient amounts of DNA from five of the 12 patients to identify the same mutations for each patient, with on-chip cf-DNA concentrations ranging between 37.3 and 472 ng/mL. We have demonstrated that the DEP collected cf-DNA can be PCR amplified directly on the chip itself, through the use of a custom designed thermal cycling system, which we used to detect KRAS mutations in pancreatic cancer patients [26].

DEP is further demonstrated as a unique platform for diagnostic applications as it can isolate nanoparticles not just from plasma, but from whole blood. The analysis of whole blood reduces labor and processing time by avoiding the centrifugation steps necessary to produce plasma or serum. The major challenge in working directly with whole blood is the high concentration of red blood cells that can obscure optical analysis of biomarker content. The rupture of those cells needs to be avoided to prevent release of their hemoglobin content, which can interfere with biomarker quantification. As described in Section 35.1.3, the size of red blood cells causes them to respond differently to the DEP force compared to nanoparticles. Red blood cells undergo negative DEP that pushes them to be loosely held in the low field region in between electrodes while at the same time the nanoparticles are undergoing positive DEP and being collected around the electrode edges. Washing removes the loosely held red blood cells and the positive DEP forces hold the nanoparticles in place. From these complex whole blood samples, the chip platform has demonstrated the isolation of cf-DNA from CLL patient whole blood or even T7 bacteriophage (55 nm in diameter) spiked into whole blood samples [22].

Larger organelles like mitochondria (500–600 nm in diameter) isolated from Jurkat cells can be recovered by DEP from high conductance buffer and analyzed by fluorescent labeling (Figure 35.3E) [22]. Mitochondria are multi-shelled particles, which are more complex particles to recover due to the electric field interacting with the surface membrane and the internal membranes as well. The DEP system is a versatile platform to collect and analyze a wide range of different particles from whole blood applicable to clinical diagnostics.

Traditional EV isolation by ultracentrifugation from 1 mL of plasma can take up to 14 hours for isolation to recover detectable EV-based biomarker signals [29], while DEP takes only 10 min to recover and 6 min to wash in order to isolate EVs from 20 μL of sample. This process is highly automatable and can be run in parallel to increase throughput, both necessary characteristics needed for clinical translation of EV-based diagnostics.

35.4.4 DEP-Based Cancer Detection Strategies Using Reporter Systems

In addition to the submicron-sized biological particles described previously, the DEP technique can also be used to collect bacteria in the high field region around the electrodes. Cancer detection is not limited to biomarkers directly derived from tumor cells but can also be expanded to detect particles or organisms that change their characteristics in the presence of a tumor. Such is the case for detecting oral carcinoma (OC), which has a five-year survival rate of 40% due to the current standard visual inspection diagnostic method followed by tissue biopsy and histopathological evaluation of lesions [23]. The oral cavity allows for the use of novel diagnostic methods that can be administered as a mouth wash and then spit out for analysis, such as in the application of tumor reporting bacteria.

Developments in synthetic biology have enabled modification of commensal bacteria, such as *Escherichia coli* strain 1917 Nissle (*E. coli*), already FDA approved for therapeutic applications, to produce GFP fluorescence in the presence of lactate, which is constitutively released by OC lesions. The lactate is diluted into saliva, and detecting these low concentrations in bulk saliva samples requires specialized sensitive equipment for techniques such as mass spectrometry. The bacteria on the other hand can get close to the tumor itself to sense lactate at the highest possible concentration and report a simple fluorescence readout. The challenge is to recover the bacteria from saliva quickly and efficiently.

We show that these modified bacteria can be collected by DEP from artificial saliva media allowing for fluorescence quantification around the electrode edges. This fluorescence intensity corresponds to the amount of lactate the bacteria had been exposed to (Figure 35.3F) [23]. The concentration of the bacteria around the electrode edges results in a signal amplification allowing quantification of bacteria fluorescence in the presence of lactate at a concentration as low as 0.1 mM [23]. This method of diagnosis could increase the sensitivity of testing and can be readily modified to create a point-of-care diagnostic tool, significantly reducing the need for invasive techniques and the time needed for testing.

35.4.5 Multiplexing DEP-Based Biomarker Detection

The heterogeneity of different cancer-driving mechanisms results in heterogeneity in biomarker expression between cancer patients. It is important to detect a panel of different biomarkers to account for as much of this heterogeneity as possible.

As previously discussed, the DEP microarray device uses a label-free method for isolation of many nanoparticle types, based on their dielectric properties, that carry multiple biomarkers. We have developed cyclic-immunofluorescence (cyclic-IF) techniques that can be used with the DEP chips to analyze up to six proteins in a single DEP collected sample (Figure 35.4) [30]. Cyclic-IF has gained traction as a method to increase multiplexing of biomarker staining in tissue analysis by first staining and then bleaching the fluorescence away to allow for a subsequent immunofluorescence staining round using different antibodies with the same fluorescent colors as before. Here we achieve photobleaching by incubating in a basic solution followed by exposure to high intensity UV light [30]. This has applications for different cancer types, such as prostate cancer, where multi-biomarker panels are necessary for the complex staging of the disease. A typical cyclic-IF protocol is used on tissue biopsies obtained by invasive surgeries, but the electrokinetic microelectrode array used here enables

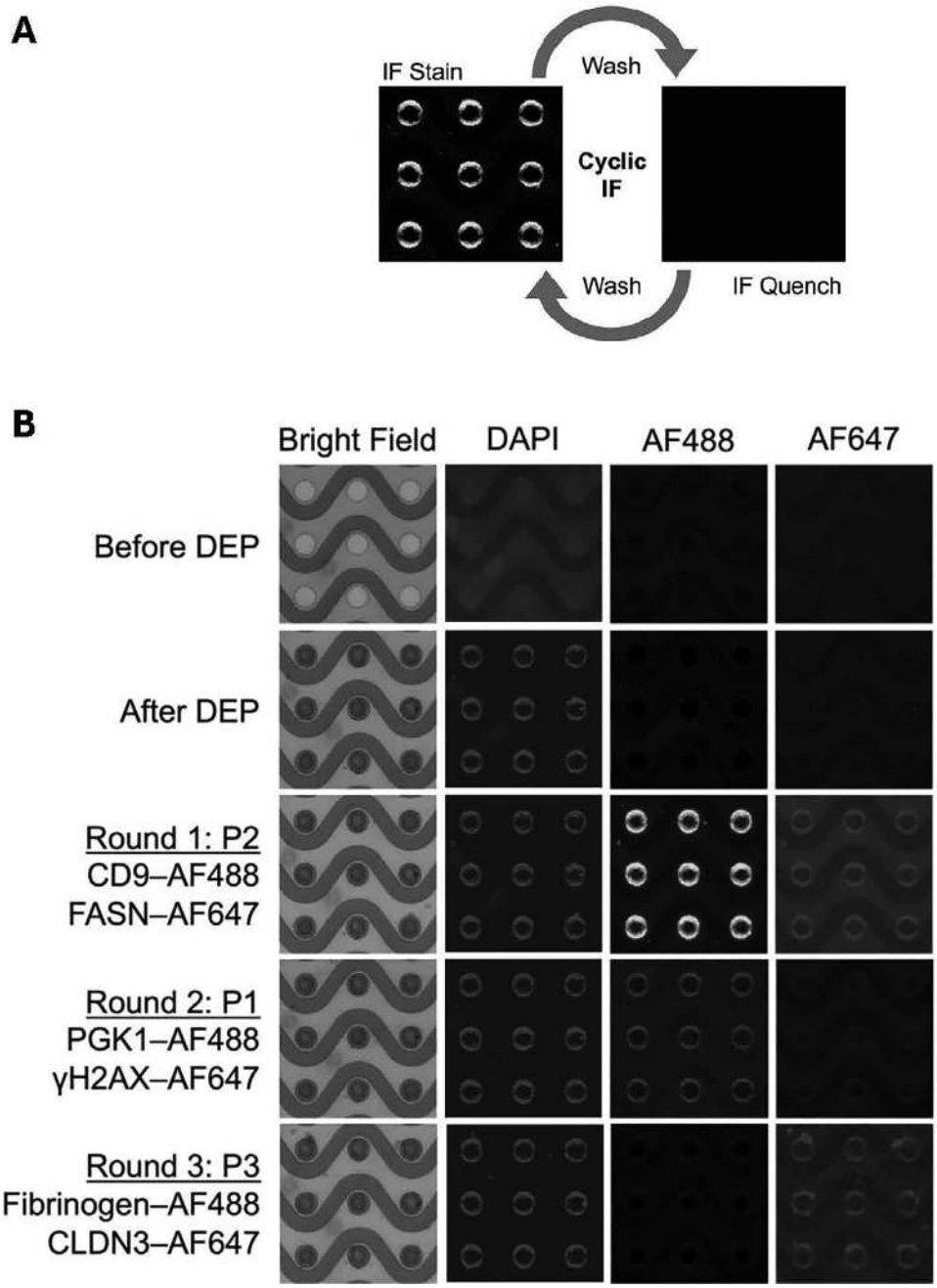


FIGURE 35.4 *The process of conducting cyclic immunofluorescence on the DEP chip.* (A) A schematic showing the cyclic-IF process of immunostaining a set of biomarkers, imaging, and then bleaching the dyes to eliminate the fluorescence thereby getting the sample ready for a subsequent round of immunostaining. (Reproduced and adapted under the terms of the Creative Commons Attribution [30]. Copyright 2022, The Authors. Electrophoresis published by Wiley-VCH GmbH.) (B) A series of different biomarkers were analyzed using the cyclic-IF technique on a single set of DEP collected EVs. (Reproduced and adapted under the terms of the Creative Commons Attribution [30]. Copyright 2022, The Authors. Electrophoresis published by Wiley-VCH GmbH.)

similar analysis of biomarker panels through low volume, less invasive liquid biopsy.

35.5 Discussion

Clinical liquid biopsy diagnostics will benefit greatly from the addition of the valuable biomarkers carried by nanoparticles released into circulation by tumors. The challenge preventing the clinical translation of nanoparticle-based diagnostics is the time-consuming and labor-intensive traditional methods used to recover the nanoparticles from plasma for analysis. This creates a critical unmet need to simplify nanoparticle-based diagnostics. Our work developing DEP techniques to recover EVs and other nanoparticles directly from undiluted patient plasma addresses this challenge by making EV recovery quick

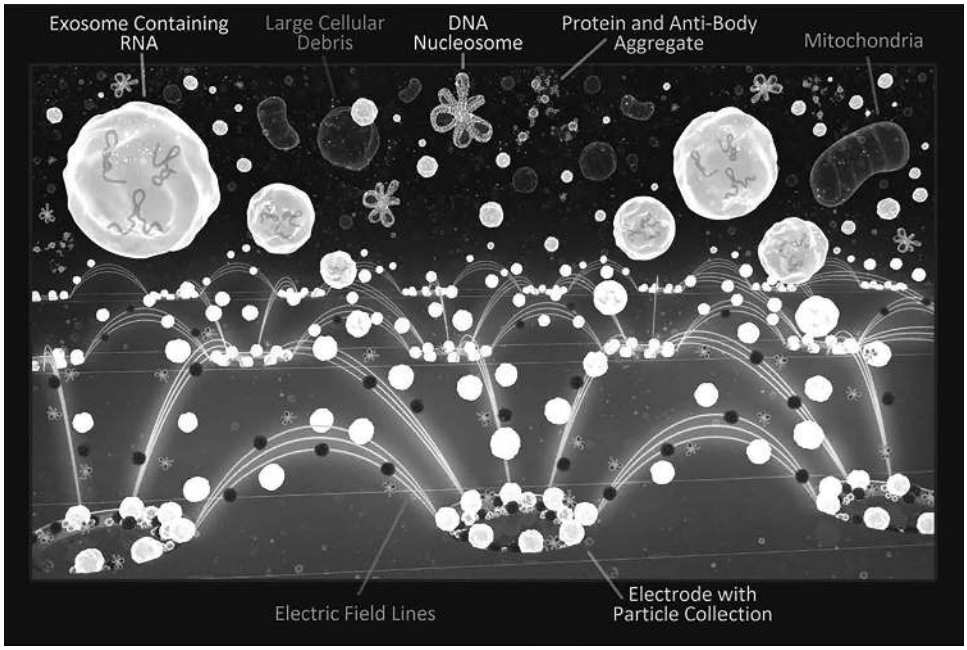


FIGURE 35.5 *Collection of various types of particles from plasma*. The electric field gradients are represented by the curved lines that draw a select subset of circulating nanoparticles to the electrode edge for collection and purification from the bulk plasma sample. (Adapted with permission from [11]. Copyright 2017 American Chemical Society.)

and highly automatable. We have shown that DEP can collect a range of different nanoparticles from high-conducting buffer and undiluted plasma including EVs, cf-DNA nanoparticles, bacteria, synthetic drug delivery nanoparticles, and organelles (Figures 35.3 and 35.5).

We have also shown the ability to detect cancer-related biomarkers on these DEP collected nanoparticles and their ability to differentiate between patients with pancreatic cancer and healthy controls (Figure 35.3G and H). We have also developed methods to multiplex biomarker detection by using cyclic-IF techniques directly on the chip (Figure 35.4).

Together this work shows the exciting promise of DEP-based technology for bringing EV-based diagnostics into the clinical setting.

35.6 Future Trends

New avenues of research in the application of DEP-based nanoparticle collection to the field of liquid biopsy diagnostics include the expansion of the technique to additional body fluid types. Various body fluids, especially those that diseased tissues are in direct contact with or secrete directly into, can be enriched for different types of nanoparticles and biomarkers relative to the blood or plasma, making direct sampling of these fluids a potential way to increase the signal-to-noise ratio for biomarkers of interest. This requires the DEP collection parameters to be optimized, based on the conductivity and viscosity of each fluid type, to maximize collection.

New research will also include expanding the number of biomarkers in these diagnostic panels to improve overall disease detection performance. Since DEP can simultaneously collect a range of different kinds of nanoparticles carrying different types of biomarkers, it is possible to collect biomarkers that span the genomic and proteomic levels of biological information from a single fluid sample on a single device. Efficiently combining different types of biomarkers, which are independent from one another in their expression levels, into a single panel can help detect disease at higher levels of sensitivity and specificity compared to a single biomarker due to the heterogeneity in biomarker expression levels observed across the patient population.

Acknowledgments

Figure 35.1A was created with BioRender.com.

REFERENCES

1. Jeppesen DK, Zhang Q, Franklin JL, Coffey RJ. Extracellular vesicles and nanoparticles: Emerging complexities. *Trends Cell Biol*. 2023; 33(8): 667–681.
2. Greten FR, Grivennikov SI. Inflammation and cancer: Triggers, mechanisms, and consequences. *Immunity*. 2019; 51(1): 27–41.
3. Li X-F, O'Donoghue JA. Hypoxia in microscopic tumors. *Cancer Lett*. 2008; 264(2): 172–180.
4. Bockhorn M, Jain RK, Munn LL. Active versus passive mechanisms in metastasis: Do cancer cells crawl into vessels, or are they pushed? *Lancet Oncol*. 2007; 8(5): 444–448.
5. Stejskal P, Goodarzi H, Srovnal J, Hajduch M, van 't Veer LJ, Magbanua MJM. Circulating tumor nucleic acids:

Biology, release mechanisms, and clinical relevance. *Mol Cancer*. 2023; 22(1): 15.
6. Chen Y, Wu FH, Wu PQ, Xing HY, Ma T. The role of the tumor microbiome in tumor development and its treatment. *Front Immunol*. 2022; 13: 935846.
7. Aytar Celik P, Derkus B, Erdogan K, Barut D, Blaise Manga E, Yildirim Y, Pecha S, Cabuk A. Bacterial membrane vesicle functions, laboratory methods, and applications. *Biotechnol Adv*. 2022; 54: 107869.
8. Luna R, Heineck D, Hinestrosa JP, Dobrovolskaia I, Hamilton S, Malakian A, Gustafson KT, Huynh KT, Kim S, Ware J, Stimson E, Ross C, Schutt CE, Ibsen SD. Enhancement of dielectrophoresis-based particle collection from high conducting fluids due to partial electrode insulation. *Electrophoresis*. 2023; 44(15–16): 1234–1246.
9. Kalra H, Adda CG, Liem M, Ang CS, Mechler A, Simpson RJ, Hulett MD, Mathivanan S. Comparative proteomics evaluation of plasma exosome isolation techniques and assessment of the stability of exosomes in normal human blood plasma. *Proteomics*. 2013; 13(22): 3354–3364.
10. Ibsen S, Sonnenberg A, Schutt C, Mukthavaram R, Yeh Y, Ortac I, Manouchehri S, Kesari S, Esener S, Heller MJ. Recovery of drug delivery nanoparticles from human plasma using an electrokinetic platform technology. *Small*. 2015; 11(38): 5088–5096.
11. Ibsen SD, Wright J, Lewis JM, Kim S, Ko SY, Ong J, Manouchehri S, Vyas A, Akers J, Chen CC, Carter BS, Esener SC, Heller MJ. Rapid isolation and detection of exosomes and associated biomarkers from plasma. *ACS Nano*. 2017; 11(7): 6641–6651.
12. Gramse G, Dols-Perez A, Edwards MA, Fumagalli L, Gomila G. Nanoscale measurement of the dielectric constant of supported lipid bilayers in aqueous solutions with electrostatic force microscopy. *Biophys J*. 2013; 104(6): 1257–1262.
13. Pegtel DM, Gould SJ. Exosomes. *Annu Rev Biochem*. 2019; 88: 487–514.
14. Pitera JW, Falta M, van Gunsteren WF. Dielectric properties of proteins from simulation: The effects of solvent, ligands, pH, and temperature. *Biophys J*. 2001; 80(6): 2546–2555.
15. Li L, Li C, Zhang Z, Alexov E. On the dielectric "constant" of proteins: Smooth dielectric function for macromolecular modeling and its implementation in DelPhi. *J Chem Theory Comput*. 2013; 9(4): 2126–2136.
16. Sarno B, Heineck D, Heller MJ, Ibsen SD. Dielectrophoresis: Developments and applications from 2010 to 2020. *Electrophoresis*. 2021; 42(5): 539–564.
17. Pethig R. Protein dielectrophoresis: A tale of two clausius-mossottis—or something else? *Micromachines*. 2022; 13(2): 261.
18. Pethig R. Review article-dielectrophoresis: Status of the theory, technology, and applications. *Biomicrofluidics*. 2010; 4(2).
19. Sarno B, Heineck D, Heller MJ, Ibsen SD. Dielectrophoresis: Developments and applications from 2010 to 2020. *Electrophoresis*. 2021; 42: 539–564.
20. Luna R, Heineck DP, Bucher E, Heiser L, Ibsen SD. Theoretical and experimental analysis of negative dielectrophoresis-induced particle trajectories. *Electrophoresis*. 2022; 43(12): 1366–1377.
21. Gustafson KT, Huynh KT, Heineck D, Bueno J, Modestino A, Kim S, Gower A, Armstrong R, Schutt CE, Ibsen SD. Automated fluorescence quantification of extracellular vesicles collected from blood plasma using dielectrophoresis. *Lab Chip*. 2021; 21(7): 1318–1332.
22. Sonnenberg A, Marciniak JY, McCanna J, Krishnan R, Rassenti L, Kipps TJ, Heller MJ. Dielectrophoretic isolation and detection of cfc-DNA nanoparticulate biomarkers and virus from blood. *Electrophoresis*. 2013; 34(7): 1076–1084.
23. Hamilton S, Shea D, Ibsen S, Brasino M. On-chip dielectrophoretic recovery and detection of a lactate sensing probiotic from model human saliva. *Electrophoresis*. 2023; 44(3–4): 442–449.
24. Lewis JM, Vyas AD, Qiu Y, Messer KS, White R, Heller MJ. Integrated analysis of exosomal protein biomarkers on alternating current electrokinetic chips enables rapid detection of pancreatic cancer in patient blood. *ACS Nano*. 2018; 12(4): 3311–3320.
25. Lotvall J, Hill AF, Hochberg F, Buzas EI, Di Vizio D, Gardiner C, Gho YS, Kurochkin IV, Mathivanan S, Quesenberry P, Sahoo S, Tahara H, Wauben MH, Witwer KW, Thery C. Minimal experimental requirements for definition of extracellular vesicles and their functions: A position statement from the international society for extracellular vesicles. *J Extracell Vesicles*. 2014; 3: 26913.
26. Hamilton S, Evans-Dutson S, Mira JLM, Heller MJ, Ibsen SD. A single microfluidic device approach to direct isolation, purification, and amplification of cfDNA from undiluted plasma. *Sens. Actuators B Chem*. 2024: 136374.
27. Ware J, Shea D, Lim JY, Malakian A, Armstrong R, Pethig R, Ibsen S. Collection of serum albumin aggregate nanoparticles from human plasma by dielectrophoresis. *Electrophoresis*. 2024; 45(19–20): 1748–1763. doi: 10.1002/elps.202400046. Epub 2024 Jul 30. PMID: 39081085; PMCID: PMC11695576.
28. Manouchehri S, Ibsen S, Wright J, Rassenti L, Ghia EM, Widhopf GF, 2nd, Kipps TJ, Heller MJ. Dielectrophoretic recovery of DNA from plasma for the identification of chronic lymphocytic leukemia point mutations. *Int J Hematol Oncol*. 2016; 5(1): 27–35.
29. Baranyai T, Herczeg K, Onódi Z, Voszka I, Módos K, Marton N, Nagy G, Mäger I, Wood MJ, El Andaloussi S. Isolation of exosomes from blood plasma: Qualitative and quantitative comparison of ultracentrifugation and size exclusion chromatography methods. *PLoS One*. 2015; 10(12): e0145686.
30. Gustafson KT, Sayar Z, Le H, Gustafson SL, Gower A, Modestino A, Ibsen S, Heller MJ, Esener S, Eksi SE. cycDEP: Cyclic immunofluorescence profiling of particles collected using dielectrophoresis. *Electrophoresis*. 2022; 43(16–17): 1784–1798.

36

Conditional Reprogramming: A Living Biomarker and Phenotypic Screening Drug Platform for Urological Cancer

Danyal Daneshdoust, Abdul Qawee Rani, Peng Wang, Shang-jui Wang, Akshay Sood, Debasish Sundi, Qi-en Wang, Zaibo Li, Anil V. Parwani, Nagireddy Putluri, Cheryl Lee, Jenny Li, and Xuefeng Liu

36.1 Introduction

36.1.1 Urological Cancer Clinical Background

Urological cancers refer to malignancies of the prostate and urothelial (including bladder, ureter, and urethra), kidney, testicular, and penile cancers. It has been estimated that 496,160 new cases and 70,020 deaths from urological cancers will occur in the United States in 2025 [1]. According to recent statistics, prostate cancer is the second most prevalent malignancy in males across the world. Over the past few decades, the treatment of advanced prostate cancer has significantly improved due to the use of androgen deprivation therapy and more potent inhibitors of the androgen receptor pathway. However, the acquisition of cellular plasticity inevitably leads to treatment resistance [2, 3].

In comparison to other cancer types, bladder cancer survival rates have not demonstrated significant advancements over the past three decades. The efficacy of particular bladder cancer treatment strategies varies based on a patient's clinical characteristics, associated risk factors, and clinical stages. The five-year recurrence rate for patients with non-muscle invasive bladder cancer is high, ranging from 60% to 70% [4], while the prognosis for those with muscle invasive bladder cancer is relatively poor, with a 65% chance of death within the five-year follow-up period [5].

Given the significant burden of urological cancers, using advanced approaches to comprehend tumor biology and treatment outcomes in urological cancer is still the ultimate goal pursued by researchers worldwide. Unfortunately, our understanding of the underlying mechanisms is limited due to the scarcity of suitable models. Consequently, it is imperative to enhance the development of more robust models and broaden the utilization of existing ones to facilitate the advancement of innovative therapeutic strategies. We present here a novel cell culture technology called conditional reprogramming (CR). We discovered the CR method in 2012, and in 2017 developed comprehensive protocols that were published in *Nature Protocols* [6, 7]. This technology development and its applications have been supported by several NCI grants.

36.1.2 Conditional Reprogramming Technology

Conditional reprogramming is a novel and robust cell culture technique that allows rapid, efficient, and continuous expansion of normal/benign and precancerous/cancerous cells derived from patients. CR technology is co-culturing epithelial cells with a Rho kinase (ROCK) inhibitor, Y-27632, in the presence of irradiated Swiss-3T3-J2 mouse fibroblasts (feeder cells). The CR method can rapidly convert normal and tumor cells to a state of "reprogrammed stem-like" in which the cells are highly proliferative and maintain their original karyotypes. Importantly, the removal of these conditions restores the capacity for cell differentiation and the phenotype will be completely reversible [6–8]. Hence, this method is called "conditional reprogramming." From a variety of tissues, such as fine-needle aspiration (FNA), core biopsies, surgical specimens, and patient-derived xenograft tissues, the CR technique can effectively generate huge numbers of primary cells [9]. Briefly, the tissue samples obtained from cancer patients are first assessed histologically by a pathologist to determine the composition of the specimen (to analyze the ratio of benign and malignant cells). The samples are first mechanically diced then dispersed into single cells by enzymatic digestion and plated in a medium containing irradiated Swiss-3T3-J2 mouse fibroblasts (feeder cells) and ROCK inhibitor. Epithelial colonies are readily observed in two days and usually proliferate rapidly to reach confluence in approximately five days. CR technology enables the establishment and culture of $2*10^6$ epithelial cells in only five to six days and continues to passage for at least 110 days and 100 population doublings until the experiment was terminated [7, 10, 11].

Conditionally reprogrammed cells (CRCs) is the term used to describe cells generated using this method. Subsequently, a sequencing analysis should be performed on both the CRCs and their parental tissues to validate the derivation of the resultant cells [12, 13].

The genetic and histological properties of the parent tissue can be recapitulated in CRCs, allowing them to maintain a highly proliferative state [14].

An important feature is that these cultures can be used for establishing xenografts [8], patient-derived xenograft cell

DOI: 10.1201/9781003449942-41

lines [15], cell cultures from PDXs, and organoid cultures. CRCs maintain cell lineage commitment and the cell heterogeneity found in a biopsy [6, 16].

Ease of execution, genotype stability, exponential growth, and high success rate in a single model system are some of the reported advantages of CR technology. These properties qualified CR as an outstanding *in vitro* model compared to other models such as organoids and patient-derived xenografts [9]. Modeling disease, precision medicine, drug discovery, regenerative medicine, and noninvasive diagnosis are some of the applications of CR technique [12, 17, 18].

Clinical research applications of CR technology have been investigated in breast cancer [19], lung cancer [20], prostate cancer [21], bladder cancer [22], gastric cancer [23], and liver cancer [24]. The conditional reprogramming system can be successfully applied to various epithelial tissues such as prostate, kidney, skin, lung, and several others [7]. Furthermore, CR technology has shown its potential in rare cancers such as respiratory papillomatosis [25] and salivary cancer [26].

A study investigating the use of CRCs to determine therapy for recurrent respiratory papillomatosis was published in the *New England Journal of Medicine* in 2012 [25]. A 24-year-old patient with recurrent respiratory papillomatosis had taken several chemotherapy agents and undergone over 350 laryngeal ablation surgeries, but all strategies had failed. The CR technique was approved for the patient's paired normal and tumor cells to be cultured for drug screening in order to control the progressive and chemoresistant disease. Consequently, the researchers found that the mutant HPV-11 genomes in the lung tumor CRCs and the laryngeal tumor CRCs were of different sizes, and vorinostat was found to be a useful treatment. Consequently, the tumor sizes had stabilized after three months of vorinostat treatment. The aforementioned case indicates that the CR method holds significant promise for advancing individualized precise medicine [25]. Moreover, the use of the system is not limited to humans. It can be applied to many kinds of mammals such as mice, rats, ferrets, horses, cows, and dogs [22, 27].

36.1.3 Comparison of the Technology and Approaches to Existing Ones

Conventionally, cancer cell lines have broadly served as efficient models for cancer research, precision medicine, regenerative medicine, and preclinical studies [28–30]. Despite their significant contributions to cancer research, these models have several drawbacks. Depending on the original tissue [7], only about 10% of cell lines can be successfully developed into immortalized or tumorous cell lines, and most primary cells, especially normal cells, are difficult to culture because of their short lifespan [29]. Furthermore, there have been notable mutations in the genes of the established cell lines, which makes it difficult to accurately recapitulate the complex characteristics of the original tissue and maintain the tumor heterogeneity [31]. Therefore, marked differences between the *in vivo* tumor microenvironment and the *in vitro* culture increase worries that these cell lines do not accurately mimic tumors in the patients.

Preclinical research on novel therapeutics has benefited greatly from the use of animal models, particularly the recently developed engineered mouse models and primary patient-derived xenografts, which have also contributed significantly to our understanding of the biology and etiology of cancer. In general, genetically modified animals, xenograft models, syngeneic models, chemical carcinogenesis models, and other models are used in animal cancer research [32, 33]. While animal models have the potential to overcome the limitations of cell lines and better mimic human disease and therapeutic response, heterogeneity exists between the different species. Hence, the application of animal models might be limited by low throughput, high costs, and technical difficulties; in addition, differences in specific species cause incorrect recapitulation of biological and therapeutic responses.

Recent progress in biotechnology has led to significant changes in the development of cancer models. Patient-derived models (PDMs) have emerged as a particularly promising approach, as they maintain consistent genetic characteristics with their parental tumors. Different types of PDMs such as organoids, induced pluripotent stem cells (iPSCs), patient-derived xenografts (PDXs), and conditionally reprogrammed cells have been widely used in cancer research due to their ability to better replicate the complexity of human tumors [34–38]. PDMs have different uses in cancer research depending on the context and methods employed. A comparison of these models (Table 36.1) is presented in this chapter, along with a detailed discussion of the conditional reprogramming technology.

This chapter provides a comprehensive review of the current state and potential applications of the CR method in urological cancer research.

36.2 Materials

36.2.1 Reagents: All Reagents Are Listed in Table 36.2

36.2.2 Preparation of the Reagents

100x hydrocortisone: dissolve 25 mg in 5 mL cold 100% ethanol to make a 5 mg/mL solution. Add 0.8 mL of 5 mg/mL solution to 100mL HBES containing 5% fetal bovine serum (FBS). Filter sterilize and store at –20 °C in 10 mL aliquots.

100x cholera toxin (ICN): add 1.2 mL of sterile water to 1 mg vial, making 10 μM solution. Dilute 50 μL of 10 μM solution into 50 mL HBES containing 0.1% bovine serum albumin. Filter sterilize and store at 4 °C in 10 mL aliquots.

100x insulin (Sigma): prepare immediately before use and never freeze! Dissolve 12.5 mg insulin in 25 mL of 0.005 M HCL. Filter sterilize with a syringe filter pre-wet with FBS.

100x adenine (Sigma): dissolve 121mg adenine in 50 mL of 0.05 M HCl by stirring for 1 h. Filter sterilize and store at –20 °C in 10 mL aliquots.

100x epidermal growth factor (EGF; R&D Systems): dissolve 100 μg in 10 mL sterile water. Add 90 mL HBES containing 0.1% bovine serum albumin. Filter sterilize and store at –20 °C in 10 mL aliquots.

Tissue-culture dishes (Falcon/Becton Dickinson; Franklin Lakes, NJ).

TABLE 36.1

Comparison of Patient-Derived Models

Methods	Advantages	Shortcomings	References
2D culture	1. Inexpensive technique 2. Easy to manipulate genetically 3. Suitable for high-throughput drug screens in a short amount of time at a low cost	1. Loss of tumor heterogeneity 2. Genetic drift between different laboratories (for cell lines) 3. Lack of microenvironment 4. Not suitable for low-grade tumor establishment 5. Limitation of cell-cell and cell-extracellular matrix interactions	[28–30]
Organoids	1. 3D culturing 2. Can generate both healthy and tumor organoids 3. Maintain tumor heterogeneity 4. Possibility to co-culture tumor organoids with elements of the microenvironment (pathogens [bacteria] and immune cells)	1. Lack of microenvironment (immune cells, vasculature, and microbiota) 2. Dependent on stem cells 3. Lack of protocol and medium standardization 4. Overgrowth of nonmalignant cells	[34, 35, 38–43]
iPSC	1. Pluripotent differentiation 2. Can combine with gene editing and 3D organoids	1. Slow and inefficient procedure 2. Difficult to reprogram cancer cells 3. Safety issues	[44, 45]
PDXs	1. *In vivo* model 2. Direct engraftment from human tumor 3. Maintain histological, genomic, and transcriptomic features of tissue of origin 4. Recapitulate the natural environment of the tumor 5. Humanized mice model with reconstituted human immune systems	1. Expensive technique 2. Resource and time consuming 3. Not suitable for high-throughput drug screening 4. Rely on interactions with a mouse microenvironment 5. Only tumor models 6. Challenging to be reproducible on a large scale	[34, 37]
CRCs	1. A wide range of specimen sources 2. Paired normal and tumor cells culturing 3. Cost saving and rapid expansion (1–10 days) 4. Can maintain original karyotype and tumor heterogeneity 5. High-throughput drug screening 6. Gene profiling analyses 7. Suitable for low-grade tumor establishment	1. Contamination with feeder cells 2. Overgrowth of benign cells 3. Lack of stromal components	[6, 7, 9, 12, 22]

TABLE 36.2

Reagents

Name	manufacturer	Cat #
Trypsin 0.05%	Thermo Fisher Scientific	25300120
Trypsin 0.25%	Thermo Fisher Scientific	25200114
EtOH	Thermo Fisher Scientific	AC61510010
Clorox® germicidal bleach, regular	VWR	89501–620
Kanamycin	Thermo Fisher Scientific	15160054
Nystatin	Sigma	N6261
Collagenase/Hyaluronidase	Stemcell Technologies	NC-9694308
hydrocortisone	Sigma	H-0888
EGF	Thermo Fisher Scientific	Cat #AF-100-15-1MG
insulin	Sigma	I-5500
cholera toxin	VWR	80055–160
DMEM	Thermo Fisher Scientific	11965–092
fetal bovine serum	Thermo Fisher Scientific	16140–071
glutamine	Thermo Fisher Scientific	25030–081
pen/strep	Thermo Fisher Scientific	15140–122
F-12 NUTRIENT MIX (1X)	Thermo Fisher Scientific	11765–054
GENTAMICIN (10 mg/ml)	Thermo Fisher Scientific	15710–064
FUNGIZONE (250 µg/ml)	Thermo Fisher Scientific	BP264550
PBS	Thermo Fisher Scientific	14190250
Dispase	Thermo Fisher Scientific	354235
Y-27632	Enzo	270–333M025
DMSO	Thermo Fisher Scientific	BP-231–100
1X phosphate buffered saline (PBS)	Thermo Fisher Scientific	10010023
Trypsin-EDTA (0.25%), phenol red	Thermo Fisher Scientific	25200056
Trypsin-EDTA (0.05%), phenol red	Thermo Fisher Scientific	25300054
UltraPure™ 0.5M EDTA, pH 8.0	Thermo Fisher Scientific	15575020
Ham's F-12 Nutrient Mix	Thermo Fisher Scientific	11765054
DMEM, high glucose	Thermo Fisher Scientific	11965092

Swiss-3T3-J2 mouse fibroblasts

3T3 cell medium:500mL DMEM with high glucose, 50 mL bovine calf serum, 5 mL Pen/Strep.

F medium: 375mL F12 medium, 125 mL DMEM with high glucose, 25 mL FBS, 5 mL 100x hydrocortisone, 5 mL 100x insulin, 5 mL 100x cholera toxin, 5 mL 100x adenine, 5 mL, Pen/Strep. 5 mL 100x EGF.

36.2.2.1 Cryopreservation Medium

Cryopreservation medium: 90% of fetal bovine serum (FBS) (Gibco, cat. no.16140–071) with 10% of dimethyl sulfoxide (DMSO) (Fisher Scientific, cat. no. BP-231–100) and 10 mM Y–27632. The medium was stored at −80 °C for up to six months.

36.2.2.2 Conditional Reprogramming Medium (Complete F Medium)

500mL of complete F medium (1x):125 ml of F12 nutrient medium (Gibco, cat. no. 11765–054) were added to 373ml of complete DMEM (Gibco, cat. no. 11965–092). The following reagents were added to the mixture: 0.5 ml of insulin (Sigma-Aldrich, cat. no. I-5500), 0.5 ml of hydrocortisone/epidermal growth factor (EGF) mix (Life Technologies, cat. no. PHG0313), 0.5 ml of gentamicin (Gibco, cat. no. 15710–064), 0.5 ml of amphotericin B (Fisher Scientific, cat. no. BP264550), 4.3 ml of cholera toxin (Sigma-Aldrich, cat. no. C-3012), and 0.5 ROCK inhibitor Y-27632 (Enzo, cat. no. 270–333M025). The medium was filter-sterilized using a 0.2 mm sterile filter and stored at 4 °C for up to two months. The final concentrations: insulin, 5 mg/ml, hydrocortisone, 25 ng/ml, EGF, 0.125 ng/ml, gentamicin, 10 mg/ml, amphotericin B, 250, cholera toxin, 0.1 nM, ROCK inhibitor Y-27632, 10 mM.

36.2.3 Equipment

Equipment and supplies are listed in Table 36.3.

TABLE 36.3

Equipment and Supplies

Laminar flow hood with vacuum
Centrifuge (Beckman Coulter Allegra X15-R with SX4750 swinging bucket rotor or equivalent)
Cell culture incubator (37 °C, 95% humidity, 5% CO_2, Heracell, Thermo Scientific)
Inverted phase contrast microscope (4× and 10× objectives, Zeiss or equivalent)
Freezer – 80 °C (Thermal Scientific)
Water bath, 37 °C
Isopropanol cell-freezing container
Cryovials, 1.2 ml (USA Scientific, cat. no. 1412–9100 or equivalent)
Six-well Nunclon Delta Surface tissue culture dishes (Thermo Scientific, cat #. 140675 or equivalent)
P1000 Pipetman and sterile tips with filter
Pipet-aid and sterile 5 and 10 ml plastic disposable pipettes
Sterile filter units, 500 ml (Millipore, Stericup cat. #: SCGPU05RE)
Conical tubes, 15 and 50 ml (Falcon, cat #: 352097 and 352098 or equivalent)
Eppendorf tube
Gauge needle (20.5)
EVOS fl digital inverted microscope (Advanced Microscopy Group, Bothell)

36.3 Methods (Step by Step SOP)

36.3.1 Fresh Tissue Handling, Transportation, and Shipping

1. Need to have an IRB approval.
2. Observe specimen sterility.
3. Cut fresh specimen/s to pieces (5–10 mm in maximum dimension) at surgery or immediately when received by the pathologist.
4. Dissect the tumor lesions from surrounding normal tissues, as much as possible.

   Caution: If sending paired normal and tumor samples from the same subject, place each place sample in a separate, clearly labeled container. The matching normal tissue should be obtained as far from the tumor lesion as possible.
5. Place specimens a container with sample medium.

   Caution: Place each specimen in a separate, clearly labeled container. The corresponding "twin" specimens should be sent for histology.
6. Close the lid of the container tightly.
7. Package samples and ship overnight, priority delivery. Ship on ice bags (4 °C).

   Caution: *Not* on −20 °C ice bags. *Never* freeze or fix the samples prior to shipping.

36.3.2 Cryopreservative Tissue

1. Same as prior steps 1–4.
2. Bring containers with specimens to tissue culture hood.
3. Use a sterile scalpel for cutting the tissue fragments. cut the fragment into cubes of about 1–2 mm size.
4. Place the cubes in 1.0 ml cryopreservative medium (following list) for 10–30 minutes at room temperature.
5. Store the vials at –80 °C or LN. Use dry ice for transportation.

   Cryopreservative medium: 90% serum, 10% DMSO, 5 μM Y-27632.

36.3.3 Isolation of Primary Cells

1. Dilute 1 part 10x collagenase/hyaluronidase mixture with 9 parts F medium (DMEM/F12) and place into a 14 mL or 50 mL centrifuge tube. Approximately 2–5 mL of the F/collagenase/hyaluronidase/FBS solution will be required for every –two to three mouse prostates. The volume of dissociation mix for human samples will be dependent on the size of the sample (typically ten times more solution than the volume of the sample).
2. Rinse resect tissues (e.g., prostates) with 95%–100% ethanol very quickly and 1x PBS twice, and transfer to a sterile Petri dish containing cold PBS. (Using a dissecting microscope, a fine set of forceps, and scissors, remove residual amounts of fat from tissue).

3. Transfer the tissue to the tube containing 3 ml of DMEM/F12/collagenase/hyaluronidase/FBS and 1 ml of dispase and incubate for 1–3 hours at 37 °C.
4. After dissociation, centrifuge the cells at 1000 rpm for five minutes with the brake on and discard the supernatant.
5. Resuspend the pellet in 2–5 mL of 0.25% Trypsin-EDTA and leave on ice for 1 hour or RT for 5 min.
6. Add 10 mL DMEM with 10% FBS and filter the cell suspension through a 40–70 µm cell strainer into a new 50 mL centrifuge tube. Transfer filter-through to new 15 ml tube and centrifuge at 1000 rpm for 5 minutes with the brake on and discard the supernatant.
7. Resuspend cells in F medium and put them in flask (5ml for T-25 and 10 ml for T-75).

36.3.4 Feeder Layer Tissue Culture and Treatment

Preparing the Swiss 3T3 cells for use as a feeder layer:

36.3.4.1 Irradiate Feeder Cells

1. Start with the T175 flask of J2 cells at lower passage number and at less than 100% density (in DMEM media).
2. Rinse cells with Ca++/Mg2+ free 1x PBS.
3. Trypsin cells with 1 ml of 0.05% trypsin for 20–40 sec.
4. Add 9 ml of complete DMEM and stop trypsinization.
5. Pellet cells with 1000rm at 4 °C.
6. Resuspend all cell pellets in 10 ml of DMEM.
7. Irradiate the cell suspension with 3000 Rad (30 Gray) with Cesium-137 irradiator (JL Shepherd Mark).
8. Plate the irradiated cells to flasks. (These irradiated cells can be plate in T-75 or T-25 with DMEM/F12 medium for immediate use, or DMEM for next day [up to three days] use, or storage at 4 °C for 3 days and long term storage at –80 °C or in LN.)

36.3.5 Co-Culture and Separation of Feeders

1. Take the flask with irradiated cells and replace DMEM media with F media. (Total amount of media in flask with cell resuspension is 15 ml for T75 and 30 ml for T175.)
2. Add primary tumor cell suspension to the flask with irradiated feeder cells.
3. Place the flask in the incubator.
4. Change media in flask with fresh F media one day after plating.
5. Change media every day or every other day, depending on cell growth and media condition.
6. Split the cells when they are more than 75% dense.

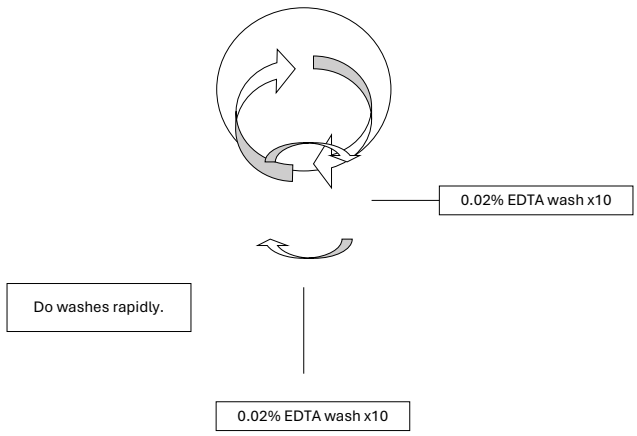
You may use irradiated feeder cells for at least 3 days after plating them. Usually they may last up to 7–8 days.

36.3.6 Differential Trypsinization

1. Co-cultures were rinsed with PBS and incubated with 0.05% trypsin solution at room temperature for 30 seconds to 1 minute.
2. Close monitor by phase microscopy. When the feeder cells rounded up and began to detach from the substrate, the cultures were gently tapped and the detached cells removed by aspiration. Epithelial colonies remained tightly adherent.
3. The epithelial cells were rinsed again with PBS and then re-trypsinized at 37 °C for 3–5 minutes.
4. Gentle pipetting to disperse the epithelial cells into a single-cell suspension.
5. The cells were pelleted through PBS containing 10% serum (to neutralize trypsin).
6. After centrifugation, the cell pellets were resuspended in F medium for passaging or resuspended in cryopreservative medium for freezing.

36.3.7 Remove Feeder Cells with EDTA

1. Once the epithelial cells on the plate are 70%–75% confluent, aspirate off medium and wash once with pre-warmed PBS (cells that are too confluent initiate differentiation).
2. Aspirate off PBS and wash with pre-warmed 0.02% EDTA (0.68mM) in 1x PBS. This wash should be performed fairly rapidly and is done to remove the Swiss 3T3 cells, but leave behind the epithelial cells. The EDTA wash is done ten times by pipetting 10–12 mLs directly onto the cells on the outer perimeter of the plate, while rotating the plate. Then using this same EDTA again, do the same for the inner half of the plate cells.
3. Diagram of 0.02% EDTA in 1x PBS washes:



4. Aspirate off the EDTA from the previous washes, and repeat the 0.02% EDTA washes for the inner and outer halves of the plate. This makes a total of 40 washes. This EDTA wash protocol is for human keratinocytes; other cells may be less strongly adherent and the number of washes may have to be modified.

5. Aspirate off the EDTA and wash the plate with pre-warmed PBS x3 still keeping the plate tilted. On the last PBS wash look at the cells under the microscope and verify that all of the Swiss 3T3 are washed away, if not, more EDTA washes should be performed.
6. Trypsinize the epithelial cells using their normal protocol (for keratinocytes trypsin 2 minutes RT, aspirate off the trypsin and then incubate for 5 min at 37 °C, tap plate, resuspend in DMEM/F12). Epithelial cells should be plated back onto irradiated feeder layers in low density 250 K for a 100 mm plate.

36.3.8 Isolate Cells from Urine Samples

1. Collect urine samples from the clean, mid-stream urine of patients.
2. 50 mL of urine is collected in sterile tubes with penicillin/streptomycin mix (final concentration: 200 U/mL penicillin and 200 mg/mL streptomycin).
3. Centrifuge at 100 rpm for 10 min. The supernatant will be aspirated, and the pellet will be washed with PBS three times before being added to culture plates.

36.4 Results

36.4.1 Urine CRCs as a Functional Platform for BC Biology and Treatment

Our previous work evaluated drug responses via CR culture of bladder cancer cells, which was taken of urine samples. The success rate of urine CRCs cultures was high at 83.3% (50 of 60 cases); specifically, low-grade bladder cancer was at 75.0% (9 of 12 cases) and high-grade bladder cancer was at 85.4% (41 of 48 cases). Interestingly we obtained 100% success rate of BC cell cultures after we optimized collection conditions as follows. Additionally, we did not observe bias for the success rates of CRC urine cultures with respect to pathology group, disease status, age, or gender from all bladder cancer cases. We also reported CR cells had 79.7%–82.6% of genetic variation similar to the original tumors. Thirteen patients underwent drug sensitivity tests, which revealed varied responses to conventional drugs like gemcitabine, cisplatin, pirarubicin, and epirubicin. However, no clinical susceptibility data was provided for other drugs such as paclitaxel, lapatinib, bortezomib, and docetaxel, which showed significant inhibition of urine culture growth [46]. Thus, urine BC cells can be a promising tool to assess drug responses and sensitivity by a minimal invasive method rather than invasive procedures (Figure 36.1).

36.4.2 Urine CRCs as a Living Biomarker for BC Patients' Response

In order to study whether urine BC cells can predict clinical response, we collected urine samples before and after treatment. Thus, we were able to establish urine CR BC cells from Patients 1 and 2 before treatment (surgery and/or intravenous chemotherapy, gemcitabine + cisplatin); and from urine samples after treatment; however, we only established UCCC from Patient 1, not Patient 2, after two cycles of chemotherapy. This outcome was consistent with clinical follow-up showing that Patient 1's tumor soon relapsed after treatment. However, the tumor in Patient 2 showed no recurrence after the first and second cycles of chemotherapy, correlating with our urine BC cell culture results. Our data also showed that response of CR BC cells generated from patients' urine samples prior to treatment correlated with response of the BC patients. The CR BC cells of one patient showed relatively low sensitivity to pirarubicin, which was consistent with clinical follow-up. This patient's tumor soon relapsed three months later after intravesical treatment with pirarubicin (40 mg) followed by undergoing transurethral resection. Another patient showed relative sensitivity to gemcitabine and cisplatin. In the absence of surgery, his CT obtained six months later showed stable disease (SD) after treatment with gemcitabine (1000 mg/m2/intravenous) and cisplatin (70 mg/m2/intravenous). These clinical results correlated well with the drug-sensitivity profiles of the corresponding CR cells in response to gemcitabine and cisplatin (Figure 36.2). Thus, urine BC cells can be used to test drug sensitivity, and their response seemed to correlate with the patients' response.

36.4.3 Establishment of the First Patient-Derived NEPC Cell Line

NEPC is a lethal subtype. The study of NEPC development has been hampered by a lack of clinically relevant models. The previous study reported a unique and first-in-field patient-derived xenograft (PDX) model of adenocarcinoma (LTL331)-to-NEPC (LTL331R) transdifferentiation [47]. Recently, we applied CR culture to establish a LTL331 PDX-derived cancer cell line named LTL331_CR_Cell (Figure 36.3) [48]. These cells retain the same genomic mutations as the LTL331 parental tumor. They can be continuously propagated *in vitro* and can be genetically manipulated. Androgen deprivation treatment on LTL331_CR_Cells had no effect on cell proliferation. Transcriptomic analyses comparing the LTL331_CR_Cell to its parental tumor revealed a profound downregulation of the androgen response pathway and an upregulation of stem and basal cell marker genes. When grafted under the renal capsules of male NOD/SCID mice, LTL331_CR_Cells spontaneously gave rise to NEPC tumors. This is evidenced by the histological expression of the NE marker CD56 and the loss of adenocarcinoma markers such as PSA. This NEPC cell line provides a novel research tool of NEPC and allows for the investigation of mechanisms underlying NEPC development by enabling gene manipulations *ex vivo* and subsequent functional evaluations *in vivo.*

36.5 Discussion

36.5.1 CR Technology in Cancer Detection, Diagnosis, and Other Applications

Using tissue biopsies is currently the gold standard for genetic profiling of malignancies. Tissue biopsies are invasive procedures and also have several drawbacks, such as limitation

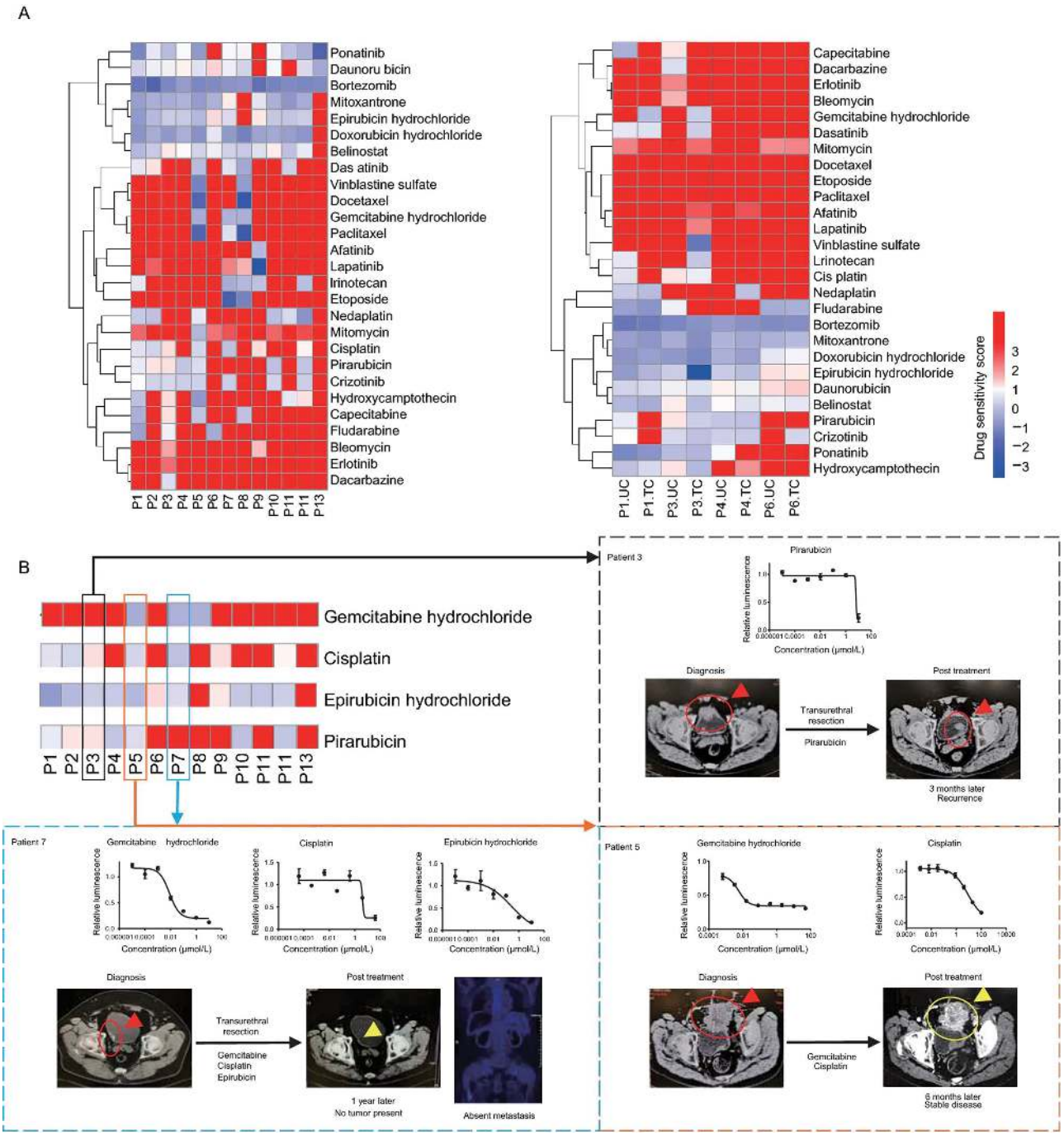


FIGURE 36.1 *Correlation of drug responses of CRCs and clinical outcomes of corresponding drug-treated patients*. (A) The heatmap of drug sensitivity scores (DSS) in urine CRCs derived from patient 1–13 and in urine and corresponding tumor CRCs of patients 1, 3, 4, and 6. DSS means log10 transformed fold-change of IC50 in comparison to 5637 cell line. Red represents higher DSS and blue represents lower DSS. The DSS was clustered by heatmap in R package ($n = 3$ independent experiments, the data indicates the mean). TC = tumor CRC. (B) The top left panel shows parts of heatmap of drug sensitivities of urine CRCs. The solid borders highlight patients 3, 5, and 7. Within the dotted borders, the patients' dose-response curves for the indicated drugs are exhibited in the upper row, and the CT and PET-CT of the patients before and after treatment are shown in the lower row. The type of surgical treatment is shown on the black arrow and the drugs used are under the black arrow. Red circles and arrows indicate primary or recurrence tumors. Yellow circles and arrows represent response [46]. (This is an open access article distributed under the terms of the Creative Commons CC BY license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.)

in sample preparation, sensitivity and accuracy, patient risk, and procedural expenses. As a result, the technique is incompatible with clinical longitudinal monitoring. Additionally, a key shortcoming of tissue biopsies is their inability to detect intra-tumoral and intermetastatic genetic variability, which affects the test's accuracy.

Liquid biopsies, as a revolutionary tool, hold great promise to overcome these existing sampling limitations. Unlike

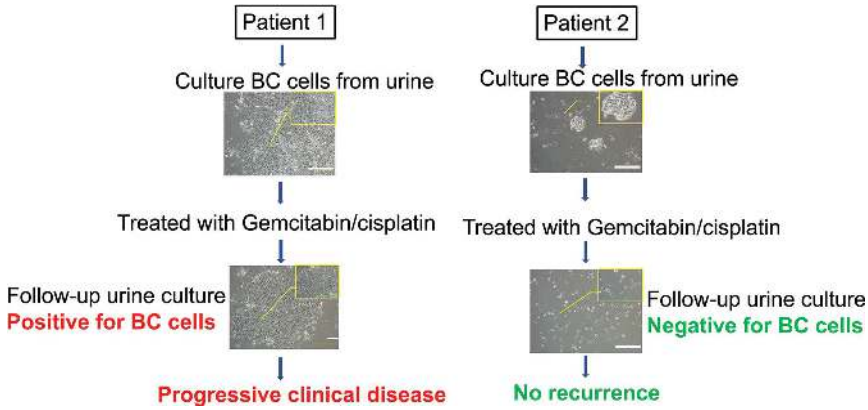


FIGURE 36.2 *Clinical relevance of urine CR BC cells from Patients 1 and 2*. Patient 1 accepted tumor resection following chemotherapy but soon relapsed (red). Patient 2's tumor was sensitive to chemotherapy. Phase contract images of urine CR cell conditions before and after chemotherapy treatment in Patients 1 and 2. We failed to establish urine CR BC cells after the first and second cycles of chemotherapy in Patient 2 since this patient was disease-free (green). Red color indicated positive CR BC culture and active disease (non-responder, or BC recurrence). Green color indicated negative CR BC culture and disease free (responder, or no recurrence).

traditional biopsies that require invasive procedures, liquid biopsy harnesses the power of a simple blood draw or urine sample to detect and analyze various biomarkers, such as cell-free DNA (cf-DNA), circulating tumor cells (CTCs), circulating RNAs (miRNAs/lncRNAs/mRNAs), cell-free proteins, and exosomes. For instance, CTCs from bladder cancer can be detected in the urine and serum of patients with metastatic bladder cancer, with higher levels of CTCs correlating with cancer aggressiveness [49, 50]. CTCs derived from bladder cancer can be measured by using CTC-specific proteins, such as c-MET and PD-L1 [51–53]. Increased CTC levels were also able to predict clinical outcomes, such as recurrence and survival [54]. CellSearchTM, an FDA-approved CTC assay kit, is currently being used in clinical settings for prognostic purposes in breast cancer but has low sensitivity and specificity [55]. Being able to culture cancer cells from blood and other liquid biopsies would be extremely valuable, but there are many technical issues that limit the study of CTC biology and their applications. It is technically difficult to capture and characterize viable CTCs [56–58].

In addition to CTCs, DNA, RNA, and exosome candidate biomarkers in serum and/or urine have been identified in urological cancer. Mutation of the human telomerase reverse transcriptase gene (hTERT) promoter occurs in 70% of human bladder cancer and is also associated with recurrence and poor prognosis of bladder cancer [59, 60]. In the urine specimens of bladder cancer patients, the telomerase reverse transcriptase (TERT) promoter mutations correlated with recurrence [61], while KRAS2 mutations were found in the plasma even before bladder cancer diagnosis [62]. Urinary UBE2C and hTERT mRNA were found to be potential markers for early diagnosis and prognosis of bladder cancer [63]. Urinary levels of miR–126 andmiR–146a–5p were also found to be elevated in bladder cancer and are associated with tumor grade and invasiveness [64]. Exosomes transfer biologically active molecules and can be secreted into the urine, blood, and other body fluids [65]. Hence, exosomes are essential mediators of cell–cell communication [66]. Interestingly, there is a strong association between heightened exosome levels and bladder cancer [67]. In urinary exosomes, significantly increased levels of active molecules (e.g., TACSTD2, lncRNAs–HOTAIR, HOX-AS-2, ANRIL, and linc-ROR) were found in patients with high-grade muscle invasive bladder cancer (MIBC) [68]. However, none of the aforementioned markers are found in a majority of bladder cancer and urine samples of patients and thus have not been suggested for clinic use. Further, they cannot be used for functional analyses to predict clinical outcomes.

These circulating molecules have emerged as noninvasive biomarkers for different malignancies, especially for urological cancers. One of its primary applications lies in the early detection and monitoring of cancer. By analyzing these circulating components, liquid biopsy can identify the presence of cancer at an early, more treatable stage, allowing for timely interventions and personalized treatment plans. This noninvasive approach is not only more comfortable for patients but also provides a valuable tool for oncologists to track a patient's response to treatment and detect the emergence of resistance, thus aiding in precision medicine.

In the case of urothelial carcinoma, however, urine cytology and endoscopy with biopsy continue to be the gold standards for diagnosis despite the ongoing identification of noninvasive markers based on urine and peripheral blood. The major reason for this is that these two tests can offer morphological and visual information for the early detection of urinary cancers, and the pathological data can serve as a useful guide for tumor staging and grading. But even with their strength, these two analyses are not without drawbacks. For instance, urine cytology, as a specific tool, has limited sensitivity when it comes to detecting low-grade tumors [69]. Despite being highly effective in diagnosing bladder cancer, biopsy and cystoscopy are invasive procedures that are unable to accurately identify small, flat tumors. Thus, there is still a need to develop advanced noninvasive diagnostic tools that have high sensitivity and specificity.

The most practical source of liquid biopsies is urine, which has been thoroughly studied for biomarkers and cytology in

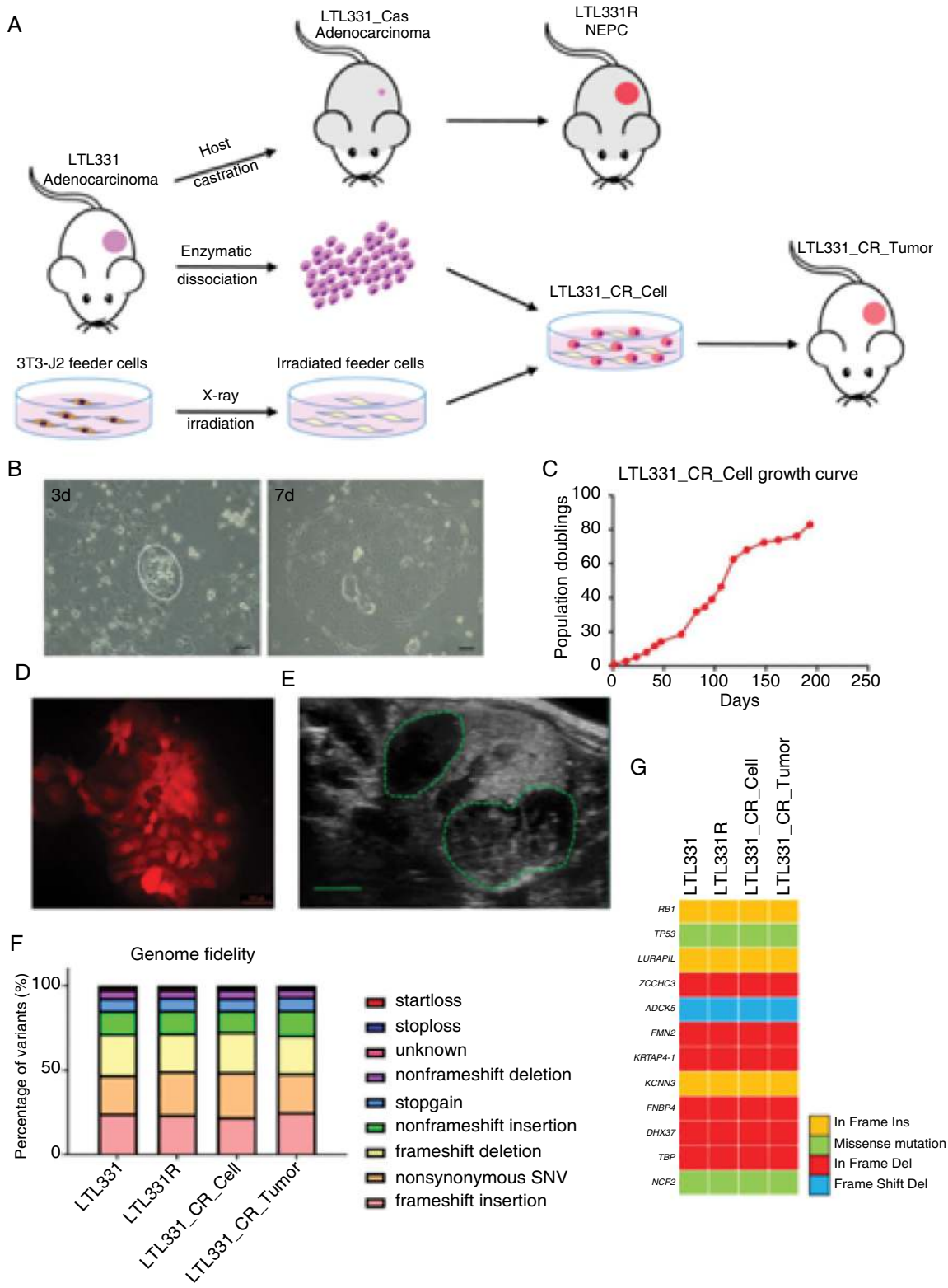


FIGURE 36.3 *Establishment, application, and molecular characterization of CR NEPC cell line from the LTL331 PDX* [48]. (A) Schematic showing the conditional reprogramming culture establishment process. (B) Representative images of LTL331_CR_Cells. (C) Growth curve of LTL331_CR_Cell. The CR cells were serially passaged and the cell numbers were recorded at each passage until the culture was terminated. (D) Representative image of LTL331_CR_Cell following genetic manipulation. Cells were transduced with lentivirus to express the mCherry fluorescent protein. (E) Representative image of LTL331_CR_Tumor growing under the renal capsule of a mouse. (F–G) Mutation analysis of expressed transcripts across the LTL331-derived models. Mutation type and frequency was analyzed with RNA-seq data. The percentage of each mutation type is stacked into one bar, with different types indicated by the corresponding colors (F). Representative gene mutations are shown (G).

the clinical diagnosis of bladder cancer [56]. Since the CR technique enables the isolation and culture of both patient-derived tumor and normal cells without alteration in karyotype, it is feasible to use CR as a noninvasive approach to diagnose urothelial carcinomas if tumor and/or normal urothelial cells from urine can be successfully isolated and cultured *in vitro*. Using the CR technology, Jiang and colleagues have successfully established bladder cancer cells from patients' urine samples. The overall success rate of urine CRC cultures was 83.3% (50 of 60 cases); specifically, low-grade bladder cancer was 75.0% (9 of 12 cases), and high-grade bladder cancer was 85.4% (41 of 48 cases). Interestingly, they obtained a 100% success rate of bladder cancer cell cultures after they optimized collection conditions. Additionally, they did not observe bias for the success rates of CRC urine cultures with respect to pathology group, disease status, age, or gender from all bladder cancer cases. They also reported that CR cells had 79.7%–82.6% of genetic variation similar to the original tumor. Subsequently, the urine CRCs were used for drug sensitivity tests and compared with clinical consistency in urine CRC-based drug testing.

Recently, Wang and colleagues established 35 primary prostate patient-derived cells (ten samples of benign prostate hyperplasia [BPH] and 25 samples of prostate tumors) to capture specific features of prostate cancer patients, such as distributions of mRNA, protein, and surface proteins, and responses to therapy. Prostate cancer epithelial cells were found to be viable and proliferating over the long term, as evidenced by their 100% primary culture success rate. Employing primary cell lines for prostate cancer, they carried out a thorough multi-omics analysis, quantifying over 1800 mutations, 17,000 transcripts, 7000 proteins, 2000 cell-surface proteins, and responses to 33 FDA-approved drugs. They found that the disease nature of patients in the prostate cancer model repository is correlated with transcriptome expressions, protein/surface protein abundance, and genomic abnormalities using integrated multi-omics analysis. The resulting data could be used to predict drug sensitivity relevant to precision medicine [70]. Using matched CR prostate normal and tumor cells, Pollock and colleagues [71] assessed the anticancer effect of natural product strigolactone analogues. They found that strigolactone selectively killed prostate cancer cells. CRCs derived from prostate cancer patients were helpful in determining that p53-induced apoptotic and autophagy signaling pathways to mediate cell death in prostate-cancer cells was the possible mechanism of VMY-1– 103 (VMY), which shows rapid and effective antitumor activities. Timofeeva and colleagues [21] established matched normal and tumor cultures from patient specimens obtained during prostatectomy; the result was CR cells that proliferated *in vitro* indefinitely while maintaining stable karyotypes. Prostate cancer could be studied *in vitro* because tumor CR cells and normal cells had distinct exome sequences and different gene expression patterns, despite having very similar morphologies. It has been demonstrated that Y-27632 and J2 feeders can be used to quickly isolate a large number of prostate cancer cells from primary tissues while preserving the tumor's original characteristics [6]. The aforementioned studies have demonstrated the potential of CR in establishing patient-derived urological cancer cells, which can be employed to establish tumor and normal cells as a means of diagnosis, monitoring, and detection of anticancer effective therapies.

36.5.2 Challenges and Prospects for the Clinical Setting

The employment of CR technology in urological cancer research exhibits significant potential; however, it is crucial to recognize its limitations. One limitation is the absence of stromal components, such as matrix elements, vascular immune cells, and endothelial cells. This limitation hinders the comprehensive analysis of how stromal cells influence tumor cell growth and how tumor cells respond to drugs [72].

Differentiating between normal and tumor epithelial cells can be a challenge since normal cells can sometimes proliferate faster than tumor cells, even surpassing tumor cells in cultures. However, adjustments to the conventional CR procedure can allow for the *in vitro* selective growth of tumor cells [72]. CR technology has great potential for use in urological cancer research, despite these drawbacks. Biomedical researchers must select a model that is appropriate for their particular research question, as there is no one-size-fits-all model. Scientists frequently combine technologies at various levels, ranging from molecules to cells, organs, and populations, to achieve their research objectives.

36.6 Future Trends

The conditional reprogramming system has shown great promise in breast cancer research, but it is important to acknowledge and address its limitations and challenges as we discussed earlier. There are several future trends of the CR technology in basic and translational cancer research: (1) Selection protocols are needed after initial and rapid expansion of primary cultures; thus the combined protocols will allow to selectively expand normal or tumor cells or heterogenous tumors based on the primary research purposes. Indeed, cell surface markers based on flowcytometry and serum containing medium with ROCK inhibitor have been used for this purpose for advanced lung cancer cell lines, breast cancer cells with rare mutations, and mouse bladder cancer cell lines. (2) A combination of CR, organoids, and PDX models is needed for a variety of types of human cancer from cancer initiation to progressions and recurrences. This combination is also important for the generation of living biobanks, anti-drug discovery, and drug repurposes *in vitro* and *in vivo*, since any single method has its unique advantage and disadvantages. (3) Simplified or optimized CR protocols, especially conditioned medium and ready-to-go irradiated feeder cells, are required for clinical lab settings. (4) Preanalytical variability of biospecimen including procurement, process, storage, and transportation can affect the reproducibility of clinical assays. These preanalytical factors need to be studied and optimized for future clinical assays. (5) Finally, genetic manipulation of the patient-derived CR cells is needed for basic cancer biology studies.

Acknowledgment

This work was partially supported by the grants R33CA177466, R01CA222148, R33CA258016, R01CA276474, U01CA278927 from the National Institutes of Health and startup funding from The Ohio State University Comprehensive Cancer Center.

Conflicts of Interest

Several patents for conditional reprogramming technology have been awarded to Georgetown University by the US Patent Office. The license for this technology has been given to a Maryland-based startup company for commercialization. The inventor, X.L., and Georgetown University receive potential royalties and payments from the company. CR media and CR cells have been distributed by Propagenix, Stem Cell Technology, Fisher Scientific, ATCC, and so on. Other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

1. Siegel, R.L., et al., Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 2023. 75(1): p. 10–45.
2. Carceles-Cordon, M., et al., Cellular rewiring in lethal prostate cancer: The architect of drug resistance. *Nature Reviews Urology*, 2020. 17(5): p. 292–307.
3. Quintanal-Villalonga, Á., et al., Lineage plasticity in cancer: A shared pathway of therapeutic resistance. *Nature Reviews Clinical Oncology*, 2020. 17(6): p. 360–371.
4. Chen, W., et al., Cancer statistics in China, 2015. *CA: A Cancer Journal for Clinicians*, 2016. 66(2): p. 115–132.
5. Kamat, A.M., et al., Bladder cancer. *The Lancet*, 2016. 388(10061): p. 2796–2810.
6. Liu, X., et al., ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. *The American Journal of Pathology*, 2012. 180(2): p. 599–607.
7. Liu, X., et al., Conditional reprogramming and long-term expansion of normal and tumor cells from human biospecimens. *Nature Protocols*, 2017. 12(2): p. 439–451.
8. Suprynowicz, F.A., et al., Conditionally reprogrammed cells represent a stem-like state of adult epithelial cells. *Proceedings of the National Academy of Sciences*, 2012. 109(49): p. 20035–20040.
9. Palechor-Ceron, N., et al., Conditional reprogramming for patient-derived cancer models and next-generation living biobanks. *Cells*, 2019. 8(11): p. 1327.
10. Liu, X., et al., Conditionally reprogrammed cells (CRCs): A new model for cancer research and personalized medicine. *Cancer Research*, 2014. 74(19_Supplement): p. 2972–2972.
11. Zhong, M. and L. Fu, Culture and application of conditionally reprogrammed primary tumor cells. *Gastroenterology Report*, 2020. 8(3): p. 224–233.
12. Liu, W., et al., Conditional reprogramming: Modeling urological cancer and translation to clinics. *Clinical and Translational Medicine*, 2020. 10(2): p. e95.
13. Wu, X., et al., Conditional reprogramming: Next generation cell culture. *Acta Pharmaceutica Sinica B*, 2020. 10(8): p. 1360–1381.
14. Alamri, A.M., et al., Primary cancer cell culture: Mammary-optimized vs conditional reprogramming. *Endocrine-Related Cancer*, 2016. 23(7): p. 535–554.
15. Mondal, A.M., et al., Fidelity of a PDX-CR model for bladder cancer. *Biochemical and Biophysical Research Communications*, 2019. 517(1): p. 49–56.
16. Saeed, K., et al., Comprehensive drug testing of patient-derived conditionally reprogrammed cells from castration-resistant prostate cancer. *European Urology*, 2017. 71(3): p. 319–327.
17. Stover, E.H., et al., Implementation of a prostate cancer-specific targeted sequencing panel for credentialing of patient-derived cell lines and genomic characterization of patient samples. *The Prostate*, 2022. 82(5): p. 584–597.
18. Jung, M., et al., Novel paired normal prostate and prostate cancer model cell systems derived from African American patients. *Cancer Research Communications*, 2022. 2(12): p. 1617–1625.
19. Daneshdoust, D., et al., Unlocking translational potential: Conditionally reprogrammed cells in advancing breast cancer research. *Cells*, 2023. 12(19): p. 2388.
20. Gao, B., et al., Non-malignant respiratory epithelial cells preferentially proliferate from resected non-small cell lung cancer specimens cultured under conditionally reprogrammed conditions. *Oncotarget*, 2017. 8(7): p. 11114.
21. Timofeeva, O.A., et al., Conditionally reprogrammed normal and primary tumor prostate epithelial cells: A novel patient-derived cell model for studies of human prostate cancer. *Oncotarget*, 2017. 8(14): p. 22741.
22. Daneshdoust, D., et al., Conditional reprogramming modeling of bladder cancer for clinical translation. *Cells*, 2023. 12(13): p. 1714.
23. Zhao, R., et al., Conditional cell reprogramming in modeling digestive system diseases. *Frontiers in Cell and Developmental Biology*, 2021. 9: p. 669756.
24. Su, S., et al., Long-term culture and characterization of patient-derived primary hepatocytes using conditional reprogramming. *Experimental Biology and Medicine*, 2019. 244(11): p. 857–864.
25. Yuan, H., et al., Use of reprogrammed cells to identify therapy for respiratory papillomatosis. *New England Journal of Medicine*, 2012. 367(13): p. 1220–1227.
26. Alamri, A.M., et al., Expanding primary cells from mucoepidermoid and other salivary gland neoplasms for genetic and chemosensitivity testing. *Disease Models & Mechanisms*, 2018. 11(1): p. dmm031716.
27. Saffari, P.S., et al., Most canine ameloblastomas harbor HRAS mutations, providing a novel large-animal model of RAS-driven cancer. *Oncogenesis*, 2019. 8(2): p. 11.
28. Sharma, S.V., D.A. Haber, and J. Settleman, Cell line-based platforms to evaluate the therapeutic efficacy of candidate anticancer agents. *Nature Reviews Cancer*, 2010. 10(4): p. 241–253.
29. Meijer, T.G., et al., Ex vivo tumor culture systems for functional drug testing and therapy response prediction. *Future Science OA*, 2017. 3(2): p. FSO190.
30. Gillet, J.-P., S. Varma, and M.M. Gottesman, The clinical relevance of cancer cell lines. *Journal of the National Cancer Institute*, 2013. 105(7): p. 452–458.

31. Ghani, F.I., et al., An ex-vivo culture system of ovarian cancer faithfully recapitulating the pathological features of primary tumors. *Cells*, 2019. 8(7): p. 644.
32. Oliveira, P.A., R. Arantes-Rodrigues, and C. Vasconcelos-Nobrega, Animal models of urinary bladder cancer and their application to novel drug discovery. *Expert Opinion on Drug Discovery*, 2014. 9(5): p. 485–503.
33. Reis, L.O., et al., Experimental animal model and RNA interference: A promising association for bladder cancer research. *World Journal of Urology*, 2009. 27: p. 353–361.
34. Xu, X., et al., A living biobank of matched pairs of patient-derived xenografts and organoids for cancer pharmacology. *PLoS One*, 2023. 18(1): p. e0279821.
35. Clevers, H., Modeling development and disease with organoids. *Cell*, 2016. 165(7): p. 1586–1597.
36. Rowe, R.G. and G.Q. Daley, Induced pluripotent stem cells in disease modelling and drug discovery. *Nature Reviews Genetics*, 2019. 20(7): p. 377–388.
37. Okada, S., K. Vaeteewoottacharn, and R. Kariya, Application of highly immunocompromised mice for the establishment of patient-derived xenograft (PDX) models. *Cells*, 2019. 8(8): p. 889.
38. O'Connell, L. and D.C. Winter, Organoids: Past learning and future directions. *Stem Cells and Development*, 2020. 29(5): p. 281–289.
39. Drost, J., et al., Sequential cancer mutations in cultured human intestinal stem cells. *Nature*, 2015. 521(7550): p. 43–47.
40. Li, X., et al., Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture. *Nature Medicine*, 2014. 20(7): p. 769–777.
41. Sato, T., et al., Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, 2009. 459(7244): p. 262–265.
42. Drost, J., et al., Organoid culture systems for prostate epithelial and cancer tissue. *Nature Protocols*, 2016. 11(2): p. 347–358.
43. Gao, D., et al., Organoid cultures derived from patients with advanced prostate cancer. *Cell*, 2014. 159(1): p. 176–187.
44. Xiao, X., et al., Generation of induced pluripotent stem cells with substitutes for Yamanaka's four transcription factors. *Cellular Reprogramming (Formerly" Cloning and Stem Cells")*, 2016. 18(5): p. 281–297.
45. Huangfu, D., et al., Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology*, 2008. 26(7): p. 795–797.
46. Jiang, S., et al., Continuous culture of urine-derived bladder cancer cells for precision medicine. *Protein & Cell*, 2019. 10(12): p. 902–907.
47. Wang, Y., et al., Molecular events in neuroendocrine prostate cancer development. *Nature Reviews Urology*, 2021. 18(10): p. 581–596.
48. Ci, X., et al., Conditionally reprogrammed cells from patient-derived xenograft to model neuroendocrine prostate cancer development. *Cells*, 2020. 9(6).
49. Flaig, T.W., et al., Detection of circulating tumor cells in metastatic and clinically localized urothelial carcinoma. *Urology*, 2011. 78(4): p. 863–867.
50. Naoe, M., et al., Detection of circulating urothelial cancer cells in the blood using the CellSearch system. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 2007. 109(7): p. 1439–1445.
51. Anantharaman, A., et al., Programmed death-ligand 1 (PD-L1) characterization of circulating tumor cells (CTCs) in muscle invasive and metastatic bladder cancer patients. *BMC Cancer*, 2016. 16: p. 1–11.
52. Qi, F., et al., Quantitation of rare circulating tumor cells by folate receptor α ligand-targeted PCR in bladder transitional cell carcinoma and its potential diagnostic significance. *Tumor Biology*, 2014. 35: p. 7217–7223.
53. Zhang, T., et al., Development of a novel c-met–based ctc detection platform. *Molecular Cancer Research*, 2016. 14(6): p. 539–547.
54. Zhang, Z., et al., The prognostic and diagnostic value of circulating tumor cells in bladder cancer and upper tract urothelial carcinoma: A meta-analysis of 30 published studies. *Oncotarget*, 2017. 8(35): p. 59527.
55. Alva, A., et al., Circulating tumor cells as potential biomarkers in bladder cancer. *The Journal of Urology*, 2015. 194(3): p. 790–798.
56. Di Meo, A., et al., Liquid biopsy: A step forward towards precision medicine in urologic malignancies. *Molecular Cancer*, 2017. 16: p. 1–14.
57. Contreras-Sanz, A., et al., Recent progress with next-generation biomarkers in muscle-invasive bladder cancer. *International Journal of Urology*, 2017. 24(1): p. 7–15.
58. Lodewijk, I., et al., Liquid biopsy biomarkers in bladder cancer: A current need for patient diagnosis and monitoring. *International Journal of Molecular Sciences*, 2018. 19(9): p. 2514.
59. Pietzak, E.J., et al., Next-generation sequencing of nonmuscle invasive bladder cancer reveals potential biomarkers and rational therapeutic targets. *European Urology*, 2017. 72(6): p. 952–959.
60. Cowan, M.L., et al., Detection of TERT promoter mutations in primary adenocarcinoma of the urinary bladder. *Human Pathology*, 2016. 53: p. 8–13.
61. Kinde, I., et al., TERT promoter mutations occur early in urothelial neoplasia and are biomarkers of early disease and disease recurrence in urine. *Cancer Research*, 2013. 73(24): p. 7162–7167.
62. Gormally, E., et al., TP53 and KRAS2 mutations in plasma DNA of healthy subjects and subsequent cancer occurrence: A prospective study. *Cancer Research*, 2006. 66(13): p. 6871–6876.
63. Xie, X.-Y., et al., Analysis of hTERT expression in exfoliated cells from patients with bladder transitional cell carcinomas using SYBR green real-time fluorescence quantitative PCR. *Annals of Clinical Biochemistry*, 2007. 44(6): p. 523–528.
64. Sasaki, H., et al., Expression level of urinary microRNA-146a-5p is increased in patients with bladder cancer and decreased in those after transurethral resection. *Clinical Genitourinary Cancer*, 2016. 14(5): p. e493–e499.
65. Wang, Y.-N., et al., The impact of PD-L1 N-linked glycosylation on cancer therapy and clinical diagnosis. *Journal of Biomedical Science*, 2020. 27(1): p. 1–11.
66. Yu, S., et al., Tumor-derived exosomes in cancer progression and treatment failure. *Oncotarget*, 2015. 6(35): p. 37151.
67. Nawaz, M., et al., The emerging role of extracellular vesicles as biomarkers for urogenital cancers. *Nature Reviews Urology*, 2014. 11(12): p. 688–701.
68. Berrondo, C., et al., Expression of the long non-coding RNA HOTAIR correlates with disease progression in

bladder cancer and is contained in bladder cancer patient urinary exosomes. *PLoS One*, 2016. 11(1): p. e0147236.
69. Muhammad, A.S., et al., Effectiveness of bladder tumor antigen quantitative test in the diagnosis of bladder carcinoma in a schistosoma endemic area. *Urology Annals*, 2019. 11(2): p. 143.
70. Wang, Z., et al., Integrative multi-omics and drug–response characterization of patient-derived prostate cancer primary cells. *Signal Transduction and Targeted Therapy*, 2023. 8(1): p. 175.
71. Pollock, C.B., et al., Strigolactone analogues induce apoptosis through activation of p38 and the stress response pathway in cancer cell lines and in conditionally reprogramed primary prostate cancer cells. *Oncotarget*, 2014. 5(6): p. 1683.
72. Palechor-Ceron, N., et al., Radiation induces diffusible feeder cell factor (s) that cooperate with ROCK inhibitor to conditionally reprogram and immortalize epithelial cells. *The American Journal of Pathology*, 2013. 183(6): p. 1862–1870.

37

Proteolytic Activity Signatures as Candidate Biomarkers for Thyroid Cancer

Danielle E. Dettling and Giselle M. Knudsen

37.1 Introduction

37.1.1 Proteolytic Biomarkers

A new area for the discovery of cancer diagnostic biomarkers lies in a specific branch of proteomics called "degradomics" for the analysis of proteolysis, which is catalyzed by proteases and peptidases and regulated by their endogenous inhibitors.[1, 2] Unlike proteomic measurements at the protein identity level, enzyme activities are evidenced by their product formation. Proteomics and epigenetics have already standardized the positional analysis of enzyme-catalyzed chemical marks on protein and DNA sequences, such as glycosylation, phosphorylation, acetylation, and ubiquitination post-translational modifications (PTMs),[3] as well as epigenetic marks in the form of methylation of DNA and PTMs of histones.[4, 5] Additionally, the products of enzyme activities in the form of metabolites are also an active area of research in metabolomics for the discovery of cancer biomarkers.[6, 7] In short, enzymes actuate an important niche between proteomics, epigenetics, and metabolomics.

Why monitor proteolysis (rather than other PTM enzyme activities)? In contrast to reversible and dynamic PTMs, which can be low stoichiometry, proteolysis is an irreversible process that amplifies the signal of its products and commits a cleaved protein to its fate, as in the case of caspase cascades activating cell death pathways, or in the fibrinolytic cascade catalyzed by thrombin, FactorXIIIa, and plasmin that activate blood clot formation.[8, 9] Protease function is integral to each of the hallmarks of cancer and to all stages of the cancer-immunity cycle, for example through catalysis of tumor microenvironment remodeling, regulation of antigen presentation, and processing of immune cell signaling molecules.[10–14]

A last argument that may be the most compelling is that multiple approved cancer biomarkers are themselves hypothesized, if not mechanistically proven, to be products or inhibitors of protolysis. These include the DR-70 fibrin product from fibrinogen, Pro-PSA vs free forms of PSA, CA125 MUC16 and CA15–3 or CA27.29 MUC1 shed forms of mucins, and HE4 which functions as an inhibitor of multiple proteases in fibrosis.[15–19] There is also precedent for the clinical use of a protease activity assay, for ADAMTS13, which cleaves von Willebrand factor and is used in the diagnosis of diseases of thrombosis.[20]

37.1.2 Thyroid Cancer Diagnosis

Incidence rates in the US for thyroid cancer have tripled over the period of 1975–2020, as reported in the SEER22 registries at NCI.[21, 22] Thyroid nodules (TN) also have a very high prevalence of 33%–68% in the general population.[23, 24] Most incidentally detected thyroid nodules are benign and asymptomatic, while only 2%–12% are found to be malignant.[25] This incongruence between thyroid cancer incidence and the high prevalence of benign nodules is likely due to increased detection sensitivity, which has led to the identification of more indolent or clinically insignificant cases.

The risk stratification of incidental thyroid nodules is initially based on ultrasonography, and if high-risk imaging features are found, a fine-needle aspirate biopsy (FNAB) with cytology is used to further classify the nodule.[26] FNA biopsy is estimated to be collected in more than 1,000,000 people every year in the US, with a 16%–21% annual increase.[27, 28] Although FNAB is a simple and safe procedure, cytological analysis on the aspirate is not always definitive, with approximately 30% of nodules being classified as indeterminate by cytology, a proportion that also appears to be rising, resulting in the need for repeat biopsies or diagnostic surgery.[29, 30] Retrospective studies show that 25%–41% of patients who undergo FNAB proceed to surgery, even though the majority 36%–75% of these patients ultimately will have benign nodules.[31–34] Consequently, the number of thyroidectomies performed in the United States has increased steadily.[35]

Thyroid surgery can result in permanent hypothyroidism and is associated with a 1%s–6% risk of complications. Beyond this burden to patients, the total cost of these surgeries to the American healthcare system has also steadily increased.[36] The cumulative result of thyroidectomies performed in the US, estimated at 137,000 based on HCUP data for 2020,[37] when multiplied by the median surgery cost of $8,508 per patient,[38] yields a total of $1.16 billion. Assuming 75% of these surgical cases involve benign nodules, this translates to a potential healthcare cost reduction of $874 million. Active surveillance, as the alternative for selected patients with very low-risk tumors, also has associated costs for repeated imaging and requires specialized and highly skilled radiologists to monitor disease progression.[39–41] Thus, there is significant need to correctly risk-stratify incidental thyroid nodules both to appropriately triage patients for resection and to remove from surveillance.

DOI: 10.1201/9781003449942-42

Over the past decade, genomic testing has been commercialized for use on FNA biopsy in order to risk-stratify indeterminate nodules, using three types of genomic markers: (1) somatic mutations, gene fusions, and copy number alterations; (2) gene expression alterations; and (3) microRNA expression (Afirma® GSC, ThyroSeq®, ThyGeNEXT/ThyraMIR®), improving the preoperative diagnosis of indeterminate nodules. All of these tests are based on the analysis of genomic alterations with modest cancer specificity; consequently, they are characterized by high sensitivity and low specificity, resulting in good negative predictive value (NPV) (91%–96%) and modest positive predictive value (PPV) (30%–66%) in post-validation studies,[42, 43] limiting their clinical use to opt-out only.

There remains an opportunity to improve specificity of tumor detection. Proteolytic enzymes have been identified in proteomic studies of thyroid cancer, with reported variation between cancer types.[44] In addition, proteases as well as endogenous inhibitors that regulate their activity are dysregulated in thyroid cancer compared to normal tissue, as revealed by both gene expression and proteomic studies.[45–49] Thus, proteolytic activity markers, as orthogonal species to genomic markers, were hypothesized to provide alternative and potentially highly specific information for improved diagnosis. This chapter will describe an approach for the discovery of thyroid cancer-selective proteolytic activity markers using an unbiased approach, their use as cleavable protease activity probes, and their initial evaluation for activity across a selection of benign and malignant thyroid tissues.

37.2 Materials

Fresh frozen samples of thyroid tissues were purchased from iSpecimen (Lexington, MA) and ProteoGenex (Inglewood, CA). All clinical samples were obtained following official protocols, with appropriate Institutional Review Board/Independent Ethics Committee IRB/IEC approval, which operates in accordance with the current federal regulations in addition to ICH, HIPAA, and GCP guidelines pertaining to protection of human subjects.

Cell culture media, biochemical reagents, and plasticware including serum-free DMEM and cell culture grade dPBS, 10k MWCO ultracentrifugal filter (UFC8010), and buffers were purchased from MilliporeSigma. Sequencing grade trypsin (V5111) was purchased from Promega. Analytical reagents include BCA kit (Thermo Fisher, 23227), HPLC-grade solvents and reagents: water, acetonitrile, and formic acid from SigmaAldrich, C18 zip tips (ZTC18S) from MilliporeSigma, and a UPLC column (ES900) purchased from Thermo Fisher. Custom-synthesized peptides were purchased from Alan Scientific, Inc., or Genscript.

37.3 Methods

37.3.1 Donor Tissue Samples

Fresh frozen tissues used as a discovery cohort included matched tumor and normal adjacent tissue (NAT) from $N = 12$ donors with thyroid cancer diagnoses. As a retrospective cohort for method testing, $N = 68$ additional samples were obtained from normal, benign, and thyroid cancer donors for a total of $N = 80$ donors. Investigator was blinded to the pathology diagnoses until after analysis. The diagnoses for these donor samples are summarized in Table 37.1.

TABLE 37.1

Donor Thyroid Tissue Samples

Diagnosis (N)	Cancer (52)	NAT (12)	Benign (28)
Histology	Papillary (41) Follicular (6) Medullary (5)	Papillary (12)	Goiter (10) Adenoma (2) Follicular adenoma (7) Hyperplasia (1) Hyperplastic nodules (5) Hashimoto's thyroiditis (2) Unknown (1)
Clinical Stage	I (25) II (13) III (5) IV (9)	I (8) II (3) III (1)	–
Gender	Female (32) Male (20)	Female (7) Male (5)	Female (26) Male (3)

37.3.2 Proteolytic Extract Preparation

Fresh frozen primary tissues samples were stored at −80 °C, then removed to dry ice and processed in a BSL-2 cell culture hood. Conditioned media was prepared from 100 to 200 mg tissue that was diced into 2–4 mm pieces, rinsed, and incubated for 48 h in serum-free DMEM media. The conditioned media was then centrifuged to remove debris, and the supernatant was buffer exchanged into dPBS and concentrated via ultracentrifugal filtration at 10kD MWCO. Mock FNA was prepared from 10–20 mg tissue that was mechanically dissociated in dPBS, and incubated for 1 h on wet ice prior to centrifugation to remove debris. Protein concentration of the proteolytic extract samples was determined by BCA assay, samples were flash frozen, and transferred to −80 °C for storage.

37.3.3 Proteomic Analysis

For peptide-sequencing with mass spectrometry, samples of conditioned media containing 10 μg of total protein were digested in solution with trypsin.[50] The resulting peptide samples were analyzed by LC-MS/MS using a QExactive Plus mass spectrometer (Thermo Scientific) equipped with a NanoAcuity UPLC system (Waters), using higher-energy collisional dissociation fragmentation, according to published protocols.[51] Protein identifications were output for the NAT and tumor conditioned media samples from $N = 12$ thyroid cancer patients, measured in technical duplicate. Label-free quantitation by spectral counting was used to approximate protein abundance,[52] and median normalized to the total signal obtained for each sample to allow comparison across data sets.[53] The input normalized peptide count data for each

sample was then $\log_2$-converted, and zero values were imputed from the normal distribution. An unpaired *t*-test using Welch's correction was performed between the NAT and tumor sets.

37.3.4 Protease Specificity Profiling

Protease specificity profiling was performed by multiplex substrate profiling with mass spectrometry (MSP-MS) method.[54] This method employs a physico-chemically diverse peptide library as substrates for proteases, and reactions are monitored over time with mass spectrometric detection of cleaved products. The resulting cleavages are assessed for specific cleavage in the enzyme-treated sample by comparison with results from a no-enzyme control incubation. In these reactions, proteolytic extracts were assayed at 50 µg/mL, and peptide substrates were 500nm concentration, measured at either pH 4.5 in 50 mM sodium acetate buffer or pH 7.4 in PBS, following published protocols.[50] The resulting peptide samples were analyzed with the same mass spectrometry methods as were used for proteomic analysis. Data were otherwise searched and analyzed as reported.[54]

37.3.5 Fluorescence Assays

Fluorescence assays were recorded using a Molecular Devices Flexstation 3 plate reader. Custom fluorescence resonance energy transfer (FRET) substrates containing a 5-mer cleavage sequence of an MSP-MS peptide were designed with guidance from the substrate preferences of the proteolytic enzymes identified in the sample, and synthesized commercially at > 95% purity. Substrates ALTC-03 and ALTC-04, containing MCA/DNP or FITC/DNP fluorophore-quencher pairs, respectively, were prepared as 5 mM stock solutions in DMSO, protected from light, and stored at −80 °C. Assay buffers mirrored those used for MSP-MS experiments. Substrates were prepared at 40 µM and added 10 µl/well into a 384-well plate. Protease extract samples were prepared as 1:25 v/v dilutions in assay buffer, then serially diluted 2.5-fold, and assayed in triplicate ($N = 3$). Reactions were started with addition of 10 µl diluted protease extract into each well containing substrate, with gentle mixing. Each assay plate was heated 5 min to 37 °C, followed immediately by kinetic measurement monitored over 1 h, at 1 min intervals. ALTC-03 was measured at ex310/em392nm, and ALTC-04 was measured at ex485/em520 nm. V_{max} was calculated over the linear range selected for each substrate. Specific activity was calculated as the catalytic rate normalized to the total mass of protein added to each well (V_{max}/[protein] in arbitrary units RFU/min/µl).

37.4 Results

To identify proteolytic signatures that can differentiate thyroid cancer from normal tissue activities, a discovery cohort of $N = 12$ thyroid cancer donors was selected, for whom paired tumor and normal adjacent tissue samples were available (Table 37.1). These patients were diagnosed with papillary thyroid cancer, the most prevalent form of thyroid cancer accounting for 80%–90% of cases.[21] Proteolytic extracts from conditioned media were prepared to enrich for the pericellular and secreted proteolytic activity released from these thyroid tissue samples and subjected to both a standard proteomic analysis for protein identification and protease activity profiling. Shown in Figure 37.1 is a volcano plot of the degradome proteins identified, indicating those with significant abundance differences between the NAT and tumor samples. A total of 106 active proteases or inhibitors from the degradome were identified, of which 13 were significantly upregulated in tumor and two were upregulated in NAT.

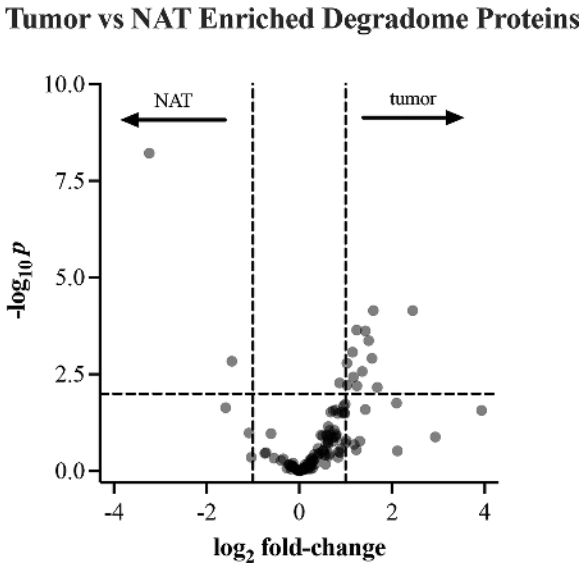


FIGURE 37.1 Differentially expressed degradome proteins identified with proteomic analysis of matched NAT and tumor tissue from $N = 12$ donors, represented in a volcano plot. Gridlines show significance cutoff for $p = 0.01$ and fold change = 2.

To profile the proteolytic activity released from these NAT vs tumor samples in a global and unbiased manner, the MSP-MS method was applied in an end-point analysis.[54] Reactions were performed under acidic and neutral pH conditions to allow for detection of a range of proteolytic activities. This library contains within it 2964 unique amide backbone bonds, each of which represent candidate cleavage motifs for proteases. Of these, the NAT and tumor samples produced a combined 484 unique peptide cleavages at pH 4.5 and 614 unique cleavages at pH 7.4. The statistically significant differential cleavages were all tumor-enriched, with a total of 74 significant cleavages detected at pH 4.5 and 32 at pH 7.4, as shown by volcano plots (Figure 37.2). Candidate cleavage sequences with the highest enrichment for cleavage in tumor vs NAT were evaluated for design into FRET reporter substrates, taking into consideration the substrate preferences of the tumor-enriched proteases identified in these samples. Candidate FRET reporter substrates from the acidic and neutral pH discovery analysis were tested against the discovery cohort, as shown for two substrates ALTC-03 and ALTC-04 with $N = 11$ donors (one donor's NAT sample had been consumed), to confirm that conversion to a FRET design would maintain tumor vs NAT selectivity (Figure 37.3). As shown, positive tumor vs NAT signal was obtained for 10/11 samples with ALTC-04, and 9/11 samples with ALTC−03.

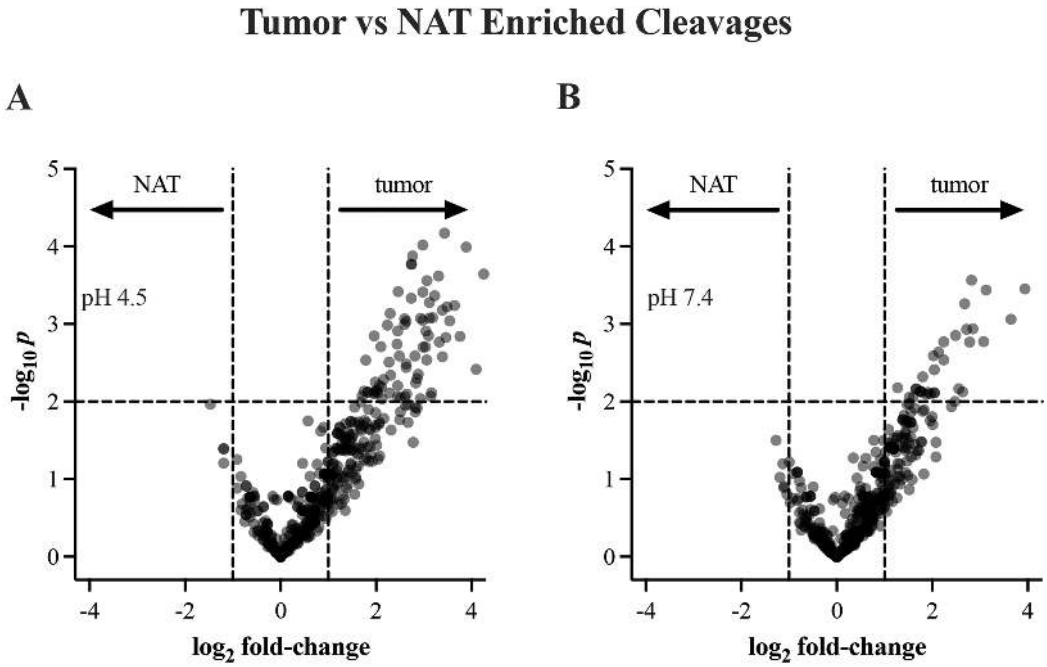


FIGURE 37.2 Differential cleavages between NAT and tumor samples are represented in a volcano plot for (A) pH 4.5 and (B) pH 7.4 conditions, identified through MSP-MS analysis of matched tumor and NAT tissue from $N = 12$ donors. Gridlines show significance cutoff for $p = 0.01$ and fold change = 2.

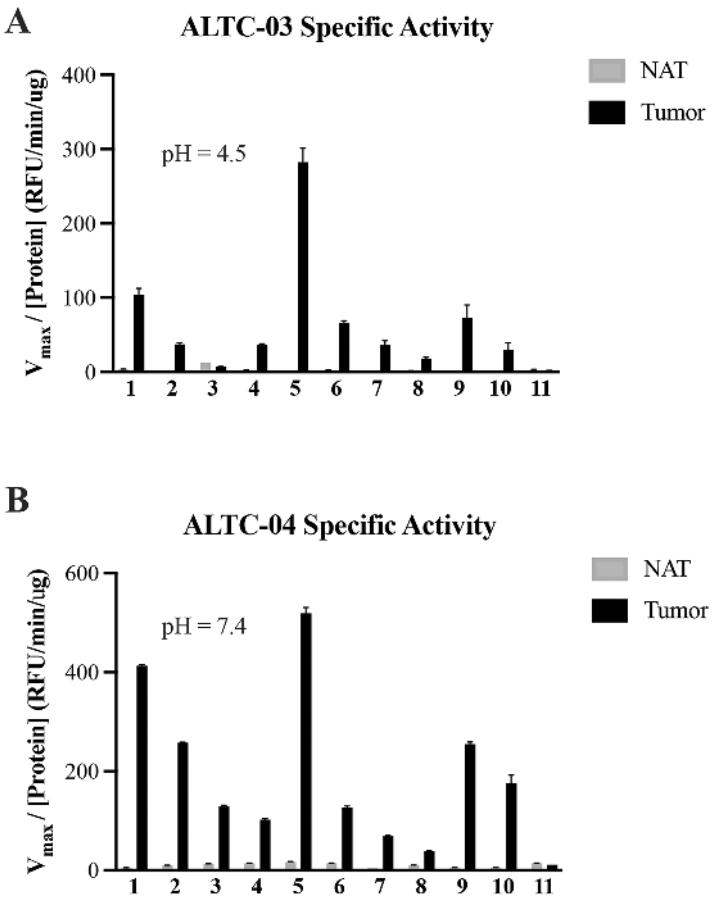


FIGURE 37.3 Specific activity (V_{max}/[protein]) measured for $N = 11$ discovery cohort NAT and tumor samples for substrates (A) ALTC-03 measured at pH 4.5 and (B) ALTC-04 measured pH 7.4. Protein extract samples from conditioned media were prepared at 1:50 dilution. Mean and standard deviation are shown for triplicate $N = 3$ measurements.

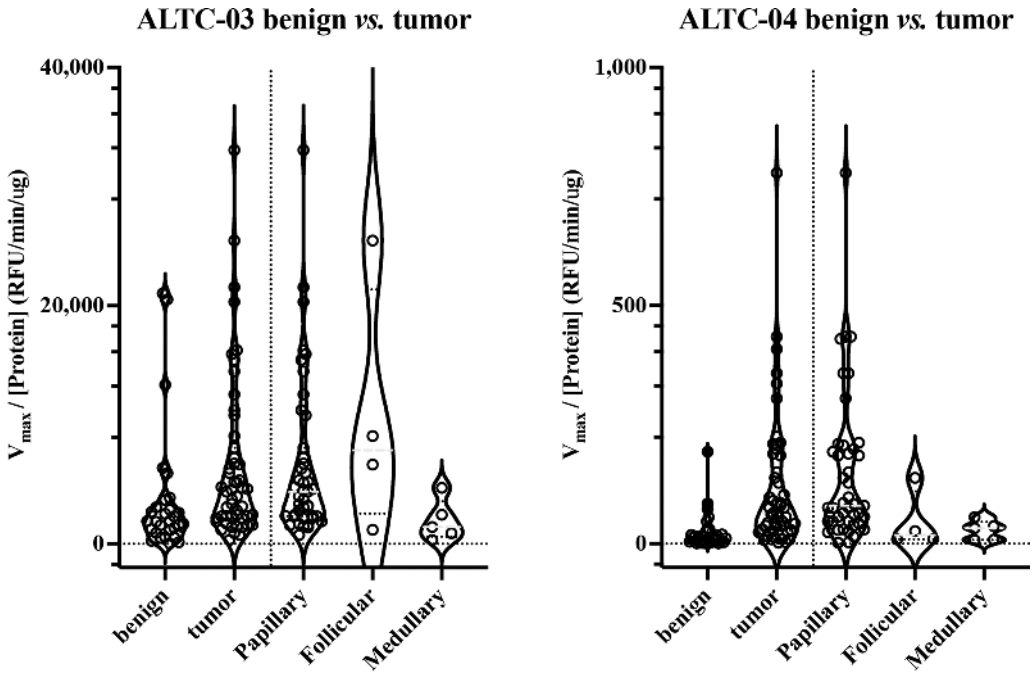


FIGURE 37.4 Violin plots for the distribution of specific activity (V_{max}/[protein]) measured for benign and tumor samples, broken out by papillary, follicular and medullary tumor subgroups, from $N = 80$ donors in a retrospective cohort for substrates (A) ALTC-03 measured at pH 4.5 and (B) ALTC-04 measured pH 7.4. Protein extract samples prepared as mock FNA were assayed at 1:1953 for ALTC-03 and 1:50 for ALTC-04 dilution.

To benchmark activities for substrates ALTC-03 and ALTC-04 across a broad range of tumor and benign thyroid tissue samples, a total of $N = 80$ donor samples, including $N = 12$ tumor samples from the discovery cohort, were all prepared using a mock FNA protocol and assayed. The specific activity score distributions for ALTC-03 and ALTC-04 are shown in Figure 37.4. Post-analysis of the distributions by subgroups revealed that the highest signal benign (likely false positive) samples had adenoma, hyperplasia, or hyperplastic nodule diagnoses. Among the tumor subtypes, positive signal was observed across follicular and medullary tumors (representing 5%–15% and 2%–3% of cases),[21] in addition to papillary tumors used to discover these substrates. Preliminarily, ALTC-04 is projected to have a better specificity for benign vs tumor, and ALTC-03 had better sensitivity in a small number of follicular and medullary tumors, while a scoring algorithm is under development.

37.5 Discussion

This chapter provides an overview of a proteomics-based approach for the discovery of candidate cancer biomarkers using two key data inputs: the identity of proteases and peptidases combined with the proteolytic activity profiles, measured in conditioned media extracts from "normal" vs tumor patient tissues. This basic approach was previously applied to the discovery and development of proteolytic activity reporter assays for detection of premalignant pancreatic cysts.[50, 55, 56] In the case of thyroid nodule diagnosis, a discovery cohort of 12 papillary thyroid cancer samples were analyzed for the enrichment of proteases and proteolytic cleavages in tumor vs normal adjacent tissue samples. This approach has the advantage that discovered substrates could be directly leveraged for the development of fluorescence reporter assays. This project has yielded promising initial proteolytic activity assays using the substrates ALTC-03 and ALTC-04, for which detailed chemistry and analytical parameters will be reported elsewhere. These substrates demonstrated differential cleavage between tumor and NAT samples from their own discovery cohort, as expected, but the true test of their broad applicability was demonstrated in screening against a cohort of $N = 80$ patient-derived samples prepared in a mock-FNA protocol. Initial performance data showed that ALTC-04 as a lead test may differentiate between benign and tumor samples, and ALTC-03 as a secondary test showed promising increased sensitivity for the detection of tumors, including from the follicular and medullary subtypes, reducing likely false negatives overall. A scoring algorithm is now under development for these substrate assays, and the next stage of the project will focus on performance in surgical tissue samples resected from patients with previously indeterminate FNAB cytology.

37.6 Future Trends

Proteases and peptidases play crucial roles in the molecular mechanisms underlying cancer and immunity,[10–14] making their activity signatures valuable for detecting molecular events in both diseased and healthy tissues. The catalytic activity of a protease can serve as a signal, amplified and detected in microliter-sized samples of patient biofluids or tissues. With calibration, proteolytic markers are proving to be sensitive

enough for use in solid and liquid biopsy samples, and a future can be easily envisioned where these cleavages can be coupled with highly sensitive spectroscopic or colorimetric detection for rapid testing at the point of care, including in surgical settings. Similar markers correlating with dysplasia or inflammation could provide early warning signs to guide patient treatment. Our immediate goal is to establish the performance of proteolytic activity signatures, laying the groundwork for evidence-based tests that support oncologists, surgeons, and pathologists in their practices, and to enable a future of rapid -omics development.

Acknowledgments

We would like to thank Francesco Caiazza for his contributions to the early stages of this project, and Prof. Kimberly S. Kirkwood at UCSF as our clinical advisor. This research was funded through NCI R43/R44 CA254649. Mass spectrometry was performed at the UCSF Mass Spectrometry Facility supported by NIH P41GM103481 and 1S10OD016229.

REFERENCES

1. López-Otín, C. & Overall, C. M. Protease degradomics: A new challenge for proteomics. *Nat. Rev. Mol. Cell Biol.* 3, 509–519 (2002).
2. Quesada, V., Ordóñez, G. R., Sánchez, L. M., Puente, X. S. & López-Otín, C. The degradome database: Mammalian proteases and diseases of proteolysis. *Nucleic Acids Res.* 37, D239–D243 (2009).
3. Pagel, O., Loroch, S., Sickmann, A. & Zahedi, R. P. Current strategies and findings in clinically relevant post-translational modification-specific proteomics. *Expert Rev. Proteomics*. 12, 235–253 (2015).
4. Ibrahim, J., Peeters, M., Van Camp, G. & Op de Beeck, K. Methylation biomarkers for early cancer detection and diagnosis: Current and future perspectives. *Eur. J. Cancer Oxf. Engl. 1990*. 178, 91–113 (2023).
5. Rothbart, S. B. & Strahl, B. D. Interpreting the language of histone and DNA modifications. *Biochim. Biophys. Acta*. 1839, 627–643 (2014).
6. Wang, W. *et al*. Cancer metabolites: Promising biomarkers for cancer liquid biopsy. *Biomark. Res.* 11, 66 (2023).
7. Suri, G. S., Kaur, G., Carbone, G. M. & Shinde, D. Metabolomics in oncology. *Cancer Rep. Hoboken NJ.* 6, e1795 (2023).
8. Julien, O. & Wells, J. A. Caspases and their substrates. *Cell Death Differ.* 24, 1380–1389 (2017).
9. Weisel, J. W. & Litvinov, R. I. Fibrin formation, structure and properties. *Subcell. Biochem.* 82, 405–456 (2017).
10. Hanahan, D. Hallmarks of cancer: New dimensions. *Cancer Discov.* 12, 31–46 (2022).
11. Mellman, I., Chen, D. S., Powles, T. & Turley, S. J. The cancer-immunity cycle: Indication, genotype, and immunotype. *Immunity.* 56, 2188–2205 (2023).
12. Vizovisek, M., Ristanovic, D., Menghini, S., Christiansen, M. G. & Schuerle, S. The tumor proteolytic landscape: A challenging frontier in cancer diagnosis and therapy. *Int. J. Mol. Sci.* 22, 2514 (2021).
13. Wyganowska-Świątkowska, M., Tarnowski, M., Murtagh, D., Skrzypczak-Jankun, E. & Jankun, J. Proteolysis is the most fundamental property of malignancy and its inhibition may be used therapeutically (review). *Int. J. Mol. Med.* 43, 15–25 (2019).
14. Habič, A., Novak, M., Majc, B., Lah Turnšek, T. & Breznik, B. Proteases regulate cancer stem cell properties and remodel their microenvironment. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 69, 775–794 (2021).
15. Adam, S. S., Key, N. S. & Greenberg, C. S. D-dimer antigen: Current concepts and future prospects. *Blood*. 113, 2878–2887 (2009).
16. Jakobsen, N. A., Hamdy, F. C. & Bryant, R. J. Novel biomarkers for the detection of prostate cancer. *J. Clin. Urol.* 9, 3–10 (2016).
17. Das, S. & Batra, S. K. Understanding the unique attributes of MUC16 (CA125): Potential implications in targeted therapy. *Cancer Res.* 75, 4669–4674 (2015).
18. Nath, S. & Mukherjee, P. MUC1: A multifaceted oncoprotein with a key role in cancer progression. *Trends Mol. Med.* 20, 332–342 (2014).
19. LeBleu, V. S. *et al*. Identification of human epididymis protein-4 as a fibroblast-derived mediator of fibrosis. *Nat. Med.* 19, 227–231 (2013).
20. Favaloro, E. J. & Pasalic, L. An overview of laboratory testing for ADAMTS13. *Methods Mol. Biol. Clifton NJ.* 2663, 481–486 (2023).
21. Surveillance Research Program, National Cancer Institute. SEER*Explorer Application. *SEER*Explorer: An interactive website for SEER cancer statistics [Internet].* https://seer.cancer.gov/statistics-network/explorer/.
22. Powers, A. E., Marcadis, A. R., Lee, M., Morris, L. G. T. & Marti, J. L. Changes in trends in thyroid cancer incidence in the United States, 1992 to 2016. *JAMA*. 322, 2440–2441 (2019).
23. Guth, S., Theune, U., Aberle, J., Galach, A. & Bamberger, C. M. Very high prevalence of thyroid nodules detected by high frequency (13 MHz) ultrasound examination. *Eur. J. Clin. Invest*. 39, 699–706 (2009).
24. Wilhelm, S. Evaluation of thyroid incidentaloma. *Surg. Clin. North Am.* 94, 485–497 (2014).
25. Fisher, S. B. & Perrier, N. D. The incidental thyroid nodule. *CA. Cancer J. Clin.* 68, 97–105 (2018).
26. Grant, E. G. *et al*. Thyroid ultrasound reporting lexicon: White paper of the ACR thyroid imaging, reporting and data system (TIRADS) committee. *J. Am. Coll. Radiol.* 12, 1272–1279 (2015).
27. Sosa, J. A., Hanna, J. W., Robinson, K. A. & Lanman, R. B. Increases in thyroid nodule fine-needle aspirations, operations, and diagnoses of thyroid cancer in the United States. *Surgery.* 154, 1420–1427 (2013).
28. Zevallos, J. P., Hartman, C. M., Kramer, J. R., Sturgis, E. M. & Chiao, E. Y. Increased thyroid cancer incidence corresponds to increased use of thyroid ultrasound and fine-needle aspiration: A study of the veterans affairs health care system. *Cancer.* 121, 741–746 (2015).
29. Durante, C. *et al*. The diagnosis and management of thyroid nodules: A review. *JAMA*. 319, 914–924 (2018).
30. Manning, A. M., Yang, H. & Steward, D. L. Thyroid ultrasound-guided fine-needle aspiration cytology results: Observed increase in indeterminate rate over the past decade. *Otolaryngol. Neck Surg.* 156(4), 611–615 (2017).

31. Hoang, J. K. *et al*. Managing incidental thyroid nodules detected on imaging: White paper of the ACR incidental thyroid findings committee. *J. Am. Coll. Radiol.* 12, 143–150 (2015).
32. Davies, L. & Welch, H. G. Current thyroid cancer trends in the United States. *JAMA Otolaryngol. Neck Surg.* 140, 317–322 (2014).
33. Brito, J. P. *et al*. Prevalence of thyroid cancer in multinodular goiter versus single nodule: A systematic review and meta-analysis. *Thyroid*. 23, 449–455 (2013).
34. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA. Cancer J. Clin*. 70, 7–30 (2020).
35. Welch, H. G. & Doherty, G. M. Saving thyroids—overtreatment of small papillary cancers. *N. Engl. J. Med*. 379, 310–312 (2018).
36. Tufano, R. P., Noureldine, S. I. & Angelos, P. Incidental thyroid nodules and thyroid cancer: Considerations before determining management. *JAMA Otolaryngol. Neck Surg*. 141, 566–572 (2015).
37. HCUP-US Home Page. https://hcup-us.ahrq.gov/.
38. Finn, C. B. *et al*. The use and costs of same-day surgery versus overnight admission for total thyroidectomy: A multistate, all-payer analysis. *Surgery.* 175, 207–214 (2024).
39. Haugen, B. R. *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*. 26, 1–133 (2016).
40. Ito, Y., Miyauchi, A. & Oda, H. Low-risk papillary microcarcinoma of the thyroid: A review of active surveillance trials. *Eur. J. Surg. Oncol*. 44, 307–315 (2018).
41. Tuttle, R. M. *et al*. Natural history and tumor volume kinetics of papillary thyroid cancers during active surveillance. *JAMA Otolaryngol. Neck Surg*. 143, 1015–1020 (2017).
42. Vargas-Salas, S. *et al*. Genetic testing for indeterminate thyroid cytology: Review and meta-analysis. *Endocr. Relat. Cancer*. 25, R163–R177 (2018).
43. Silaghi, C. A. *et al*. Thyroseq v3, afirma GSC, and microRNA panels versus previous molecular tests in the preoperative diagnosis of indeterminate thyroid nodules: A systematic review and meta-analysis. *Front. Endocrinol*. 12, (2021).
44. Gawin, M. *et al*. Proteome profiles of different types of thyroid cancers. *Mol. Cell. Endocrinol*. 472, 68–79 (2018).
45. Brown, L. M. *et al*. Quantitative and qualitative differences in protein expression between papillary thyroid carcinoma and normal thyroid tissue. *Mol. Carcinog*. 45, 613–626 (2006).
46. Jarząb, B. *et al*. Gene expression profile of papillary thyroid cancer: Sources of variability and diagnostic implications. *Cancer Res*. 65, 1587–1597 (2005).
47. Liu, C.-L. *et al*. Expression of serine peptidase inhibitor Kunitz type 1 in differentiated thyroid cancer. *Histochem. Cell Biol*. 149, 635–644 (2018).
48. Zarkesh, M. *et al*. The role of matrix metalloproteinase-9 as a prognostic biomarker in papillary thyroid cancer. *BMC Cancer.* 18, 1199 (2018).
49. Zhang, Y. *et al*. High-accuracy detection of preoperative thyroid nodules using combination of BRAF mutation and TMPRSS4 mRNA level. *Arch. Med. Res*. 49, 365–372 (2018).
50. Ivry, S. L. *et al*. Global protease activity profiling provides differential diagnosis of pancreatic cysts. *Clin. Cancer Res*. 23, 4865–4874 (2017).
51. Sampognaro, P. J. *et al*. Mutations in α-synuclein, TDP-43 and tau prolong protein half-life through diminished degradation by lysosomal proteases. *Mol. Neurodegener.* 18, 29 (2023).
52. Choi, H., Fermin, D. & Nesvizhskii, A. I. Significance analysis of spectral count data in label-free shotgun proteomics. *Mol. Cell. Proteomics MCP.* 7, 2373–2385 (2008).
53. Tyanova, S. *et al*. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods*. 13, 731–740 (2016).
54. O'Donoghue, A. J. *et al*. Global identification of peptidase specificity by multiplex substrate profiling. *Nat. Methods*. 9, 1095–1100 (2012).
55. Caiazza, F. *et al*. Accurate identification of mucinous pancreatic cystic lesions using small-volume analytes. *J. Surg. Res*. 284, 322–331 (2023).
56. Suresh, V. *et al*. Translation of a protease turnover assay for clinical discrimination of mucinous pancreatic cysts. *Diagn. Basel Switz*. 12, 1343 (2022).

38

Carbohydrate Microarrays Identify the Stage-Specific Embryonic Antigen (SSEA)-0 as a Novel Oncofetal Cancer Marker

Denong Wang

38.1 Introduction

Carbohydrates, like nucleic acids and proteins, are essential biological molecules. Owing to their intrinsic physicochemical properties, carbohydrates can generate structural diversity in a multitude of ways and are prominently displayed on the surfaces of cell membranes or on the exposed regions of macromolecules. Carbohydrate moieties are critical for host immune recognition of invading microbes or malignantly transformed tumor cells and induction of glycan-specific antibodies with protective potential. Recognition of abnormal glycosylation in virtually any cancer type has raised a great interest in the exploration of the tumor glycome for biomarker discovery [1–3].

Recently, our team made a focused effort to identify potential immunological targets of breast cancer (BCA) [4–8]. BCA is among the most prevalent cancers and accounts for the highest number of cancer-related deaths among women worldwide [9]. Since tumor metastasis is the main cause of BCA-associated death [10, 11], there is a pressing need to identify tumor-specific cell-surface biomarkers for early detection of metastatic BCA (MBCA) and to develop targeted immunotherapies to eradicate BCA metastasis.

We explored carbohydrate microarrays as a broad-spectrum immunoassay to probe the tumor biomarkers that are immunologically potent and are likely valuable in tumor detection and targeted immunotherapy [4, 7, 8, 12, 13]. We reasoned that the immune systems of many animal species can distinguish subtle changes in carbohydrate moieties displayed by tumor cells and have potential to mount specific antibody responses to the abnormally expressed tumor glycans that exhibit the "non-self" immunological characteristics. Screening anti-tumor antibodies against a large panel of carbohydrate antigens in microarrays is a practical approach to identify the anti-tumor antibodies that recognize tumor glycan targets. As summarized in the following, our effort has led to the identification of blood group precursor SSEA-0 antigen as an important immunological target of MBCA cells.

38.2 Materials

38.2.1 Anti-Tumor Antibodies

Native tumor glycoprotein antigens may express glycan-based antigenic determinants in addition to protein epitopes. Monoclonal antibodies (mAbs) elicited by native tumor antigens may recognize either glycan or protein targets and are highly valuable for tumor immunological target discovery [7, 12–15]. In this study, mAb HAE3 served as a key reagent for probing BCA-glycan markers. This mAb was produced by mouse immunization with a tumor antigen asialo-epiglycanin (EPGN), which was derived from the major sialomucin glycoprotein of murine mammary adenocarcinoma TA3 cells [14]. Interestingly, HAE3 also reacts with several human epithelial tumors in tissues [16–18]. This cross-species tumor binding profile prompted us to predict that HAE3 may recognize an unexplored tumor glycan marker that is co-expressed by both mouse- and human-derived epithelial cancers.

38.2.2 Carbohydrate Antigens

To explore the natural ligands of HAE3, we applied a collection of purified natural carbohydrate antigens for antibody screening. Key agents include blood group substance reference reagents [19], such as Cyst 9 and Cyst 14, A active; Beach phenol insoluble, B active; Hog, H active; JS phenol insoluble, H and Le^b active, and N-1 20% from the second 10%, Le^a active, and a panel of blood group precursor antigens—OG, Tij II, Beach P1, and McDon P1. A preparation of HAE3-reactive human carcinoma-associated antigen (HCA) served as a positive control for microarray screening [16].

38.3 Methods

38.3.1 Carbohydrate Microarrays

We used carbohydrate microarrays as a multiplex immunoassay to facilitate tumor glycan marker discovery. Carbohydrate microarray assays were performed as described [20]. In brief, a microarray robot (PIXSYS 5500C, Cartesian Technologies, Irvine, CA) was used to spot antigen preparations onto glass slides pre-coated with nitrocellulose polymer (FAST Slides; Schleicher & Schuell, Keene, NH). The printed microarrays were incubated at room temperature with HAE3 (IgM) antibody at 5 μg/ml in 1% (wt/vol) BSA in PBS containing 0.05% (wt/vol) NaN_3 and 0.05% (vol/vol) Tween 20. An R-phycoerythrin (R-PE)-conjugated affinity-purified F(ab') fragment of goat anti-mouse IgM secondary antibody preparation (Rockland Immunochemicals, Inc., PA) was applied at 2.0 μg/ml to reveal HAE3-specific staining signal. Fluorescence

DOI: 10.1201/9781003449942-43

intensity values for each micro-spot and its background were calculated using ScanArray Express software. SAS Institute's JMP software package (Cary, NC) was used for microarray data standardization and statistical analysis.

38.3.2 Carbohydrate-Specific Enzyme-Linked Immunosorbent Assay (ELISA)

ELISA and ELISA inhibition assays were performed as described [20, 21]. Briefly, glycoprotein antigen preparations were diluted in 0.1 M sodium bicarbonate buffer solution, pH 9.6, for coating on ELISA microplates (NUNC, MaxiSorp, Thermo Scientific, Rochester, NY) followed by blocking using 1% BSA, 1X phosphate buffered saline Tween-20 (PBST). MAb HAE3 (IgM) (2.5 μg/ml) and biotinylated peanut agglutinin (PNA) (2.0 μg/ml) were diluted in 1% BSA, PBST for the ELISA binding assay. The bound HAE3 and PNA were revealed by an alkaline phosphatase (AP)-conjugated goat anti-mouse IgM and an AP-streptavidin-conjugate (Sigma Chemical Co., St. Louis, MO), respectively. ELISA inhibition assays were performed with EPGN (1μU/ml) coated to display the native HAE3 epitope and a series of blood group reference antigens as potential inhibitors (25.0 μg/ml/each) to compete with the HAE3 (1.0 μg/ml) binding of EPGN. Percent inhibition was calculated as follows: % inhibition = ((standard A – blank A) – (A with inhibitor – A))/(standard A – blank A) x100.

38.4 Results

38.4.1 Identification of Blood Group Precursor SSEA-0 Tumor Biomarker

We scanned HAE3 using carbohydrate microarrays to explore potential tumor glycan markers [7]. In Figure 38.1, the HAE3-binding profiles are graphically presented as overlay plots with the mean values of triplicate detections of each antigen on the Y-axis and corresponding antigen preparations on the X-axis. Visual inspection of this figure readily reveals that HAE3 is strongly positive with the HCA positive control agent (1# and 2#) and with four well-characterized blood group precursor reference antigens, including Beach P1 (29#), McDon P1 (30#), Tij II (31#), and OG (32#), By contrast, HAE3 has minimal or no cross-reactivity with blood group substances A, B, O, or Lewis antigens, or the large panel of other carbohydrate antigens spotted in the same array.

Of note, OG [22] and Tij II [23] antigens were prepared by pepsin digestion, ethanol precipitation, and solubilization in 90% phenol, followed by fractional ethanol precipitation from phenol. Beach P1 [24] and McDon P1 [25] were obtained as the non-dialyzable O-cores from partially hydrolyzed blood group antigen, Beach B and McDon A1, respectively. Thus, these blood group precursor substances were prepared to eliminate peripheral glyco-epitopes, such as A, B, H, or Lewis antigen-specific epitopes, but preserve their O-glycan core structures. Selective detection of these blood group precursors from a large panel of blood group substances by HAE3 demonstrated that this antibody is specific for a shared cryptic O-core of these blood group precursor substances, which is termed as SSEA-0.

38.4.2 Target Validation Using Conventional Immunoassays

We further examined whether the observed HAE3 binding reactivity can be attributed to cross-reactivity with T/Tn glyco-epitopes that are often expressed as components of blood group precursor substances [7]. For this purpose, we tested ELISA binding of HAE3 or T-specific lectin PNA with blood group precursor Tij II (31#) and two T/Tn-positive glyco-conjugates, asialo-PSM (T) (81#) and asialo-OSM (Tn)(82#). Unlike PNA, which binds to both Tij II and asialo-PSM (T), HAE3

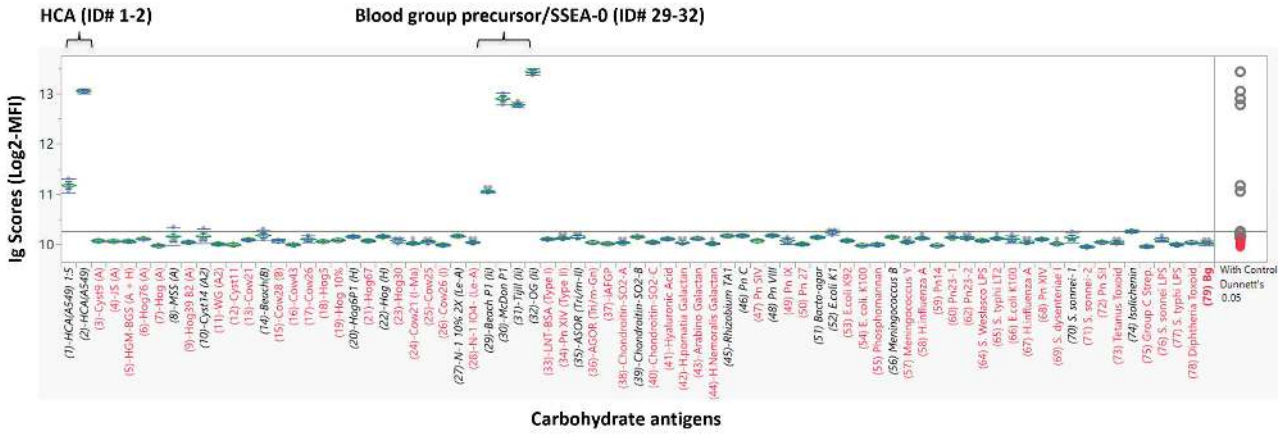


FIGURE 38.1 Carbohydrate microarray analysis of anti-tumor mAb HAE3. A collection of native carbohydrate antigens was spotted in triplicates in one to two dilutions to yield the customized microarrays for antibody screening. Levels of antigen-specific antibody signal captured by carbohydrate microarrays are presented as the standardized mean Ig-scores—the Log2-transformed mean fluorescent intensity (MFI) of triplicate detections. The comparison circles (right) are results of a Dunnett's test; the more the circles intersect, the less significant the difference, and vice versa. The antigen names on the X axis in *black* indicate the means of detection is significantly different from background microspots (Bg); *red* refers to non-significant detection as compared to the Bg.

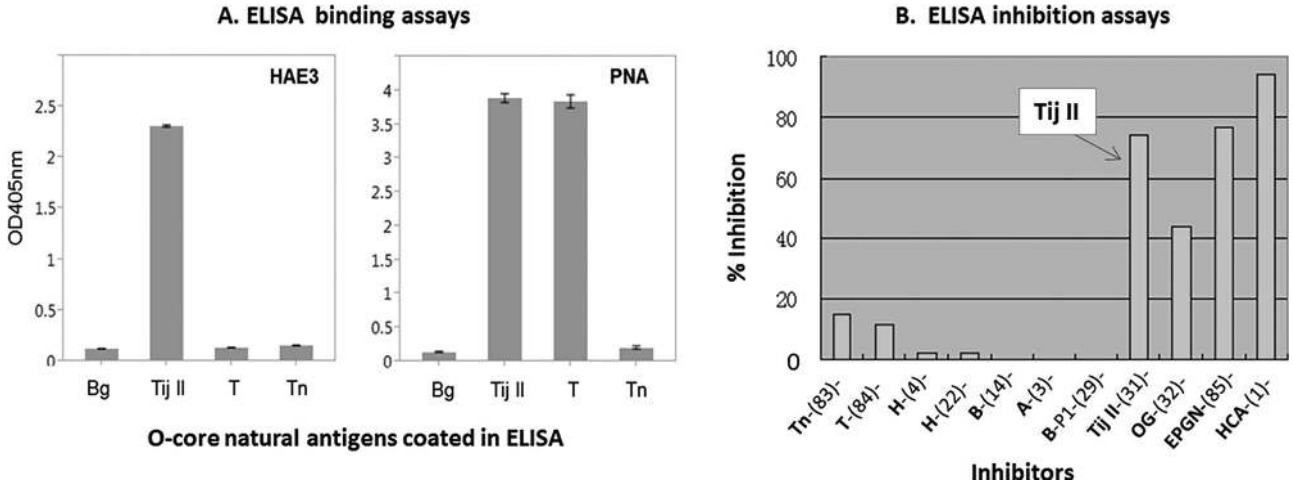


FIGURE 38.2 Carbohydrate-specific ELISA validates binding specificities of HAE3. (A) An antigen-specific ELISA distinguished HAE3-binding specificity from the T-antigen-specific PNA. ELISA plates were coated with O-core antigen, Tij (31#), T (81#), and Tn (82#) at 10 μg/ml to react with either HAE3 (5.0 μg/ml) or PNA (1.0 μg/ml). (B) ELISA inhibition assays with a series of carbohydrate antigens as competitors (25.0 μg/ml) to inhibit interaction between asialo-EPGN (1.0 μU/ml) and anti-HCA (HAE3, 1.0 μg/ml). These antigens include HCA (1#), EPGN (85#), which is the immunogen of HAE3, and blood group substances Tn (83#), T (84#), H (4# or 22#), A (3#), B (14#), Beach P1 (29#), Tij II (31#), and OG (32#). Results are shown as percent inhibition in the presence of an inhibitor. (Adapted from [7].)

specifically binds to Tij II without cross-reactivity with asialo-PSM (T) or asialo-OSM (Tn) (Figure 38.2A). Thus, HAE3-binding of Tij II is irrelevant to the native T/Tn-glyco-epitopes expressed by these glyco-conjugates. ELISA inhibition assays further demonstrated that blood group precursors Tij II (31#) and OG (32#), but not other blood group antigens, including A (3#), B (14#), H (22#), T (84#), or Tn (83#), significantly inhibited HAE3 binding to the immunogen asialo-EPGN (85#) (Figure 38.2B).

38.5 Discussion

Our carbohydrate microarray analysis identified blood group precursor substances as the natural ligands of antibody HAE3 [7]. Figure 38.3 is a schematic of blood group precursor core structure SSEA-0, which was proposed based on extensive glycan structural analyses and immunochemical studies of blood group substances [22, 26–28]. The four-branched structure in the circle represents the internal portion of the carbohydrate moiety of blood group substances. Fucosylation of these internal domains or cores blocks surface exposure of the SSEA-0/gp^{C1} cryptic epitope (s) but results in mature blood group antigens, such as A, B, O, and the Lewis (Le) series of antigens.

It is noteworthy that several SSEA antigens in the histo-blood group system have been found to be overexpressed in BCA. These include blood group I antigen [29–31], SSEA-1/Le^X [32–34], and SSEA-5/H type-I [35]. The SSEA-1/Le^X and SSEA-5/H-type I are fucosylated "matured" blood group antigens derived from the blood group precursor core SSEA-0 (Figure 38.3). The I-antigen is known as a blood group precursor polylactosamine chain with branches. However, these blood group antigens are HAE3-negative [4, 7, 13]. Thus, the HAE3-positive SSEA-0 is a previously unexplored oncofetal antigen in the histo-blood group substance repertoire.

We examined tumor cell expression of SSEA-0 using additional experimental approaches. In a flow cytometry-based cell-surface staining, we characterized SSEA-0 in a panel of BCA lines, including some triple-negative cancers (TNBCs) that lack the estrogen, progesterone, and HER2/neu receptors but have breast cancer stem cell (bCSC) features [7]. We found that BT-549 and MDA-MB-468 were strongly HAE3-positive. MDA-MB-231 was weakly positive. Hs 578T cells were, however, generally negative except for a small set of weakly positive cells. We further introduced the fiber optic array scanning technology (FAST-scan) to explore cell-surface glycan markers of breast circulating tumor cells (bCTCs) [4, 8]. These cancer cells are rare in blood but may contain CSCs to spur metastasis [36, 37]. In our clinical case study [4], the $HAE3^+$-CTCs were detected in all subjects with ~40% of bCTCs strongly SSEA-0 positive. Notably, the CTCs from a patient with TNBC and multiple sites of metastasis were predominantly SSEA-0 positive (92.5%, 37/40 CTCs). Thus, SSEA-0 is overexpressed by murine and human-derived BCAs with highly metastatic potential.

38.6 Future Trends

The potential of SSEA-0 marker and targeting antibodies for the early detection of MBCA and targeted immunotherapy against tumor metastasis remain largely unexplored. It is noteworthy the HAE3 positive antigens were also overexpressed by other human epithelial tumors, including lung, prostate, bladder, esophagus, and ovarian cancers [16–18]. Thus, the SSEA-0-targeting therapeutic strategies may be effective therapeutics for additional epithelial cancers.

Glyco-conjugates expressing SSEA-0 may be explored as tumor vaccines. The blood group precursor substances are native glycoconjugates and are often more complex in

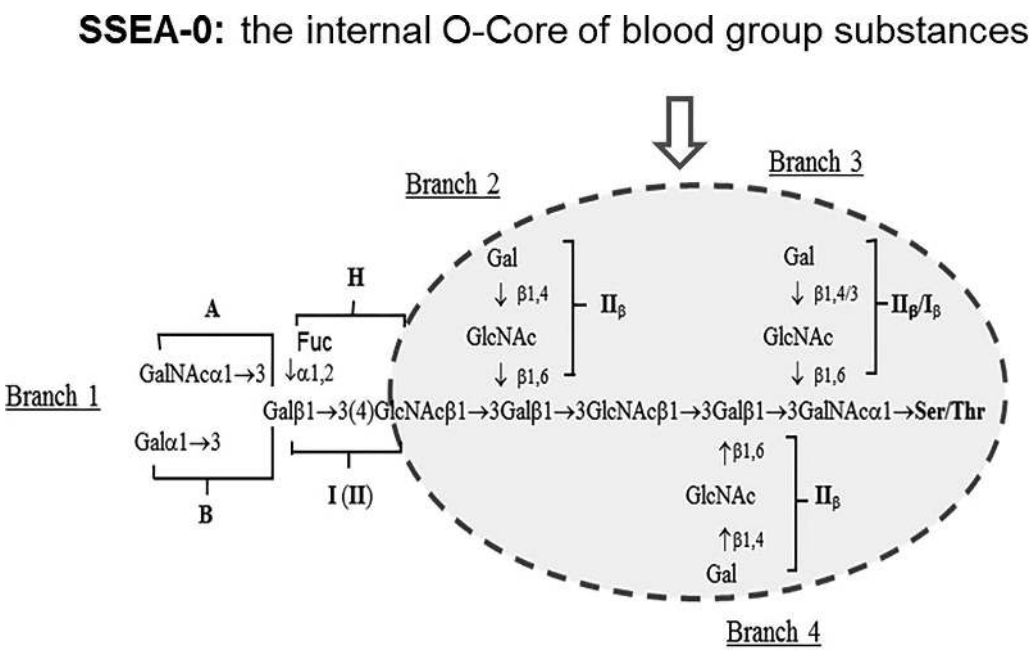


FIGURE 38.3 A schematic of SSEA-0/gp^{C1} blood group precursor antigens. Blood group precursors are characteristically composed of oligosaccharide chains with branches of IIβ (Galβ1,4GlcNAcβ1,6)- and/or Iβ (Galβ1,3GlcNAcβ1,6)-moieties without fucosylation. The four-branched structure in the circle represents the internal portion of the carbohydrate moiety of blood group substances. The terminal nonreducing β-galactoside epitopes of IIβ and/or Iβ are crucial for preserving the potent SSEA-0/gp^{C1} determinant(s). The key glycogenes involved in the biosynthesis of blood group precursor include β-1,6-N-acetylglucosaminyltransferases (GCNT1, 2, and 3) that catalyze β-1,6-branching of the O-cores and the polylactosamine extension enzyme (B3GNT2), commonly required for blood group I synthesis. (Adapted from [7].)

carbohydrate structures than the most known model O-cores [38]. As determined by Bio-Gel P-10 filtration, the sizes of the Tij II blood group precursor substance were distributed in a range of molecular weights from 7 KDa to 15 KDa, approximately [7]. Since this antigen contains only approximately 2% peptide sequences, its mass is apparently made up predominately of carbohydrates [23, 39, 40].

Chemical synthesis of such complex blood group precursors is expected to be challenging. However, it is not impossible to rationally design and produce HAE3-positive compounds by stabilizing relatively simple O-core moieties via specific structural modifications [41]. A recent glycan array screening revealed an unpredicted binding of HAE3 to a sulfated glycolipid SM1a, Galβ1–3GalNAcβ1–4(3-O-sulfate)Galβ1–4GlcCer [13], which may be viewed as an O-core-derivative. Recent development in bioconjugation offers a promising technology for rapid development of efficacious vaccines against emerging bacterial threats [42, 43], which may be further explored to advance the glycoconjugate-based tumor vaccines.

Acknowledgments

The authors acknowledge the Kabat Collection of Carbohydrate Antigens at SRI International for a panel of carbohydrate antigens that were applied in this study. This work was supported in part by US Government grants U01CA128416 (NIH), R21CA254048 (NIH), and PR210339 (CDMRP) to DW. The content is solely the responsibility of the author and does not necessarily represent the official views of the funding agents.

Conflict of Interest

There are no conflicts of interest to declare for any of the authors involved in this work.

REFERENCES

1. Hakomori S: Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens. *Advances in Cancer Research* 1989, 52:257–331.
2. Fukuda M: Possible roles of tumor-associated carbohydrate antigens. *Cancer Research* 1996, 56(10):2237–2244.
3. Dube DH, Bertozzi CR: Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nature Reviews Drug Discovery* 2005, 4(6):477–488.
4. Wang D, Liu X, Hsieh B, Bruce R, Somlo G, Huang J, Sambucetti L: Exploring glycan markers for immunotyping and precision-targeting of breast circulating tumor cells. *Arch Med Res* 2015, 46(8):642–650.
5. Wang D: Unraveling sugar chain signatures of the "seeds" of tumor metastasis. *J Proteomics Bioinform* 2017, 10(1).
6. Jiang Y, Wang D: Liquid biopsy in the OMICS era of tumor medicine. *Open Access J Biomed Eng Appl* 2018, 1(3).
7. Wang D, Tang J, Liu S, Huang J: Carbohydrate microarrays identify blood group precursor cryptic epitopes as potential immunological targets of breast cancer. *Journal of Immunology Research* 2015, 2015:510810.
8. Wang D, Wu L, Liu X: Glycan markers as potential immunological targets in circulating tumor cells. *Advances in Experimental Medicine and Biology* 2017, 994:275–284.

9. Chatterjee SK, Zetter BR: Cancer biomarkers: Knowing the present and predicting the future. *Future Oncol* 2005, 1(1):37–50.
10. Siegel RL, Miller KD, Fuchs HE, Jemal A: Cancer statistics, 2022. *CA Cancer J Clin* 2022, 72(1):7–33.
11. Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, Jemal A, Siegel RL: Breast cancer statistics, 2022. *CA Cancer J Clin* 2022, 72(6):524–541.
12. Newsom-Davis TE, Wang D, Steinman L, Chen PF, Wang LX, Simon AK, Screaton GR: Enhanced immune recognition of cryptic glycan markers in human tumors. *Cancer Research* 2009, 69(5):2018–2025.
13. Palma AS, Liu Y, Childs RA, Herbert C, Wang D, Chai W, Feizi T: The human epithelial carcinoma antigen recognized by monoclonal antibody AE3 is expressed on a sulfoglycolipid in addition to neoplastic mucins. *Biochemical and Biophysical Research Communications* 2011, 408(4): 548–552.
14. Codington JF, Sanford BH, Jeanloz RW: Glycoprotein coat of the TA3 cell. Isolation and partial characterization of a sialic acid containing glycoprotein fraction. *Biochemistry* 1972, 11(14):2559–2564.
15. Gao C, Liu Y, Zhang H, Zhang Y, Fukuda MN, Palma AS, Kozak RP, Childs RA, Nonaka M, Li Z *et al*.: Carbohydrate sequence of the prostate cancer-associated antigen F77 assigned by a mucin O-glycome designer array. *The Journal of Biological Chemistry* 2014, 289(23):16462–16477.
16. Li R, Yao JL, Bourne PA, di Sant'Agnese PA, Huang J: Frequent expression of human carcinoma-associated antigen, a mucin-type glycoprotein, in cells of prostatic carcinoma. *Arch Pathol Lab Med* 2004, 128(12):1412–1417.
17. Liang S, Yao J, Bourne PA, diSant'Agnese PA, Huang J, Lei JY: Overexpression of human carcinoma-associated antigen in esophageal adenocarcinoma and its precursor lesions. *Am J Clin Pathol* 2004, 122(5):747–751.
18. Yao JL, Bourne PA, Yang Q, Lei J, di Sant'Agnese PA, Huang J: Overexpression of human carcinoma-associated antigen in urothelial carcinoma of the bladder. *Arch Pathol Lab Med* 2004, 128(7):785–787.
19. Kabat EA: *Blood group substances, their chemistry and immunochemistry*. New York: Academic Press Inc.; 1956.
20. Wang D, Liu S, Trummer BJ, Deng C, Wang A: Carbohydrate microarrays for the recognition of cross-reactive molecular markers of microbes and host cells. *Nat Biotechnol* 2002, 20(3):275–281.
21. Wang D, Dafik L, Nolley R, Huang W, Wolfinger RD, Wang LX, Peehl DM: Anti-oligomannose antibodies as potential serum biomarkers of aggressive prostate cancer. *Drug Dev Res* 2013, 74:65–80. (Biomarkers in Drug Development and Companion Diagnostics).
22. Vicari G, Kabat EA: Structures and activities of oligosaccharides produced by alkaline degradation of a blood group substance lacking A, B, H, Le b specificities. *Biochemistry* 1970, 9(17):3414–3421.
23. Maisonrouge-McAuliffe F, Kabat EA: Immunochemical studies on blood groups. Fractionation, heterogeneity, and chemical and immunochemical properties of a blood group substance with B, I and I activities purified from human ovarian cyst fluid. *Arch Biochem Biophys* 1976, 175(1):71–80.
24. Allen PZ, Kabat EA: Immunochemical studies on blood groups. XXII. Immunochemical studies on the nondialyzable residue from partially hydrolyzed blood group A, B and O(H) substances (P1 fractions). *J Immunol* 1959, 82(4):340–357.
25. Sikder SK, Kabat EA, Roberts DD, Goldstein IJ: Immunochemical studies on the combining site of the blood group A-specific lima bean lectin. *Carbohydrate Research* 1986, 151:247–260.
26. Feizi T, Kabat EA, Vicari G, Anderson B, Marsh WL: Immunochemical studies on blood groups.XLIX. The I antigen complex: Specificity differences among anti-I sera revealed by quantitative precipitin studies; partial structure of the I determinant specific for one anti-I serum. *J Immunol* 1971, 106(6):1578–1592.
27. Feizi T, Kabat EA, Vicari G, Anderson B, Marsh WL: Immunochemical studies on blood groups. XLVII. The I antigen complex--precursors in the A, B, H, Lea, and leb blood group system--hemagglutination-inhibition studies. *The Journal of Experimental Medicine* 1971, 133(1):39–52.
28. Wu AM, Khoo KH, Yu SY, Yang Z, Kannagi R, Watkins WM: Glycomic mapping of pseudomucinous human ovarian cyst glycoproteins: Identification of Lewis and sialyl Lewis glycotopes. *Proteomics* 2007, 7(20):3699–3717.
29. Burchell J, Wang D, Taylor-Papadimitriou J: Detection of the tumour-associated antigens recognized by the monoclonal antibodies HMFG-1 and 2 in serum from patients with breast cancer. *Int J Cancer* 1984, 34(6):763–768.
30. Dube VE, Haid M, Chmiel JS, Anderson B: Serum cold agglutinin and IgM levels in breast carcinoma. *Breast Cancer Research and Treatment* 1984, 4(2):105–108.
31. Dube VE, Kallio P, Chmiel JS, Haid M, Hakim A: An immunosorbent assay for blood group I antigens in breast carcinoma. *Clin Immunol Immunopathol* 1987, 45(2): 196–207.
32. Koh YW, Lee HJ, Ahn JH, Lee JW, Gong G: Expression of Lewis X is associated with poor prognosis in triple-negative breast cancer. *Am J Clin Pathol* 2013, 139(6):746–753.
33. Renkonen J, Paavonen T, Renkonen R: Endothelial and epithelial expression of sialyl Lewis(x) and sialyl Lewis(a) in lesions of breast carcinoma. *Int J Cancer* 1997, 74(3):296–300.
34. Nakagoe T, Fukushima K, Itoyanagi N, Ikuta Y, Oka T, Nagayasu T, Ayabe H, Hara S, Ishikawa H, Minami H: Expression of ABH/Lewis-related antigens as prognostic factors in patients with breast cancer. *J Cancer Res Clin Oncol* 2002, 128(5):257–264.
35. Tang C, Lee AS, Volkmer JP, Sahoo D, Nag D, Mosley AR, Inlay MA, Ardehali R, Chavez SL, Pera RR *et al*.: An antibody against SSEA-5 glycan on human pluripotent stem cells enables removal of teratoma-forming cells. *Nat Biotechnol* 2011, 29(9):829–834.
36. Jacob K, Sollier C, Jabado N: Circulating tumor cells: Detection, molecular profiling and future prospects. *Expert Review of Proteomics* 2007, 4(6):741–756.
37. Hayashi N, Yamauchi H: Role of circulating tumor cells and disseminated tumor cells in primary breast cancer. *Breast Cancer* 2012, 19(2):110–117.
38. Brockhausen I: Mucin-type O-glycans in human colon and breast cancer: Glycodynamics and functions. *EMBO Reports* 2006, 7(6):599–604.
39. Maisonrouge McAuliffe F, Kabat EA: Immunological studies on blood groups. Structures and immunochemical

properties of oligosaccharides from two fractions of blood group substance from ovarian cyst fluid differeing in B, I, and I activities and reactivity toward concanavalin A. *Arch Biochem Biophys* 1976, 175:90–113.

40. Maisonrouge-McAuliffe F, Kabat EA: Immunochemical studies on blood groups. Heterogeneity of oligosaccharides liberated by degradation with alkaline borohydride of two human ovarian cyst fractions differing in B, I, and i activities and in reactivity toward concanavalin A. *Arch Biochem Biophys* 1976, 175(1):81–89.
41. Hirose H, Tamai H, Gao C, Imamura A, Ando H, Ishida H, Feizi T, Kiso M: Total syntheses of disulphated glycosphingolipid SB1a and the related monosulphated SM1a. *Org Biomol Chem* 2015, 13(45):11105–11117.
42. Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, Aebi M: Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli. *Proc Natl Acad Sci U S A* 2005, 102(8):3016–3021.
43. Feldman MF, Mayer Bridwell AE, Scott NE, Vinogradov E, McKee SR, Chavez SM, Twentyman J, Stallings CL, Rosen DA, Harding CM: A promising bioconjugate vaccine against hypervirulent Klebsiella pneumoniae. *Proc Natl Acad Sci U S A* 2019, 116(37):18655–18663.

39

Collective Attributes of Extracellular Vesicles as Biomarkers for Cancer Detection

Alfred Akinlalu, Tommy Gao, Sky Gao, and Dali Sun

39.1 Introduction

In the fight against cancer, there is an ongoing effort to discover biomarkers that provide accurate and reliable cancer detection.[1] Biomarker-based detection methods are significantly less invasive than traditional diagnostics and can serve as an early screening method for cancer. Most disease-causing cancers exhibit hallmark features including rapid growth rate and metastatic potential, and these functional characteristics make it possible to differentiate cancer cells from healthy cells based on their molecular signature. Since extracellular vesicles (EVs) are released into the circulatory system while still containing many of the same proteins, nucleic acids, and lipids as their parent cells,[2] they are attractive targets for these biomarker-based methods.

Current biomarker-based cancer detection methods face technological and methodological limitations, particularly in differentiating between cancer subtypes, due to the overlapping molecular signatures often present among various cancer subtypes. This may be due to inter- and intra-tumor heterogeneity, where the molecules and cells within a single tumor mass can be highly different from one another.[3] Moreover, the research prospect of validating the sensitivity and specificity of such biomarkers is lengthy and intricate. Consequently, relying solely on individual biomarkers for diagnostics often falls short of the discriminatory power needed for clinical applications. Still, biomarker-based diagnostic methods hold the advantage of being relatively cheap, rapid, and noninvasive.

Extracellular vesicles are an underexplored cancer biomarker source with promising translational capabilities that can work around some challenges of traditional biomarker-based cancer detection methods. Tumor EVs contain nucleic acids among their cargo, highlighting their involvement in the transfer of genetic components and the role they play in cancer progression and metastasis.[4, 5] While various traits of EVs derived from tumor and nontumor cells have been studied, their widespread acceptance as cancer biomarkers has been hindered by inconsistent results across different EV-based assays.[4] Existing methods for measuring EVs have demonstrated inconsistencies in sample preparation, often necessitating the use of dyes or other reagents that may alter the EVs' inherent characteristics.[4] Current approaches to cancer detection using nucleic acids often rely on single biomarkers such as circulating tumor DNA (ctDNA) and microRNA (miRNA).[6] However, these markers are inefficient in disease detection due to their lack of subtype-specificity, high false positive rate, and inconsistency in protocol. In this chapter, we introduce innovative EV-based assays developed through a comprehensive analysis of nucleic acids and protein contents in EVs from tumor and nontumor cells. This analysis led to the development of three cancer detection assays and a device, namely NPr, EVIPThT, EVIPqPCR, and elliptical dichroism (ED) spectrometer.

39.2 Materials

39.2.1 Cell Lines

Non-malignant immortalized primary ductal pancreatic cells (HPNE), and three primary pancreatic ductal adenocarcinoma (PDAC) cell lines with different genotypes (PANC-1, BxPC-3, MIA PaCa-2) were obtained from American Type Culture Collection (ATCC).

39.2.2 Chemicals

Isolation magnetic beads modified with anti-human epithelial cell adhesion molecule (EpCAM) (Invitrogen). Lysis buffer (Sigma-Aldrich). M-PER mammalian protein extraction reagent (Thermo). Bicinchoninic acid assay (Thermo). AllPrep DNA/RNA Mini Kit (Qiagen). TaqMan Universal Master mix II (Thermo).

39.2.3 Equipment

Nanodrop 8000 Spectrophotometer (Thermo). Applied Biosystems Quantstudio (Thermo).

39.3 Methods

Malignant and non-malignant pancreatic adenocarcinoma cell lines were cultured in serum-free media to collect EVs during cell starvation. Non-starvation conditions were achieved through culture in media supplemented with EV-depleted fetal bovine serum.

DOI: 10.1201/9781003449942-44

39.3.1 EV Isolation through Immunoprecipitation (EVIP)

200 µl each of serum samples from healthy individuals, patients with pancreatic ductal adenocarcinoma (PDAC), and patients with pancreatitis were used. 200 µl EpCAM was washed twice with PBS and dissolved in 400 µl PBS. The serum samples and 200 µl dissolved magnetic beads were incubated for 2 h at room temperature with gentle agitation. Followed by PBS wash, the beads were dissolved in 10 µl lysis buffer and incubated for 30 min. The collected supernatant containing IP-isolated EVs was used for further analysis.

39.3.2 EV Protein/DNA/RNA Extraction

30 µl each from 100 ml EV stock sample isolated from cells incubated with M-PER for 30 min on ice were used for lysis. The lysate was quantified with bicinchoninic acid assay. Total DNA and RNA samples were isolated with AllPrep DNA/RNA Mini Kit by dissolving 40 µl of isolated EVs in 200 µl of buffer with 1 min homogenization, then lysate was added to isolation columns while eluting RNA and DNA using 50 µl EB buffer and 30 µl nuclease-free water respectively. DNA/RNA concentrations were quantified using Nanodrop.

39.3.3 EV Nucleic-Acid-to-Protein-Ratio (NPr)

2 µl of EVs from 100 ml EV stock isolated from cells or 2 µl of human serum sample obtained from 600 ml stock were used. The NPr measurements were obtained from UV absorption of the samples at 200 nm and 260 nm using Nanodrop. Five PBS blank samples were measured between each sample measurement to avoid cross contamination and five repeat NPr readings were taken.

39.3.4 ThT-Based Assay of EV Proteins (EVIPThT), qPCR-Based Assay of EV DNA (EVIPqPCR)

For the EVIPThT, IP-isolated EVs were dissolved in 20 µl lysis buffer for 15 min. Then, 35 µl of PBS was added to the supernatant collected. Next, 50 µl of the sample mixture was obtained, and added to 50 µl ThT solution (70 µM). The resulting 100 ml solution was added to a 96-well plate and incubated for 30 min while avoiding exposure to light, after which readings were taken using the plate reader (excitation 450nm / emission 482nm). The final fluorescent intensity was normalized by PBS blank with assigned ThT concentration.

For the EVIPqPCR, following EV isolation using IP process in EVIP method, the lysate was collected, and the DNA level was evaluated by taking the absorption at 280 nm using Nanodrop. 2 µl of the DNA samples were amplified in technical replicates in 26 µl reactions by qPCR using TaqMan Universal Master mix II (10 µl) on Applied Biosystems Quantstudio 6 real-time PCR system. The TaqMan assay targeting mtDNA (MT-ND1 ID: Hs02596873_s1, FAM-MGB) and stabilized against the nDNA gene (GAPDH ID: Hs02786624_g1, VICMGB) to ensure high sensitivity. An endogenous control probe (GAPDH ID: Hs02786624_g1, VIC-MGB_PL) was applied to normalize expression results due to stronger signal expression of MT-ND1 over GAPDH. mtDNA/nDNA was evaluated using equation $2^{(ct_{mt}-ct_n+1)}$, where Ct_{mt} is the cycle threshold for mtDNA and Ct_n is the cycle threshold for nDNA.

39.3.5 EV Elliptical Dichroism (ED) Stereochemical Analysis

200 µl each of samples (tumor and nontumor), amino acids (1 mg/ml in PBS) and protein samples (1 mg/ml in PBS) was loaded into a sample cuvette covered with a lid for the measurements. The system was turned on 30 min before readings for stable readouts. Before each sample measurement, PBS was measured five times and averaged as blank, and the samples were measured five times as well. The absorbance was normalized to the blank using the equation: $A(\theta) = \log\left(\frac{I_{blank(\theta)}}{I_{sameple(\theta)}}\right)$.

In the angle sweep (AS) mode, the baseline was corrected by linear subtraction connecting the starting and ending points. In ED mode, ED value was calculated as $ED = A_l - A_r$, where A_l and A_r represent absorbance at left and right elliptical polarization ($\theta = 45°$ and $135°$), respectively.

39.3.6 Statistics

All statistics were conducted using Origin Pro software. Comparison between two groups was performed using two-tailed student's *t*-test (normally distributed parameters), unpaired two-tailed Mann-Whitney U-test (unpaired samples), a paired two-tailed Mann-Whitney U-test (paired samples). $p < 0.05$ is set to be statistically significant.

39.4 Results

39.4.1 DNA Abundance in Tumor EVs as a Cancer Biomarker

The underlying physical principle behind spectroscopy is that different biomolecules absorb light in a measurably different way due to their unique molecular structure.[7] Proteins exhibit a peak absorbance at a light wavelength of 280 nm due to the aromatic ring structure of their amino acids, while nucleic acids exhibit an absorbance at 260 nm due to their double bonds. It is known that tumor cells have a higher nucleic acid content than healthy cells and since EVs encompass biomolecules reflective of their host cell, this discrepancy can serve as a proxy biomarker.

We evaluated the nucleic acid to protein absorption ratio (NPr) of EVs from tumor and nontumor cells by measuring the ultraviolet (UV) absorbance of the EV samples at 260 nm and 280 nm and taking the ratio of the spectroscopic measurements. This statistic serves as an indirect measurement of the DNA and protein content of the EVs. The relative abundance of nucleic acid content to protein (99:1) is higher in tumor cells compared to their healthy counterparts, with the largest ratio difference in the malignant cells (MIA-PaCa-2, Figure 39.1a and b). Figure 39.1c demonstrates that the nucleic acid signal is primarily from the EVs and not from other sources in the samples. This NPr assay was then used in clinical samples of patients with PDAC and

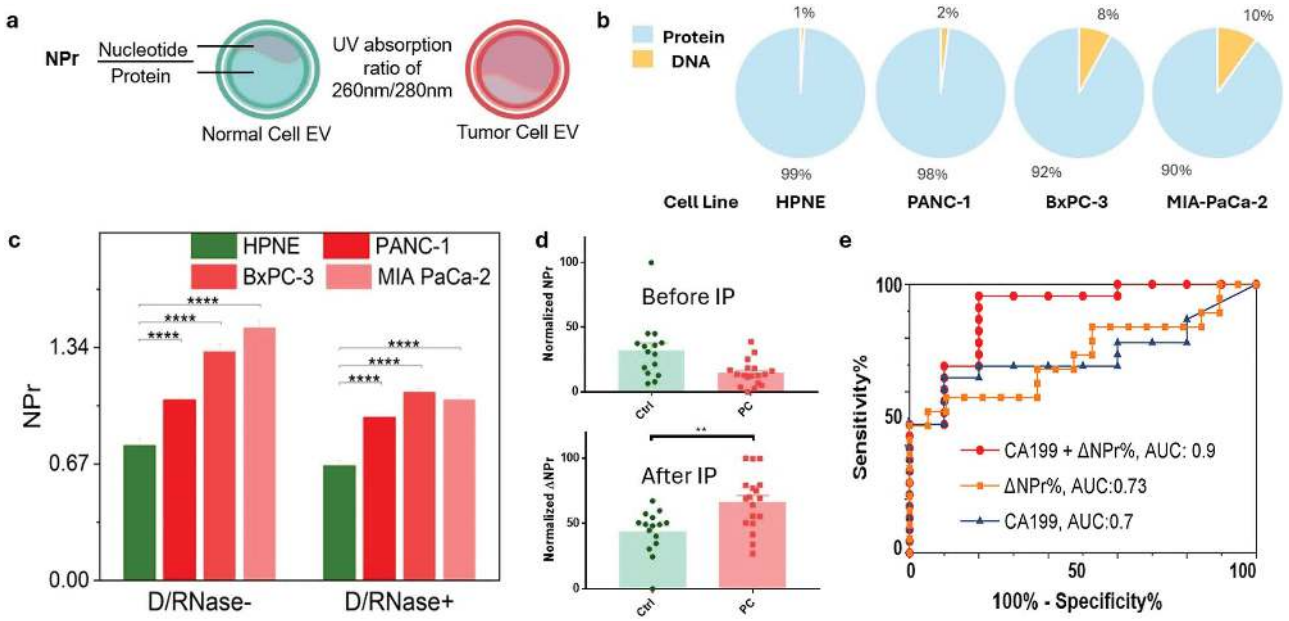


FIGURE 39.1 *Diagnostic potential of using nucleic acid-to-protein ratio for differentiating tumor and nontumor samples.* (a) Defining nucleic acid to protein ratio. (b) NPr of one nontumor and three tumor cell lines. (c) NPr of sample with and without DNA/RNase treatment indicates most nucleic acid signal is due to EV contents. (d) Change of NPr for control vs. PDAC-derived EVs pre- vs. post-enrichment through IP. (e) Combined NPr and CA19-9 yielded a higher AUC (0.90) compared with individual NPr (0.73) and individual CA19-9 (0.70).

healthy controls to assess its diagnostic utility. Our results show that the immunoprecipitation (IP) step, concentrating the EVs in the sample, statistically increases the discriminatory power of NPr (Figure 39.1d). We found that NPr can distinguish between PDAC and control patients with IP enrichment. Although NPr demonstrates a weaker correlation compared to established pancreatic cancer biomarkers like CA19–9 (carbohydrate antigen 19–9) and LIPase, it can be utilized alongside these validated biomarkers to enhance diagnostic accuracy. Combining an NPr-based assay with CA19–9 yielded a higher area under the curve (AUC) for the ROC at 0.9, in contrast to the individual AUCs of NPr (0.73) and CA19–9 (0.7) (Figure 39.1e).

39.4.2 Mitochondrial DNA Abundance in Tumor EVs as Cancer Biomarker

Building off NPr, which focused on the relative abundance of DNA over protein in tumor EVs, we aimed to investigate the DNA component further. From NPr, we observed that mitochondrial DNA (mtDNA) has a higher distribution in the nucleic acid component of NPr. Moreover, in our pancreatic cancer cells, the cancer induces oxidative stress and cellular dysfunction leading to alterations in the mtDNA to nuclear DNA (nDNA) ratio.[8,9] Utilizing fluorescent staining and confocal microscopy, we identified a significant increase in mitochondrial activity in pancreatic tumor cells compared to nontumor cells and developed an EV-based assay to take advantage of this. On examining the mtDNA/nDNA ratio at both cellular and EV levels, we observed that this ratio is independent of the total DNA level for each cell line. This underscores its cell-specific nature, showing resistance to variations in overall DNA content or technical factors that may influence DNA quantification.

For the clinical assay development, we looked at the ratio of mtDNA to nDNA and found that tumor EVs have a higher mtDNA/nDNA ratio when compared with nontumor EVs. To quantify the difference in serum samples, we combined the EV immunoprecipitation procedure with a duplexed qPCR measurement into an assay we named EVIPqPCR. EpCAM on tumor EVs was targeted to enrich the EVs through IP to increase assay power, similar to the NPr assay. qPCR was then used to measure the relative abundance of mtDNA (Figure 39.2a). To validate our results, an *in vivo* tumor growth study on PANC-1 injected mice found that EVIPqPCR could distinguish between tumor and nontumor mice serum (Figure 39.2b). Further, in a pilot study of healthy patients (n = 6), pancreatitis patients (n = 9), and PDAC patients (n = 15), EVIPqPCR significantly distinguished between them (Figure 39.2c). When compared with the currently used pancreatic cancer biomarker, CA19–9, EVIPqPCR has both a higher specificity and sensitivity (ROC-AUC: 0.88), however, with a weak correlation, indicating EVIPqPCR can be used as a complementary noninvasive screening assay (Figure 39.2d and e).

39.4.3 Protein β-Sheet Richness in Tumor EVs as a Cancer Biomarker

After looking at the DNA component of EVs, we then looked at the protein component. Using Fourier transform infrared (FT-IR) and circular dichroism (CD) spectroscopy, we found that tumor EV proteins are comparatively β-sheet rich compared to nontumor EV proteins. These findings were consistently corroborated with fluorescent microscopy, flow cytometry, and bioinformatics analysis (Figure 39.3a and b). Protein-protein interaction (PPI) analysis indicated that malignant cells

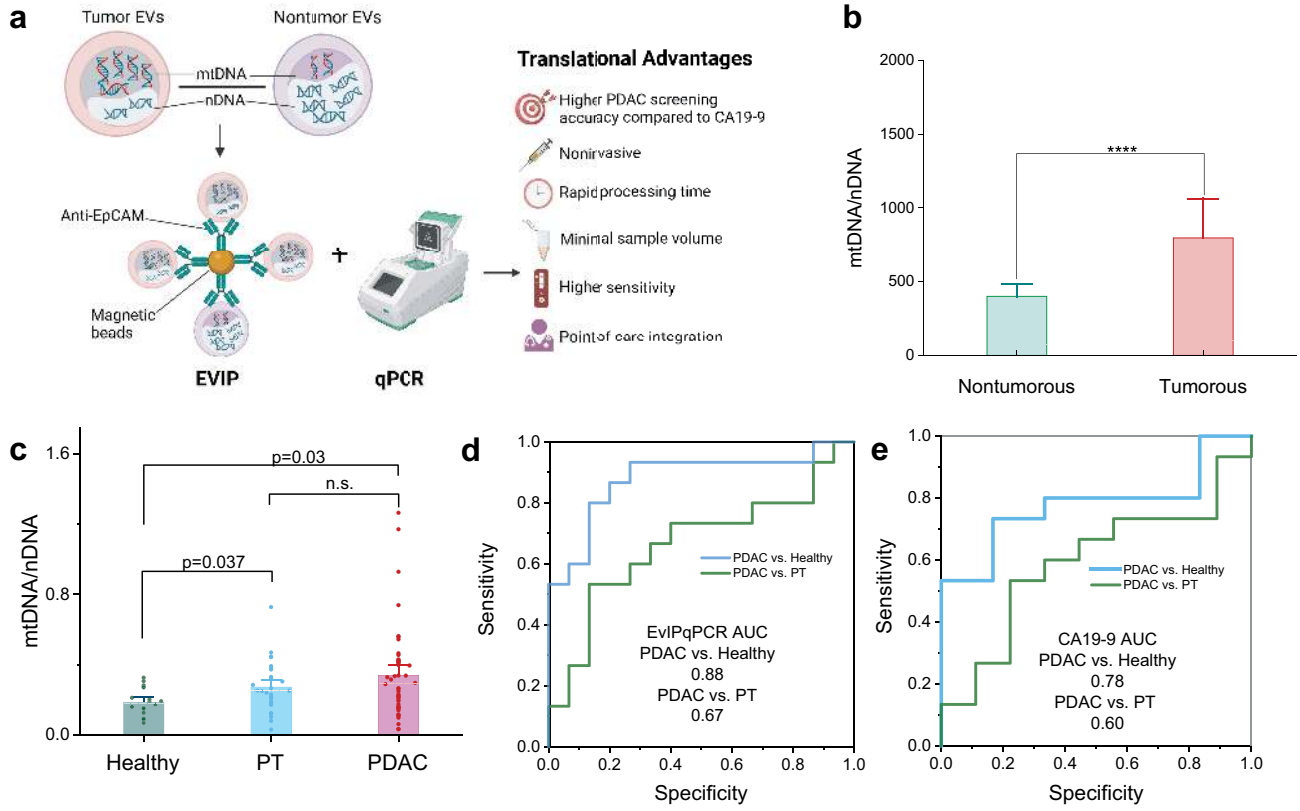


FIGURE 39.2 *mtDNA/nDNA ratio can discriminate between tumor and nontumor.* (a) PDAC detection scheme using mtDNA/nDNA ratio. PDAC-derived EVs have higher mtDNA/nDNA compared to nontumor EVs. (b) EVIPqPCR can distinguish between tumor and nontumor mice serum. (c) EVIPqPCR can differentiate between healthy patients, pancreatitis patients, and PDAC patients. (d, e) ROC curve: EVIPqPCR has higher specificity and sensitivity than CA19–9.

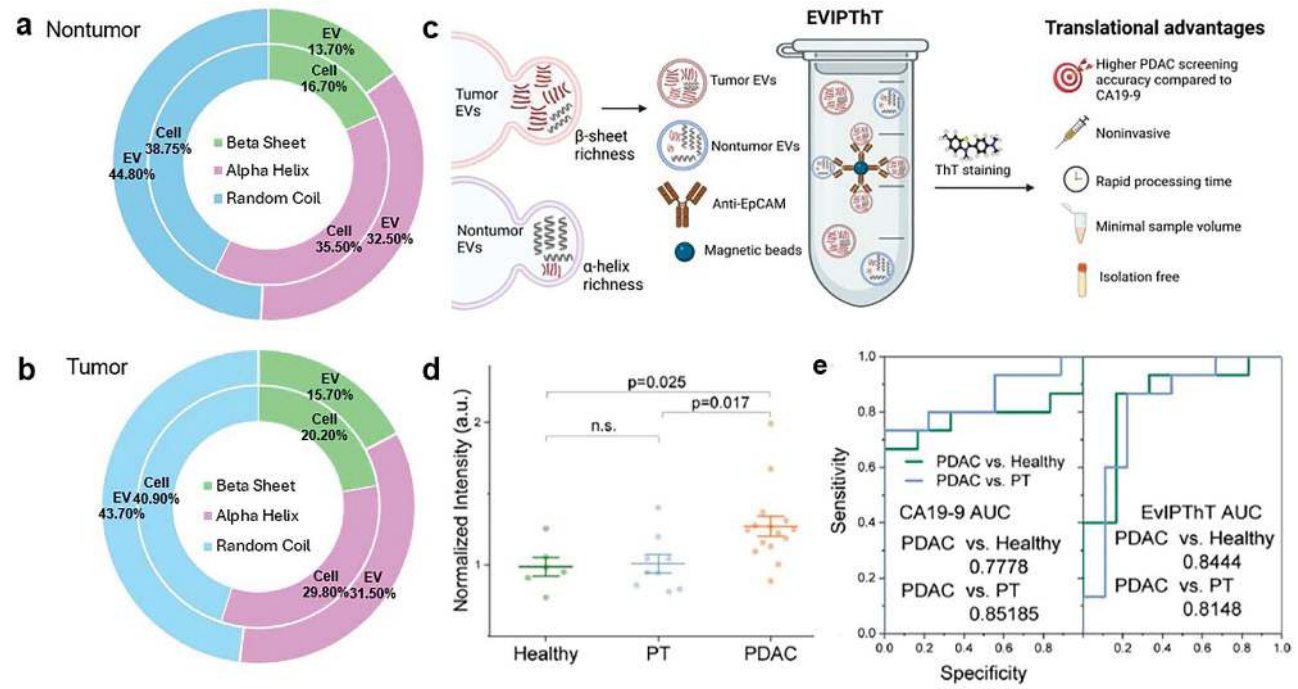


FIGURE 39.3 *β-sheet richness can distinguish between tumor and nontumor samples.* (a, b) Tumor cells and EVs are more β-sheet rich than nontumor counterparts. (c) Detection scheme for PDAC using β-sheet richness. (d) EVIPThT can distinguish between healthy patients, pancreatitis patients, and PDAC patients. (e) ROC curve: EVIPThT has higher specificity and sensitivity than CA19–9.

exhibit a more intricate interaction network than nontumor cells, demonstrating that singular protein markers are not a good indicator of cancer diagnoses, but rather that a statistic such as β-sheet richness is a better alternative as it relates to both structural information as well as intercellular functional activity in biological components.

Given the β-sheet richness observed in PDAC-derived EVs, we aimed to develop a noninvasive clinical cancer diagnostic assay devoid of complex instrumentation such as those required for CD or FT-IR analysis. To develop this, we targeted PDAC-derived EVs using the β-sheet rich feature and developed a fluorescent assay, termed EVIPThT, integrating IP enrichment with thioflavin-T (ThT) staining (Figure 39.3c). The IP method selectively enriches tumor-derived EVs by again targeting their surface antigen—EpCAM, followed with ThT staining to indicate the β-sheet richness of these EVs. ThT staining is a common method in fluorescence spectroscopy, and the protocol is simple and widely accessible. We assessed the diagnostic potential of EVIPThT by evaluating a pilot cohort of healthy controls, pancreatitis patients, and PDAC patients and found that it can discriminate between PDAC, pancreatitis, and healthy subjects (Figure 39.3d). Comparison with current biomarkers, CA19–9, and carcinoembryonic antigen (CEA) indicated the superior discriminatory power of EVIPThT (Figure 39.3e).

39.4.4 ED Spectroscopy for Detecting β-Sheet Richness in Tumor EVs

Having identified multiple collective attributes of EVs as cancer markers, a crucial next step involves developing a cost-effective and widely accessible detection device for these biomarkers. In this context, we developed an elliptical dichroism spectrometer, leveraging the birefringence of biomolecules to different polarization light. This technology can utilize noninvasive samples from bodily fluids, requires minimal to no preparation, and accurately detects tumor EV biomarkers through absorbance spectrometry.

The operating principle of polarized absorbance spectrometers is based on the interaction between a solvent sample and a light source. Light polarization has an effect due to the chiral properties of many biological molecules.[7] Differential absorbance, as a function of light polarization, is then measured using a photodetector.[10] One such device is the circular dichroism spectrometer, which is standard in analyte analysis but has some limitations. Notably, CD spectrometers are large and costly while requiring more complex and intricate parts and maintenance,[11] hindering their ability to be implemented in field hospitals or other remote areas. The ED spectrometer developed herein is a significantly lower-cost spectrometer that uses a combination of linear polarizer and rotating waveplate to generate elliptically polarized light without much loss in detection capabilities. The components of this ED spectrometer are cost-effective, avoiding the expensive components of CD spectrometers, such as high-powered light sources, polarimeters, photoelastic modulators, and photomultiplier tubes.[11] Instead, as depicted in Figure 39.4a, affordable alternatives, such as a UV deuterium lamp and a rotating waveplate, are used. The employment of a UV deuterium lamp encompasses the absorbance spectrum of peptide bonds found in the secondary structures of proteins.[12]

This ED spectrometer has two primary modes of operation: the angle sweep mode and the ED absorbance mode. The AS

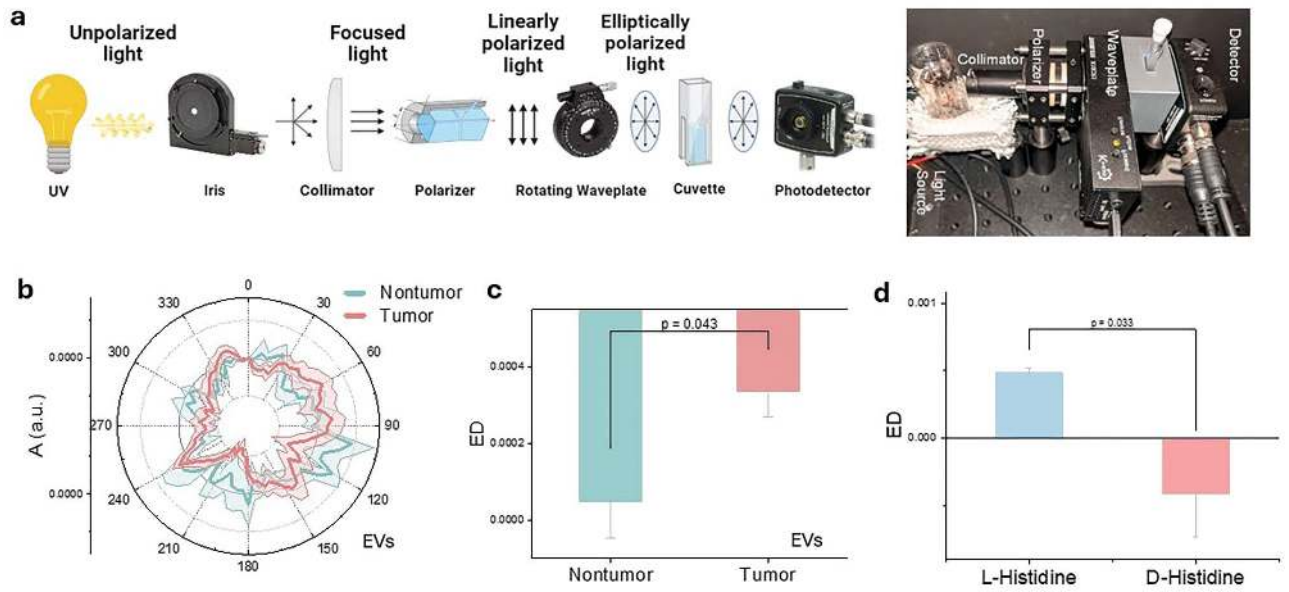


FIGURE 39.4 *Diagnostic potential of the elliptical dichroism device in differentiating tumor and chiral samples.* (a) Optics path and physical layout of components in the ED spectrometer design. (b) Absorbance sweep graph for nontumor and tumor EVs showing the difference in absorbance characteristics. (c) ED values for nontumor and tumor EVs showing discriminatory power of ED spectrometer operating in ED mode. (d) ED value distinguishing chiral isomers of amino acid histidine.

mode measures absorbance values as the waveplate rotates from 0° to 360°, while the ED mode takes the difference in absorbance at 45° and 135°.[12] As found previously, β-sheet richness is a secondary structure that serves as a distinguishing characteristic in the EVIPThT assay. We evaluated the utility of ED spectrometer in characterizing proteins' secondary structures by testing various proteins with known structural features and comparing the ED results with those from a standard CD spectrometer.[13]

We have shown that the ED spectrometer data had strong correlations with CD readouts by validating with BSA, chymotrypsin, and avidin, which are known to be alpha–helix and β-sheet dominant, respectively. Our results indicate that the ED spectrometer can distinguish different protein structures at a similar level to CD spectrometry. As far as diagnostic potential, it was found that there was significant distinguishing power between healthy and tumor cell lines using both derived cell lysates and EVs as samples, both of which showed differences in absorbance sweep. EV results of absorbance sweep and ED values are shown in Figure 39.4b and c. Furthermore, the ED spectrometer was able to distinguish chirality differences within the same molecule. L- and D- isomers of arginine and histidine were tested, and it was found that ED value could distinguish between them; histidine results are shown in Figure 39.4d.

39.5 Discussion

39.5.1 DNA Abundance in Tumor EVs as a Cancer Biomarker

The objective of utilizing an NPr-based assay is to establish a novel parameter that simplifies the distinction between nontumor and tumor EVs. NPr represents a collective characteristic that is easily and rapidly measurable using standard commercial spectrometers, making it an attractive option for cancer screening. The discriminatory power of NPr is grounded in the understanding that EVs from tumor cells contain a higher ratio of double-stranded DNA compared to nontumor EVs.[14] One roadblock that remains is that while the spectrophotometer NPr measurement itself is straightforward, isolating EVs still presents a significant methodological challenge. Without a pure EV sample, accurate measurements cannot be achieved. Furthermore, for diagnostic applications, the amount of tumor EVs in the circulatory system is minimal and tends to be masked by the vast number of nontumor EVs, even in end-stage cancer patients. These translational challenges prompted further evaluation of the nucleic acids and protein components of NPr separately and gave rise to immunoprecipitation for EV enrichment method. The results from these evaluations led to the development of additional EV-based qPCR and fluorescence assays.

39.5.2 Mitochondrial DNA Abundance in Tumor EVs as Cancer Biomarker

qPCR protocols often involve time-consuming and operator-dependent DNA extraction processes. To streamline the assay, we aimed to determine whether accurate results could be obtained without this step. DNA extraction procedures typically result in a reduction of DNA quantity in the sample. Therefore, for an assay concentrating on the ratio of mtDNA to nDNA, eliminating the extraction step is both logical and beneficial. We developed a DNA-extraction-free protocol for evaluating mtDNA/nDNA in tumor EVs and compared it with the DNA-extraction protocol[15] and found that EVIPqPCR between isolation-free and DNA-extraction methods on mice serum samples reveals the mtDNA/nDNA ratio is actually more than three times higher for the isolation-free method, indicating the isolation-free method better reflects differences between tumor and nontumor cases.

The assay's unique features, such as using EV mitochondrial DNA abundance, isolation-free qPCR, duplex probes, minimal sample volume requirement, user-friendly design, and robust discriminatory capacity, render it an effective and convenient method for noninvasive cancer screening.

39.5.3 Protein β-Sheet Richness in Tumor EVs as a Cancer Biomarker

Traditional protein-based assays have concentrated on pinpointing individual protein markers, often lacking in-depth analysis of their structural characteristics. Such an approach tends to suffer from high background noise and has been impractical for clinical diagnostics, notably due to the absence of cancer-specific EV-origin protein markers. By looking at the structural information of protein contents from tumor EVs instead, namely the **β**-sheet richness, the EVIPThT assay addresses these issues.

The EVIPThT assay demonstrates significant potential in pancreatic cancer diagnosis, offering several advantages such as high discriminatory power, not requiring EV isolation, rapid processing time (< 0.75 h), and minimal sample consumption (200 ml serum). These features make the assay compatible with clinical settings that demand high throughput and viable noninvasive cancer screening method.

39.5.4 ED Spectroscopy for Detecting β-Sheet Richness in Tumor EVs

The innovative ED technology provides several advantages over traditional CD spectroscopy, without requiring skilled technicians, long analysis time, bulky and costly instrumentation, or large sample volumes. However, the ED spectrometer is not without its limitations. Due to the lower cost parts used in the design of the ED spectrometer, data results typically have lower resolution, especially as signal-to-noise ratio decreases. To quantify the difference in resolution capability between ED and CD spectroscopy, we found that traditional CD spectrometers have an SNR of 20[16] while the ED spectrometer had an SNR of 14, primarily due to the lower-powered photodetector.[12] For future improvements of this ED spectrometer, the lower SNR value can be mitigated by taking additional replicate measurements[12] or using a more sensitive photodetector.

39.6 Future Trends

Cancer diagnoses may take weeks due to the complexity of laboratory methods. Developing low-footprint, replicable detection tools is crucial for rapid diagnostics at the point

of care. This chapter introduces four innovative methods for utilizing collective extracellular vesicle biomarkers in cancer detection: NPr, EVIPThT, EVIPqPCR, and ED spectroscopy. Initially focused on pancreatic cancer, these adaptable methods promise rapidity, sensitivity, and scalability for various cancer types. The collective analysis of tumor EV cargo demonstrates superior diagnostic capabilities over single biomarkers. For instance, the NPr method employs commercial spectrometers for quick DNA absorbance analysis in EV samples, producing reliable statistics suitable for straightforward application. EVIPqPCR and EVIPThT, using clinical-oriented fluorescence detection, bypass the need for complex EV isolation or DNA extraction, showing clinical promise. ED spectroscopy offers a cost-effective alternative to traditional CD spectroscopy for detecting β-sheet richness in tumor EVs without losing diagnostic power. Advancing these technologies for clinical use is vital for offering diagnostic options, facilitating timely therapeutic interventions, and extending applications to various cancer types. The current focus is on enhancing the assays' specificity and sensitivity through methodological and instrumental improvements, with the goal of automation for clinical use while ensuring accessibility, affordability, and user-friendliness.

REFERENCES

1. Schiffman, J. D., Fisher, P. G. & Gibbs, P. Early detection of cancer: Past, present, and future. *American Society of Clinical Oncology Educational Book* 35(1), 57–65 (2015).
2. Doyle, L. M. & Wang, M. Z. Overview of extracellular vesicles, their origin, composition, purpose, and methods for exosome isolation and analysis. *Cells* 8, (2019).
3. Horswell, S., Matthews, N., & Swanton, C. Cancer heterogeneity and "the struggle for existence": Diagnostic and analytical challenges. *Cancer Letters* 340(2), 220–226 (2013).
4. Sun, D. et al. Dye-free spectrophotometric measurement of nucleic acid-to-protein ratio for cell-selective extracellular vesicle discrimination. *Biosens Bioelectron* 179, 113058 (2021).
5. Bebelman, M. P., Smit, M. J., Pegtel, D. M. & Baglio, S. R. Biogenesis and function of extracellular vesicles in cancer. *Pharmacol Ther* 188, 1–11 (2018).
6. Vaidya, M. & Sugaya, K. DNA associated with circulating exosomes as a biomarker for glioma. *Genes (Basel)* 11, 1–13 (2020).
7. Soltani, S., Ojaghi, A. & Robles, F. E. Deep UV dispersion and absorption spectroscopy of biomolecules. *Biomedical Optics Express*, 10(2), 487–499 (2019).
8. Martinez-Useros, J., Li, W., Cabeza-Morales, M. & Garcia-Foncillas, J. Oxidative stress: A new target for pancreatic cancer prognosis and treatment. *Journal of Clinical Medicine* 6, 29 (2017).
9. Lynch, S. M. et al. Mitochondrial DNA copy number and pancreatic cancer in the alpha-tocopherol beta-carotene cancer prevention (ATBC) study. *Cancer Prev Res (Phila)* 4, 1912 (2011).
10. Bachmann, L. M. & Miller, W. G. Spectrophotometry. *Contemporary Practice in Clinical Chemistry* 119–133 (2020). doi:10.1016/B978-0-12-815499-1.00007-7.
11. Circular Dichroism Spectroscopy. JASCO Inc. (2023, June 5). https://jascoinc.com/learning-center/theory/spectroscopy/circular-dichroism-spectroscopy/.
12. Bauer, A. et al. A portable elliptical dichroism spectrometer targeting secondary structural features of tumorous protein for pancreatic cancer detection. *Biosens Bioelectron* 222, 114934 (2023).
13. Rasuleva, K. et al. β-sheet richness of the circulating tumor-derived extracellular vesicles for noninvasive pancreatic cancer screening. *ACS Sens* 6, 4489–4498 (2021).
14. Thakur, B. K. et al. Double-stranded DNA in exosomes: A novel biomarker in cancer detection. *Cell Research* 24(6), 766–769 (2014).
15. Rasuleva, K. et al. EvIPqPCR, target circulating tumorous extracellular vesicles for detection of pancreatic cancer. *Anal Chem* 95, (2023).
16. Dinitto, J. M. & Kenney, J. M. Noise characterization in circular dichroism spectroscopy. *Appl Spectrosc* 66, 180–187 (2012).

40

Using a Gamified Mobile App Approach to Train Tobacco Control Program Implementers in Schools in India

Eve M. Nagler, Chuck Sigmund, Smita P. Warke, Priyanka Ghosh, Paromita Mehta, Leah C. Jones, and Mangesh Pednekar

40.1 Introduction

Tobacco use is a global public health crisis. Tobacco, a leading preventable cause of disease and death, will kill approximately eight million people annually worldwide by 2030, with 80% of deaths occurring in low- and middle-income countries (LMICs) (1, 2). India is the world's second-largest consumer of tobacco products: 29% of people aged 15 and over use smokeless tobacco and/or smoke (3). India has the highest oral cancer rate globally (4, 5) and low cessation rates (6–8). Each year, approximately 1.2 million people in India die from tobacco-related causes (9, 10). Reducing tobacco-related cancers and mortality in India and other LMICs requires a scale-up of tobacco control evidence-based interventions (EBIs) (11, 12)—such as the *Tobacco-Free Teachers, Tobacco-Free Society* (TFT-TFS) program (13)—that leverage the power of community institutions and leaders. Wider-scale EBI implementation can bolster government efforts and reduce tobacco use, ultimately reducing cancer risk and improving public health (14).

EBI scale-up relies on training program implementers, often in person (15, 16). Yet logistical challenges limit the potential of scaling in-person training models (17). Further, variations in training styles and expertise can compromise fidelity (18, 19). Smartphone-based mobile training may bridge some of these gaps (20) when scaling up tobacco control EBIs, particularly in rural and resource-constrained settings.

Defined as "learning across multiple contexts, through social and content interactions, using personal electronic devices," (21) mobile learning has grown exponentially in the past decade (22). Mobile learning on smartphones offers advantages not found in traditional face-to-face educational settings, including ready access to current content in audio and video formats (23); facilitated communication between and among instructors and participants (24–27); real-time data feedback and tracking; and information sharing among participants and with community members (23, 28). Digital content can be updated more quickly and less expensively than printed materials and can be instantly shared with large audiences—a benefit particularly relevant to underserved communities (22, 28). Mobile learning has been used in many government sectors in India, including education, but has not been tested in the implementation of school-based tobacco control EBIs.

Researchers from the Healis-Sekhsaria Institute for Public Health (Healis) in India and from Harvard T.H. Chan School of Public Health (Harvard) and Dana-Farber Cancer Institute (DFCI) in the United States previously demonstrated the success of TFT-TFS in a cluster-randomized study in Bihar, India. Tobacco use cessation rates among teachers in intervention schools were double those of teachers in control schools (13), and 100% of schools adopted school tobacco control policies (29). TFT-TFS engages tobacco users and non-users around six themes designed to resonate with teachers' personal experiences and centers teachers as school and community role models. We tested a cascade in-person training of trainers model, leveraging Bihar Department of Education (DOE) infrastructure (30). However, program scalability faced resource limitations and logistical challenges. These constraints inspired the development of a smartphone-based TFT-TFS training format that could reach a broader geographic area with increased efficiency and data collection capabilities.

Grounded in the ubiquitous adoption of smartphones across India, our Healis/Harvard/DFCI research team is undertaking a new study (2021–2025) examining the utility of a gamified smartphone-based training app for TFT-TFS program implementers, that is, school principals,* in two Madhya Pradesh (MP) districts. Our study aims to offer insight into scaling up tobacco control EBIs using smartphone technology in schools across India and in other LMICs—a critical step toward reducing tobacco-related cancers and other illnesses. Our study will also inform the application of smartphone-based training for other health-related EBIs in resource-constrained areas.

This chapter provides an overview of our study and explains the gamified TFT-TFS smartphone training app we are testing in MP schools. We describe the formative research to develop the app; the app design; testing of the app; and contextual factors likely to affect program implementation. Finally, we conclude with preliminary findings and reflections on how this technology might be adapted to train people who implement cancer and other non-communicable disease (NCD) control programs in low-resource settings.

* In MP, the program implementer—whether the principal or person designated by the principal to implement the program—is referred to as the Program-In-Charge. For clarity, throughout this paper we refer simply to the "principal."

DOI: 10.1201/9781003449942-45

40.2 Materials

40.2.1 Key Elements of the Training App

40.2.1.1 The TFT-TFS Program

TFT-TFS is a school-based tobacco control program for teachers. The principal conducts the program during work hours for six months across an academic year. With the goal of creating a tobacco-free school environment, TFT-TFS centers on *six themes*, which the principal delivers one per month (following one initial orientation session): (i) Teachers as Role Models, (ii) Health Effects of Tobacco, (iii) Motivation to Quit Tobacco Use, (iv) Skills to Quit Tobacco, (v) Dealing with Withdrawal, and (vi) Maintenance and Celebration. Teachers are encouraged to share program information and to help others quit. TFT-TFS has *four core program components* that incorporate the six monthly themes:

1. Group discussions with teachers, facilitated by the principal.
2. Cessation support via a self-help quit booklet and referrals to government resources.
3. Hanging six theme-based tobacco control posters.
4. Displaying and implementing a school tobacco control policy.

The program also encourages schools to prominently display other tobacco control information, such as "No Tobacco" signs. The essential role of teachers in TFT-TFS—voluntary, but encouraged by the DOE—extends the program's reach and impact within the community. Teacher-led "bonus activities" include motivating community members to quit. Teachers are also encouraged to share information on the harmful effects of tobacco; cessation resources, such as the National Tobacco Quit Helpline (1800-112-356); and the TFT-TFS quit booklet.

40.2.1.2 Gamification

Utilizing game elements in a non-game environment (31), or gamification, increases motivation to complete tasks (31, 32). Researchers have identified more than 200 game "mechanics," or elements employed to engage end users (33). Common game mechanics used in the TFT-TFS smartphone training app include *points* to reward participants for completing modules, quizzes, and activities; *leaderboards* to foster competition; *challenges* to participants to apply what they learned in real-world settings; and *badges* rewarding participants for achieving milestones, for example completing a module. Gamification can significantly enhance learning outcomes by increasing motivation, engagement, and retention (34). In the context of health behavior change, gamification has been shown to encourage participation and program adherence (35).

40.2.1.3 Digital Storytelling

Digital storytelling is a pedagogical approach that (36) presents consistent themes, characters, and story arcs throughout content that continually reinforces key messaging and familiarizes the participant with the digital environment. Digital storytelling allows learners to synthesize the material into their understanding of their daily activities, draw conclusions about how to apply their knowledge, and identify thematic relationships across course materials. As a result, digital storytelling is a powerful and successful mechanism for creating long-term behavior change. Digital storytelling has been shown to increase participant engagement, reduce time to mastery, and dramatically reduce abandonment rates (37).

40.2.2 Characteristics of Study Participants

Study participants include principals and teachers from districts in MP representing urban, rural, and tribal populations. Participants work in MP state government high schools and higher secondary schools, are between the ages of 20 and 64, and are approximately 60% male and 40% female. India's second largest state, MP has a population of more than 72 million (38). Approximately 34% of all adults (50% males, 17% females) (3) and 21% of school personnel (28% males, 8% females) (39) in MP smoke tobacco and/or use smokeless tobacco. The two study districts have populations of 2,185,793 (51% male, 49% female) and 1,563,715 (52% male, 48% female), respectively (40).

40.3 Methods for TFT-TFS Training App Development

40.3.1 Formative Research

We conducted formative research with mixed groups of principals and teachers to explore and describe mobile usage patterns, technological familiarity, and awareness and perceptions of gamification and other game mechanics. In addition, we explored their preferences for delivering the TFT-TFS program via a self-explanatory smartphone-based training app.

Formative research findings suggested that almost all principals and teachers interviewed had Android-based smartphones. They were well-acquainted with using various smartphone apps for teaching, since COVID-19 required distance learning. For the app-based TFT-TFS training, they preferred shorter and animated audio-visuals in Hindi. Further, they recommended the use of a leaderboard to motivate and maintain healthy competition among schools.

We also conducted the Reiss Motivational Profile with six school principals in MP to reveal what engages and disengages them (41). Principals ranked high on curiosity, honor, idealism, order, and status, which helped us develop strategies for reaching and motivating them, and which pointed to the need for a gamified smartphone training app. The TFT-TFS training program thus employs a combination of gamification mechanics and digital storytelling to immerse and involve participants in the content.

40.3.2 Kite Theme

Based on our formative research findings, we selected a kite-flying motif for our app design and gamification strategy. In Indian society, kite flying symbolizes communal harmony

through friendly competition. Colorful kites symbolize the festive spirit, their height depicts achievement, and their flight depicts freedom. The ritual of preparing and flying a kite takes skill, and skill takes time to build; thus, the kite is symbolic of the schools' efforts to make their environment tobacco-free—and fosters friendly competition with other schools. In the app, each school is represented by a kite. The app employs the idea of school kites, segmented into six parts corresponding to the six TFT-TFS monthly themes. A school's performance is gauged by gains in both colors and height.

40.3.3 Cultural Relevance

All in-person TFT-TFS training content was converted into training videos created in Vyond or Camtasia, which allowed us to design a creative, customizable, and immersive environment. We created characters and stories aligning with local MP culture and representation. We further ensured cultural and linguistic relevance to app users by delivering all content in Hindi.

40.3.4 Technology

The application uses a combination of user interface (UI) frameworks, including Flutter, React.js, Laravel, and MySQL, deployed across various platforms to suit different program components. Based on user preference, we developed an Android app, which is hosted on the Google Play Store. The admin panel is hosted on Amazon Web Service (AWS) Simplify; its application program interface (API) server and database are hosted on AWS EC2; its push notifications are triggered from Firebase; and the data is stored on AWS S3. This setup ensured robust performance monitoring and management of the initiative across its diverse user base.

40.3.5 App User Roles

A database of districts, schools, their principals and teachers, and their respective mobile numbers was created to manage all potential users and their roles. *Principals* have access to the whole app training, including gamified content, core activities, and program tracking functions. *Teachers* use the app only to log completion of their bonus activities.

40.3.6 Detailed Description of the TFT-TFS Training App

Once the TFT-TFS training app is downloaded and installed, the principal logs in with their mobile number, which is verified through a one-time password (OTP) sent to their device. The login screen (see Figure 40.1) features the image of "*Aadarsh Vidyalaya*" (The Ideal School), emphasizing TFT-TFS's aim to help schools create a tobacco-free environment. The app is divided into three main sections: (1) the "Principal's Journey" provides access to gamified training content and the four core components for each theme; (2) the "Teachers' Journey" helps teachers track their bonus activities (e.g., sharing tobacco information and helping others quit); and (3) the "Dashboard" displays a school's progress in comparison to other schools.

FIGURE 40.1 Login screen with the idyllic school image.

40.3.6.1 Principal's Journey

The principal's journey begins with a brief call-to-action video (in English with Hindi subtitles) featuring Dr. Eve Nagler, the study's principal investigator, who welcomes principals to TFT-TFS and acknowledges their significance as society's role models. Videos of participating district officials also endorse and emphasize the importance of schools participating in this program.

At the end of the call-to-action video, the principal lands on the home page. They then watch a navigation video and learn how to unlock themes using the map, navigate the app using the drop-down menu, and locate the dashboard and settings, as well as access the navigation video for future reference.

The home page features a map of MP (see Figure 40.2) with the themes plotted on it to simplify navigation. The Orientation module at the beginning of the journey map provides an overview of the program, participant roles, app features, game mechanics, scoring, navigation, and insights into tobacco use in India in animated videos. To track how well principals understand the Orientation content, they are asked to take a short quiz. This same process of verifying participant's knowledge, that is, with quizzes of three to five questions each, is repeated after each informational video throughout the app. Upon completion of the Orientation module, the school is awarded an outline of a kite divided into six segments, with each segment representing a TFT-TFS theme.

Each theme includes training videos, quizzes, and activities to implement the theme's four core components. The themes unlock sequentially at the start of each month, contingent on the principals' completion of tasks in the prior theme. TFT-TFS administrators (in our case, the study team) have the flexibility to unlock themes as needed. These themes are easily identifiable by their unique colors and icons, which are mirrored on the school kite. Once a school completes all four core components in Theme 1, the school kite gains the color blue. Similarly, after completion of Theme 2, the school kite gains orange, and so on, until the kite is filled in with the six colors representing the six themes.

Each theme's module has its own landing page. Content is organized in a carousel-style display (see Figure 40.3) that allows the learner to swipe and see viewed and unviewed content within the theme. Users can freely navigate back to previous content or activities at any time. To begin a specific task, users select "Get Started," with each task type distinguished by its icon (e.g., a training video has a "Play button" icon, a knowledge check has a "Magnifying glass" icon). A progress bar offers principals a visual indication of their progress within the theme.

All themes are similarly structured. The principal first views training videos that explain the theme. For instance, Theme 1, titled *Shikshak ek Aadarsh* (Teacher as a Role Model), features

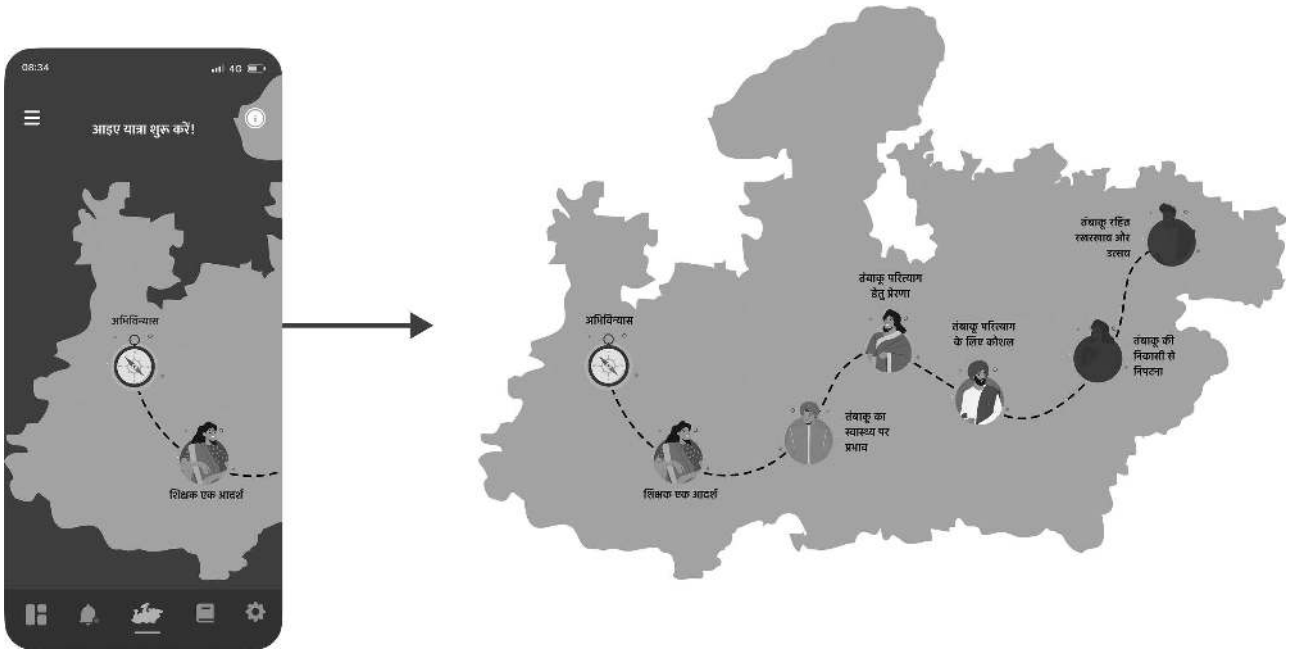


FIGURE 40.2 App home page and the full map of Madhya Pradesh when swiped.

FIGURE 40.3 Theme layout with carousel display.

a 10-minute video narrating the story of a teacher who raises awareness about tobacco's harmful effects after a student drops out due to family circumstances related to tobacco use. The video is followed by a three-question quiz to assess the principal's understanding of the content.

After completing this quiz, principals download a guide to facilitate the Theme 1 group discussion with teachers. This app also contains instructions for implementing the theme's other three core program components (i.e., distributing the quit booklet, hanging the poster, and displaying the school tobacco policy). Additionally, the app guides principals in encouraging teachers to participate in bonus activities.

After implementing each program component, the principal is prompted to fill out a completion form. For example, after conducting the group discussion, the principal records the discussion date, time, and number of teacher participants. The principal also uploads photos of component completion (e.g., photo of the Theme poster on the wall). There is a completion form for each program component. Progress within a theme is displayed in the four tabs on the dashboard, as described later.

The app delivers weekly push notifications to motivate principals to complete their monthly activities. Within the Settings tab of the app, the Help page allows participants to contact the program team for assistance with app navigation or other technical difficulties.

40.3.6.2 Teachers' Journey

When teachers log in to the app, they watch the TFT-TFS program overview video outlining the six themes and core program components. A demonstration video illustrates how the themes are connected to the game of kite flying and how schools can score points. Upon completion of these videos, teachers begin to conduct bonus activities with students, family, friends, and community members and record their actions. Teachers fill out a unique form for the bonus activities they conduct for each theme.

How schools score points: Implementation of core program components, alongside teacher bonus activities, earns schools points, determining the school's standing in the kite flying game. Points range from two to six per activity, with a maximum of 24 points per theme achievable for the school. If a single teacher completes the activity, the school scores two points, while if 100% of the teachers at the school complete the activity, the school scores a maximum of six points. So, the combination of core and bonus activities within each theme offers schools opportunities to score in the game, with greater activity completion indicating greater implementation within the school and community.

40.3.6.3 Dashboard

The dashboard (see Figure 40.4) is divided into four tabs that track the progress of every school: School's Kite Page, Bonus Activities Page, Leaderboard, and Scoreboard. The School's Kite Page offers a summary of program component completion per theme, displaying the earned colors on their kite. The Bonus Activities Page displays all bonus activities; by tapping on "Get Details," principals can view who has completed each activity. The Leaderboard displays the top five schools in the district, represented by their kites flying highest in the sky; tapping on a school kite reveals additional scoring details. Finally, the Scoreboard lists all district schools, ranking them based on score.

For TFT-TFS administrators, there is an Admin Panel of data collected on the app, which shows the progress of each school implementing the TFT-TFS program.

40.3.7 Testing the App in a Comparative Effectiveness Trial: Study Design Overview

Our study will evaluate whether the implementation and effectiveness of the TFT-TFS program when delivered through smartphone-based training is non-inferior (i.e., as good or superior) to traditional in-person training. We are conducting a two-arm comparative effectiveness trial using a cluster-randomized design. We randomly assigned 200 high schools and higher secondary schools from the two districts in MP to receive either traditional in-person training (n = 90) or training via smartphone (n = 110) for the TFT-TFS program. Once trained, principals in both groups implement TFT-TFS within their schools for six months. We will compare the training arms

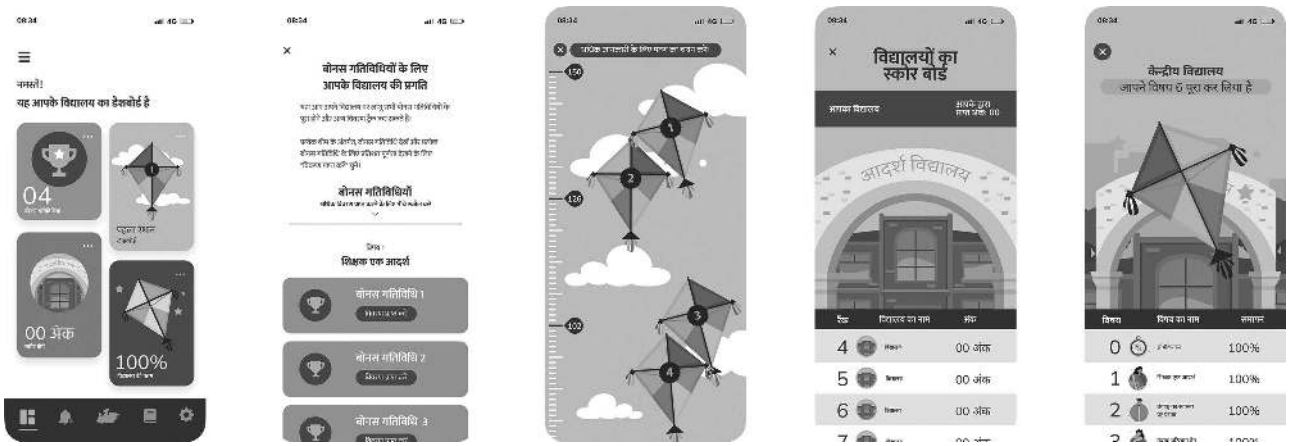


FIGURE 40.4 Dashboard with bonus activities, leaderboard, and school kite.

in terms of program implementation, teacher participation, cost, and effectiveness (tobacco use cessation among teachers). Post-implementation, we will use mixed methods to assess factors contributing to outcomes, guided by the Consolidated Framework for Implementation Research (CFIR).

40.3.8 Identifying Factors Affecting Program Implementation

After the post-implementation assessment, we will conduct a mixed-methods evaluation to identify factors underlying observed variations in TFT-TFS program implementation for each arm. Based on our prior research (42–44), we selected factors expected to influence implementation of TFT-TFS that are relevant to the DOE and that would inform future scale-up efforts. These include factors within four CFIR domains related to (1) the TFT-TFS program, and smartphone training app features and functioning; (2) DOE leadership support for TFT-TFS and the smartphone training app; (3) compatibility of the TFT-TFS program and its training app with school processes and how TFT-TFS supports existing school tobacco control efforts; and (4) principals' ease of using and perceived effectiveness of the TFT-TFS smartphone app. This will deepen our understanding of contextual factors associated with using the TFT-TFS smartphone training app.

In each training arm, we selected schools who (1) implemented the program monthly; (2) implemented multiple themes in a month; and (3) did not implement the program. Principals participated in a qualitative semi-structured interview to explore factors related to the CFIR constructs. We also asked open-ended questions about their implementation experience, their perspectives on strategies employed to promote TFT-TFS program tracking on smartphones, and their reactions to trends identified in the quantitative analysis. Additionally, principals proposed solutions to implementation challenges and made recommendations for future use of this technology.

40.3.9 Mixing Qualitative and Quantitative Data

We will present side-by-side comparisons of findings in each arm. Table cells will contain the quantitative/qualitative results for each CFIR construct and mean number of core program components implemented, teacher participation, cost, and effectiveness. Together with principals and district-level representatives, we will synthesize principals' solutions and recommendations and present them to the state DOE.

40.4 Results

We are gathering implementation data and expect to have final results in 2025. Preliminary findings indicate that 96/110 (87%) of the principals assigned to the smartphone training arm completed the Orientation session on the TFT-TFS training app. This indicates their ability to download the app, log in, interact with content, and complete quizzes.

Comments from principals in a TFT-TFS WhatsApp group also suggest they enjoy using the app, particularly the healthy competition the training app created. One male principal commented after completing Theme 4, "My school came at 2nd position this time, but I assure you that next theme onwards we will come in the 1st position." Another expressed, "My school's kite is at the first position since the beginning of this program. Our teachers are very motivated with this. Thank you for bringing this program to us." Principals trained also commented that having audio-visuals in Hindi was essential. Finally, a female principal expressed her overall satisfaction: "The app and the program are good, and the leaderboard is quite a motivating factor."

40.5 Discussion

This novel application of an existing technology that is already ubiquitous in Indian schools and communities will demonstrate the feasibility of smartphone-based training and its effect on EBI implementation in a school setting to improve tobacco control in LMICs. We expect to provide a strong scientific evaluation of the use of smartphones as a strategy to rapidly scale up training of program implementers to implement tobacco control EBIs in LMICs. Our results may also help designers tailor the training app for cancer control and address the growing burden of other NCDs.

Despite these contributions, when we consider the use of smartphone technology for cancer control, we must recognize advantages and limitations. On the one hand, the smartphone training app provides a scalable, low-cost, and engaging training solution for tobacco control program implementers in low-resource settings. End users can learn anytime/anywhere and repeat course content for greater retention. Still, although app users can interact virtually with other participants and trainers, they lack face-to-face communication, opportunities for role plays, and hands-on experiences (45, 46). Learners may also feel less motivated to implement a program if they are not expected to report on their progress in person. The app is also dependent on Internet connectivity.

40.6 Future Trends

40.6.1 Future Directions

An important next step in our investigation of this technology for TFT-TFS training would be to explore more deeply how teachers use smartphone technology to share tobacco control information. Further, we could explore strategies to use gamification to enhance the user experience and interaction with smartphone training and to improve data collection, particularly in resource-constrained settings.

40.6.2 Potential Applications

There is increased use of mobile apps for cancer control (47), but there have not been ample randomized control trials with sufficient sample sizes to test the effectiveness of these initiatives (48). The smartphone technology used to train TFT-TFS implementers might be adapted to train, at scale, implementers

of other government health programs that address modifiable risk factors for cancer and other NCDs. For example, community health workers in LMICs are instrumental to public health initiatives, including encouraging patients to receive routine cancer screenings at government health centers. The app could be tailored to train community health workers in behavior change communication strategies, for example, or to record information from home visits and community outreach activities on the app.

The overall concept, design, and architecture of the smartphone app can be easily modified to address a broad variety of issues and audiences. Designers would need to analyze the technology needs and preferences of the intended audience and identify ways to make the storyline and delivery culturally and linguistically relevant. They would also need to modify the app content for the language and health literacy of the user population (48). Collecting formative data enables designers to create a thematic approach that is relevant and engaging and to select the appropriate game mechanics to reinforce the desired user behaviors. For instance, if the app were adapted to train implementers of an adolescent diabetes program in Mexico, designers could use a *luchador* (brightly costumed Mexican wrestler) theme. Designers could create a digital story in which the user takes on the luchador persona and engages in a series of "matches," as they progress through the training app.

40.7 Funding

This study is supported by the National Cancer Institute of the National Institutes of Health, Division of Cancer Control & Population Sciences (DCCPS), Grant Number: 1R01CA248910–01A1.

REFERENCES

1. World Health Organization. *WHO Report on the Global Tobacco Epidemic, 2008: The MPOWER Package*. Geneva: World Health Organization; 2008.
2. World Health Organization. *WHO Report on the Global Tobacco Epidemic, 2011, Executive Summary [Internet]*. Geneva; 2011 [cited 2016 Aug 5]. http://apps.who.int/iris/bitstream/10665/70680/1/WHO_NMH_TFI_11.3_eng.pdf.
3. Tata Institute of Social Sciences (TISS) M, Ministry of Health and Family Welfare G of India. *Global Adult Tobacco Survey GATS 2 India 2016–17*. New Delhi, India: Ministry of Health and Family Welfare, Government of India; 2018.
4. Gupta PC, Ray CS, Murti PR, Sinha DN. Rising incidence of oral cancer in Ahmedabad city. *Indian J Cancer*. 2014/12/20 ed. 2014 Dec;51 Suppl 1:S67–72.
5. Reddy KS, Gupta PC. *Report on Tobacco Control in India*. New Dehli, India: Ministry of Health and Family Welfare, Government of India; 2004 Nov. Report No. 1.
6. Jindal SK, Aggarwal AN, Chaudhry K, Chhabra SK, D'Souza GA, Gupta D, et al. Tobacco smoking in India: Prevalence, quit-rates and respiratory morbidity. *Indian J Chest Allied Sci*. 2006/02/18 ed. 2006 Jan;48(1):37–42.
7. International Institute for Population Sciences. *Global Adult Tobacco Survey (GATS) India 2009–2010*. New Delhi: Ministry of Health and Family Welfare, Government of India; 2010.
8. Thankappan KR. Tobacco cessation in India: A priority health intervention. *Indian J Med Res*. 2014/07/25 ed. 2014 Apr;139(4):484–6.
9. Sinha DN, Palipudi KM, Gupta PC, Singhal S, Ramasundarahettige C, Jha P, et al. Smokeless tobacco use: A meta-analysis of risk and attributable mortality estimates for India. *Indian J Cancer*. 2014/12/20 ed. 2014 Dec;51 Suppl 1:S73–7.
10. Jha P, Jacob B, Gajalakshmi V, Gupta PC, Dhingra N, Kumar R, et al. A nationally representative case-control study of smoking and death in India. *N Engl J Med*. 2008/02/15 ed. 2008 Mar 13;358(11):1137–47.
11. Siddiqi K, Newell J, Robinson M. Getting evidence into practice: What works in developing countries? *Int J Qual Health Care*. 2005/05/06 ed. 2005 Oct;17(5):447–54.
12. McMichael C, Waters E, Volmink J. Evidence-based public health: What does it offer developing countries? *J Public Health Oxf*. 2005/04/12 ed. 2005 Jun;27(2):215–21.
13. Sorensen G, Pednekar MS, Sinha DN, Stoddard AM, Nagler E, Aghi MB, et al. Effects of a tobacco control intervention for teachers in India: Results of the Bihar school teachers study. *Am J Public Health*. 2013/09/14 ed. 2013 Nov;103(11):2035–40.
14. Gupta PC, Lando HA, Pednekar MS, Narake SS, Nagler EM, Pawar PS, et al. Improvement in prevalence of tobacco use among teachers in Bihar after COTPA. *Indian J Cancer*. 2014/12/20 ed. 2014 Dec;51 Suppl 1:S19–23.
15. Fernandes PT, Noronha AL, Sander JW, Bell GS, Li LM. Training the trainers and disseminating information: A strategy to educate health professionals on epilepsy. *Arq Neuropsiquiatr*. 2007/06/28 ed. 2007 Jun;65 Suppl 1:14–22.
16. Mormina M, Pinder S. A conceptual framework for training of trainers (ToT) interventions in global health. *Glob Health*. 2018 Dec;14(1):100.
17. Orfaly RA, Frances JC, Campbell P, Whittemore B, Joly B, Koh H. Train-the-trainer as an educational model in public health preparedness. *J Public Health Manag Pr*. 2005/10/06 ed. 2005 Nov;Suppl:S123–7.
18. Hayes D. Cascade training and teachers' professional development. *ELT J*. 2000 Apr;54(2):135–45.
19. Baron N. The 'TOT': A global approach for the training of trainers for psychosocial and mental health interventions in countries affected by war, violence and natural disasters. *Intervention*. 2006 Jul;4(2):108–25.
20. Government of India. *DIKSHA National Digital Infrastructure for Our Teachers [Internet]*. 2020 [cited 2024 Mar 24]. https://www.india.gov.in/spotlight/diksha-national-digital-infrastructure-teachers.
21. Crompton H. *A Historical Overview of Mobile Learning: Toward Learner-Centered Education*. London: Routledge; 2013. (Berg ZL, Muilenburg L, editors. Handbook of Mobile Learning).
22. Baker A, Dede C, Evans J. *The 8 Essentials for Mobie Learning Success in Education*. San Diego, CA: Qualcomm Wireless Reach; 2018.
23. Curran V, Matthews L, Fleet L, Simmons K, Gustafson DL, Wetsch L. A review of digital, social, and mobile technologies in health professional education. *J Contin Educ Health Prof*. 2017/08/24 ed. 2017 Summer;37(3):195–206.

24. Hardyman W, Bullock A, Brown A, Carter-Ingram S, Stacey M. Mobile technology supporting trainee doctors' workplace learning and patient care: An evaluation. *BMC Med Educ*. 2013/01/23 ed. 2013 Jan 21;13:6.
25. Lall P, Rees R, Law GCY, Dunleavy G, Cotic Z, Car J. Influences on the implementation of mobile learning for medical and nursing education: Qualitative systematic review by the digital health education collaboration. *J Med Internet Res*. 2019/03/01 ed. 2019 Feb 28;21(2):e12895.
26. Alegria DA, Boscardin C, Poncelet A, Mayfield C, Wamsley M. Using tablets to support self-regulated learning in a longitudinal integrated clerkship. *Med Educ Online*. 2014/03/22 ed. 2014;19:23638.
27. Doherty I, Sharma N, Harbutt D. Contemporary and future eLearning trends in medical education. *Med Teach*. 2014/08/27 ed. 2015 Jan;37(1):1–3.
28. Anikeeva O, Bywood P. Social media in primary health care: Opportunities to enhance education, communication and collaboration among professionals in rural and remote locations: did you know? Practical practice pointers. *Aust J Rural Health*. 2013/04/17 ed. 2013 Apr; 21(2):132–4.
29. Pawar PS, Nagler EM, Gupta PC, Stoddard AM, Lando HA, Shulman L, et al. Tracking intervention delivery in the "tobacco-free teachers/tobacco-free society" program, Bihar, India. *Health Educ Res*. 2015/09/06 ed. 2015 Oct;30(5):731–41.
30. Nagler EM, Pednekar M, Sinha DN, Stoddard AM, Narake S, Adhikari K, et al. Implementation of an evidence-based tobacco control intervention for school teachers in India: Evaluating the effects of a capacity-building strategy. *Implement Res Pract*. 2023 Jan 1;4. https://doi.org/10.1177/26334895231159428.
31. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: Defining "gamification". In: *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments [Internet]*. New York, NY, USA: Association for Computing Machinery; 2011 [cited 2024 Apr 18]. p. 9–15. (MindTrek '11). https://doi.org/10.1145/2181037.2181040.
32. Werbach K, Hunter D. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton School Press; 2012.
33. Peters J, Cornetti M. *Deliberate Fun: A Purposeful Application of Game Mechanics to Learning Experiences*. Sententia Publishing; 2019.
34. Sailer M, Hense JU, Mayr SK, Mandl H. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Comput Hum Behav*. 2017 Apr 1;69:371–80.
35. Cugelman B. Gamification: What it is and why it matters to digital health behavior change developers. *JMIR Serious Games*. 2013 Dec 12;1(1):e3.
36. Smeda N, Dakich E, Sharda N. *Developing a Framework for Advancing e-Learning through Digital Storytelling*. 2010 [cited 2024 Apr 18]. p. 169–76. https://www.iadisportal.org/digital-library/developing-a-framework-for-advancing-e-learning-through-digital-storytelling.
37. Hunter LP, Hunter LA. Storytelling as an educational strategy for midwifery students. *J Midwifery Womens Health*. 2006;51(4):273–8.
38. Government of India. *Census India—2011 [Internet]*. 2011 [cited 2024 Feb 27]. https://www.censusindia.co.in.
39. Sinha DN, Gupta PC, Gangadharan P. Tobacco use among students and school personnel in India. *Asian Pac J Cancer Prev*. 2007/12/28 ed. 2007 Jul;8(3):417–21.
40. Office of the Registrar General & Census Commissioner India. *Census Data 2011 [Internet]*. New Delhi, India; 2011 [cited 2022 Feb 22]. https://censusindia.gov.in/2011-common/censusdata2011.html.
41. Reiss S. *The Reiss Motivation Profile: What Motivates You?* IDS Publishing Corporation; 2013.
42. Karuppusamy B, Paulson W, Chellappan S, Behera SK, Balabaskaran Nina P. District-level epidemiology, hot spots and sociodemographic determinants of tobacco use in Indian men and women: Analysis of national family health survey-4 (2015–16). *Public Health*. 2021 May;194:127–34.
43. Aghi M, Nagler E, Lando H, Pednekar M, Gupta P, Sorensen G. Training lay interventionists to support tobacco cessation among teachers in India. *Int J Health Promot Educ*. 2016/01/01 ed. 2016;54(6):304–17.
44. Nagler EM, Aghi M, Rathore A, Lando H, Pednekar MS, Gupta PC, et al. Factors associated with successful tobacco use cessation among teachers in Bihar state, India: A mixed-method study. *Health Educ Res*. 2020 Feb 1;35(1):60–73.
45. Knowles MS, Holton EF, Swanson RA, Robinson PA. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. 9th ed. London: Routledge/Taylor & Francis Group; 2020.
46. Vella J. *On Teaching and Learning: Putting the Principles and Practices of Dialogue Education into Action*. San Francisco, CA: Jossey-Bass; 2008.
47. Houghton LC, Howland RE, McDonald JA. Mobilizing breast cancer prevention research through smartphone apps: A systematic review of the literature. *Front Public Health*. 2019;7:298.
48. Coughlin S, Thind H, Liu B, Champagne N, Jacobs M, Massey RI. Mobile phone apps for preventing cancer through educational and behavioral interventions: State of the art and remaining challenges. *JMIR MHealth UHealth*. 2016 May 30;4(2):e69.

41

Parallel Reaction Monitoring (PRM) Quantitative Analysis of Glycopeptide Biomarkers from Patient Serum

Komal Abhange, Yu Lin, Jie Zhang, Jianhui Zhu, Amit G. Singal, Neehar D. Parikh, and David M. Lubman

41.1 Introduction

41.1.1 Hepatocellular Carcinoma

Currently we are using PRM technology to study biomarkers for hepatocellular carcinoma (HCC), the fourth most common cause of cancer-related death worldwide (1). Over 80% of HCC develops in patients who have underlying cirrhosis. HCC is one of the few cancers with a rising mortality in the Western world, including the United States, with a disproportionate burden among racial and ethnic minority patients and those of low socioeconomic status (2).

Global professional society guidelines recommend HCC surveillance for early-stage HCC detection (3, 4). Curative treatments can be generally applied to patients who have early-stage HCC, resulting in a five-year survival rate greater than 70%; whereas patients diagnosed at late stage have a five-year survival rate less than 5% (5). The increasing incidence of cirrhosis is due to both the peak in hepatitis C–related complications and the emergence of metabolic dysfunction–associated steatotic liver disease (MASLD) and alcohol-related liver disease (6–9). MASLD typically occurs in patients with obesity, diabetes, and metabolic syndrome, thus in parallel with the obesity epidemic, the rise in MASLD incidence has contributed to the growth of HCC incidence in the United States (10). Therefore, with shifting epidemiology of HCC, early detection strategies should be tailored to emerging etiologies of liver disease, especially MASLD, for effective therapeutic treatment and improvement in survival.

41.1.2 Current Methods in Early Detection

Only about 20%–30% of patients with HCC are eligible for curative therapy with surgical resection, orthotopic liver transplantation, and percutaneous ablative techniques (11). The impact of these therapeutic interventions on the overall survival of patients with HCC underscored the importance of detecting HCC at earlier stages (12–14). HCC surveillance has been associated with significant improvements in early detection, curative treatment, and reduced mortality across several cohort studies (15). Improved methods for surveillance are needed to increase early detection and to identify patients that can be offered curative therapies. Currently recommended strategies for HCC early detection involves imaging ultrasound every six months (16); however, the sensitivity of ultrasound to detect HCC at an early stage is below 50% when used alone (17). Further, ultrasound visualization and performance are likely worse in patients with non-viral liver disease, an increasingly common population seen in clinical practice (18, 19). Finally, effectiveness in clinical practice is further impeded by low implementation given patient and provider barriers to imaging-based surveillance (20–22).

Serum markers are used as a complementary method in patients with cirrhosis or small tumors (23–25). Serum alpha-fetoprotein (AFP) is most widely used as a complementary HCC detection marker. Although AFP has suboptimal test performance when used alone, it appears to be complementary to ultrasound when used in combination (17, 26). Des-gamma-carboxy prothrombin (DCP) is widely used as an alternative marker for AFP in Japan (27). However, the performance of DCP varies by etiology of liver disease. Finally, AFP-L3 has also been evaluated as an early detection marker but has limited utility (28). Biomarker panels, such as GALAD, have promising performance in early phase data but require validation in larger cohort studies, particularly with contemporary cirrhosis etiologies (29, 30). Biomarkers with improved performance characteristics for the early detection of HCC are therefore needed to improve the sensitivity for detection over existing markers, with the possibility of development of disease specific markers that have been optimized for the specific etiology of disease (31–42).

41.1.3 Site-Specific Glycopeptides Serve as Biomarkers for Cancer Detection

Unique changes in the glycosylation of proteins that involve structural changes in glycan groups may be important as serum biomarkers for early cancer detection (43–70). These changes in glycan structure may be associated with distinct proteins/peptides at specific sites (64–66, 69, 70). The idea of monitoring both glycan structure and its site specificity and occupancy provide a marker fingerprint that could be more specific to particular organ sites and disease states. Moreover, there may be isomers of these structures where subtle differences provide key information as markers of cancer. These structures may include multiple isoforms of fucosylated and sialylated glycans where novel methods are required for resolving and analyzing these structures (71). Such glycan and glycopeptide structural changes can be detected and monitored quantitatively using mass spectrometry (MS), which has proven to be a means to

DOI: 10.1201/9781003449942-46

study even minor changes in structure that may be significant as biomarkers. This capability has been made possible by the advent of high resolution mass analysis (46–48, 64–70), tandem mass spectrometry (MS/MS) measurements, and novel software and databases that have recently been developed for such glycoprotein analysis (46–63). In addition, these glycopeptide markers can be assayed without initial isolation of the protein from serum, where the peptide portion of the marker can be identified with a specific protein.

41.1.4 Technology-Parallel Reaction Monitoring as a Quantitative Assay for Glycopeptides

The technology developed in our work is based on targeted mass spectrometry analysis of glycopeptides using PRMs. A schematic showing the basics of PRM in mass spectrometry is shown in Figure 41.1. In PRMs the various glycopeptides are separated by high-performance liquid chromatography (HPLC) and then injected into the mass spectrometer using an electrospray source via electrospray ionization (ESI). As they elute from the HPLC, the target ions are selected by the first mass filter, then fragmented in a collision gas (higher energy collisional dissociation [HCD]) and then all product ions are monitored with high resolution using an orbitrap-MS device. In PRMs, all product ion transitions and possible product ions are analyzed together with high resolution and mass accuracy. PRMs are a well-developed method for analysis of peptides using an internal standard synthesized with an isotopic label (72–80). The PRM method does not require antibodies and is a highly accurate method for quantitation. The lack of a need for antibodies makes it applicable to any protein where often good antibodies are not available, especially for low-abundance proteins. The method also enhances sensitivity for detection based on the use of targeted detection where background ions are eliminated from the analysis. In addition, the technique can be multiplexed so that large numbers of target markers can be evaluated simultaneously based on initial mass spectrometer target selection, and the method can be automated to run large numbers of samples.

The PRM-MS method is used on a regular basis in various assays in our laboratory with an orbitrap fusion lumos mass spectrometer. This instrument is especially suited for the proposed work because it will allow us to perform the proposed LC-HCD-PRM-MS method so that we can monitor the peptide backbone and glycosylation site and the glycan structure. The stepped HCD method will be used for fragmentation and structural analysis as found to provide excellent results in our previous work (71, 81–84). Essentially, we have a mass spectrometry version of an ELISA for detection and quantitation of glycopeptides, where ELISAs cannot readily be developed for glycopeptides and glycoproteins. The issue for this work is that standards for glycopeptides for quantitation are generally not readily available. We have thus developed methods to normalize the signals for this method for glycopeptides so that accurate quantitative work can be performed.

In this chapter, we will describe our work on the use of PRM-MS technology in biomarker confirmation and validation in patient serum. The study herein describes an application to glycopeptides specifically obtained from haptoglobin initially isolated from patient serum. The biomarker work involves an initial discovery phase performed by mass spectrometry–based techniques where our target marker glycopeptides are identified, and the best performing markers are selected for further study. These glycopeptide markers can then be analyzed by PRM-MS where the time on the HPLC separation gradient is known and the mass of the glycopeptides can be monitored in the mass spectrometry assay (74, 75). This can be performed for many target glycopeptides so that our assays have been multiplexed for several markers to develop the optimal assay. The markers can then be combined with other AFP or various clinical variables to develop an assay with the best performance in terms of the sensitivity/specificity for early detection.

41.2 Materials

41.2.1 Reagents

Sequencing-grade trypsin and GluC were purchased from Promega (Madison, WI, USA). The 7KDa MWCO Zeba Spin Desalting Columns were purchased from Thermo Scientific (Rockford, IL, USA). Haptoglobin (Hp) standard protein was purchased from Abcam. Other reagents were from Sigma (St. Louis, MO, USA).

41.2.2 Serum Samples

Serum was used in our studies. Serum samples were obtained by consenting patients from UT Southwestern Medical Center, Dallas, Texas, and the University of Michigan Hospital, Ann Arbor, MI. These included samples of MASLD cirrhosis, early-stage MASLD HCC, and late-stage MASLD HCC. Samples were aliquoted and stored at −80 °C, without prior

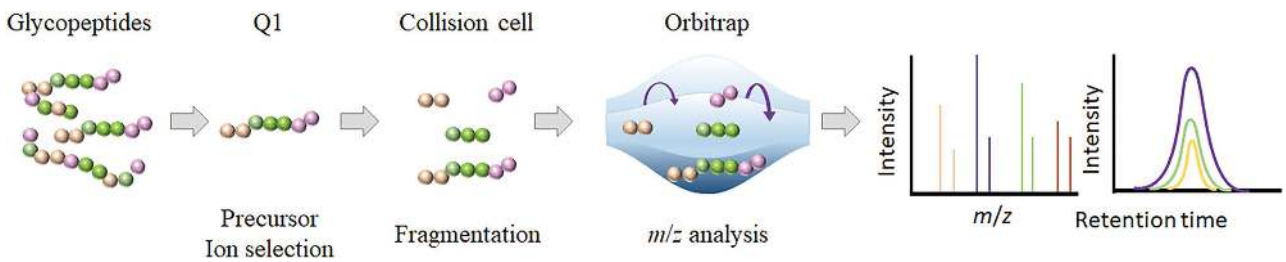


FIGURE 41.1 A schematic showing the basics of PRM in mass spectrometry.

thaw cycles. For these glycoprotein studies it is essential not to freeze thaw the samples. Samples were approved by the IRBs at each institution. As noted previously (85, 86), clinical annotations of samples were obtained from a prospectively maintained database of treatment-naïve patients. MASLD cirrhosis and MASLD HCC were defined clinically by the presence of the metabolic syndrome including obesity, diabetes, dyslipidemia, and absence of other causes of chronic liver disease such as viral hepatitis and alcohol abuse (87). HCC was defined per criteria from the American Association for the Study of Liver Disease, that is, radiographically or histologically (88, 89). The clinical features of patients in one of our studies are summarized in Table 41.1. These clinical parameters are important for further combinations with the glycopeptide markers. The serum samples were classified as NASH (also known as MASLD) cirrhosis and NASH HCC. Early-stage disease was defined by Milan criteria, which is the most common criteria for liver transplantation in the United States. The tumor, node, and metastasis (TNM) score indicates the stage of disease. The analysis was performed using R 4.0.5.

TABLE 41.1

Clinical Characteristics of Individual Patients with NASH[a] for Investigation

Variable	Cirrhosis	Early HCC	Late HCC	P-Value
N	46	32	17	
Gender (M/F)	13/33	15/17	11/6	0.026§
Age	61.5 [25, 84]	70.7 [60, 91]	64 [43, 78]	< 0.001
Laboratory				
AFP	3 [1.4, 10]	5 [2.0, 310.4]	101.9 [4, 60500]	< 0.001
Total_bilirubin	0.7 [0.2, 4.1]	0.7 [0.2, 4.7]	1.6 [0.3, 3.1]	0.121
INR	1.1 [0.9, 2.6]	1.1 [0.9, 2.5]	1.1 [1.0, 1.9]	0.527
Creatinine	0.94 [0.48, 7.4]	0.95 [0.59, 3.69]	0.83 [0.41, 7.47]	0.249
Score				
MELD_score	8 [0,24]	8.5 [1, 21]	8 [1, 30]	0.775
CTP score	6 [5, 10]	5 [1, 10]	6 [2, 11]	0.058
Ascites (%)				0.559§
1. None	35 (76.1)	23 (71.9)	11 (64.7)	
2. Mild	9 (19.6)	9 (28.1)	5 (29.4)	
3. Severe	2 (4.3)	0 (0.0)	1 (5.9)	
TNM				
I	0	21	0	
II	0	11	0	
III	0	0	11	
IV	0	0	6	
AFP < 20ng/ml	46 (100.0)	26 (81.2)	5 (29.4)	< 0.001§
max_diameter	NA [NA, NA]	3.05 [1.3, 12.0]	14.75 [8.5, 18.5]	< 0.001

Note: This table has been reproduced with permission from ACS (82). [a]Values are presented as median with the range [Min, Max]. P-values with "§" are obtained from Fisher's exact test; all others are obtained from Kruskal-Wallis Test.

AFP, TBili, ALT, AST, INR, and creatinine values and MELD and CTP scores were provided by the UT Southwestern Medical Center. Values are presented as median with the interquartile range (IQR). AFP: alpha-fetoprotein; TBili: total bilirubin; ALT: alanine aminotransferase; AST: aspartate aminotransferase; INR: international normalized ratio; MELD: model for end stage liver disease; CTP: Child-Turcotte-Pugh.

41.2.3 LC-Stepped HCD-DDA-MS/MS and LC-Stepped-HCD-PRM-MS/MS

To obtain the parameters of the targeted precursors such as the retention time, charges, and ratio of mass to charges, a survey scan in DDA mode was required before running the PRM analysis. The dried glycopeptides were dissolved in distilled water containing 0.1% formic acid (FA) and then analyzed on an orbitrap fusion lumos tribrid mass spectrometer (Thermo) coupled with a Dionex UPLC system with conditions as described in our prior work (58, 71, 90). The mass spectrometer was set as data dependent mode with the MS1 scan range set as m/z 400–2000 and MS1 data acquired in the orbitrap (120k resolution, 4e5 AGC, 100 ms injection time) followed by stepped HCD-MS/MS acquisition with the stepped collision energies of 31.5%, 35%, and 38.5%. When the LC-Stepped HCD-PRM-MS/MS was performed, the elution linear gradient was like that used in the DDA detection mode mentioned earlier. Two differences between DDA-MS/MS and PRM-MS/MS involve the stepped collision energies (19%, 26%, 33%) which were set to fragment the glycopeptides, and the PRM analysis required predefined precursor ions. As a method of targeted quantitative analysis, the sensitivity of the detection of PRM is improved compared to the DDA detection mode. The targeted precursor ions from Hp include potential glycopeptide markers with mono- and bi-fucosylated glycans at sites N184, N207, and N241 and those with fully sialylated bi- and triantennary glycan motifs.

41.2.4 Software and Data Analysis: Byonic-Assignment of Glycopeptides

All mass spectra were searched with Byonic (Protein Metrics, San Carlos, CA) incorporated in Proteome Discoverer 2.1 (Thermo) against the UniProt human haptoglobin database (P00738) (50). A UniProt human protein database containing the complete human proteome with haptoglobin removed was also applied in data searching to evaluate the assignments of glycopeptides. A human N-glycan database containing 179 human N-glycans was employed for data searching, which combines a Byonic database of 164 mammalian N-glycans but no multiple fucoses and 15 N-glycans with multiple fucoses reported in the literature for serum Hp in HCC and liver cirrhosis patients (91–93).

Trypsin and GluC were selected as enzymes with a maximum of two missed cleavages allowed. The search was performed using the following parameters: (1) static modification, carbamidomethyl of C; (2) dynamic modifications, oxidation of M, deamidation of N and Q, and glycan modifications; (3) maximum missed cleavages, 2; (4) precursor ion mass tolerance 10 ppm; (5) fragment ion mass tolerance 0.01 Da. Results were filtered at 1% FDR and confidence threshold of Byonic score >100, and further validation was performed manually. The manual check criteria include the retention time and the presence of oxonium ions, for example m/z204.09 for HexNAc, 292.10 for NeuAc, 274.09 for NeuAcH2O, 366.13 for HexHexNAc, 512.20 for HexHexNAcFuc, and 657.23 for HexNAcHexNeuAc. The monoisotopic peak and

diagnostic oxonium ion were further checked in the case of 2Fuc − 1.02 = 1NeuAc to eliminate the false assignment of NeuAc to 2Fuc. In the main text, glycans were described as AxGxFxSx according to the Oxford glycan nomenclature, that is, Ax, number (x) of antenna; Gx, number (x) of linked galactose on antenna; Fx, number (x) of fucose; Sx, number (x) of sialic acids. For example, A3G3F2S2 represents the triantennary trigalactosylated bifucosylated bisialylated glycan (90).

41.2.5 Byologic-Quantification of Glycopeptides

Automated quantitative analysis was performed using Byologic (Protein Metrics Inc.), which uses as input both MS1 raw data and Byonic search results. Byologic can compare the quantitative results among samples so all the raw data and Byonic search results from 30 MS runs were uploaded into one Byologic project for comparison. Byologic incorporates the list of glycopeptides identified, the MS/MS spectrum, XIC, and precursor isotope plot in one single window, allowing one to gather complete information for quantitation analysis. The isotope plot of the precursor ion is one of the features of Protein Metrics software. Individual scans are used to find the isotopes within a given tolerance window (i.e., 18 ppm), and the isotope m/z apex is used to determine which m/z value to use for XIC creation. The use of a single isotope has previously been empirically justified by demonstrating more consistent XIC behavior and thus was used in our prior work (90). The peak area of the XIC of a given glycopeptide was automatically integrated and normalized against that of total glycopeptides of Hp identified in each MS run.

The abundance of a site-specific glycoform was represented by the sum of all glycopeptides containing the same glycan at the glycosite, where the peptide sequences may differ in the case of miscleavage or deamidation. The relative quantitation of each site-specific glycopeptide was achieved at two levels: the total glycopeptide level and its corresponding glycosite level. Each sample was analyzed with two injections, and the average normalized abundance of a given N-glycopeptide from the two replicates was used to measure its relative levels between disease groups.

41.2.6 Skyline for Quantitative Analysis of Glycopeptide from PRM

The Skyline platform was used to quantitatively analyze the selected glycopeptides from PRM results. Like DDA analysis, oxonium ions, including HexNAc, NeuAc, HexNAcHex, HexHexNAcFuc, and HexNAcHexNeuAc and other possible b/y ions, were used for glycopeptide identification, while the Y1 ion (peptide+HexNAc) was used for quantification (58). Skyline analysis requires peptide settings, transition settings, and an.ms2 file converted from the survey scan raw data. A fasta file (P00738 from UniProt human database) of haptoglobin was uploaded as the background protein database. Peptide settings are required to create a library (.ssl file) with the parameters taken from a DDA survey scan, such as the glycopeptide sequence, scan number, retention time, charges, and so on. The transition settings and the procedures for quantification are as described in our recent publication (58, 71).

41.2.7 Statistical Analysis and Software

The abundance of site-specific glycopeptides in early HCC, cirrhosis, and healthy controls was compared using GraphPad Prism 7 (La Jolla, CA). The reproducibility analysis of four replicates of a normal serum sample was performed with the Perseus software as in our earlier study (90). The heatmap was generated with HemI 1.0.39. The significance of the difference in site-specific glycopeptides between groups was evaluated using Tukey's multiple comparisons test in ANOVA. The threshold $p < 0.05$ was used to determine the differentially expressed glycopeptides. The scatter plots of the differentially expressed site-specific N-glycopeptides between early HCC and cirrhosis were generated with GraphPad Prism 7 (La Jolla, CA) (90).

41.3 Methods

41.3.1 Glycopeptides Preparation

Figure 41.2 illustrates the experimental workflow of the study. Each of these steps will be further discussed in detail.

41.3.1.1 Removal of Serum High-Abundance Proteins

High-abundance proteins such as albumin, immunoglobulins, and transferrin make up a substantial part of serum, and they can hinder the accurate measurement of low-abundance proteins due to their high complexity and wide dynamic range. These high-abundance proteins therefore pose a challenge for the detection of low-abundance proteins, such as cytokines, that are typically found at ng/mL levels in serum. Thus, removing low-abundance proteins is essential for the study of serum.

There are several immunoaffinity-based commercial kits, reagents for precipitating proteins as well as nanocomposite-based methods for HAPs depletion. Here, we introduce a commonly used immunoaffinity-based high-abundance proteins depletion method (Cat. 5188–6560, Multiple Affinity Removal Spin Cartridge Human 14, Agilent).

Step 1. Draw 4 mL of Buffer A into the syringe then dispense Buffer A slowly through spin cartridge to prepare resin and to remove any trapped air bubbles.

Step 2. Dilute 10 μL of serum sample with Buffer A to a final volume of 200 μL. Add 200 μL of diluted serum sample to top of resin bed, then centrifuge for 1 minute at 100 × *g*. Collect the flow-through fraction in the collection tube F1.

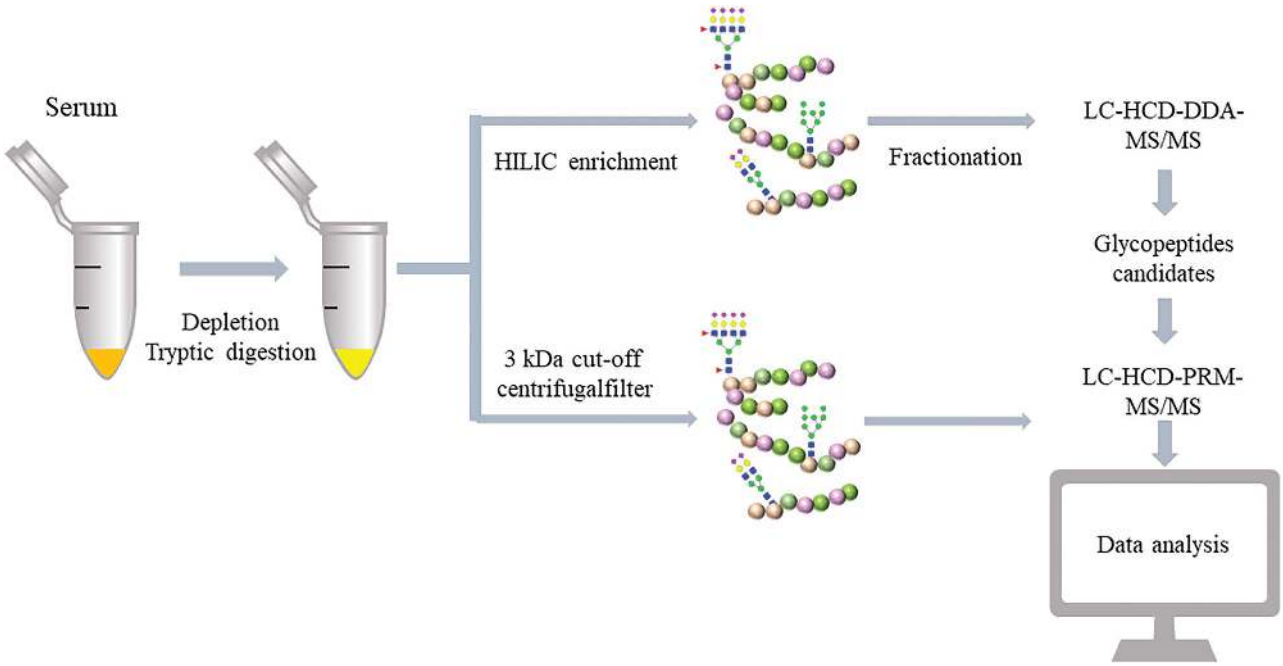


FIGURE 41.2 Workflow of quantitative LC-Stepped HCD analysis of intact N-glycopeptides derived from serum of patients between different groups.

Step 3. Add 400 µL of Buffer A to top of resin bed and centrifuge for 2.5 minutes at 100 × *g*. Collect the flow-through into the same F1 collection tube used in step 2.

Step 4. Perform second wash with Buffer A and collect F2 flow-through fraction (repeat step 3).

Step 5. Load 2.5 mL of Buffer B to the cartridge to elute the bound fraction, and use additional 4 mL of Buffer A to re-equilibrate the cartridge with Buffer A. Then, the spin cartridge is ready for processing the next sample.

Step 6. Combine the F1 and F2 fractions for the next step such as removal of the buffer or buffer exchange by filters.

41.3.1.2 Filter Processes

The proteins collected earlier were then processed for desalting using an Ultra-4 centrifugal filter (10 kDa MWCO, Millipore) and then dried down in a SpeedVac concentrator. This step allows for buffer exchange, protein concentration, or desalting to be achieved. Then, the proteins are usually digested by enzymes.

41.3.1.3 Double Enzymatic Digestion

Step 1. Dissolve purified serum Hp (~3 µg) in 50 mM NH_4HCO_3.

Step 2. Reduce with 20 mM dithiothreitol (DTT) and alkylate with 50 mM iodoacetamide (IAA).

Step 3. Add 0.2 µg of trypsin (Promega, Madison, WI) and incubate at 37 °C overnight.

Step 4. Quench the reaction at 95 °C for 10 min.

Step 5. Add 0.2 µg of GluC (Promega, Madison, WI) and incubate at 37 °C for 12 h.

Step 6. Dry down the sample.

41.3.1.4 Hydrophilic Interaction Liquid Chromatography (HILIC) Enrichment for Glycopeptides

The glycopeptides enriched by HILIC have been widely used. Some commercial tips with HILIC resin are time-effective for glycopeptides preparation. Here, we introduce the enrichment of glycopeptides by HILIC TopTip (Glygen, Columbia, MD).

Step. 1. Activate the HILIC TopTips with 20 µL water three times.

Step 2. Equilibrate the HILIC TopTips with the binding buffer (85% acetonitrile containing 15 mM ammonium acetate, pH 3.5).

Step 3. Resuspend the peptides in binding buffer and load onto the HILIC TopTip, where the flow-through was collected and reloaded, three times to ensure complete binding.

Step 4. After glycopeptide capture, wash the HILIC TopTip three times with 10 µL of 85% ACN containing 15 mM ammonium acetate, pH3.5.

Step 5. Elute the bound glycopeptides with 10 µL of water twice and collect.

Optional step: a 3kDa MWCO centrifugal filter can be used as an alternative to enrich the glycopeptides.

41.3.2 Glycopeptides Analyzed by Mass Spectrometer

41.3.2.1 LC-Stepped HCD-DDA-MS/MS

To obtain the retention time of the targeted precursors, a survey scan in DDA mode was performed before running the PRM analysis.

Instrument: orbitrap fusion lumos tribrid mass spectrometer (Thermo), Dionex UPLC system

Reagents: formic acid, acetonitrile (ACN).

Step 1. The dried fractions of glycopeptides are dissolved in 0.1% formic acid (FA).

Step 2. The binary solvent system was composed of 0.1% FA in H_2O (A) and 80% ACN/0.1% FA (B).

Step 3. Samples were separated using a 75 μm × 50 cm column (C18, 2 μm, 100 Å; Acclaim PepMap RSLC, Thermo) under a 90 min linear gradient from 2% to 40% B at a flow rate of 300 nL/min.

Step 4.1. The MS instrument was operated in data dependent mode.

Step 4.2. The MS1 scans (m/z 400−1800) were acquired in the orbitrap (120k resolution, 4×10^5 AGC, 100 ms injection time).

Step 4.3. Orbitrap was set at 60k resolution, 2×10^5 AGC, 250 ms injection time.

Step 5. The stepped collision energies were set as 31.5%, 35%, and 38.5%.

Step 6. Data acquisition.

41.3.2.2 LC-Stepped-HCD-PRM-MS/MS

The LC-MS system was similar to that used in the DDA detection mode already mentioned.

Step 1. The gradient of the binary solvent used a 90 min linear gradient from 2% to 40% B at a flow rate of 300 nL/min.

Step 2. The stepped collision energy CE was set as 19%, 26%, and 33%.

Step 3. Targeted precursor ions were defined in the PRM mass table.

Step 4. Data acquisition.

41.3.2.3 Data Analysis

For the PRM results, we used the Skyline software for quantification of the selected glycopeptides. Similar to DDA analysis, oxonium ions GlcNAc, NeuAc, GlcNAc-Hex, HexHexNAcFuc, and HexNAcHexNeuAc and other possible b/y ions were used for glycopeptides identification, while the Y1 ion (peptide+HexNAc) was used for quantification. For Skyline analysis, peptide settings and transition settings are required.

Step 1. Of the peptide settings, a library (.ssl file) was created with the parameters retrieved from a survey scan before PRM detection, including glycopeptide sequence, scan number, retention time, charges, and an .ms2 file converted from the survey scan raw data.

Step 2. A fasta file of the targeted proteins was uploaded as the background protein database.

Step 3. Of the transition settings, the precursor charges were set as +2 to +5, the ion charges were set as +1 and +2, and the ion types were set as y and b ions.

Step 4. The ion match tolerance was set as 0.05 m/z.

Step 5. The peptide backbone with one HexNAc moiety (pep+HexNAc) was defined as Y1 ion for quantitation.

Step 6. The integral peak areas of the Y1 ions were exported from Skyline in .csv format, which were used for glycopeptide quantification and data normalization. Relative abundance of each glycopeptide was calculated by normalizing its peak area to the sum of peak areas of all targeted glycopeptides in each sample.

Step7. The relative abundances of the target glycopeptides were compared between different disease state groups and the scatter plots were created by statistical software.

41.3.2.4 Application

Herein, serum haptoglobin (Hp) will be used as an example to further illustrate the application of PRM in mass spectrometry quantitative protein methodology. Serum Hp containing four glycosites at N184, N207, N211, and N241 is an abundant serum glycoprotein secreted into the bloodstream primarily by the liver, and these glycopeptides are promising biomarkers for liver cancer prognosis.

Step 1. Haptoglobin purification.

An in-house antibody-immobilized HPLC column was used to purify the Hp protein from 20 μL of each individual patient serum.

Step 2. Enzymatic digestion, glycopeptide enrichment.

Enzymatic digestion was performed with trypsin and GluC. After that, the glycopeptides were enriched using HILIC TopTips.

Step 2.1. Purified serum Hp (~3 μg) was dissolved in 50 mM NH_4HCO_3, reduced with 20 mM dithiothreitol (DTT), and alkylated with 50 mM iodoacetamide (IAA).

Step 2.2. 0.2 μg of trypsin (Promega, Madison, WI) was added and incubated at 37 °C overnight. Then, quenching of the reaction at 95 °C for 10 min.

Step 2.3. 0.2 μg of GluC (Promega, Madison, WI) was added and incubated at 37 °C for 12 h.

Step 2.4. The sample was dried down and then reconstituted in 15 μL of 85% acetonitrile containing 15 mM ammonium acetate, pH 3.5, where glycopeptides were enriched using HILIC TopTips (Glygen, Columbia, MD).

Step 3. LC-Stepped HCD-PRM-MS/MS.

The enriched glycopeptides were dissolved in 0.1% formic acid (FA) and analyzed on the orbitrap fusion lumos tribrid mass spectrometer (Thermo) coupled with a Dionex UPLC system. A binary solvent system was composed of 0.1% FA in H_2O (A) and 80% CH_3CN/0.1% FA (B).

Step 3.1. Samples were separated on a 75 μm × 50 cm column (C18, 2 μm, 100 Å; Thermo) under a 90 min linear gradient from 2% to 40% B at a flow rate of 300 nL/min.

Step 3.2. The MS instrument was operated in PRM mode. The MS1 scans (m/z 400–2000) were acquired in the orbitrap (120 K resolution, 4e5 AGC, 100 ms injection time) followed by stepped HCD-MS/MS acquisition with the stepped collision energies of 31.5%, 35%, and 38.5%.

Step 3.3. A list of precursor ions of the target glycopeptides.

Step 3.4. Data acquisition.

Step 4. Data interpretation and relative quantitation.

All spectra from DDA results were searched with Byonic software (Protein Metrics), and a UniProt human Hp database (P00738) was used for data searching.

Step 4.1. The theoretical m/z of the oxonium ions in glycopeptides from HCD-MS were used to check the fragment ions.

Step 4.2. The Skyline platform was used to quantitatively analyze the selected glycopeptides from PRM results.

Step 4.3. Oxonium ions, including HexNAc, NeuAc, HexNAc-Hex, HexHexNAcFuc, and HexNAcHexNeuAc, and other possible b/y ions, were used for glycopeptide identification, while the Y1 ion (peptide+HexNAc) was used for quantification.

Step 4.4. A fasta file (P00738 from UniProt human database) of haptoglobin was uploaded as the background protein database.

Step 4.5. A library (.ssl file) was created with the parameters taken from a DDA survey scan, such as the glycopeptide sequence, scan number, retention time, charges, and so on.

Step 4.6. Peptide settings and transition settings were based on survey scan raw data.

Step 4.7. Data were exported for further analysis.

41.4 Results

41.4.1 Glycopeptides Identified as Potential Biomarker Candidates by LC-Stepped-HCD-PRM-MS/MS

To confirm the differential expression of glycopeptides for diagnostic potential in the detection of NASH HCC, LC-Stepped-HCD-PRM-MS/MS was performed for targeted quantitative analysis of the selected glycopeptide candidates among different liver disease states, including 46 cirrhosis, 32 early-stage HCC, and 17 late-stage HCC (82). The selection of the targeted glycopeptides for PRM-MS was based on our previous report on the site-specific glycopeptides of Hp showing a significant difference during the progression from NASH cirrhosis to HCC by differential LC-DDA-MS/MS analysis (52). The glycopeptide marker candidates were predominantly found with fucosylated or fully sialylated glycan motifs.

In our current study, 49 precursor ions originating from 32 site-specific glycopeptides, including mono- and bi-fucosylated glycoforms at sites N184, N207, and N241, were selected for PRM-MS analysis among different liver disease states. Both charge states 3 and 4 of the precursor ion were included for PRM-MS in some glycopeptides based on the MS profile observed in the DDA mode. Based on the relative abundance of each glycopeptide quantified, the expression of six glycopeptides showed statistically significant differences between patients with HCC and those with cirrhosis as shown in Figure 41.3. The relative abundance of each marker was normalized by their peak, and the normalized value was used for the ROC analysis.

These six glycopeptides involved two glycosites—N184 and N241. The two glycopeptides on N184 were N184_A2G2S2 (Figure 41.3a) and N184_A3G3F1S3 (Figure 41.3b). There are four glycopeptides derived from the same peptide backbone containing N241 including the bi-antennary glycoform N241_A2G2F1S2 (Figure 41.3c), tri-antennary glycoform N241_A3G3F1S3 (Figure 41.3d), tetra-antennary glycoform N241_A4G4F1S3 (Figure 41.3e), and N241_A4G4F2S4 (Figure 41.3f). All four glycopeptides were increased significantly during the progression from cirrhosis to HCC. It is notable that the non-fucosylated glycopeptide N184_A2G2S2 was decreased during the progression of HCC, whereas the other five fucosylated glycan forms at site N184 and N241 were elevated in HCC compared to cirrhosis, indicating the complexity of fucosylation and sialylation in HCC as previously reported (52, 92). The expression of these six glycopeptides also presented statistically significant differences between patients with early-stage HCC and those with cirrhosis (Figure 41.3g–l).

41.4.2 Diagnostic Performance of Site-Specific N-Glycopeptides in All HCC and Early HCC

We performed the receiver operating characteristic (ROC) analysis for Hp site-specific N-glycopeptides in differentiating all HCCs and early HCCs from cirrhosis patients, respectively. The estimated AUC as well as its 95% confidence interval (CI) were summarized for each individual marker in Table 41.2. AFP had an estimated AUC of 0.790 (95% CI: 0.697, 0.872), and the sensitivity was 52.3% at 90% specificity when comparing all HCC vs. cirrhosis. The six glycopeptides derived from two glycosites N184 and N241 discussed previously were all expressed statistically significantly different between cirrhosis and all HCC groups. The estimated AUCs of two glycopeptides involving the site N184 were 0.776 (95% CI: 0.676, 0.866) and 0.719 (95% CI: 0.606, 0.820), respectively. Although these two glycopeptides were significantly different between cirrhosis and HCC, they did not yield a better AUC than AFP. The four glycopeptides involving glycosite N241 included one bi-antennary glycopeptide N241_A2G2F1S2, one tri-antennary

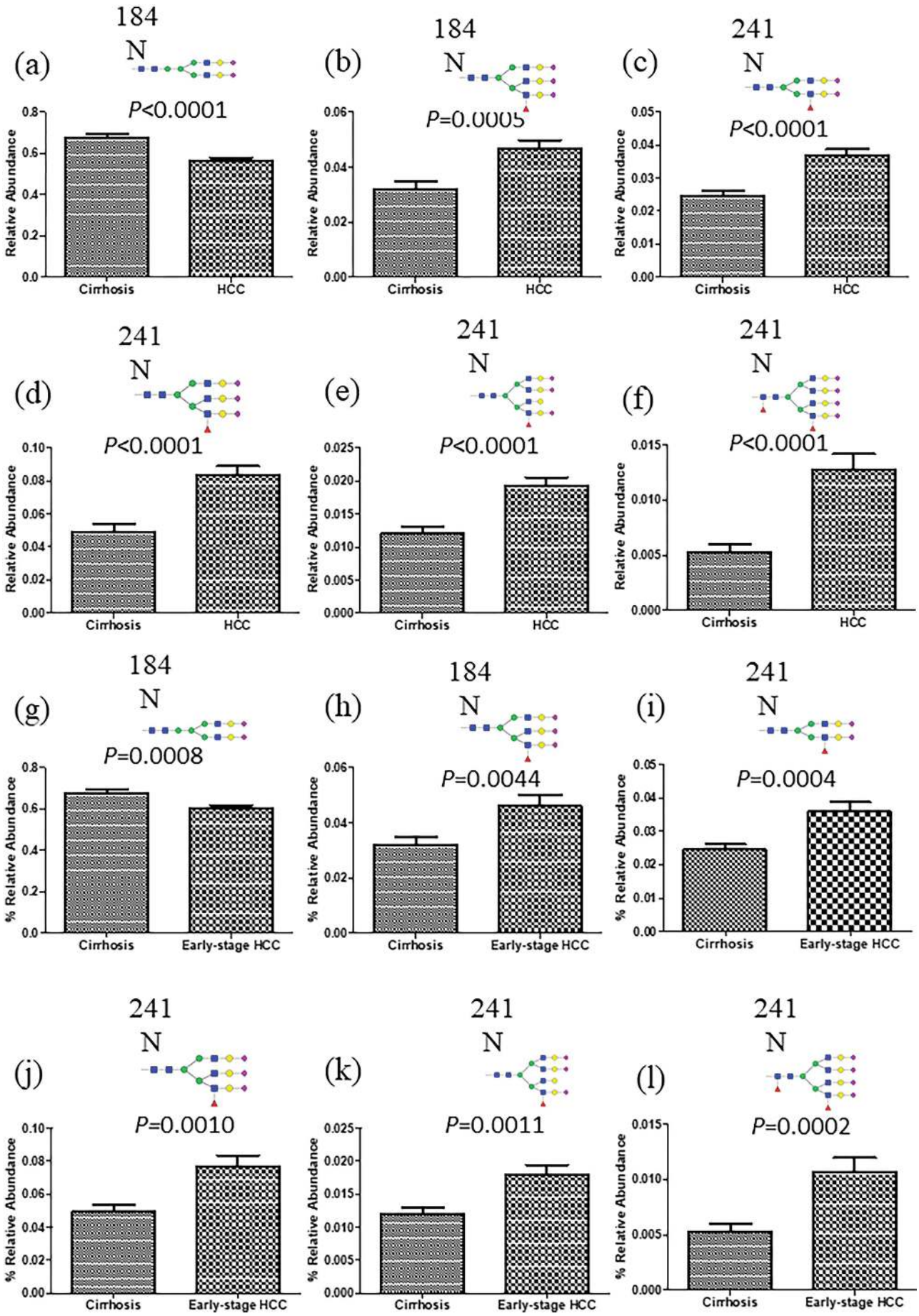


FIGURE 41.3 Histograms of the six differentially expressed site-specific N-glycopeptides between HCC and cirrhosis (a–f) as well as between early-stage HCC and cirrhosis (g–l). (Figure 41.3 reproduced with permission from ACS [83].)

TABLE 41.2

The Estimated AUC of Individual Markers for All HCCs and Early-Stage HCCs, Respectively

	Cirrhosis (n = 46) vs all HCC (n = 49)			Cirrhosis (n = 46) vs early HCC (n = 32)		
Marker	AUC	95% CI	Sensitivity at 90% Specificity	AUC	95%CI	Sensitivity at 90% Specificity
AFP	0.790	(0.697, 0.872)	52.3%	0.715	(0.592, 0.827)	39.5%
N184_A2G2S2	0.776	(0.676, 0.866)	40.8%	0.703	(0.579, 0.819)	21.9%
N184_A3G3F1S3	0.719	(0.606, 0.820)	24.5%	0.682	(0.552, 0.806)	21.9%
N241_A2G2F1S2	0.765	(0.665, 0.853)	34.7%	0.732	(0.617, 0.832)	37.5%
N241_A3G3F1S3	0.796	(0.693, 0.883)	46.9%	0.729	(0.611, 0.836)	31.3%
N241_A4G4F1S3	0.758	(0.657, 0.853)	32.7%	0.704	(0.579, 0.818)	25.0%
N241_A4G4F2S4	0.793	(0.700, 0.874)	44.9%	0.757	(0.643, 0.861)	34.4%

Note: Table 41.2 has been reproduced with permission from ACS (82).

glycopeptide N241_A3G3F1S3, and two tetra-antennary glycopeptides N241_A4G4F1S3 and N241_A4G4F2S4. N241_A3G3F1S3 and N241_A4G4F2S4 had similar estimated AUC as AFP, 0.796 (95% CI: 0.693, 0.883), and 0.793 (95% CI: 0.700, 0.874), respectively.

When comparing early-stage HCCs to cirrhosis, AFP had an estimated AUC of 0.715 (95% CI: 0.592, 0.827). It had a sensitivity of 39.5% at 90% specificity. Like the results of all HCCs together, the results from glycopeptides involving site N184 did not perform better than that of AFP alone. However, the glycopeptides N241_A2G2F1S2, N241_A3G3F1S3, and N241_A4G4F2S4 showed slightly better AUC, and their AUCs were 0.732 (95% CI: 0.617, 0.832), 0.729 (95% CI: 0.611, 0.836) and 0.757 (95% CI: 0.643, 0.861), respectively.

41.4.3 Diagnostic Performance of Combinatorial Analysis of Hp N-Glycopeptide Markers with AFP

To improve the discrimination performance of AFP, we sought the best combination panels. Considering six candidate glycopeptide markers, as well as age and gender, using AFP as the anchor marker, we built the 2-marker and 3-marker panels. The marker panels with the highest AUC were selected as the best panel for future validation. Table 41.3 summarizes the results of the best 2- and 3-marker panels. The P values for comparing each selected marker panel with the panel using AFP alone were calculated based on the bootstrapping method.

All the 2-marker combination models showed AUC greater than 0.841 in distinguishing all HCC samples from cirrhosis and better than the AUC of 0.790 using AFP alone. N241_A4G4F2S4 combined with AFP had the best estimated AUC in all 2-marker combinations with an estimated AUC of 0.898 (95% CI:0.835, 0.951), which was significantly better than AFP alone (P = 0.0481) (82). This best 2-marker panel had a sensitivity of 63.3% at 90% specificity, whereas AFP alone had a sensitivity of 52.3%. When using three-fold cross-validation, it had an AUC of 0.904 (95% CI: 0.844, 0.960). The other three glycopeptides on site 241, involving N241_A2G2S1F2, N241_A3G3F1S3, and N241_A4G4F1S3, achieved AUC values of 0.870 (82), 0.869, and 0.841respectively. The sensitivity for each at 90% specificity was 59.2%, 57.1%, and 61.2%. At site N184, the AUC for the combination for glycopeptide N184_A3G3F1S3 + AFP was estimated to be 0.843 (95% CI: 0.763, 0.917). And the panel of N184_A2G2S2 + AFP had an AUC of 0.869 (95% CI: 0.795, 0.931). Notably, this glycopeptide's expression is decreased in distinguishing between HCCs to cirrhosis, where there is no Sialyl Lewis epitope.

In the comparison of early-stage HCC vs. cirrhosis, the panel of AFP+N241_A4G4F2S4 had the best estimated AUC of 0.845 (95% CI: 0.757, 0.919), which was marginally significantly better than the AUC of 0.715 (95% CI: 0.592, 0.827) when using AFP alone (P = 0.0691) (82). It had sensitivity of 50.0% at 90% specificity, whereas AFP alone had a sensitivity of 39.5%. Its AUC was estimated to be 0.823 (95% CI: 0.750, 0.923) in three-fold cross-validation.

Like the results from all HCCs, in early-stage HCC, the AUC values of the other three glycopeptides N241_A2G2F1S2, N241_A3G3F1S3, and N241_A4G4F1S3 were larger than 0.762, and they had better sensitivity at 90% specificity than AFP alone in distinguishing early-stage HCCs from cirrhosis. At site N184, the AUC value of N184_A2G2S2 combined with AFP was 0.806, and the AUC of N184_A3G3F1S3 + AFP was 0.777. It is notable that the glycopeptide N241_A4G4F2S4 combined with AFP provided the best AUC values among these glycopeptides in distinguishing either all HCCs or early-stage HCC from cirrhosis.

Since both age and gender showed significant differences between cirrhosis and HCC samples, we also tested the performance of the 2-marker panels of AFP+age and AFP+gender. When differentiating all HCC, the panels of AFP+age and AFP+gender had the estimated AUC being 0.846 (95% CI: 0.760, 0.916) and 0.811 (95% CI: 0.714, 0.890), respectively. When differentiating early HCC, the panels of AFP+age and AFP+gender had the estimated AUC of 0.832 (95% CI: 0.734, 0.906) and 0.761 (95% CI: 0.650, 0.858), respectively. None of them had significant improvement compared to AFP alone.

41.4.4 Diagnostic Performance of Combinatorial Analysis Using 3-Marker Panels

When considering 3-marker panels and using AFP as the anchor marker, the best panel was selected based on the estimated AUC. The performance of the best panel was

summarized in Table 41.3, and Figure 41.4 showed the ROCs. In the results for all HCC vs. cirrhosis, the combination of AFP+N241_A2G2F1S2+N241_A4G4F2S4 performed the best with the AUC estimated to be 0.928 (95% CI: 0.877, 0.970), with 65.3% sensitivity at 90% specificity. It was statistically better than the panel using AFP alone ($P = 0.0083$). When using three-fold cross-validation, it had AUC of 0.923 (95% CI: 0.874, 0.975). In the comparison of early-stage HCCs vs. cirrhosis, the combination of AFP+age+N241_A2G2F1S2 performed the best with an AUC value achieved of 0.902 (95% CI: 0.829, 0.961), with 71.9% sensitivity at 90% specificity. It was statistically better than the panel using AFP alone ($P = 0.0048$). When using three-fold cross-validation, it had an AUC of 0.885 (95% CI: 0.861, 0.967).

41.5 Discussion

In our work, we have explored the use of PRM-MS as a liquid biopsy for biomarker analysis from patient serum. The significant advantage of PRM-MS is that it is a targeted quantitative analysis as compared to DDA-MS, which scans over the entire mass range. The increased sensitivity of PRM is due to its targeted nature where the mass spectrometer can spend

TABLE 41.3

The Best 2-Marker and 3-Marker Panels in Differentiating All HCCs and Early-Stage HCCs, Respectively

Cirrhosis (n = 46) vs all HCC (n = 49)

	Best Model	AUC	95% CI	Sensitivity at 90% Specificity	P
	AFP	0.79	(0.697, 0.872)	52.30%	NA
2-marker	AFP+N241_A4G4F2S4	0.898	(0.835, 0.951)	63.30%	0.0481
3-marker	AFP+N241_A2G2F1S2+N241_A4G4F2S4	0.928	(0.877, 0.970)	65.30%	0.0083

Cirrhosis (n = 46) vs early HCC (n = 32)

	Best Model	AUC	95% CI	Sensitivity at 90% Specificity	P
	AFP	0.715	(0.592, 0.827)	39.50%	NA
2-marker	AFP+N241_A4G4F2S4	0.845	(0.757, 0.919)	50.00%	0.0691
3-marker	AFP+age+N241_A2G2F1S2	0.902	(0.829, 0.961)	71.90%	0.0048

Note: Table 41.3 has been reproduced with permission from ACS (82). P-values were based on the bootstrapping method of 1000 bootstrapping samples; NA: Not applicable.

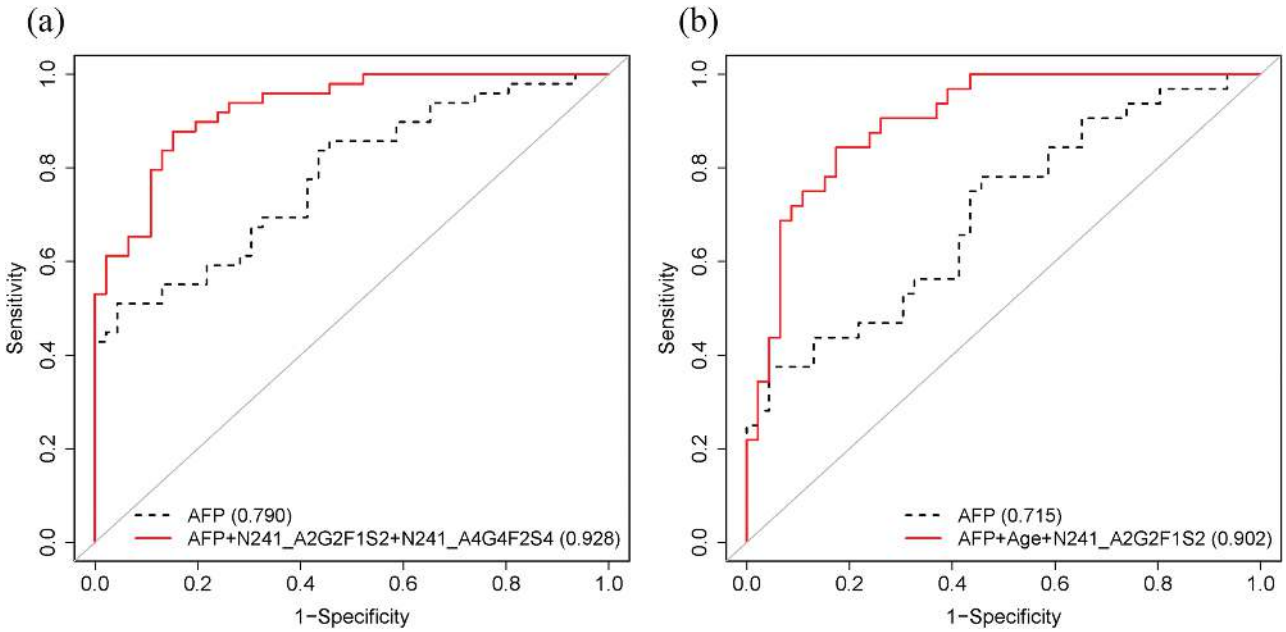


FIGURE 41.4 Receiver operating characteristic (ROC) curves of the best 3-marker panel to differentiate all HCCs (a), and early HCCs (b) from cirrhosis patients. Black dashed line, AFP only; red solid line, the combination of Hp N-glycopeptide and AFP. (Figure 41.4 reproduced with permission from ACS [83].)

longer times collecting the target peptide ions, which is termed dwell time, compared to DDA-MS. Another important aspect of PRM-MS is its multiplexing ability where large numbers of markers can be run together. PRM allows all fragment ions of a predefined precursor to be measured in parallel, where a full MS2 product ion spectrum of a specified precursor ion can be acquired, and all detectable product ions can be simultaneously monitored at high accuracy and resolving power (72, 81, 82). In addition, the PRM-MS can provide a wide dynamic range for quantitative analysis for target precursors. These aspects of PRM allow accurate quantification of a larger number of target precursors in an anticipated elution time interval. Also, an important aspect of PRM-MS is that antibodies are not required for these assays. This is especially important in these studies where antibodies are not readily available for these glycosylated peptides.

In the results of the PRM-MS confirmation set that we showed herein, we found that six fucosylated glycopeptides from serum Hp could differentiate HCC from cirrhosis in patients with MASLD using LC-Stepped HCD-PRM-MS/MS ($P < 0.05$). These six glycopeptides involved two glycopeptides at site N184 and four glycopeptides at site N241. At site N184, the expression of the glycopeptide N184_A2G2S2 without Sialyl Lewis epitopes decreased in distinguishing all HCCs from cirrhosis. Another glycopeptide involving site N184, N184_A3G3F1S3, and the four glycopeptides with N241 glycans with Sialyl Lewis epitopes were expressed differentially during disease progression.

The ROC curves of all these glycopeptides when combined with AFP performed better than AFP alone. Of these fucosylated glycopeptides, a tetra-antennary N241_A4G4F2S4 provided the best performance with an AUC value of 0.898 (95% CI: 0.835, 0.951, sensitivity: 63.3%, specificity: 90%) for the comparison of all HCCs versus cirrhosis. When considering 3-marker panels, the AUC value of the combination of AFP+N241_A2G2F1S2+N241_A4G4F2S4 performed the best in all HCCs. These results demonstrated that the bifucosylated tetra-antennary glycopeptides N241_A4G4F2S4 which bear the Sialyl Lewis (SLe) epitope may potentially play a role in distinguishing HCC from cirrhosis. The bifucosylated tetra-antennary glycopeptide at site N241 contains both a core-fucosylation and an outer-arm fucosylation. These markers are promising for early detection of HCC from this initial confirmation set, but further work on validation using a larger blinded set will be required in the future.

41.6 Future Trends

In conclusion, it should be noted that several companies are setting up PRM-MS assays for biomarker monitoring. Such mass spec–based assays are already in use for small molecule analysis, but for peptide and large molecule analysis the industry still needs development. Although mass specs are expensive devices, the mass specs for the clinical laboratory are generally stripped-down versions of complex research-grade instruments. It should be noted that clinical labs make so much profit on these assays that the instrument rapidly pays for itself. The ability to perform multiplex analysis on large numbers of markers is a genuine advantage of this method, and the ability to automate it to run large number of samples is a significant advantage. Also, the ability to eliminate antibodies is essential where antibodies are not readily available for these glycopeptides, but ELISAs depend on well-validated antibodies.

The PRM technique, when used to analyze small peptide sets (usually less than 50 peptides) in a complex mixture, enhanced the detection and quantification capabilities compared to selected reaction monitoring (SRM). However, implementing PRM for analyzing large numbers of peptides requires adjusting MS acquisition parameters, and thus impacts the data quality and, consequently, the overall experiment output. Gallien and colleagues (94) developed the internal standard triggered-parallel reaction monitoring (IS-PRM) technique by incorporating an internal standard and dynamically adjusting data acquisition parameters, which allows for precise quantification of endogenous peptide sequences, ensuring accurate mass spectrometry results even when analyzing a moderate-to-large number of peptide data. The IS-PRM method demonstrates exceptional performance when analyzing extensive peptide sets (up to 600 peptides) in complex samples, achieving remarkable quantification limits in the low amol range. We speculate that PRM-MS will find broad applicability for early detection, prognosis, and therapeutic monitoring for various cancers based on a minimally invasive assay using patient serum. It will also have applications to numerous other illnesses that can be monitored through blood biomarkers, including Alzheimer's and mild cognitive impairment and potentially also Parkinson's and other related conditions (95).

Acknowledgments

Funding sources: We acknowledge the support of this work from the National Cancer Institute under grants 1R01 CA160254 (DML), U01 CA225753 (DML), R50 CA221808 (JZ), and 5U01CA27188 (NP and AS). DML acknowledges support under the Maud T. Lane Professorship. We also acknowledge the support of this work from Protein Metrics Inc. for the use of the Byonic/Byologic software. Further we knowledge ACS publications for granting copyrights permission to reuse our work (82, 90). Please note that further permission related to the material excerpted should be directed to the ACS.

REFERENCES

1. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR, A global view of hepatocellular carcinoma: Trends, risk, prevention and management. *Nat Rev Gastroenterol Hepatol* 2019;16(10):589–604.
2. Rich NE, Hester C, Odewole M, Murphy CC, Parikh ND, Marrero JA, Yopp AC, Singal AG, Racial and ethnic differences in presentation and outcomes of hepatocellular carcinoma. *Clin Gastroenterol Hepatol* 2019 Feb;17(3):551–559.e1.
3. Singal AG, Llovet JM, Yarchoan M, Mehta N, Heimbach JK, Dawson LA, Jou JH, Kulik LM, Agopian VG, Marrero JA, Mendiratta-Lala M, Brown DB, Rilling WS, Goyal L,

Wei AC, Taddei TH, AASLD practice guidance on prevention, diagnosis, and treatment of hepatocellular carcinoma. *Hepatology* 2023 Dec 1;78(6):1922–1965.
4. European Association for the Study of the Liver. Electronic address EEE, european association for the study of the L. EASL clinical practice guidelines: Management of hepatocellular carcinoma. *J Hepatol* 2018;69:182–236.
5. Parikh ND, Mehta AS, Singal AG, Block T, Marrero JA, Lok AS, Biomarkers for the early detection of hepatocellular carcinoma. *Cancer Epidemiol Biomarkers Prev* 2020 Dec;29(12):2495–2503.
6. Fattovich G, Stroffolini T, Zagni I, Donato F, Hepatocellular carcinoma in cirrhosis: Incidence and risk factors. *Gastroenterology* 2004;127(5 Suppl 1): S35–S50.
7. Bacon BR, Farahvash MJ, Janney CG, Neuschwander-Tetri BA, nonalcoholic steatohepatitis: An expanded clinical entity. *Gastroenterology* 1994;107(4):1103–1109.
8. Mori S, Yamasaki T, Sakaida I, Takami T, Sakaguchi E, Kimura T, Kurokawa F, Maeyama S, Okita K, Hepatocellular carcinoma with nonalcoholic steatohepatitis. *J Gastroenterol* 2004;39(4):391–396.
9. Kamangar F, Dores GM, Anderson WF, Patterns of cancer incidence, mortality, and prevalence across five continents: Defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006;24(14):2137–2150.
10. Huang DQ, Singal AG, Kono Y, Tan DJH, El-Serag HB, Loomba R, Changing global epidemiology of liver cancer from 2010 to 2019: NASH is the fastest growing cause of liver cancer. *Cell Metabolism* 2022;34(7):969–977.e2.
11. Bruix J, Sherman M, Practice Guidelines Committee, American Association for the Study of Liver Diseases, Management of hepatocellular carcinoma. *Hepatology* 2005;42(5):1208–1236.
12. Chen C, Schmilovitz-Weiss H, Liu XE, Pappo O, Halpern M, Sulkes J, Braun M, Cohen M, Barak N, Tur-Kaspa R, Vanhooren V, Van Vlierberghe H, Libert C, Contreras R, Ben-Ari Z, Serum protein N-glycans profiling for the discovery of potential biomarkers for nonalcoholic steatohepatitis. *J Proteome Res* 2009;8(2):463–470.
13. Blomme B, Fitzpatrick E, Quaglia A, De Bruyne R, Dhawan A, Van Vlierberghe H, Serum protein N-glycosylation in paediatric non-alcoholic fatty liver disease. *Pediatr Obes* 2012;7(2):165–173.
14. Masuzaki R, Omata M, Chapter 4: Screening program in high-risk populations. In *Hepatocellular Carcinoma*, edited by McMasters KM and Vauthey JN. Springer Nature, 2010, pp. 55–68. https://doi.org/10.1007/978-1-60327-522-4_4.
15. Singal AG, Zhang E, Narasimman M, Rich NE, Waljee AK, Hoshida Y, Yang JD, Reig M, Cabibbo G, Nahon P, Parikh ND, Marrero JA. HCC surveillance improves early detection, curative treatment receipt, and survival in patients with cirrhosis: A meta-analysis. *J Hepatol* 2022 Jul;77(1):128–139.
16. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, De Notariis S, Roda E, Bernardi M, Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: Influence of HBsAg and anti-HCV status. *J Hepatol* 2001;34(4):570–575.
17. Tzartzeva K, Obi J, Rich NE, Parikh ND, Marrero JA, Yopp A, Waljee AK, Singal AG, Surveillance imaging and alpha fetoprotein for early detection of hepatocellular carcinoma in patients with cirrhosis: A meta-analysis. *Gastroenterology* 2018 May;154(6):1706–1718.e1.
18. Chong N, Schoenberger H, Yekkaluri S, Fetzer DT, Rich NE, Yokoo T, Gopal P, Manwaring C, Quirk L, Singal AG, Association between ultrasound quality and test performance for HCC surveillance in patients with cirrhosis: A retrospective cohort study. *Aliment Pharmacol Ther* 2022 Mar;55(6):683–690.
19. Schoenberger H, Chong N, Fetzer DT, Rich NE, Yokoo T, Khatri G, Olivares J, Parikh ND, Yopp AC, Marrero JA, Singal AG, Dynamic changes in ultrasound quality for hepatocellular carcinoma screening in patients with cirrhosis. *Clin Gastroenterol Hepatol* 2022 Jul;20(7):1561–1569.e4.
20. Wolf E, Rich NE, Marrero JA, Parikh ND, Singal AG, Use of hepatocellular carcinoma surveillance in patients with cirrhosis: A systematic review and meta-analysis. *Hepatology* 2021 Feb;73(2):713–725.
21. Singal AG, Tiro JA, Murphy CC, Blackwell JM, Kramer JR, Khan A, Liu Y, Zhang S, Phillips JL, Hernaez R, Patient-reported barriers are associated with receipt of hepatocellular carcinoma surveillance in a multicenter cohort of patients with cirrhosis. *Clin Gastroenterol Hepatol* 2021 May;19(5):987–995.e1.
22. Simmons OL, Feng Y, Parikh ND, Singal AG, Primary care provider practice patterns and barriers to hepatocellular carcinoma surveillance. *Clin Gastroenterol Hepatol* 2019 Mar;17(4):766–773.
23. Marrero JA, Henley KS, The role of serum biomarkers in hepatocellular carcinoma surveillance. *Gastroenterol Hepatol (N Y)* 2011;7(12):821–823.
24. Singal AG, Conjeevaram HS, Volk ML, Fu S, Fontana RJ, Askari F, Su GL, Lok AS, Marrero JA, Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. *Cancer Epidemiol Biomarkers Prev* 2012;21(5):793–799.
25. Parikh ND, Tayob N, Singal AG, Blood-based biomarkers for hepatocellular carcinoma screening: Approaching the end of the ultrasound era? *J Hepatol* 2023 Jan;78(1):207–216.
26. Flores A, Marrero JA, Emerging trends in hepatocellular carcinoma: Focus on diagnosis and therapeutics. *Clin Med Insights Oncol* 2014;8:71–76.
27. Lin Yu, Lubman David M, The role of N-glycosylation in cancer. *Acta Pharmaceutica Sinica B* 2023. https://doi.org/10.1016/j.apsb.2023.10.014; ISSN 2211-3835.
28. Marrero JA, Feng Z, Wang Y, Nguyen MH, Befeler AS, Roberts LR, Reddy KR, Harnois D, Llovet JM, Normolle D, Dalhgren J, Chia D, Lok AS, Wagner PD, Srivastava S, Schwartz M, Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology* 2009;137(1):110–118.
29. Singal AG, Tayob N, Mehta A, Marrero JA, El-Serag H, Jin Q, Saenz de Viteri C, Fobar A, Parikh ND, GALAD demonstrates high sensitivity for HCC surveillance in a cohort of patients with cirrhosis. *Hepatology* 2022 Mar;75(3):541–549.
30. Wechsler JB, Bolton SM, Amsden K, Wershil BK, Hirano I, Kagalwalla AF, Eosinophilic esophagitis reference score accurately identifies disease activity and treatment

effects in children. *Clin Gastroenterol Hepatol* 2018 Jul;16(7):1056–1063.
31. Fattovich G, Stroffolini T, Zagni I, Donato F, Hepatocellular carcinoma in cirrhosis: Incidence and risk factors. *Gastroenterology* 2004;127(5 Suppl 1): S35–S50.
32. Bacon BR, Farahvash MJ, Janney CG, Neuschwander-Tetri BA, Nonalcoholic steatohepatitis: An expanded clinical entity. *Gastroenterology* 1994;107(4):1103–1109.
33. Mori S, Yamasaki T, Sakaida I, Takami T, Sakaguchi E, Kimura T, Kurokawa F, Maeyama S, Okita K, Hepatocellular carcinoma with nonalcoholic steatohepatitis. *J Gastroenterol* 2004;39(4):391–396.
34. Kamangar F, Dores GM, Anderson WF, Patterns of cancer incidence, mortality, and prevalence across five continents: Defining priorities to reduce cancer disparities in different geographic regions of the world. *J Clin Oncol* 2006;24(14):2137–2150.
35. Bruix J, Sherman M, Practice Guidelines Committee, American Association for the Study of Liver Diseases, Management of hepatocellular carcinoma. *Hepatology* 2005;42(5):1208–1236.
36. Chen C, Schmilovitz-Weiss H, Liu XE, Pappo O, Halpern M, Sulkes J, Braun M, Cohen M, Barak N, Tur-Kaspa R, Vanhooren V, Van Vlierberghe H, Libert C, Contreras R, Ben-Ari Z, Serum protein N-glycans profiling for the discovery of potential biomarkers for nonalcoholic steatohepatitis. *J Proteome Res* 2009;8(2):463–470.
37. Blomme B, Fitzpatrick E, Quaglia A, De Bruyne R, Dhawan A, Van Vlierberghe H, Serum protein N-glycosylation in paediatric non-alcoholic fatty liver disease. *Pediatr Obes* 2012;7(2):165–173.
38. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, De Notariis S, Roda E, Bernardi M, Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: Influence of HBsAg and anti-HCV status. *J Hepatol* 2001;34(4):570–575.
39. Marrero JA, Henley KS, The role of serum biomarkers in hepatocellular carcinoma surveillance. *Gastroenterol Hepatol (N Y)* 2011;7(12):821–823.
40. Singal AG, Conjeevaram HS, Volk ML, Fu S, Fontana RJ, Askari F, Su GL, Lok AS, Marrero JA, Effectiveness of hepatocellular carcinoma surveillance in patients with cirrhosis. *Cancer Epidemiol Biomarkers Prev* 2012;21(5):793–799.
41. Flores A, Marrero JA, Emerging trends in hepatocellular carcinoma: Focus on diagnosis and therapeutics. *Clin Med Insights Oncol* 2014;8:71–76.
42. Marrero JA, Feng Z, Wang Y, Nguyen MH, Befeler AS, Roberts LR, Reddy KR, Harnois D, Llovet JM, Normolle D, Dalhgren J, Chia D, Lok AS, Wagner PD, Srivastava S, Schwartz M, Alpha-fetoprotein, des-gamma carboxyprothrombin, and lectin-bound alpha-fetoprotein in early hepatocellular carcinoma. *Gastroenterology* 2009;137(1):110–118.
43. Pan S, Tamura Y, Chen R, May D, McIntosh MW, Brentnall TA, Large-scale quantitative glycoproteomics analysis of site-specific glycosylation occupancy. *Mol Biosyst* 2012;8(11):2850–2856.
44. Drake RR, Glycosylation and cancer: Moving glycomics to the forefront. *Adv Cancer Res* 2015;126:1–10.
45. Naitoh A, Aoyagi Y, Asakura H, Highly enhanced fucosylation of serum glycoproteins in patients with hepatocellular carcinoma. *J Gastroenterol Hepatol* 1999;14(5):436–445.
46. Zhu J, Warner E, Parikh ND, Lubman DM, Glycoproteomic markers of hepatocellular carcinoma-mass spectrometry based approaches. *Mass Spectrom Rev* 2019;38(3):265–290.
47. Narimatsu H, Kaji H, Vakhrushev SY, Clausen H, Zhang H, Noro E, Togayachi A, Nagai-Okatani C, Kuno A, Zou X, Cheng L, Tao SC, Sun Y, Current technologies for complex glycoproteomics and their applications to biology/disease-driven glycoproteomics. *J Proteome Res* 2018;17(12):4097–4112.
48. Suttapitugsakul S, Sun F, Wu R, Recent advances in glycoproteomic analysis by mass spectrometry. *Anal Chem* 2020;92(1):267–291.
49. Liu MQ, Zeng WF, Fang P, Cao WQ, Liu C, Yan GQ, Zhang Y, Peng C, Wu JQ, Zhang XJ, Tu HJ, Chi H, Sun RX, Cao Y, Dong MQ, Jiang BY, Huang JM, Shen HL, Wong CCL, He SM, Yang PY, pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nat Commun* 2017;8(1):438.
50. Bern M, Kil YJ, Becker C, Byonic: Advanced peptide and protein identification software. *Curr Protoc Bioinformatics* 2012;40(1):13–20.
51. Park GW, Kim JY, Hwang H, Lee JY, Ahn YH, Lee HK, Ji ES, Kim KH, Jeong HK, Yun KN, Kim YS, Ko JH, An HJ, Kim JH, Paik YK, Yoo JS, Integrated GlycoProteome analyzer (I-GPA) for automated identification and quantitation of site-specific N-glycosylation. *Sci Rep* 2016;6:21175.
52. Zhu J, Chen Z, Zhang J, An M, Wu J, Yu Q, Skilton SJ, Bern M, Ilker Sen K, Li L, Lubman DM, Differential quantitative determination of site-specific intact N-glycopeptides in serum haptoglobin between hepatocellular carcinoma and cirrhosis using LC-EThcD-MS/MS. *J Proteome Res* 2019;18(1):359–371.
53. Yu Q, Wang B, Chen Z, Urabe G, Glover MS, Shi X, Guo LW, Kent KC, Li L, Electron-transfer/higher-energy collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization. *J Am Soc Mass Spectrom* 2017;28(9):1751–1764.
54. Walsh I, Zhao S, Campbell M, Taron CH, Rudd PM, Quantitative profiling of glycans and glycopeptides: An informatics' perspective. *Curr Opin Struct Biol* 2016;40:70–80.
55. Sun S, Hu Y, Ao M, Shah P, Chen J, Yang W, Jia X, Tian Y, Thomas S, Zhang H, N-GlycositeAtlas: A database resource for mass spectrometry-based human N-linked glycoprotein and glycosylation site mapping. *Clin Proteomics* 2019;16:35.
56. Yin H, Lin Z, Nie S, Wu J, Tan Z, Zhu J, Dai J, Feng Z, Marrero J, Lubman DM, Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma. *J Proteome Res* 2014;13(6):2887–2896.
57. Yin H, Tan Z, Wu J, Zhu J, Shedden KA, Marrero J, Lubman DM, Mass-selected site-specific core-fucosylation of serum proteins in hepatocellular carcinoma. *J Proteome Res* 2015;14(11):4876–4884.
58. Lin Y, Zhu J, Pan L, Zhang J, Tan Z, Olivares J, Singal AG, Parikh ND, Lubman DM, A panel of glycopeptides as candidate biomarkers for early diagnosis of NASH hepatocellular carcinoma using a stepped HCD method and PRM evaluation. *J Proteome Res* 2021 Jun 4;20(6):3278–3289.
59. Liu MQ, Zeng WF, Fang P, Cao WQ, Liu C, Yan GQ, Zhang Y, Peng C, Wu JQ, Zhang XJ, Tu HJ, Chi H, Sun RX, Cao Y, Dong MQ, Jiang BY, Huang JM, Shen HL,

Wong CCL, He SM, Yang PY, pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification. *Nat Commun* 2017;8(1):438.

60. Bern M, Kil YJ, Becker C, Byonic: Advanced peptide and protein identification software. *Curr Protoc Bioinformatics* 2012;40(1):13–20.
61. Park GW, Kim JY, Hwang H, Lee JY, Ahn YH, Lee HK, Ji ES, Kim KH, Jeong HK, Yun KN, Kim YS, Ko JH, An HJ, Kim JH, Paik YK, Yoo JS, Integrated GlycoProteome analyzer (I-GPA) for automated identification and quantitation of site-specific N-glycosylation. *Sci Rep* 2016;6:21175.
62. Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM, GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res* 2008;7(4):1650–1659.
63. Mayampurath A, Yu CY, Song E, Balan J, Mechref Y, Tang H, Computational framework for identification of intact glycopeptides in complex samples. *Anal Chem* 2014;86(1):453–463.
64. Wu Y, Mechref Y, Klouckova I, Mayampurath A, Novotny MV, Tang H, Mapping site-specific protein N-glycosylations through liquid chromatography/mass spectrometry and targeted tandem mass spectrometry. *Rapid Commun Mass Spectrom* 2010;24(7):965–972.
65. Zhu J, Chen Z, Zhang J, An M, Wu J, Yu Q, Skilton SJ, Bern M, Ilker Sen K, Li L, Lubman DM, Differential quantitative determination of site-specific intact N-glycopeptides in serum haptoglobin between hepatocellular carcinoma and cirrhosis using LC-EThcD-MS/MS. *J Proteome Res* 2019;18(1):359–371.
66. Yu Q, Wang B, Chen Z, Urabe G, Glover MS, Shi X, Guo LW, Kent KC, Li L, Electron-transfer/higher-energy collision dissociation (EThcD)-enabled intact glycopeptide/glycoproteome characterization. *J Am Soc Mass Spectrom* 2017;28(9):1751–1764.
67. Walsh I, Zhao S, Campbell M, Taron CH, Rudd PM, Quantitative profiling of glycans and glycopeptides: An informatics' perspective. *Curr Opin Struct Biol* 2016;40:70–80.
68. Sun S, Hu Y, Ao M, Shah P, Chen J, Yang W, Jia X, Tian Y, Thomas S, Zhang H, N-GlycositeAtlas: A database resource for mass spectrometry-based human N-linked glycoprotein and glycosylation site mapping. *Clin Proteomics* 2019;16:35.
69. Yin H, Lin Z, Nie S, Wu J, Tan Z, Zhu J, Dai J, Feng Z, Marrero J, Lubman DM, Mass-selected site-specific core-fucosylation of ceruloplasmin in alcohol-related hepatocellular carcinoma. *J Proteome Res* 2014;13(6):2887–2896.
70. Yin H, Tan Z, Wu J, Zhu J, Shedden KA, Marrero J, Lubman DM, Mass-selected site-specific core-fucosylation of serum proteins in hepatocellular carcinoma. *J Proteome Res* 2015;14(11):4876–4884.
71. Lin Y, Zhang J, Arroyo A, Singal AG, Parikh ND, Lubman DM, A fucosylated glycopeptide as a candidate biomarker for early diagnosis of NASH hepatocellular carcinoma using a stepped HCD method and PRM evaluation. *Front Oncol* 2022 Mar 17;12:818001.
72. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ, Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* 2012;11(11):1475–1488.
73. Wilson JP, Raghavan AS, Yang YY, Charron G, Hang HC, Proteomic analysis of fatty-acylated proteins in mammalian cells with chemical reporters reveals S-acylation of histone H3 variants. *Mol Cell Proteomics* 2011;10(3):M110–001198.
74. Meng Z, Veenstra TD, Targeted mass spectrometry approaches for protein biomarker verification. *J Proteomics* 2011;74(12):2650–2659.

75 Shi T, Su D, Liu T, Tang K, Camp DG, 2nd, Qian WJ, Smith RD, Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. *Proteomics* 2012;12(8):1074–1092.

76. Kuzyk MA, Smith D, Yang J, Cross TJ, Jackson AM, Hardie DB, Anderson NL, Borchers CH, Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma. *Mol Cell Proteomics* 2009;8(8):1860–1877.
77. Parker CE, Borchers CH, Mass spectrometry based biomarker discovery, verification, and validation–quality assurance and control of protein biomarker assays. *Mol Oncol* 2014;8(4):840–858.
78. Carr SA, Anderson L, Protein quantitation through targeted mass spectrometry: The way out of biomarker purgatory? *Clin Chem* 2008;54(11):1749–1752.
79. Ahn YH, Shin PM, Kim YS, Oh NR, Ji ES, Kim KH, Lee YJ, Kim SH, Yoo JS, Quantitative analysis of aberrant protein glycosylation in liver cancer plasma by AAL-enrichment and MRM mass spectrometry. *Analyst* 2013;138(21):6454–6462.
80. Ahn YH, Kim YS, Ji ES, Lee JY, Jung JA, Ko JH, Yoo JS, Comparative quantitation of aberrant glycoforms by lectin-based glycoprotein enrichment coupled with multiple-reaction monitoring mass spectrometry. *Anal Chem* 2010;82(11):4441–4447.
81. Yin H, Zhu J, Wang M, Yao ZP, Lubman DM, Quantitative analysis of alpha-1-antitrypsin glycosylation isoforms in HCC patients using LC-HCD-PRM-MS. *Anal Chem* 2020;92(12):8201–8208.
82. Lin Y, Zhu J, Zhang J, Dai J, Liu S, Arroyo A, Rose M, Singal AG, Parikh ND, Lubman DM, Glycopeptides with Sialyl Lewis antigen in serum haptoglobin as candidate biomarkers for nonalcoholic steatohepatitis hepatocellular carcinoma using a higher-energy collision-induced dissociation parallel reaction monitoring-mass spectrometry method. *ACS Omega* 2022 Jun 22;7(26):22850–22860. https://pubs.acs.org/doi/10.1021/acsomega.2c02600.
83. Pan L, Lin Y, Zhu J, Zhang J, Tan Z, Lubman DM, Large scale screening and quantitative analysis of site-specific N-glycopeptides from human serum in early alzheimer's disease using LC-HCD-PRM-MS. *J Proteomics Bioinform* 2022;15.
84. Gutierrez Reyes CD, Huang Y, Atashi M, Zhang J, Zhu J, Liu S, Parikh ND, Singal AG, Dai J, Lubman DM, Mechref Y, PRM-MS quantitative analysis of isomeric *N*-glycopeptides derived from human serum haptoglobin of patients with cirrhosis and hepatocellular carcinoma. *Metabolites* 2021 Aug 23;11(8):563.
85. Rich NE, Murphy CC, Yopp AC, Tiro J, Marrero JA, Singal AG, Sex disparities in presentation and prognosis of 1110 patients with hepatocellular carcinoma. *Aliment Pharmacol Ther* 2020;52(4):701–709.
86. Rich NE, Hester C, Odewole M, Murphy CC, Parikh ND, Marrero JA, Yopp AC, Singal AG, Racial and ethnic

differences in presentation and outcomes of hepatocellular carcinoma. *Clin Gastroenterol Hepatol* 2019;17(3):551–559.e1.
87. Hester CA, Rich NE, Singal AG, Yopp AC, Comparative analysis of nonalcoholic steatohepatitis- versus viral hepatitis- and alcohol-related liver disease-related hepatocellular carcinoma. *Journal of the National Comprehensive Cancer Network* 2019;17(4):322–329.
88. van der Pol CB, Lim CS, Sirlin CB, McGrath TA, Salameh JP, Bashir MR, Tang A, Singal AG, Costa AF, Fowler K, McInnes MDF, Accuracy of the liver imaging reporting and data system in computed tomography and magnetic resonance image analysis of hepatocellular carcinoma or overall malignancy-a ystematic review. *Gastroenterology* 2019;156(4):976–986.
89. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, Roberts LR, Heimbach JK, Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the american association for the study of liver diseases. *Hepatology* 2018;68(2):723–750.
90. Zhu J, Huang J, Zhang J, Chen Z, Lin Y, Grigorean G, Li L, Liu S, Singal AG, Parikh ND, Lubman DM, Glycopeptide biomarkers in serum haptoglobin for hepatocellular carcinoma detection in patients with nonalcoholic steatohepatitis. *J Proteome Res* 2020;19:3452–3466. https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00270.
91. Huang Y, Zhou S, Zhu J, Lubman DM, Mechref Y, LCMS/MS isomeric profiling of permethylated N-glycans derived from serum haptoglobin of hepatocellular carcinoma (HCC) and cirrhotic patients. *Electrophoresis* 2017;38(17):2160–2167.
92. Sanda M, Pompach P, Brnakova Z, Wu J, Makambi K, Goldman R, Quantitative liquid chromatography-mass spectrometrymultiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol. Cell. Proteomics* 2013;12(5):1294–1305.
93. Chandler KB, Pompach P, Goldman R, Edwards N, Exploring site-specific N-glycosylation microheterogeneity of haptoglobin using glycopeptide CID tandem mass spectra and glycan database search. *J Proteome Res* 2013;12(8):3652–3666.
94. Gallien S, Kim SY, Domon B, Large-scale targeted proteomics using internal standard triggered-parallel reaction monitoring (IS-PRM). *Mol Cell Proteomics* 2015;14(6):1630–1644.
95. Gutierrez Reyes CD, Atashi M, Fowowe M, Onigbinde S, Daramola O, Lubman DM, Mechref Y, Differential expression of N-glycopeptides derived from serum glycoproteins in mild cognitive impairment (MCI) patients. *Proteomics* 2024:e2300620.

42

Point-of-Care Technologies for Molecular Subtyping of Breast Cancer in Low- and Middle-Income Countries

Wenting Gao, Iftak Hussain, Sally Adebamowo, David Erickson, and Clement Adebamowo

42.1 Introduction

42.1.1 Breast Cancer in Low- and Middle-Income Countries

In 2020, 2.3 million women were diagnosed with breast cancer, with 685,000 deaths, making it the commonest cancer globally.[1, 2] About half of these cases (53%) occurred in low- and middle-income countries (LMICs) where the five-year survival rate ranged from 12% to 53% compared to the survival rate in high-income countries (HICs) where it was at least 85% in 2020. The global incidence of breast cancer has been rising by 1.44% annually since the 1990s except in North America, where the incidence has stabilized and may be falling.[3] The global annual percentage change in breast cancer mortality was 0.23 but this ranged from an increase of 0.56 in Africa and East Asia/Pacific to a decrease of 1.75 in North America and 0.55 in Europe/Central Asia.[4] These global disparities in mortality are due to several factors including limited screening and early detection capabilities, inability to afford care, poor facilities, and lack of trained personnel.[5]

Breast cancer is subtyped using the patterns of steroid hormone receptors—estrogen receptor (ER) and progesterone receptor (PR), cell proliferation marker (Ki67), and overexpression of human epidermal growth factor receptor 2 (HER2) protein (Figure 42.1). The prevalence of different subtypes of breast cancer varies across populations, largely driven by demographic patterns and the distribution of risk factors. The ability to accurately detect and distinguish various subtypes of breast cancer has profound implications as it guides treatment decisions and prognostication. It also guides epidemiological research and clinical trials to better understand disease patterns, discover new biomarkers and treatments, and ultimately contribute to better health outcomes.[6, 7] Subtypes such as hormone receptor-positive, HER2-positive, and triple-negative breast cancer each have unique treatment algorithms and clinical outcomes.[8, 9] Nevertheless, properly characterizing the molecular subtypes of breast cancer LMICs has been challenging because of inadequate resources and personnel, poor infrastructure, and high cost. The results of immunohistochemical subtyping are often negatively affected by pre-analytic, analytic, and post-analytic factors including reporting errors leading to low accuracy of the results.[10–12] It is therefore crucial to develop novel strategies for molecular subtyping of breast cancer in LMICs.[13]

In HICs, the gold standard for molecular subtyping of breast cancer is immunohistochemistry of surgically excised specimens or core-needle biopsies by highly trained pathologists using sophisticated equipment, often in centralized laboratories. In contrast, LMICs have a limited number of well-trained pathologists and laboratory technicians. Africa has one pathologist per 1,000,000 people compared to 40–50 pathologists per 1,000,000 people in the US.[14, 15] LMIC pathologists are often generalists with little consistent exposure to breast pathology, which reduces their diagnostic accuracy. Chronic and persistent underinvestment in diagnostic services negatively impacts the quality of laboratory diagnostics, which results in poor treatment and outcomes for breast cancer patients in LMICs.[16, 17] These constraints are unlikely to be resolved soon. Proposed solutions, such as training new personnel and providing equipment, require time and rely on additional funding that may not be available.[15] Although telepathology has been suggested as a potential solution, it requires substantial capital investments, training, and coordination across multiple healthcare systems and often across time zones.[18] Furthermore, telepathology does not address the underlying infrastructural limitations and lack of trained laboratory personnel needed to make the slides for telepathology consultations. Therefore, there is an urgent need to develop affordable and easily implementable breast cancer molecular subtyping technologies that are adaptable and accessible to cancer care providers and patients in LMICs.

42.1.2 Point-of-Care Technologies for Breast Cancer Subtyping

Recent years have witnessed rapid advancements in the development of novel genomics and proteomics methods for subtyping breast cancer, but none of these have replaced the standard methods, and few of the approaches are likely to meet the needs of LMICs. To be useful and adopted in LMICs, novel breast cancer subtyping technologies must meet certain performance criteria. The World Health Organization (WHO) Sexually Transmitted Diseases Diagnostics Initiative (SDI) developed the "ASSURED" criteria (affordable, sensitive, specific, user-friendly, rapid, and robust, equipment-free, and deliverable to end users), which was modified to REASSURED with inclusion of mHealth. These criteria are used as benchmarks to evaluate the potential of novel technologies and decide if they address disease control needs.[19] They can be usefully applied to evaluate point-of-care devices for non-communicable

DOI: 10.1201/9781003449942-47

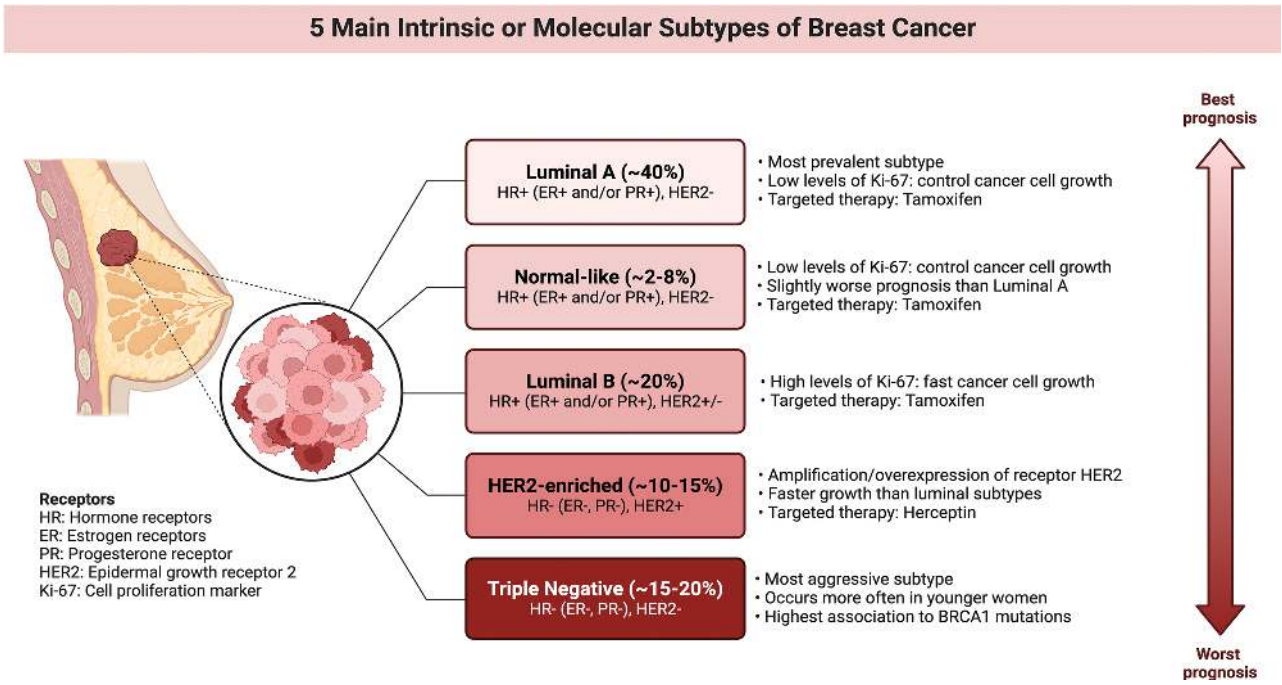


FIGURE 42.1 The prevalence and characteristics of molecular subtypes of breast cancer. Reproduced with permission from BioRender.com.

diseases including the molecular subtyping of breast cancer. Such devices can harness recent developments in optical, electrochemical, and micro-nuclear magnetic resonance technologies. The resulting tests must have minimal hands-on requirements to account for shortage of trained personnel and generate validated objective results that do not require highly trained specialists for interpretation. The tests must be affordable, generate results rapidly, preferably be implemented on generic platforms, and at the point of care to reduce risk of loss to follow-up.

There have been multiple attempts at developing molecular subtyping tests for breast cancer that meet most, if not all, the REASSURED criteria. One approach is the use of a chip-based diagnostic magnetic resonance (DMR) system that leverages molecular interactions between magnetic nanoparticle and target markers, including HER2 extracted from fine-needle aspirate (FNA) samples.[20,21] Another FNA-based approach engages a digital imaging platform and artificial intelligence diffraction (AIDA) that uses a dual-color-staining holographic method to detect multiple membrane and intracellular markers on cells, thus allowing their quantitative molecular profiling and detection of cellular heterogeneity.[22] Automated image cytometry system (CytoPAN) is yet another approach that uses multiple fluorescence signals for rapid diagnosis,[23] while EpiView-D4 is a multimodal cellphone-based platform featuring an optical microscope and an immunodiagnostic chip that quantifies the expression of protein biomarkers and provides an assay readout.[24] These FNA-based approaches are limited because breast cancer diagnosis needs to be done first, and this requires tissue biopsies to demonstrate invasiveness and differentiate invasive from ductal or lobular carcinoma *in situ* (DCIS or LCIS). Performing FNA for molecular subtyping after excisional biopsy may not be feasible as there may be no residual tissue left, tissue scaring may confound the test procedures, and the process may require multiple hospital visits that increases patients' costs, inconvenience, and risk of loss to follow-up.

Another approach to molecular subtyping of breast cancer is the use of quantitative PCR analysis to evaluate gene expression in breast biopsy samples.[25,26] Several research groups have developed qPCR-based molecular subtyping tests. While some of the resulting tests focused on single biomarkers (e.g., HER2 gene overexpression), others were multiplex tests.[27–29] The most developmentally advanced of these platforms is the GeneXpert STRAT4® test which has been tested in many countries and demonstrates excellent concordance with immunohistochemistry (IHC).[30,31] The concordance between some of these qPCR platforms and pathologist-reported IHCs, particularly in LMIC institutions where internal quality control and external quality assurance of pathology departments is less rigorous, is suboptimal.[28] Most of the qPCR-based tests require sophisticated equipment and more advanced skills than what is currently available in most LMIC institutions. Other qPCR-based tests require expensive, proprietary platforms whose availability may be limited due to competing needs or infrastructural limitations.[32] Other limitations include lack of validation data, complexity, low sensitivity, and low specificity.[33]

42.1.3 Lateral Flow Immunoassay for Breast Cancer Subtyping

Lateral flow immunoassay (LFIA) has emerged as a popular analytical tool for on-site detection of target substances, particularly in LMICs, because it meets many of WHO's REASSURED criteria for diagnostic tests.[19,34] LFIAs can be

employed to rapidly detect the presence of target analytes, such as proteins, pathogens, and nucleic acids, in various sample types including blood, saliva, and urine. Their design is conducive to point-of-care application without the need for sophisticated laboratory infrastructure, making them highly suitable for use in LMICs.[35–38] Despite their widespread use and success in infectious disease diagnostics, there are challenges with applying LFIAs to the complex domain of breast cancer molecular subtyping.[39–41] Breast cancer molecular subtyping demands a high level of sensitivity and specificity due to the heterogeneous nature of the disease, complexity of the tissues, and subtle variations in the levels of the biomarkers across different samples.[42, 43] LFIAs are typically less sensitive than laboratory-based assays such as enzyme-linked immunosorbent assays (ELISAs), which can limit their effectiveness in accurately quantifying low-abundance biomarkers such as those required to distinguish between breast cancer subtypes.[44, 45] Moreover, multiplexing is essential for comprehensive breast cancer subtyping. The development of multiplex LFIAs is technically challenging because it requires concurrent detection of several analytes without cross-reactivity or signal interference. The integration of advanced materials, such as nanoparticles and quantum dots, and innovative design strategies are subjects of ongoing research to overcome these limitations.[34, 36, 37, 46]

In addressing the challenges of meeting the REASSURED criteria by using current LFIAs in molecular subtyping of breast cancer, the development of an advanced point-of-care diagnostic platform is imperative. We therefore engineered a microfluidics-based LFIA for molecular subtyping and paired it with a portable fluorescence imaging platform capable of quantitatively measuring the biomarkers in tissue samples with high precision called PROVIDS to create the PoCBreCa toolkit.

42.2 Materials

42.2.1 Chemicals

We obtained primary antibodies—Human total ER (DYC5715), Human total ErBb2 (R&D Systems), and Human total PR (DYC5415)—from R&D Systems, Minneapolis, MN, and Europium-streptavidin (Eu-SA) chelate nanoparticles (ab270228) from Abcam, Waltham, MA. We used Whatman FF170HP Membrane Cards (Cytiva) as the foundation of our test strips, and obtained CF5 sample and absorbent pads from GE Healthcare, Chicago, IL, alongside glass fiber conjugate pad from MilliporeSigma, Burlington, MA.

42.2.2 Biological

We obtained formalin-fixed, paraffin-embedded (FFPE) blocks of adenocarcinoma of the breast expressing different ER, PR, and HER2 receptors (CB544787, CB551526, and CB555588) from Origene Technologies, Rockville, MD. Donkey anti-goat and donkey anti-sheep IgG antibodies (BAF109 and BAF016) from Biotechne, Minneapolis, MN, were used as controls.

42.3 Methods

42.3.1 Fluorometric Multiplex LFIA Design

We designed the assay component of the platform to enable specific and sensitive detection of three human breast cancer biomarkers simultaneously. The components of the multiplex LFIA test strip (Figure 42.2) were a sample pad for dispensing sample mixture, a conjugate pad for holding Eu-SA:bioAb complex, a nitrocellulose membrane for immobilizing capture reagents, and an absorbent pad for wicking excess flow. Individually assembled PR, ER, and HER2 LFIA were placed 5 mm apart and parallel with each other. Shared sample and absorbent pads were assembled at the beginning and end of the assay to enable multiplexing.

We selected 200 nm Europium-streptavidin (Eu-SA) chelate nanoparticles (Abcam, ab270228) as our fluorophore conjugates for the detection of biotinylated anti-PR, anti-ER, and anti-HER2 detection antibodies. The surface treatment of the nanoparticles makes the conjugate resistant to aggregation, while the extremely broad Stokes shift of the Eu chelate particle allows the assay to reach a higher sensitivity in any immunoassay, preventing nonspecific fluorescence interference.

To enable fluorophore-based multiplexing, a mixture of Eu-SA chelate nanoparticles and biotinylated antibody (bioAb) complex was pre-dried on the conjugate pad for each individual test strip. When reconstituted by the addition of mixed sample and running buffer, the functionalized Eu-SA:bioAb complex can bind to its corresponding human total PR, ER, or HER2 proteins present in the sample mixture. As this mixture migrates up the nitrocellulose membrane, Eu-SA:bioAb:antigen complexes bind to the immobilized capture antibodies on the membrane in sandwich format, which contributes to the development of fluorometric signals at test lines. Unbounded Eu-SA:bioAb complexes bind to a mixture of donkey anti-sheep and donkey anti-goat IgG antibodies at control line, which was designed not only as a negative control of the assay, but to confirm conjugates functionality. Fluorophore coding of test and control lines for various breast cancer subtypes is shown in Figure 42.2.

42.3.2 Singleplex LFIA Assembly and Preparation

The construction of a single LFIA test strip involved the strategic arrangement of several critical components: a sample pad, a glass fiber conjugate pad, a nitrocellulose membrane, and an absorbent pad. For the foundation of the test strip, we used Whatman FF170HP Membrane Cards (Cytiva) that incorporate a nitrocellulose membrane integrated with adhesive sections to secure additional components. This membrane offers an average capillary flow rate of 45 s/cm, the lowest rate available from the manufacturer, to enhance the sensitivity and efficiency of reagent use by prolonging the interaction between the capture reagents and the target analyte.

We used the Human total ER (R&D Systems, DYC5715), Human total ErBb2 (R&D Systems, DYC1129), and Human total PR (R&D Systems, DYC5415) DuoSet IC ELISA kits as spiked buffer samples for the development of singleplex LFIAs.

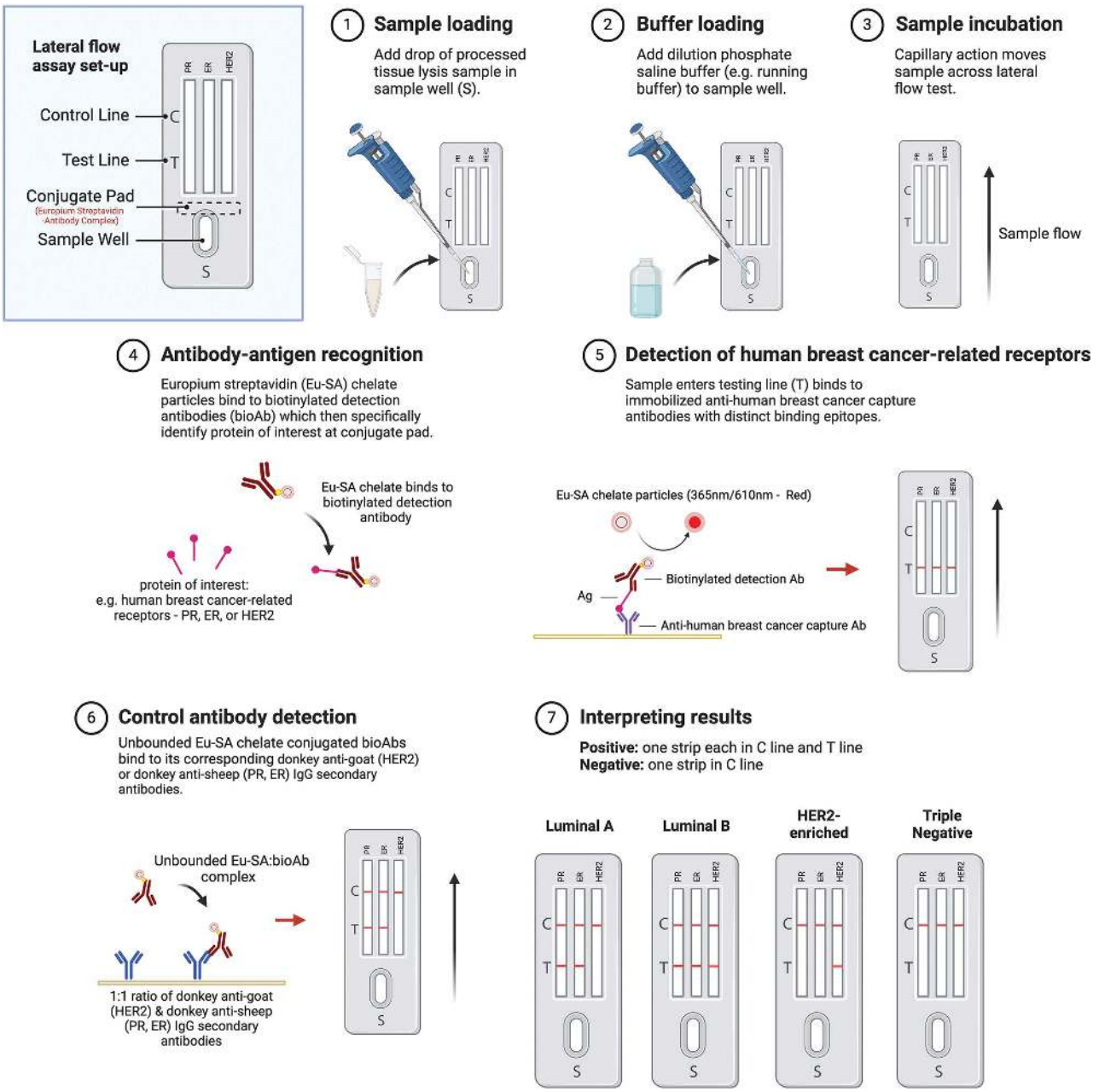


FIGURE 42.2 Architecture of multiplex LFIA. Eu-SA:bioAb conjugates are able to bind to human total PR, ER, and HER2 proteins present in the sample, respectively. Eu-SA:bioAb labeled PR, ER, or HER2 proteins are captured by the first test line (shown in red at T line for illustrative purposes), and unbound Eu-SA:bioAb are captured by the secondary Ab mixture at the control line (shown in red at C line for illustrative purposes). Examples of test strip appearance corresponding to various diagnostic scenarios were provided as user's references. Created with BioRender.com.

In the assembly process, we deposited the lines for control and test reagents 3 mm apart on the nitrocellulose membrane using a Kinematic Automated Matrix 1600 dispensing module set to dispense at 6.7 mL/min. The test line was impregnated with capture antibodies specific for human PR, ER, or HER2 proteins at a concentration of 0.1 mg/mL. Conversely, the control line is formed by a 1:1 ratio of donkey anti-goat and donkey anti-sheep IgG antibodies (Biotechne, BAF109 & BAF016) at the same concentration. These capture agents were diluted in a buffer solution composed of 1 × PBS and 2% sucrose.

Post-dispensing, the membrane cards were subjected to a drying phase at 37 °C for a minimum of two hours. Subsequently, sample and absorbent pads (GE Healthcare, CF5), alongside with glass fiber conjugate pad (MilliporeSigma), were affixed to the polystyrene adhesive backing, positioned adjacent to the nitrocellulose membrane, with a deliberate 2 mm overlap to facilitate the seamless transfer of fluids during assay running. The assembled membrane cards were then cut into strips of 4 mm width using Kinematic Matrix 2360 Programmable Shear device.

42.3.3 Development of Singleplex Quantitative LFIAs of Human PR, ER, and HER2 Proteins

We prepared 0.5% (w/v) Eu-SA chelate nanoparticles following manufacturer's protocol and stored it at 4 °C in a light-free environment until use. Upon assay running, we thoroughly diluted it to 0.05% (w/v) Eu-SA chelate nanoparticles using conjugation buffer (50 mM borate, 10% sucrose, 1% BSA, 0.5% Tween-20, and 0.02% sodium azide), in which stabilized and preserved the bioactivity of the conjugate (e.g., Eu-SA labeled antibodies), ensuring its optimal interaction with the target analyte.

For each singleplex LFIA dipstick, we then mixed 1 uL of diluted biotinylated anti-PR, anti-ER, or anti-HER2 detection antibodies (0.024 mg/mL, 0.024 mg/mL, and 0.01 mg/mL, respectively following manufacture's protocols) and 2 uL of conjugate buffer diluted 0.05% (w/v) Eu-SA chelate nanoparticles in a 96-well half area plates, allowing thorough binding of the analyte with Eu-SA:bioAb complexes. LFIAs in quadruplets were then each exposed to 25 uL of spiked buffer solutions with different concentrations of biomarkers (3.9 pmol to 1000 pmol), alongside with 30 uL of running buffer (75 mM Tris, 450 mM NaCl, 1% BSA, 1% Tween-20, and 0.02% sodium azide), which facilitated the movement of the sample and reagents along the test strip, maintaining the appropriate conditions for antigen-antibody reactions to occur effectively. Fluorescent signals were obtained using PROVIDS reader system 20 minutes after assay initiation.

42.3.4 Development of PROVIDS Reader System

We developed a portable rapid on-site versatile image-based detection system (PROVIDS), which marks a significant milestone in enhancing the accessibility and efficiency of LFIA analysis. Figure 42.3a shows a prototype image of the PROVIDS system. It comprises an imaging module to capture fluorescence signals from the LFIA device, a Raspberry Pi 4

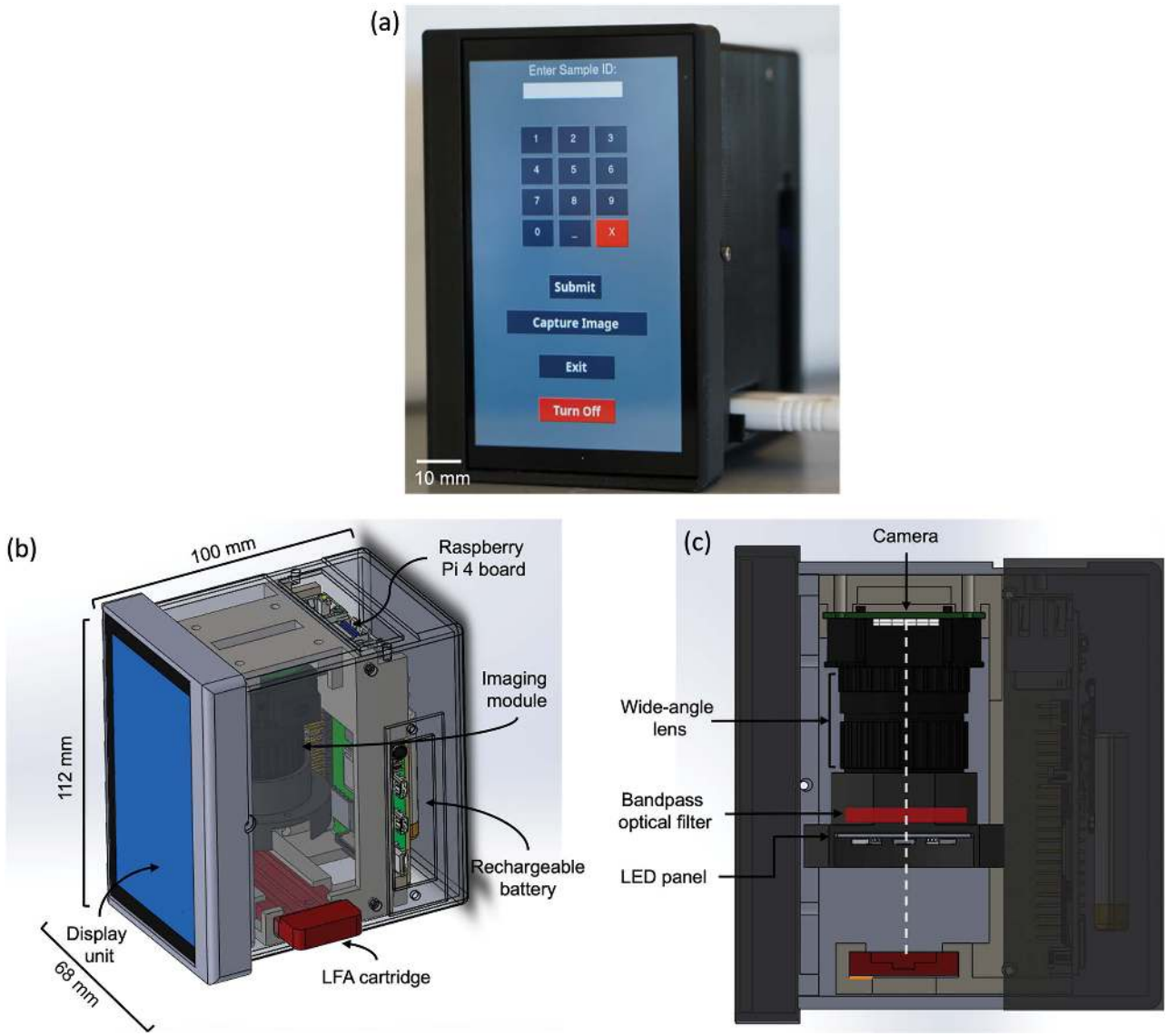


FIGURE 42.3 PROVIDS hardware device for detection of fluorescent signal. (a) Photo image of the PROVIDS reader system. (b) 3D rendering of the system. (c) Schematic illustration of the imaging module.

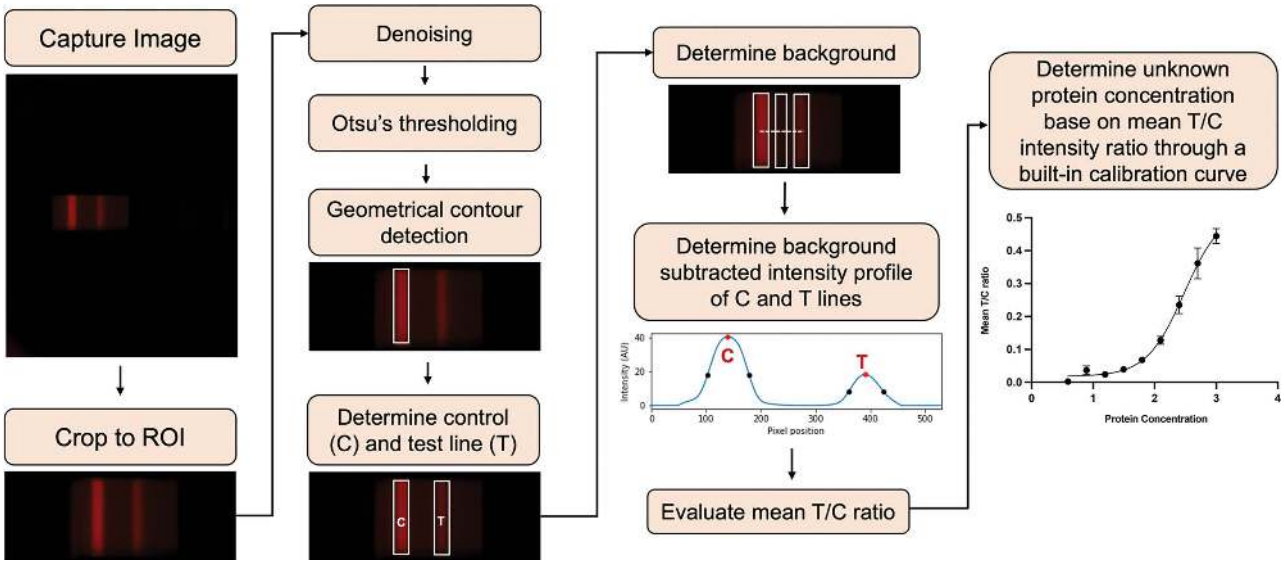


FIGURE 42.4 PROVIDS in-built LFIA image processing pipeline. After capturing the image of LFIA, the in-built pipeline automatically detects control (C) and test (T) line fluorescence intensity and processes the T/C intensity ratio in determining the unknown concentration of the investigated protein.

board to control the camera and to perform post-capture computations, an LCD screen to display the results, a 3D-printed cartridge to hold the LFIA dipstick, and various 3D-printed components to assemble the system. The optical design of the system was optimized to provide a compact and low-cost platform for analysis of spatially multiplexed LFIA devices (Figure 42.3b). As shown in Figure 42.3c, the imaging module was designed to capture three LFIAs seamlessly.

Six UV-LEDs of peak emission wavelength 365 nm matching the excitation wavelength of Eu chelate microspheres were used as a light source, and a band-pass optical filter with peak emission wavelength 610 nm with a bandwidth of 10 nm was used as the emission filter. A Raspberry Pi high-quality camera (4056 × 3040 pixels, Sony IMX477 sensor) captures the fluorescence signals passing through the emission filter with a mounted adjustable lens. A 5000 mAh rechargeable battery was integrated to enhance its autonomy and portability, offering over nine hours of operational capacity. We employed a 3D printer (Raise3D Pro3) to construct the device, utilizing black polylactic acid (PLA) filament for its fabrication. The choice of black PLA filament not only contributes to the system's aesthetic but, importantly, serves to block external light, thereby enhancing the precision of fluorescence signal capture by minimizing potential light interference. The compact dimensions (100 mm in length × 68 mm in breadth × 112 mm in height) and lightweight design (~290 grams) enhance its portability. A 32 GB microSD card is embedded for efficient storage and retrieval of LFIA images, ensuring the system's capacity for extensive field use without the immediate need for data transfer.

42.3.5 Image Processing Pipeline

We designed a Python application called LFIA Analyzer to perform all the necessary data processing steps, which includes capturing LFIA images, required image processing, and determining biomarker concentration. As shown in Figure 42.4, high-resolution images of the LFIA are acquired under controlled lighting conditions to ensure uniform illumination across all samples.

After loading the captured image file, the application automatically crops the image to the region of interest (ROI) for subsequent image processing to detect the control and test line intensity. To mitigate the impact of digital noise and enhance the signal-to-noise ratio, we applied a Savitzky–Golay filter to the image. This filter was selected for its ability to preserve the essential features of the signal, such as height and width of the assay lines, while smoothing out noise. The filtered image underwent a thresholding process using Otsu's method to distinguish between the line regions and the background more distinctly. Contours within the thresholded images were identified. A custom algorithm was applied to these contours to discern potential control line based on predefined color and geometric criteria, specifically targeting line width and height parameters relevant to our assay design. We targeted the control line for segmentation since its intensity is constant, considering the currently implemented assay format. After determining pixel position of control line, we determined the background and test line.

For each identified line, we determined a midpoint and extracted a horizontal intensity profile across the strip at this position. This profile captures the variation in pixel intensity across the assay, reflecting the presence and concentration of target analytes. The profiles were further processed using a custom implementation of the Hampel filter to reduce the influence of outliers, thereby smoothing the data for peak detection. Utilizing peak detection algorithms, we identified significant peaks for control (I_c) and test line (I_t) and the corresponding background intensity (I_b). The magnitude and relative positions of these peaks were analyzed to quantify the assay results. The analysis included calculating the ratio of

peak intensities, which serves as an indicator of the presence and concentration of the target analyte. For comprehensive analysis and interpretation, we generated visual representations of the processed images, highlighting the detected lines and the corresponding intensity profiles. These visualizations included overlays of the original images with detected lines, plots of the smoothed and corrected intensity profiles, and annotations indicating significant peaks and their quantitative analysis. These visual aids were crucial for validating the analytical process and interpreting the results.

42.4 Results

42.4.1 Singleplex LFIA for Individual Human Total PR, ER, and HER2 Recombinant Protein Detection and Quantification

We employed a singleplex LFIA to target and quantify each of the biomarkers required for molecular subtyping of breast cancer separately, thereby simplifying the test and potentially increasing its specificity and ease of use in priming the PROVIDS reader. We established a mean test-to-control (T/C) ratio curve, a metric integral to LFIA quantification for each biomarker, ensuring reliable and repeatable measurements (Figure 42.5). The limits of detection (LOD) and quantification (LOQ) for the assays were rigorously defined in adherence to analytical standards, with the LOD defined as mean + 3σ of the blank (σ is standard deviation) and the LOQ defined as mean + 10σ of the blank. These parameters were derived from their respective standard curves, generated from a series of known concentrations ranging from 0 to 1000 pmol.

Upon validation, the singleplex LFIA showed a high level of sensitivity and specificity. The assay achieved an LOD of 21.97 pmol and an LOQ of 110.917 pmol for detection of PR, an LOD of 14.19 pmol and an LOQ of 108.64 pmol for ER, and an LOD of 48.89 pmol and an LOQ of 198.15 pmol for HER2. The PROVIDS reader successfully identified each biomarker in the buffer samples, which were precisely spiked with predetermined concentrations of recombinant proteins to mimic varying clinical scenarios. It demonstrated exceptional accuracy in discriminating among a wide range of concentrations,

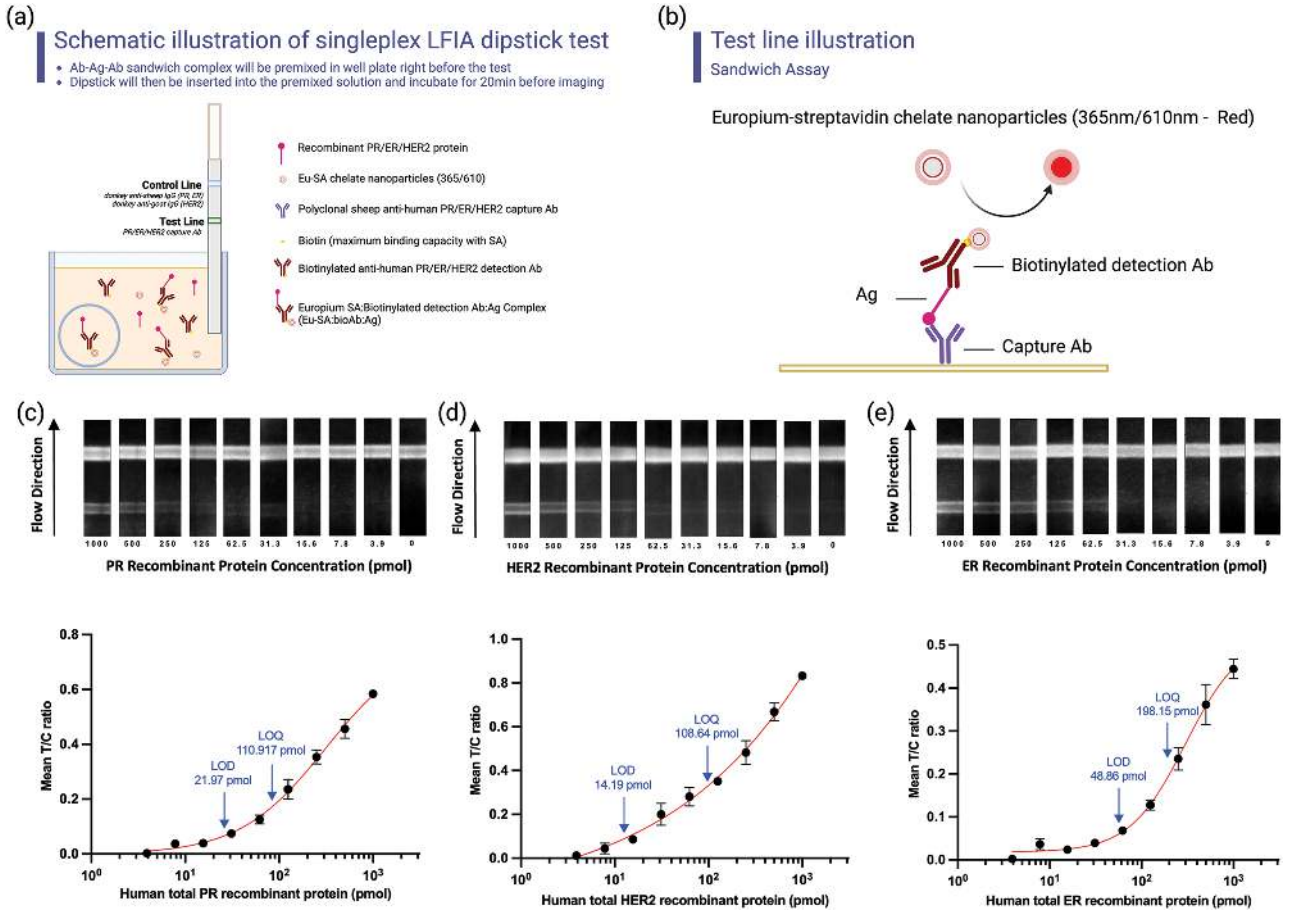


FIGURE 42.5 Quantitative singleplex LFIAs of human breast cancer biomarkers. (a) Schematic illustration of individual LFIA strips of comprising capture antibody test line and a donkey IgG control line. (b) Schematic illustration of the sandwich architecture of reagents at test line. (c) Fluorescence images of human total PR LFIAs. PROVIDS built-in code then processed images and generated dose-dependent fluorescence intensities mean T/C ratio curve of human total PR LFIAs. (d) Fluorescence images of human total HER2 LFIAs. PROVIDS built-in code then processed images and generated dose-dependent fluorescence intensities mean T/C ratio curve of human total HER2 LFIAs. (e) Fluorescence images of human total ER LFIAs. PROVIDS built-in code then processed images and generated dose-dependent fluorescence intensities mean T/C ratio curve of human total ER. Data are mean ± s.d.; n = 2 × 2 repeated tests. Schematic was created with BioRender.com.

from the lower limits of detection up to concentrations typical of clinical samples. These results underscore the potential of PoCBreCa as a reliable POC toolkit for breast cancer subtyping utilizing multiplex LFIA, offering a simplified yet precise alternative to existing diagnostic methods.

42.5 Discussion and Conclusions

In this study, we developed and demonstrated the potential of PoCBreCa toolkit as a pioneering advancement in PoC diagnostics for molecular subtyping of breast cancer, which would be particularly useful in LMICs. Our tests of the toolkit demonstrate its exceptional accuracy in detecting a broad spectrum of biomarker concentrations, from the lower detection limits to those commonly observed in clinical samples. This toolkit therefore offers a rapid and accurate alternative to conventional diagnostic methodologies. Central to our platform's innovation is the incorporation of europium (Eu), a lanthanide fluorophore, which is characterized by its exceptional luminescent properties. Notably, europium exhibits a Stokes shift of approximately 245 nm, a remarkable separation between its excitation peak (around 365 nm) and emission peak (around 610 nm). This vast shift plays a crucial role in diminishing background fluorescence, thereby significantly enhancing the clarity and definition of the assay's fluorescence signal. Furthermore, the integration of europium-streptavidin conjugates in our assays markedly enhanced the sensitivity and specificity of our LFIA platform. The strong affinity of streptavidin for biotinylated antibodies facilitates the formation of a stable complex, ensuring that the europium's luminescence is maximally utilized for the detection of target analytes. This coupling increased the LOD of our assays down to picogram levels, a significant improvement over conventional LFIA systems. More importantly, the specificity of the assay is preserved due to the high-fidelity binding between streptavidin and biotin, reducing the likelihood of nonspecific binding and false-positive results. Besides, the robust chemical and photostability of europium translates into sustained luminescence, leading to an amplification of the assay's sensitivity—essential for detecting low-abundance biomarkers.

The PoCBreCa toolkit combines immunochemistry, microfluidics, LFIA, and the PROVIDS reading system to create a low-cost, rapid, objective, sensitive, specific, and user-friendly test kit for molecular subtyping of breast cancer. The kit will substantially improve access to breast cancer testing in parts of the world where the incidence of breast cancer is still rising and where there are limited resources for diagnosis and management. The kit overcomes many of the limitations of alternative approaches to resolving the challenge of limited diagnostics in LMIC, enables more personalized medicine for breast cancer treatment, and provides physicians and patients with prognostic information. Given its low cost, limited infrastructural requirements, and usability by personnel with limited training, the PoCBreCa toolkit has great potential to save patients and healthcare systems in LMICs huge amounts of money while significantly improving care.

Despite these advancements, one of the primary challenges of current study is transitioning from singleplex to multiplex testing without losing accuracy or increasing cross-reactivity. Future steps in the PoCBreCa toolkit's evolution will concentrate on refining its multiplex capacity to enable the concurrent detection of multiple biomarkers efficiently. This development will involve rigorous optimization of assay conditions to prevent cross-reactivity and the integration of robust controls to ensure consistent performance.

42.6 Future Trends

Rapid testing using microfluidics, immunochemistry, and LFIA is potentially a feasible long-term solution to the challenges of molecular subtyping of breast cancer in LMICs. Further work will enhance the toolkit's user interface and data management capabilities, allowing for seamless integration into various healthcare settings. Additional research on the components of the immunochemistry and reporting systems and inclusion of AI algorithms and machine learning technologies will further improve the sensitivity, specificity, and validity of the results generated by the kit. By overcoming these challenges, the PoCBreCa toolkit will become an indispensable asset for breast cancer diagnostics worldwide, contributing significantly to personalized and timely cancer care.

Acknowledgments

This work was funded by NCI/NIH grant "Point of Care Diagnostic Test for Molecular Subtyping of Breast Cancer (PoCBreCa) Study" Award number R21CA255835. Additional support for CAA and SNA was provided by funds through the Maryland Department of Health's Cigarette Restitution Fund Program (CH-649-CRF) and the University of Maryland Greenebaum Cancer Center Support Grant (P30CA134274). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or the Maryland Department of Health.

Figures 42.1, 42.2, and 42.5 were created using BioRender.com (Agreement No. SL26MR60EW, YZ26MR60GP, GZ26MR60IU, and NR26MR60KH).

Author Contributions

CAA, SNA, and DE conceived, designed, and obtained funding for this project. W.G. and I.H. designed and performed most of the experiments and wrote the first draft of this chapter. CAA and DE supervised experiments, discussed and provided feedback. CAA, SNA, and DE revised and provided critical feedback to the chapter and all authors approved the final version.

Declaration of Interests

The authors declare that the study was conducted in the absence of any commercial or financial relationships that be construed as a potential conflict of interest.

REFERENCES

1. Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, Vignat J, Gralow JR, Cardoso F, Siesling S, Soerjomataram I. Current and future burden of breast cancer: Global statistics for 2020 and 2040. *Breast*. 2022; 66:15–23.
2. Wilkinson L, Gathani T. Understanding breast cancer as a global health concern. *Brit J Radiol*. 2022;95: 20211033–20211036.
3. Lei SY, Zheng RS, Zhang SW, Wang SM, Chen R, Sun KX, Zeng HM, Zhou JC, Wei WQ. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. *Cancer Commun*. 2021;41:1183–1194.
4. Lima SM, Kehm RD, Terry MB. Global breast cancer incidence and mortality trends by region, age-groups, and fertility patterns. *EClinicalMedicine*. 2021;38:100985–100993.
5. Sharma R. Breast cancer incidence, mortality and mortality-to-incidence ratio (MIR) are associated with human development, 1990–2016: Evidence from global Burden of disease study 2016. *Breast Cancer-Tokyo*. 2019;26:428–445.
6. Waks AG, Winer EP. Breast cancer treatment: A review. *JAMA*. 2019;321:288–300.
7. Haque R, Ahmed SA, Inzhakova G, Shi JX, Avila C, Polikoff J, Bernstein L, Enger SM, Press MF. Impact of breast cancer subtypes and treatment on survival: An analysis spanning two decades. *Cancer Epidem Biomar*. 2012;21:1848–1855.
8. Francies FZ, Hull R, Khanyile R, Dlamini Z. Breast cancer in low-middle income countries: Abnormality in splicing and lack of targeted treatment options. *Am J Cancer Res*. 2020;10:1568–1591.
9. da Costa Vieira RA, Biller G, Uemura G, Ruiz CA, Curado MP. Breast cancer screening in developing countries. *Clinics (Sao Paulo)*. 2017;72:244–253.
10. Adebamowo CA, Famooto A, Ogundiran TO, Aniagwu T, Nkwodimmah C, Akang EE. Immunohistochemical and molecular subtypes of breast cancer in Nigeria. *Breast Cancer Res Tr*. 2008;110:183–188.
11. Akarolo-Anthony SN, Ogundiran TO, Adebamowo CA. Emerging breast cancer epidemic: Evidence from Africa. *Breast Cancer Res*. 2010;12:1–4.
12. McCormack VA, Joffe M, van den Berg E, Broeze N, Silva ID, Romieu I, Jacobson JS, Neugut AI, Schüz J, Cubasch H. Breast cancer receptor status and stage at diagnosis in over 1,200 consecutive public hospital patients in Soweto, South Africa: A case series. *Breast Cancer Res*. 2013;15:1–13.
13. Coleman WB, Andersta CK. Discerning clinical responses in breast cancer based on molecular signatures. *Am J Pathol*. 2017;187:2199–2207.
14. Metter DM, Colgan TJ, Leung ST, Timmons CF, Park JY. Trends in the US and Canadian pathologist workforces from 2007 to 2017. *JAMA Netw Open*. 2019;2:e194337.
15. Fleming KA, Naidoo M, Wilson M, Flanigan J, Horton S, Kuti M, Looi LM, Price C, Ru K, Ghafur A, Wang J, Lago N. An essential pathology package for low- and middle-income countries. *Am J Clin Pathol*. 2017;147:15–32.
16. Shyyan R, Masood S, Badwe RA, Errico KM, Liberman L, Ozmen V, Stalsberg H, Vargas H, Vass L, Pathology GSD. Breast cancer in limited-resource countries: Diagnosis and pathology. *Breast J*. 2006;12:S27–S37.
17. Masood S, Vass L, Ibarra JA, Ljung BM, Stalsberg H, Eniu A, Carlson RW, Anderson BO, Patholo BHGI. Breast pathology guideline implementation in low- and middle-income countries. *Cancer-Am Cancer Soc*. 2008;113:2297–2304.
18. Martei YM, Pace LE, Brock JE, Shulman LN. Breast cancer in low-and middle-income countries: Why we need pathology capability to solve this challenge. *Clinics in Laboratory Medicine*. 2018;38:161–173.
19. Land KJ, Boeras DI, Chen XS, Ramsay AR, Peeling RW. REASSURED diagnostics to inform disease control strategies, strengthen health systems and improve patient outcomes. *Nat Microbiol*. 2019;4:46–54.
20. Lee H, Sun E, Ham D, Weissleder R. Chip-NMR biosensor for detection and molecular analysis of cells. *Nat Med*. 2008;14:869–874.
21. Haun JB, Castro CM, Wang R, Peterson VM, Marinelli BS, Lee H, Weissleder R. Micro-NMR for rapid molecular analysis of human tumor samples. *Sci Transl Med*. 2011;3:1–13.
22. Min J, Im H, Allen M, McFarland PJ, Degani I, Yu H, Normandin E, Pathania D, Patel JM, Castro CM, Weissleder R, Lee H. Computational optics enables breast cancer profiling in point-of-care settings. *ACS Nano*. 2018;12:9081–9090.
23. Min J, Chin LK, Oh J, Landeros C, Vinegoni C, Lee J, Lee SJ, Park JY, Liu AQ, Castro CM, Lee H, Im H, Weissleder R. CytoPAN-portable cellular analyses for rapid point-of-care cancer diagnosis. *Sci Transl Med*. 2020;12:1–14.
24. Joh DY, Heggestad JT, Zhang S, Anderson GR, Bhattacharyya J, Wardell SE, Wall SA, Cheng AB, Albarghouthi F, Liu J, Oshima S, Hucknall AM, Hyslop T, Hall AHS, Wood KC, Shelley Hwang E, Strickland KC, Wei Q, Chilkoti A. Cellphone enabled point-of-care assessment of breast tumor cytology and molecular HER2 expression from fine-needle aspirates. *NPJ Breast Cancer*. 2021;7:1–14.
25. Lehmann-Che J, Amira-Bouhidel F, Turpin E, Antoine M, Soliman H, Legres L, Bocquet C, Bernoud R, Flandre E, Varna M, de Roquancourt A, Plassa LF, Giacchetti S, Espie M, de Bazelaire C, Cahen-Doidy L, Bourstyn E, Janin A, de The H, Bertheau P. Immunohistochemical and molecular analyses of HER2 status in breast cancers are highly concordant and complementary approaches. *Br J Cancer*. 2011;104:1739–1746.
26. Sinn HP, Schneeweiss A, Keller M, Schlombs K, Laible M, Seitz J, Lakis S, Veltrup E, Altevogt P, Eidt S, Wirtz RM, Marme F. Comparison of immunohistochemistry with PCR for assessment of ER, PR, and Ki-67 and prediction of pathological complete response in breast cancer. *BMC Cancer*. 2017;17:124.
27. Di Cosimo S, Pizzamiglio S, Ciniselli CM, Duroni V, Cappelletti V, De Cecco L, De Marco C, Silvestri M, De Santis MC, Vingiani A, Paolini B, Orlandi R, Iorio MV, Pruneri G, Verderio P. A gene expression-based classifier for HER2-low breast cancer. *Sci Rep*. 2024;14:2628.
28. Laible M, Schlombs K, Kaiser K, Veltrup E, Herlein S, Lakis S, Stohr R, Eidt S, Hartmann A, Wirtz RM, Sahin U. Technical validation of an RT-qPCR in vitro diagnostic test system for the determination of breast cancer molecular subtypes by quantification of ERBB2, ESR1, PGR and MKI67 mRNA levels from formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer*. 2016;16:398.

29. Suryavanshi M, Mehta A, Jaipuria J, Kumar D, Vishwakarma G, Panigrahi MK, Verma H, Saifi M, Sharma S, Tandon S, Doval DC, Das BC. Clinical utility of RT-PCR in assessing HER 2 gene expression versus traditional IHC and FISH in breast cancer patients. *Breast Cancer-Tokyo*. 2018; 25:416–430.
30. Borges PCC, Spencer HB, Barbosa C, Costa V, Furtado A, Leal MC, Lopes C, Ferreira D, Carvalho AL, Dos-Santos-Silva I, Santos LL. XPERT® breast cancer STRAT4 as an alternative method of identifying breast cancer phenotype in Cape Verde (preliminary results). *Ecancermedicalscienc*. 2023;17:1–14.
31. Wu NC, Wong W, Ho KE, Chu VC, Rizo A, Davenport S, Kelly D, Makar R, Jassem J, Duchnowska R, Biernat W, Radecka B, Fujita T, Klein JL, Stonecypher M, Ohta S, Juhl H, Weidler JM, Bates M, Press MF. Comparison of central laboratory assessments of ER, PR, HER2, and Ki67 by IHC/FISH and the corresponding mRNAs (ESR1, PGR, ERBB2, and MKi67) by RT-qPCR on an automated, broadly deployed diagnostic platform. *Breast Cancer Res Tr*. 2018;172:327–338.
32. Olufemi A. *How machine shortage slows down Nigeria's fight against TB*. Accessed March 28, 2024. https://www.premiumtimesng.com/health/health-investigations/669126-special-report-how-machine-shortage-slows-down-nigerias-fight-against-tb.html.
33. Li ZY, Bao QJ, Liu CY, Li Y, Yang YH, Liu ML. Recent advances in microfluidics-based bioNMR analysis. *Lab on a Chip*. 2023;23:1213–1225.
34. Otoo JA, Schlappi TS. REASSURED multiplex diagnostics: A critical review and forecast. *Biosensors-Basel*. 2022; 12:1–14.
35. Khelifa L, Hu Y, Jiang N, Yetisen AK. Lateral flow assays for hormone detection. *Lab Chip*. 2022;22:2451–2475.
36. Majdinasab M, Badea M, Marty JL. Aptamer-based lateral flow assays: Current trends in clinical diagnostic rapid tests. *Pharmaceuticals (Basel)*. 2022;15:1–27.
37. Kuswandi B, Hidayat MA, Noviana E. Based sensors for rapid important biomarkers detection. *Biosens Bioelectron X*. 2022;12:100246.
38. Naghdi T, Ardalan S, Asghari Adib Z, Sharifi AR, Golmohammadi H. Moving toward smart biomedical sensing. *Biosens Bioelectron*. 2023;223:115009.
39. Mahmoodpour M, Kiasari BA, Karimi M, Abroshan A, Shamshirian D, Hosseinalizadeh H, Delavari A, Mirzei H. Paper-based biosensors as point-of-care diagnostic devices for the detection of cancers: A review of innovative techniques and clinical applications. *Front Oncol*. 2023;13:1–16.
40. Ince B, Sezgintuerk MK. Lateral flow assays for viruses diagnosis: Up-to-date technology and future prospects. *Trac-Trend Anal Chem*. 2022;157:1–19.
41. Boehringer HR, O'Farrell BJ. Lateral flow assays in infectious disease diagnosis. *Clin Chem*. 2022;68:52–58.
42. de Oliveira KA, Sengupta S, Yadav AK, Clarke R. The complex nature of heterogeneity and its roles in breast cancer biology and therapeutic responsiveness. *Front Endocrinol*. 2023;14:1–13.
43. Mavrommati I, Johnson F, Echeverria G, Natrajan R. Subclonal heterogeneity and evolution in breast cancer. *Npj Breast Cancer*. 2021;7:1–9.
44. Coleman WB, Tsongalis GJ. *Diagnostic molecular pathology: A guide to applied molecular testing*. Academic Press; 2023.
45. Yersal O, Barutca S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World J Clin Oncol*. 2014;5:412–424.
46. Di Nardo F, Chiarello M, Cavalera S, Baggiani C, Anfossi L. Ten years of lateral flow immunoassay technique applications: Trends, challenges and future perspectives. *Sensors (Basel)*. 2021;21:1–33.

43

Development of LLM for Prostate Cancer—The Need for Domain-Tailored Training

Amara Tariq, Man Luo, Aisha Urooj, Avisha Das, Jiwoong Jeong, Shubham Trivedi, Haidar Abdul-Muhsin, Umer Ghaffar, Nathan Yu, Bhavik Patel, and Imon Banerjee

43.1 Introduction

A generic large language model (LLM) excels in its flexibility and versatility across diverse topics. A generic LLM (e.g., GPT4) is trained on a wide range of topics and text from diverse sources for a capturing broad understanding of human language and generating human-like text across various domains without specializing in any specific field. However, our hypothesis posits that it could encounter difficulties when handling complex clinical notes and adhering to clinical guidelines. This challenge arises from its potential to generate generalized responses to domain-specific questions, where specialized knowledge and precise terminology are critical for informed decision-making. In this work, we present our strategy for developing domain-tailored large language models which cover the vocabulary of the target domain and train on reliable sources of clinical information. Prostate cancer was used as a use-case target domain for this work. During diagnosis, treatment planning, and post-treatment disease management, vast amounts of clinical information is generated and recorded by the care team about prostate cancer patients. This is usually in free-text form, including clinical notes and radiology and pathology reports. Large language modeling provides a unique opportunity to process this information for developing reliable agents for interactive knowledge sharing in a discreet way.[1–3] A domain-specific LLM is trained specifically on text and data from a particular domain that is optimized to understand the specific jargon, nuances, and context of its trained domain, making it more accurate and contextually appropriate for tasks within that domain.

We argue that generic LLMs are at a disadvantage when processing complex clinical notes and compliant with clinical guidelines as they may generate generic responses to domain-specific questions where specialized knowledge and proper terminology are crucial for decision-making. Domain-specific LLMs typically provide more accurate results within their domain due to their specialized training, whereas a generic LLM might lack the depth of understanding required for nuanced domain-specific tasks. Generic LLMs also suffer from hallucination, which can be especially dangerous in sensitive domains like clinical medicine.[3–6] Research effort has been put into domain-specific LLM development for broad medicine domains such as MedPALM[7] (540B), MedAlpaca[8] (13B), and BioGPT[9] (1.5B). Given that these models are intended to cover complete domains of medicine and biology, the models are growing in sizes reaching on the order of 100 billions of parameters to cover such vast fields since medicine includes subspecialties, with different domain-specific knowledge such as oncology, preventive medicine, and emergency medicine where the syntactic language and treatments vocabulary differ significantly.

Moreover, popular generic LLMs, for example GPT[10] and LLaMA,[11] have been trained on text crawled from the world wide web without verifying the reliability of the information source, which makes the LLM vulnerable to data poisoning attacks.[12] Training of specialized LLMs like BioGPT and MedPALM have been focused on domain-specific knowledge corpora such as millions of scientific abstracts from PubMed. However, patient records are arguably the most critical source of information to capture the domain of any disease and include several distinct documents that include oncology notes, nursing notes, ED reports, discharge summaries, and radiology and pathology reports. While GPT2 and LLaMA, and even medicine-focused BioGPT, have been trained on extremely large amounts of information, their training data did not include such patient-specific documents because of patient privacy requirements set by the Health Insurance Portability and Accountability Act (HIPAA). We argue that this puts them at an inherent disadvantage in terms of understanding the clinical decision-making process and information dissemination to the patients themselves. While research papers and expert treatment guidelines can impart knowledge about the latest treatment options and their general pros and cons, clinical notes actually describe how physicians and patients may decide on a treatment option and their following outcomes. We argue that the inclusion of clinical notes and reports is critical when building a language understanding and generation tool for any particular disease.

To comprehend the gain of domain-specific LLM over generic LLM in the oncology domain, we targeted the prostate cancer specialty, which is the most commonly diagnosed non-skin cancer among men in the United States with a five-year survival rate of about 100% for local or regional cancer.[13, 14] The diagnosis of prostate cancer is often followed by immediate decline in mental and physical health.[15, 16] Limited knowledge regarding the disease, embarrassment around physical examination such as rectal exam, and discussion of sexual

DOI: 10.1201/9781003449942-48

symptoms with healthcare professionals, potentially exacerbated depending on the provider gender,[17, 18] are known causes of poor outcome among older patients and/or those belonging ethnic minorities.[19–21] These uncertainties and the anxiety related to sexual symptoms could lead patients to search for information through online search engines (often backed up by generic LLM) in which the patient may end up with incomplete, contradictory, misleading, and/or inaccurate information.[22] This, in turn, could cause further delay in treatment and ultimately result in poor outcomes.

We present our work on the development of a domain-specific LLM to overcome the shortcomings of general purpose LLMs when applied to sensitive domains of prostate cancer for medical decision-making and information dissemination. We also devised a two-phase domain-specific self-supervised training methodology with clinical notes to ensure that the model learned the nuances of the chosen prostate cancer domain. We evaluated the proposed LLM against similarly sized general purpose LLM (GPT2-small) as well as larger domain-specific LLM (BioGPT) on two generic tasks: (i) clinical information mask prediction and (ii) question-answering, and two clinical tasks: (iii) treatment suggestion and (iv) guideline compliance analysis. We also measured the performance in a user study with three independent raters who evaluated the generated responses in terms of completeness, correctness, and relevance.

43.2 Materials

With the approval of the internal review board (IRB) at Mayo Clinic, we collected ~1.8M clinical notes and 23,665 radiology and pathology notes concerning more than 15k patients treated for prostate cancer at all three major sites of Mayo Clinic (Rochester, Arizona, Florida). Detailed characterization of these patients is provided in Table 43.1. We developed rule-based natural language processing (NLP) techniques to gather information about cancer characteristics of these patients such as lesion size and Gleason score. These patients were also part of an enterprise-wide cancer registry that recorded their clinical outcome information. Model development was supported by Lambda machines with four A6000 GPUS with 50 MB memory each.

TABLE 43.1

Cohort Characteristics for the Prostate Cancer Patients

Characteristics	Value
Cohort for self-supervised pretraining	
Patients	15,341
Total number of clinic notes	1,852,213
Clinical terms	Identified by UMLS parsing: 193936 After term-frequency and document-frequency based filtering: 88,743
Radiology reports	23,665
Follow-up	mean: 2.8+/−1.8 years max: 6.6 years
Age	66.7+/−8.6 years
Race	Caucasian: 14,213 (93%) African American: 548 (4%) Asian: 171 (1%)
Gleason Score (aggressive vs. non-aggressive)	6 or lower: 17%; 7 or higher: 65%
PSA	PSA ≤ 10: 57%; 10 < PSA < 20: 13%; PSA ≥ 20: 10%
Tumor size	mean: 366.3 mm, min: 1.0mm, max: 990 mm

Note: Earliest inclusion date is 2017, thus the follow-up is limited.

43.3 Methodology

Figure 43.1 shows the overall framework for domain-specific LLM development and application, and in the following subsections, we describe several steps involved in the development and training of a domain-specific LLM.

43.3.1 Text Pre-Processing

Clinical notes often contain identifiable protected health information (PHI) about both patients and providers, for example name, date of birth, and phone number. Large language models display behaviors like memorization that leaves open the possibility of unmasking identifiable patient information used during the training phase. To ensure patient and provider privacy was not violated even by sharing the weights of our domain-specific language model, we anonymized all clinical notes prior to their use in training of the proposed LLM using a pretrained BiLSTM model with CRF output layer.[23] Information like patient and provider names, dates, and location were replaced with special tokens like "[NAME]," "[DATE]," and "[LOC]." Only imaging finding sections from radiology and pathology information were used for LLM training. These sections are devoid of any patient information. Electronic signatures of providers were dropped using regular expression to ensure provider privacy. To improve the quality of training data for broadening the model learning space to capture rare events, textual dataset cleaning steps also included filtering of repetitive sentences using fuzzy matching techniques to filter out sentences that are too similar to other sentences within the clinical notes for one patient.

43.3.2 Clinical Information Marking

During the training, instead of only oncology-specific notes, we included various types of clinic notes to capture the clinical history and patient timeline. To ensure that domain-specific LLM was focused on cancer treatment, side effects, outcomes, and symptom-related information, we identified domain-specific information from free-text clinical notes by parsing them using standardized medical vocabularies and ontologies leveraging the Unified Medical Language System (UMLS) metathesaurus made available by the National Institute of Health.[24] Text parsing tools designed on top of UMLS identify medical concepts from free text and map them to concept unique identifiers (CUI). Concepts are categorized into 127 "semtypes." We manually selected 36 semtypes relevant to cancer treatment, symptoms, side effects, clinical status, and anatomical details for oncology. Filtered text from

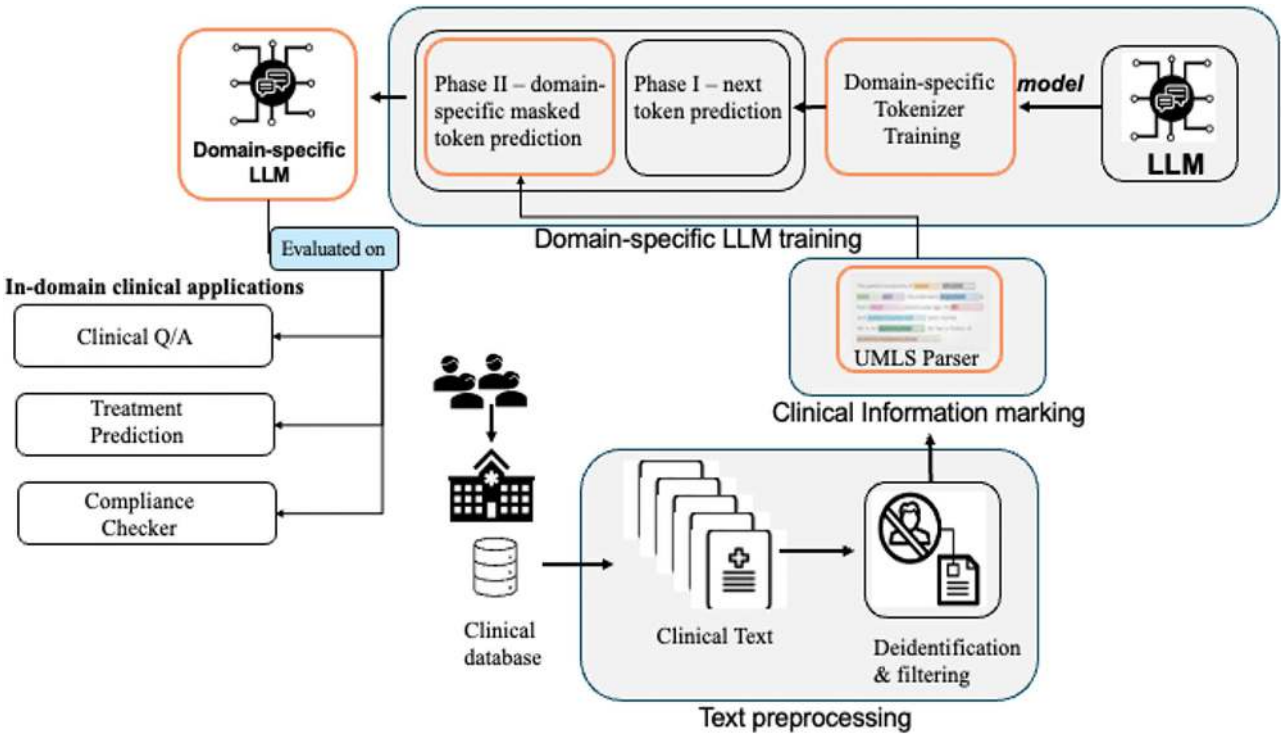


FIGURE 43.1 Domain-specific LLM training and application framework; highlighted (orange-colored) modules indicate domain-specific modifications to LLM design framework.

clinical notes was divided into sentences and passed through the UMLS parser one by one, and relevant medical entities and their location in the sentences were stored. We used a Python parser for UMLS designed on a type of medical named entity extractor for parsing.[25] Our processing indicated that 84% of the sentences from clinical notes included important medical concepts. Medical concepts have varying distribution in the data with 33% concepts with one occurrence and 0.3% with greater than 10,000 occurrences. A total of 193,936 concepts were identified. We computed term frequency–inverse document frequency (tfIdf) and used this measure to filter out very generic and too rare terms, leaving ~88K terms.

43.3.3 Domain-Specific Tokenizer

Modern LLMs rely on word-piece tokenizers that meet the challenge of out-of-vocabulary tokens by chopping up the word into smaller foundational pieces; for example, "carcinoma" may be chopped into "carcinoma" as ["car," "##cin," "##oma"], in which the word lost its actual meaning and was replaced by some non-relevant junk words, such as "car." The words that were replaced by the junk pieces may not play the original role in deriving the contextual representation of the sentence or the paragraph, even when analyzed by the powerful transformer models. A passage full of relevant medical information containing N words may be divided in M tokens when M>>N. In extreme cases, M may exceed the length of the context window causing the model to ignore some information. On the other hand, a domain-specific tokenizer may be able to preserve boundaries of medical concepts and entities, thus shortening the length of the relevant text and allowing for better attention within the textual context. Moreover, they often fail to tokenize domain-specific compound words, such as hormone therapy, and they are brittle to noise, meaning that even minor typos may affect the representation of a word.

Therefore, we decided to train a domain-specific word-piece tokenizer with a vocabulary size of 50 thousand on the cleaned deidentified text, which was similar in size to popular tokenizers such GPT-2 tokenizer (vocab size of 50,257).

43.3.4 Domain-Specific LLM

We based our LLM on a popular 12-layer masked self-attention decoder-based architecture similar to the one used by GPT-2.[10] The model was composed of about 124M trainable parameters and trained in a self-supervised fashion. Simple language-based self-supervision pretraining tasks like the prediction of the next token seem enough for general purpose models like GPT to gain fundamental understanding of natural language. We built upon this line of research by introducing a two-phase domain-specific self-supervision methodology.

43.3.5 Phase I—General Purpose Language Understanding

Under phase I, our LLM was trained on free text data of prostate cancer patients for the task of next token prediction. A domain-specific tokenizer was used to generate token sequences out of all sentences of reports and notes. Given a sequence of tokens $\left[x_1, x_2, \ldots . x_n, \ldots x_N, x \in V\right]$, the model was

trained to maximize the probability of the next word in the sequence, out of all vocabulary tokens V, conditioned on the previous words, that is, $P(x_n \mid x_{n-1}, x_{n-2}, ...x_1)$. The maximum number of previous tokens considered are based on the context length, which was set to 128 in our experiments. In terms of training loss for backpropagation, we relied on cross-entropy loss computation with every predicted token treated as predicted class label while vocabulary set used as class labels' set:

$$l_n = -log\left(\frac{p_{x_n}}{\sum_v^{|V|} p_v}\right)$$

$loss = \frac{\sum_n^N l_n}{N}$, where N is the length of the sequence.

This loss computation encourages the model to understand the linguistic structure of the text of the domain of prostate cancer and be able to predict the next word by considering the previous words within context. We named the model trained under this task only as PCa-LLM.

43.3.6 Phase II—Domain-Specific Language Understanding

In phase II of training, we aimed for the model to learn the nuance of clinical language specific to prostate cancer domain. We marked clinical terms in free-text data of prostate cancer patients using the UMLS parser. In phase II of training, we used masked clinical token prediction as a self-supervised training task. Each text snippet may contain more than one clinical term. Terms were chosen with 50% probability for masking. Note that one term may be split into multiple tokens. In our training scheme, all tokens belonging to that term were masked and the model was trained to predict masked tokens. For example, the sentence "*common side effects of hormone therapy may include shrinkage of testicles*" included clinical terms of "hormone therapy" and "shrinkage of testicles." If "hormone therapy" was selected randomly for marking, all tokens belonging to it ({"hormone," "therapy"}) would be masked. In this case, the model received "*common side effects of <mask> <mask> may include shrinkage of testicles*" and was tasked with predicting "hormone" and "therapy."

In terms of loss computation, we again utilized cross-entropy. However, the loss was computed only on the masked tokens as described by the following equation.

$$l_n = -log\left(\frac{p_{x_n}}{\sum_v^{|V|} p_v}\right) \quad if\ x_n\ is\ masked \quad else\ 0$$

$$loss = \sum_n^N \frac{l_n}{\sum_n^N 1.\{if\ x_n\ is\ masked\}}$$

We argue that this two-phase training scheme allowed the LLM to focus on domain-specific language in addition to learning correlation between treatments, side effects, symptoms, and cancer characteristics. Such correlations can be essential for domain-specific downstream tasks like patient education and treatment recommendation. For such downstream tasks, the model must be able to predict clinical terms using the context provided in the input text. This phase of training essentially taught the model to do the same. We call the model that underwent both phases of training PCa-MLM.

43.3.7 In-Domain Clinical Applications

43.3.7.1 Masked Clinical Information Prediction

This task has a similar objective to phase II of training and basically evaluates the model performance for the masked word prediction to test the linguistic understanding of the domain. We randomly selected clinical terms and masked them in a set of held-out sentences (text not used during the training phase). The model was tasked with retrieving masked terms from the vocabulary of its tokenizer. We evaluated the model through an evaluation metric for information retrieval systems, that is, recall@K where K was varied between 1 and 10, based on ranked retrieval results generated by the model probability.

43.3.7.2 Clinical Question-Answering

One of the most popular downstream applications for any LLM is question-answering that can allow the users to interact with the LLM in an interactive and flexible manner. With expert-oncologist's guidance, we manually curated questions from treatment guidelines set by the American Cancer Society for prostate cancer. We browsed through each section of the guidelines and defined question-answer pairs of the following broad categories.

- Treatment pathway recommended for cancer at different stages/risk category
- Treatment pathway recommended for cancer at different stages/risk category under certain conditions (older patient, patients in poor health, patients with certain comorbidities)
- Side effects of treatment
- Long term risks of treatment
- Medication used in treatment
- Comparative pros and cons of different treatment

About 300 question-answer pairs were manually curated. We then employed an open-source general purpose LLM—LLaMA (13B)[11]—within a generative framework to paraphrase each question four different ways. After filtering for nonsensical paraphrasing such as empty strings, we were left with 1222 QA pairs. Randomly selected unique 35 QA pairs were kept in a holdout set for evaluation purposes after dropping paraphrases of the same questions.

43.3.7.3 Treatment Category Prediction

We also evaluated our model on the prediction task for recommended and actually administered treatment categories from urology consultation notes using only the free-text describing

TABLE 43.2

Curated Cohort Description for Treatment Category Prediction

Prostate Cancer Treatment Prediction Cohort	
Patients	448
Recommended treatment category distribution	Curative treatment: 155 Surveillance: 40 Surgery: 40 Radiation: 33 Chemohormonal:30 Focal therapy: 22
Administered treatment category distribution	Surgery: 130 Radiation: 100 Surveillance: 45 Chemohormonal: 30 Focal therapy: 15

patient attributes and laboratory test results and radiology imaging findings performed before the appointment with the urologist. Such prediction can guide the patient to seek appointments with relevant medical specialists, such as a surgeon versus a radiation oncologist, and can avoid unnecessary medical appointments, saving physicians' and patients' time and efforts.

We curated information through a manual chart review about both recommended treatment and the treatment administered to the patient, taking into account their choice. This information was collected for 448 patients described in Table 43.2. A random sample of 20% of the patients was held out for evaluation purposes. Based on the clinical workflow, several competing treatment categories were identified for the models to predict, such as curative vs. non-curative (recommended and administered) treatment, radiation vs. surgery (administered), and chemotherapy vs. all other treatments (recommended and administered). In this dataset, a recommendation of curative treatment indicated that the patient can choose between radiation and surgery while some patients were specifically recommended either radiation or surgery.

We parsed the text of urology consultation notes to exclude "Assessment/Plan" sections to avoid leaking any information from the urologist. Hence, the model predictions were purely based on information (clinical history, labs, and demographics) collected before the appointment. Limited range question-answering models were built on top of the domain-specific and generic LLMs to answer (Yes/No) questions about treatment categories in the context of parsed information from urology consultation notes. We also visualized the model attention to investigate critical information as identified by the model.

43.3.7.4 Clinical Guideline Compliance

One of the primary obstacles hindering the integration of generative language models into the field of medicine is their propensity to generate erroneous or "hallucinated" information and not being compliant with the current guideline. To assess the utility of domain-specific vs. generic LLM for clinical guideline compliance, we curated a dataset with "compliant" and "non-compliant" statements regarding the treatment of prostate cancer. Compliant statements were gathered from the answers in our question-answering dataset that is curated based on current guidelines set by the American Cancer Society. Clinical experts provided one "non-compliant" statement (against the guideline) for every "compliant" statement. For example, "cabazitaxel is often the next chemo drug prescribed (along with a steroid) if docetaxel does not work" is a compliant statement, whereas comparative non-compliant statement is "lupron is often the next chemo drug tried (along with a steroid) if docetaxel does not work." A total of 632 statements (1:1 ratio of compliant to non-compliant statements) was curated and 20% of this data was held out for evaluation purposes. We built text classification models using both domain-specific LLM and generic LLM (GPT-2) and performed few-shot training to identify whether the input text statement is compliant or non-compliant.

43.4 Results

We selected two generic models as comparative baselines for our model—GPT2-small and BioGPT. Selected version of GPT2 was similarly sized as our model but lacked the advantages of domain-specific vocabulary, cancer-specific training data, and targeted fine-tuning tasks. BioGPT was about three times bigger and was designed to cover a vast domain of medicine as a whole, and thus procure a clinical domain-specific vocabulary; however, lacking the domain-tailored self-supervised pretraining. Comparison with this model was meant to establish benefits of fine-grained domain definition (cancer vs. generic medicine).

43.4.1 Domain Vocabulary Coverage

To measure the vocabulary coverage for the clinical domain, we evaluated the general purpose GPT-2 tokenizer, medicine-focused BioGPT tokenizer, and our tokenizer on clinical terms extracted by the standard UMLS parser. The GPT2 tokenizer tended to divide common medical concepts like "prostatectomy" (GPT2 tokenizer: 2 tokens; our tokenizer: 1 token), "diabetes mellitus" (GPT2 tokenizer: 4 tokens; our tokenizer: 2 tokens), and so on. The BioGPT tokenizer showed better clinical vocabulary coverage for up to 3 tokens than our tokenizer. However, it still was unable to retain prostate cancer–related words as a single token like "genitourinary" (GPT2: 5 tokens; BioGPT: 5 tokens; ours: 1 token) and "erectile dysfunction" (GPT2: 5 tokens; BioGPT: 3 tokens; ours: 2 token). Figure 43.2 shows distribution of medical concepts divided into token sets of varying length by the three different tokenizers. Even though GPT2 coverage is minimal for the clinical terms with the lowest number of tokens ≤1, BioGPT and our model have similar coverage of the common clinical terms—not specific for prostate cancer domain.

43.4.2 Masked Clinical Information Prediction

For masked clinical information retrieval, each model produced a probability estimate for every tokens in its vocabulary for prediction as replacement for the masked token. Hence, a

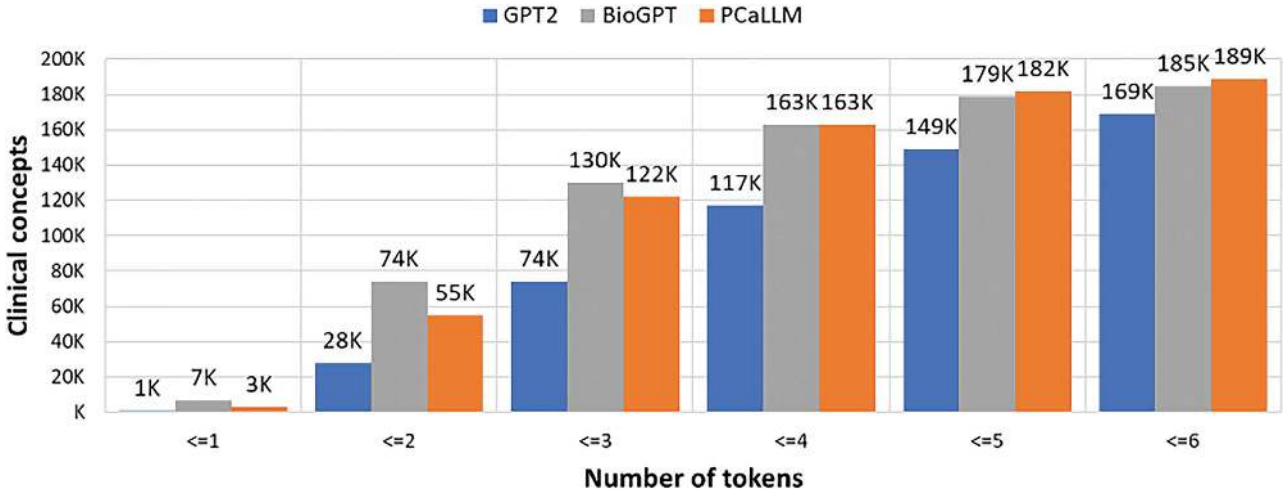


FIGURE 43.2 Distribution of clinical concepts across number of tokens formed by three different tokenizers.

TABLE 43.3

Clinical Information Retrieval Performance for Different LLMs and Proposed Domain-Specific Model: Phase I Training—PCa-LLM and Phase II Training—PCa-MLM

Model	*Recall@10*	*Recall@5*	*Recall@1*
GPT-2	0.017	0.010	~0.000
BioGPT	0.017	0.010	0.003
PCa-LLM	0.228	0.173	0.076
PCa-MLM	0.305	0.239	0.111

ranked list of all tokens in vocabulary (sorted in descending order of probability estimate) was available for each masked token. Each model was evaluated to see if the correct token was found in top K tokens for each masked token (recall@K). The performance was aggregated for all masked tokens (Table 43.3). We evaluated two phases of our model—*Phase I*: PCa-LLM and *Phase II*: PCa-MLM. Both versions significantly outperformed both comparative models with significant margin at each K, even when the clinical vocabulary coverage for BioGPT was similar to the domain-specific model, which highlights the fact that the domain-tailored model was able to capture superior understanding of the domain compared to generic models.

43.4.3 Question-Answering

For evaluation of the question-answering task, we relied on a user study where three independent clinical experts were asked to evaluate the responses of each model on the basis of three criteria.

Correctness: if the generated response was correct according to the clinical guidelines.

Completeness: if the generated response was complete and covered all aspects of the ground truth answer.

Relevance: if the generated response was relevant to the question despite being incorrect or incomplete based on the clinical guideline.

TABLE 43.4

Manual Evaluation Study for Question-Answering Performance of Different LLMs

Model	*Criterion*	*Expert 1*	*Expert 2*	*Expert 3*	*Median*
GPT2	*Correctness*	16 (46%)	7 (20%)	4 (11%)	7 (20%)
	Completeness	12 (34%)	6 (17%)	3 (9%)	6 (17%)
	Relevance	31 (89%)	17 (49%)	9 (26%)	17 (49%)
BioGPT	*Correctness*	23 (66%)	21 (60%)	23 (66%)	23 (66%)
	Completeness	14 (40%)	20 (57%)	14 (40%)	14 (40%)
	Relevance	35 (100%)	26 (74%)	27 (77%)	27 (77%)
PCa-MLM	*Correctness*	27 (77%)	27 (77%)	18 (51%)	27 (77%)
	Completeness	14 (40%)	23 (66%)	11 (31%)	14 (40%)
	Relevance	35 (100%)	32 (91%)	29 (83%)	32 (91%)

Note: Represents the count and (%).

While correctness and completeness evaluate the accuracy of the response, the third criteria was introduced to assess the model's comprehension of the question. Models, especially general purpose models—GPT2, may generate generic responses if they are unable to understand the granularity of the question.

After deidentification of the model names, all the experts were asked to provide a 1/0 response for each question for all criteria; 1 indicated that the criteria were met (complete, correct, or relevant) while 0 indicated that the generated response did not meet the criteria (incomplete, incorrect, or irrelevant). Responses in each category by every user were summed over 35 questions (Table 43.4). By aggregate opinion of users (median scores), our model outperformed both comparative models based on median value for all the criteria that are rated higher for our model. Expert 3 rated BioGPT responses higher for correctness and completeness while still rating our model higher for relevance based on the generated responses being more relevant for the prostate cancer domain. Table 43.5 shows sample answers generated by all three models, where the generic BioGPT model tended to generate common symptoms and side effects that are often not relevant for prostate cancer, for example allergic reaction and nausea.

TABLE 43.5

Sample Answers Generated by Different LLMs for the Same Questions Related to Prostate Cancer

Question 1	What treatment is recommended if the PSA level is greater than 20 and prostate cancer with a Gleason score of 8 to 10 is found in lymph nodes?
Ground truth answer	hormone therapy, possibly along with external radiation
GPT-2 response	it means that when people are given drugs that are similar to what they are given in clinical trials
BioGPT response	hormone therapy, possibly along with external radiation therapy, may be recommended. This is a type of focal therapy
PCaMLM response	hormone therapy, possibly along with external radiation therapy
Question 2	What is ablative treatment?
Ground truth answer	heat or cold or other methods to destroy (ablate) prostate tissue, rather than removing it with surgery or treating it with radiation. Methods include high intensity focused ultrasound (HIFU) and cryotherapy.
GPT-2 response	surgery (TURP) may be recommended to relieve symptoms like bleeding and urinary obstruction
BioGPT response	Cryotherapy (also called cryosurgery or cryoablation) is the use of very cold temperatures to freeze
PCaMLM response	HIFU uses highly focused ultrasound beams to heat and destroy prostate tissue and shrink prostate over time
Question 3	What are the side effects of bisphosphonates
Ground truth answer	flu-like symptoms and bone or joint pain, kidney problems, so patients with poor kidney function might not be able to be treated with these medicines
GPT-2 response	Corticosteroids are drugs that target PD-1, a checkpoint protein on immune system cells
BioGPT response	denosumab (Xgeva), a bisphosphonate-like zoled
PCaMLM response	pain, swelling in the legs or feet, nausea, and diarrhea. These drugs can also cause kidney damage.
Question 4	What are common side effects of chemotherapy drugs?
Ground truth answer	Hair loss, Mouth sores, Loss of appetite, Nausea and vomiting, Diarrhea, Increased chance of infections (from having too few white blood cells), Easy bruising or bleeding (from having too few blood platelets), Fatigue (from having too few red blood cells)
GPT-2 response	doctor might recommend external radiation therapy to the prostate bed (the area where the prostate was before it was removed)
BioGPT response	allergic reaction, nerve damage, leukemia, and increased risk of blood clots in the legs or lungs.
PCaMLM response	fatigue, nausea, vomiting, and diarrhea.
Question 5	What are common side effects of brachytherapy?
Ground truth answer	bowel problems, urinary problems, erection problems
GPT-2 response	nausea, diarrhea, and feeling weak or tired.
BioGPT response	(1) bowel problems, (2) urinary problems, (3) erection problems
PCaMLM response	Pain in the treatment area, Blood in the urine, An increased risk of a urinary tract infection (UTI)

TABLE 43.6

Comparative Performance of Generic (GPT2), Medicine Specific (BioGPT), and Domain-Specific LLM (PCa-MLM) for Treatment Category Prediction Task

Treatment Category Labels	***PCaMLM (AUROC)***	***GPT-2 (AUROC)***	***BioGPT (AUROC)***
Curative vs. non-curative (recommended)	0.78	0.65	0.48
Curative vs. non-curative (administered)	0.77	0.74	0.71
Radiation vs. surgery (administered)	0.59	0.59	0.53
Chemohormonal vs. other treatments (recommended)	0.99	0.99	0.98
Chemohormonal vs. other treatments (administered)	0.99	0.99	0.99

43.4.4 Treatment Prediction

Table 43.6 shows results of all models for various competing treatment categories using the same clinical text as context. We observed that the domain-specific model (PCaMLM) outperformed the generic LLM (GPT-2) and clinical LLM (BioGPT) for treatment category prediction on three out of five cases whereas all models achieved near-perfect performance in the remaining two tasks. This alludes to superior understanding of clinical text by domain-specific model and often able to predict complicated decisions (0.78 for curative vs. non-curative).

43.4.5 Compliance Detection

For the compliance detection task, we trained the modes in a few-shot in-context learning strategy with a limited number of training samples. We argue that this scenario simulates a real-world application where new treatments and medications are introduced and approved for cancer management and other treatments and medications are removed from practice. In such a scenario, the latest compliance detection model can be fine-tuned with a small number of statements. The PCaMLM-based compliance checker outperformed both comparative models when very few training samples were made available (Table 43.7). Three times larger, BioGPT was able to close the performance gap, eventually outperforming our model when fine-tuning was performed with more than 100 samples. GPT-2 showed uneven performance gain as the number of training samples were increased. We hypothesize that rather than semantics, GPT-2 focuses on linguistic style to identify compliant versus non-compliant sentences. Therefore, it fails to show consistent performance gain by increasing the number of training samples. The PCaMLM model shows consistent performance over a wide range of training samples, indicating that the model

TABLE 43.7

Comparative Performance of Generic and Domain-Specific LLM for Compliance Detection Task

No. of Training Examples	*PCaMLM (AUROC)*	*GPT-2 (AUROC)*	*BioGPT (AUROC)*
10	0.53	0.36	0.39
25	0.53	0.30	0.49
50	0.54	0.48	0.51
100	0.58	0.37	0.57
150	0.59	0.55	0.64

relies on its domain knowledge obtained during pretraining to perform this task. On the other hand, BioGPT relies heavily on the availability of training data to perform this task.

43.5 Discussion

Even though general purpose huge LLMs have demonstrated amazing capabilities to generalize to a wide variety of domains as zero-shot frameworks, their performance still suffers in terms of relevance or specificity to a sensitive domain and may "*hallucinate*" while being forced to generate a predefined set of critical information.[4, 6, 26] Privacy concerns make it impossible to train such models with protected healthcare data that may be critical for application of these LLMs in clinical workflow, particularly for novice end users. We present a framework for *design and development of domain tailored LLM with prostate cancer chosen as a use-case using deidentified clinical narratives.* Our framework includes domain-specific customization steps for tokenizer design, data curation, and self-supervised two-phases training. We argue that development of domain-specific LLMs provides a systematic solution to the challenges faced in terms of deployment of huge general purpose LLMs in sensitive domains. While some research has been done to specialize LLMs for sensitive domains like medicine and biology, this line of research usually focuses on the curation of large domain-specific datasets.[7, 9] BioGPT and MedPALM are the two popular biomedical LLMs developed to support this line of research. These LLMs still cover vast fields of medicine and biology, and their training has been focused on public datasets. These LLMs have absorbed the knowledge of vast sources of medicine from published literature but have had little exposure to the real clinical data and thus patient-specific treatment decisions. In comparison, our framework focuses on training the model on clinical narratives including clinical notes and radiology and pathology reports. Thus, the models trained under our framework will be exposed to unique characteristics of patient-physician exchange for the particular targeted domain and able to learn the nuance of the treatment management.

We trained a small sized LLM (124M parameters) for the particular use case of prostate cancer, which not only outperformed a similarly sized general purpose LLM (GPT-2 small) but also outperformed the three times larger specialized LLM (BioGPT) for four downstream tasks. The specialized tokenizer allows the LLM to understand and focus on domain-specific terms instead of dividing them up into semantically imperfect tokens. The GPT2 tokenizer chops up clinical terms into a larger number of tokens while our specialized tokenizer and BioGPT tokenizer can comprehend most of the clinical terms as a whole. As these tokenizers were designed for the fields of medicine and biology, they tend to describe medicine-related terms in fewer tokens. BioGPT tokenizer covers more clinical terms within three tokens splitting than our tokenizer. However, clinical terms that are more specific to prostate cancer (i.e., found more commonly in prostate cancer–related clinical text) are better handled by our tokenizer than the BioGPT tokenizer. These terms include "erectile dysfunction," "primary malignant neoplasm," "prostatic hypertrophy," "cystolitholapaxy," "cystourethroscopy," and others. In addition, several prostate cancer–related drug names are part of our specialized vocabulary, including "abiraterone" and "docetaxel." We plan to make our model publicly available after ensuring no identifiable patient or physician information is part of the model's vocabulary.

For the task of masked clinical term prediction, it is intuitive that PCa-MLM (phase I and II of training) outperformed PCa-LLM (phase I training only) as phase II training was performed using a very similar self-supervision task. However, even PCa-LLM (phase I training only) outperformed GPT-2 and BioGPT by wide margins even though PCa-LLM has been trained using generic self-supervision tasks of next token prediction. This speaks to the benefits of specialized domain-specific tokenizers. The BioGPT tokenizer displayed somewhat better coverage of generic clinical vocabulary than our tokenizer (as shown in Figure 43.2), but the model still performed extremely poorly for prediction of domain-specific (prostate cancer–related) clinical information retrieval.

One of the main applications of our specialized LLM can be patient education as an interactive question-answering framework. Therefore, we tested the performance of the model on a question-answering task. Note that these questions were curated from the website of the American Cancer Society. Since this data is publicly available, it is likely to be included in the training data of GPT2 and BioGPT's training corpus. On the other hand, our model was only trained on clinical data generated within the hospital. It was never trained explicitly on the textual guidelines. Despite this, our model outperformed GPT2 by a wide margin on all three evaluation criteria. GPT2 often generated generic or irrelevant responses. Even when answering questions about side effects (Question 5), it described generic side effects like "feeling weak and tired," which was not very specific to the treatment in question. BioGPT does not suffer from this problem and answers were mostly relevant even when it failed to produce the correct answer. For example, it generated names of bisphosphonate drugs when the question was about side effects of bisphosphonates (Question 3). In answer to Question 1, it generated a treatment option which is available for cancer (focal therapy) but was not part of the ground truth answer. PCaMLM and BioGPT responses differed from each other for Questions 2 and 4, but were still relevant and somewhat appropriate for the question.

For the clinical downstream task of treatment category prediction, domain-specific models outperformed generic LLM and medicine specific BioGPT for three out of five categories

while performing comparably on the remaining two. Treatment category set of radiation and surgery is an interesting case where the models produced poor performance since patients are recommended curative treatment courses and are able to choose between radiation and surgery based on their personal preference, which cannot be derived from the clinical history recorded urology consultation notes. Even in this case, PCaMLM was able to predict better than three-times larger BioGPT. Task of differentiating patients recommended for chemohormonal treatment seems to be an easy task because only patients with metastases were recommended for this treatment in our cohort, hence mention of metastasis became a clear indication of chemohormonal treatment. We visualized the model's attentions on the input text while answering questions about treatment (Figure 43.3a), which clearly indicates that the model paid attention to "skeletal metastasis" when generating an affirmative answer to the chemohormonal treatment question. Clinical history and Gleason score were considered important while predicting curative treatment (Figure 43.3b). PCaMLM also demonstrated superior performance for the compliance detection task when very few training samples were available (few-shot learning scenario). Such a compliance checker model can be an essential tool in curtailing the trend of hallucinated information generation through LLM. This can also be used to suppress information about out-of-date treatment and medications that may have been learnt by the model during pretraining on large clinical notes' set spanning many years.

Our experiments clearly establish the benefits of developing and employing domain-tailored LLMs for sensitive domains like cancer, especially for building interactive knowledge dissemination tools. While the initial building blocks (e.g., Q&A, compliance check, prediction) are already in place, we plan to expand upon this line of research by developing patient- and physician-facing chatbots on top of our domain-specific LLM for the use case of prostate cancer. This line of research will provide a solution to reliability and specificity issues faced by a general purpose chatbot like ChatGPT in the field of cancer.

43.6 Future Trends

Large language modeling provides a tremendous opportunity to achieve human-like performance for applications based on automated understanding and generation of natural language. However, cost of development of such models in terms of computational resources and training data is exorbitant. Still, LLMs like GPT-4 and LLaMa have proven their merit as zero-shot learners for a variety of downstream tasks, thus justifying the effort and resources required to build these models. On the other hand, several research studies have documented the limitations of these models such as hallucinating critical information and generating vague responses. These limitations are more critical when these models are applied in specialized and sensitive domains like medicine for tasks that require specific information like physician assistance or patient education.

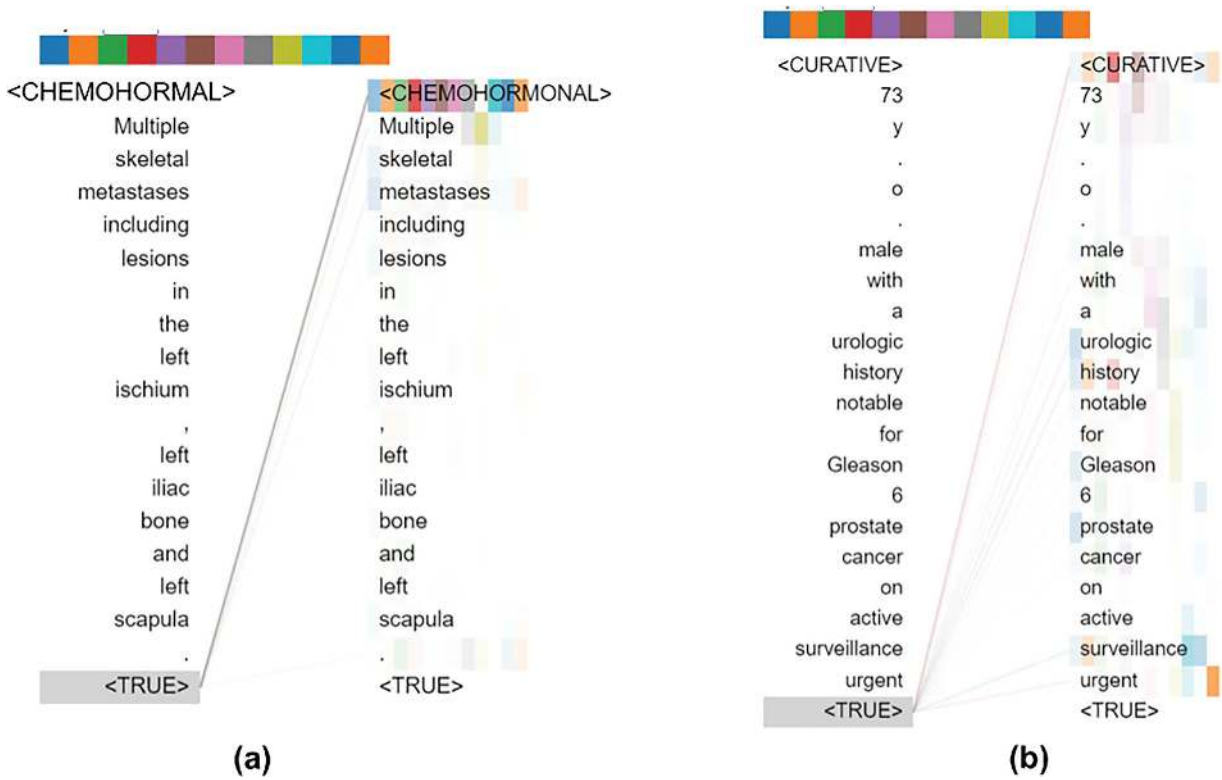


FIGURE 43.3 Visualization of model attentions from all attention heads on different parts of the input text while answering questions about different treatment categories. Twelve different colors correspond to 12 attention heads of the model. Darker color indicates higher attention: (a) patient with chemohormonal treatment, (b) patient recommended curative treatment.

Such limitations of general purpose LLMs have spurred a competing line of research that focuses on domain-tailored fine-tuning or from-scratch development of LLMs for specific target domains. Proposed LLM for prostate cancer is in line with this research trend. Our experiments clearly establish advantages of domain-specific smaller LLMs over huge generic LLMs in terms of manageable computational cost and training data requirements, reliability in language generation and effectiveness for downstream tasks. Given these advantages, domain-specific LLM development provides a sustainable research avenue for automating language-based tasks of sensitive domains like medicine.

REFERENCES

1. Harris E. Large language models answer medical questions accurately, but can't match clinicians' knowledge. *JAMA*. 2023;330(9):792–794.
2. Omiye JA, Gui H, Rezaei SJ, Zou J, Daneshjou R. Large language models in medicine: The potentials and pitfalls: A narrative review. *Annals of Internal Medicine*. 2024;177(2):210–220.
3. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nature Medicine*. 2023;29(8):1930–1940.
4. Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*. 2023;15(2).
5. Chen S, Kann BH, Foote MB, et al. Use of artificial intelligence chatbots for cancer treatment information. *JAMA Oncology*. 2023;9(10):1459–1462.
6. Sallam M. *ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns* (Vol. 11). MDPI; 2023:887.
7. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. *Nature*. 2023;620(7972):172–180.
8. Han T, Adams LC, Papaioannou JM, et al. MedAlpaca–an open-source collection of medical conversational AI models and training data. *arXiv preprint arXiv:230408247*. 2023.
9. Luo R, Sun L, Xia Y, et al. BioGPT: Generative pre-trained transformer for biomedical text generation and mining. *Briefings in Bioinformatics*. 2022;23(6):bbac409.
10. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. *OpenAI Blog*. 2019;1(8):9.
11. Touvron H, Lavril T, Izacard G, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:230213971*. 2023.
12. Das A, Tariq A, Batalini F, Dhara B, Banerjee I. Exposing vulnerabilities in clinical LLMs through data poisoning attacks: Case study in breast cancer. *medRxiv*: the preprint sever for health scince 2024 March 21.
13. Brawley OW. Trends in prostate cancer in the United States. *Journal of the National Cancer Institute Monographs*. 2012;2012(45):152–156.
14. Henley SJ, Ward EM, Scott S, et al. Annual report to the nation on the status of cancer, part I: National cancer statistics. *Cancer*. 2020;126(10):2225–2249.
15. Petros NG, Alvarsson-Hjort J, Hadlaczky G, et al. Predictors of the use of a mental health–focused eHealth system in patients with breast and prostate cancer: Bayesian structural equation modeling analysis of a prospective study. *JMIR Cancer*. 2023;9:e49775.
16. Crump C, Stattin P, Brooks JD, et al. Risks of alcohol and drug use disorders in prostate cancer survivors: A national cohort study. *JNCI Cancer Spectrum*. 2023;7(4):pkad046.
17. Fink M, Klein K, Sayers K, et al. Objective data reveals gender preferences for patients' primary care physician. *Journal of Primary Care & Community Health*. 2020;11. doi: 10.1177/2150132720967221.
18. Steinkohl F, Luger AK, Gruber L, et al. Acceptance of female urologists among patients with suspected prostate disease. *Translational Andrology and Urology*. 2021;10(7):2938.
19. Ravi P, Karakiewicz PI, Roghmann F, et al. *Mental health outcomes in elderly men with prostate cancer* (Vol. 32). Elsevier; 2014:1333–1340.
20. Shaw B, Walter FM, Hamilton W, Martins T. Symptom appraisal and help seeking in males with symptoms of possible prostate cancer: A qualitative study with an ethnically diverse sample in London. *British Journal of General Practice*. 2023;73(732):e502–e510.
21. Wray RJ, McClure S, Vijaykumar S, et al. Changing the conversation about prostate cancer among African Americans: Results of formative research. *Ethnicity & Health*. 2009;14(1):27–43.
22. Muzii B, Di Bello F, Carraturo F, et al. Mental health of prostate cancer patients: Content review on YouTubeTM. *International Journal of Environmental Research and Public Health*. 2023;20(6):4721.
23. Trienes J, Trieschnigg D, Seifert C, Hiemstra D. Comparing rule-based, feature-based and deep neural methods for de-identification of dutch medical records. *arXiv preprint arXiv:200105714*. 2020.
24. Bodenreider O. The unified medical language system (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*. 2004;32(suppl_1):D267–D270.
25. Soldaini L, Goharian N. *Quickumls: A fast, unsupervised approach for medical concept extraction*. InMedIR workshop, sigir; 2016 Jul 17:1–4.
26. Kojima T, Gu SS, Reid M, Matsuo Y, Iwasawa Y. Large language models are zero-shot reasoners. *Advances in Neural Information Processing Systems*. 2022;35:22199–22213.

44

Biosample and Method Selection for Marker Studies

Angela J. Yoon

44.1 Introduction

Biomarkers are utilized in the biomedical field to screen for disease (screening markers), assess treatment response (response markers), and identify individuals at high risk for disease-specific mortality (prognostic markers) [1]. The appropriate selection of biospecimens and methods is crucial to identifying top marker candidates. The following sections describe examples of the types of samples and the techniques used for detecting cancer biomarkers. We also discuss the statistical analysis and validation procedures to bring biomarker discovery from benchtop to bedside.

Biomarkers for screening a large number of populations to identify those who would benefit from further diagnostic testing should be minimally invasive and cost-effective [2]. Our goal was to devise a high-throughput assay that measures the expression of the signature biomarkers for radiation exposure in human tissue samples obtained by noninvasive methods. In this study, the markers were preselected: activated histone 2AX (γH2AX) and activated checkpoint kinase 2 (pChk2), DNA damage response molecules. These markers are known to be produced in irradiated cells, which can serve as screening markers of radiation exposure.

A pilot clinical trial assessed the efficacy of oral cancer chemopreventive agent (–)-epigallocatechin-3-gallate (EGCG), the most biologically active component in green tea extract. The participants consisted of those with biopsy-confirmed oral precancer [3]. To avoid additional study-related tissue biopsies, exfoliative buccal cells were collected in the form of "swish-and-spit" mouthwash. EGCG was utilized as a topical agent (EGCG power in solution) to hold in the mouth. It was beneficial as it maximizes local exposure of the oral mucosa to the agent but minimizes systemic side effects. Tissue microarray was generated from exfoliative buccal cells and protein expression was assessed using immunohistochemistry (IHC). In another immunotherapy study, tissues collected from the biopsy and surgery were compared for the change in immune phenotype change using quantitative multiplex immunofluorescence (qmIF) [4]. Hence, the marker of treatment response can be qualitatively assessed using IHC, or by quantifying tumor cell count using qmIF.

For early-stage oral cancer, there is no existing risk-stratification modality beyond conventional TNM staging system to identify patients at high risk for cancer-specific mortality. In this study, we aimed to discover a panel of prognostic biomarkers by comparing the microRNA (miRNA) expression in tumor tissues of patients who survived versus those who died of the disease within the first five years of initial surgical treatment [5].

44.2 Materials

44.2.1 Buccal Cells

A total of 100 healthy individuals, 55 males and 45 females, who received routine dental radiographs (full mouth series) were enrolled. The dental radiographs consist of 18 F-speed intraoral films, equivalent to the radiation dose of approximately 2.34 cGy. We collected exfoliative oral cells from consenting participants before and 20 minutes after exposure to ionizing radiation of the radiographs. As a comparison, oral cell samples were collected sequentially from ten non-irradiated subjects 20 min apart. The exfoliative oral epithelial cells were collected using a mouthwash technique. Each participant was asked to rigorously rinse with 25 ml commercial mouthwash solution containing 15% alcohol for 40 s and to expectorate into a cup. The samples were transferred into a 15 ml tube within 4 h of collection and centrifuged (Thermo IEC) at 1000 rpm for 10 min. The pellet containing the oral cells was resuspended in 100–200 µl phosphate buffered saline (PBS). An aliquot of 10 µl was taken and diluted with 90 µl PBS. Five smear slides were prepared by placing 10–20 µl (~1000 cells) on the blank glass slide. The slides were checked for an adequate number of cells and allowed to air dry. To fix the cells, 70% cold acetone was applied for 10 min. The smear slides were stored at −80 °C until immunohistochemical assay was performed.

44.2.2 Cell-Based Tissue Microarray and Tissue Section

At the baseline and endpoint visits, the exfoliative oral epithelial cells were collected by rinsing the mouth with 25 ml of a commercial mouthwash solution. Within four hours of collection, the sample was transferred into a 5 ml tube and centrifuged (Thermo IEC) at 2700 rpm for 10 min. The pellet was then transferred to a 1.7 ml microcentrifuge tube, and 95% formalin was added to fix the epithelial cells. After 24 hours, formalin-fixed cells were embedded in paraffin.

The H&E stained slides were prepared from the paraffin-embedded oral cell blocks to identify three areas containing densely localized cells. The identified areas were marked on the

DOI: 10.1201/9781003449942-49

corresponding cell block for tissue microarray (TMA) construction. Three tissue cores (2 mm diameter) from each sample were arrayed in the recipient paraffin block, yielding a total of 42 cores. Sections (4 μm) were generated from the TMA block for the H&E slide and immunohistochemical analysis. Each 2 mm diameter TMA core contained approximately 60 cells, which is adequate for immunohistochemical analysis. These ~60 epithelial cells made up 40% of each TMA core, and the remaining 60% consisted of keratotic debris and inflammatory cells.

In another immunotherapy study, the change in tumor cellularity and immune profile was assessed using quantitative multiplex immunofluorescence in the formalin-fixed paraffin-embedded (FFPE) biopsy and surgical tissues [4]. Unlike the clinical trial assessing the changes in the marker expression by the tumor cells described earlier, intra-tumoral immune composition is critical data to demonstrate the efficacy of the immunotherapy. For this reason, tissue sections, instead of tissue microarrays, were utilized. Both biopsy and surgical tissues were obtained as standard of care for diagnosis and treatment, respectively, not necessitating invasive tissue collection for the study.

44.2.3 Total RNA from Formalin-Fixed Paraffin-Embedded Tissues

A total of 568 early-stage oral squamous cell carcinoma (OSCC) patients who had surgery only and also with available five-year clinical outcomes data were identified. For each subject, archived FFPE tissue blocks were retrieved. In case the subject had recurrent and/or second primary OSCC, the initial OSCC surgical tissue sample was utilized for the analysis. Ten 10 μm sections were obtained from archived FFPE tumor tissue samples for all subjects. For each sample, a representative section was stained with H&E and reviewed by a pathologist to identify regions containing > 90% malignant epithelial cells for macrodissection. Total RNA was isolated from tissues using RNeasy FFPE kits (Qiagen Inc., Valencia, CA) following the manufacturer's protocol, yield was quantitated by Nanodrop, and samples were stored at −80 °C.

44.3 Methods

44.3.1 Exfoliative Buccal Cell Collection

The expression levels of pChk2 and γH2AX were assessed by IHC assay. The smear slides fixed in cold acetone were washed in PBS thrice and incubated in 0.3% H_2O_2 with PBS for 10 min. A blocking solution was then applied for 20 min. The slides were incubated overnight with anti-pChk2 (1:100, Cell Signaling, Danvers, MA) and for one hour with anti-γH2AX (1:50, Upstate Biotechnology, Lake Placid, NY). The primary antibodies were then labeled with 50 ul of horseradish peroxidase (HRP) conjugated secondary antibody (Envision+RSystem Labelled Polymer-HRP, Anti-rabbit-HRP, DakoCytomation, Denmark) and washed in PBS. Color was developed using Dako DAB for 10 min. Oral epithelial tissues with known pChk2 and γH2AX expression were used as the positive and negative controls with each batch.

The relative intensity of nuclear staining in 25–30 cells from an average of five to six randomly selected high-power fields (400x magnification) was quantified using the Cell Measurement Program software package (Cell Analysis System CAS 200 optical microscope, Becton Dickinson, San Jose, CA). Nuclear intensity was assessed based on the object average absorbance multiplied by 1000. The average staining intensity (mean, median) and standard deviation (SD) were calculated for two markers before and after the radiation exposure. The ANOVA test and the paired *t*-test were conducted to assess the difference in marker expression levels. Pearson correlation coefficient was used to evaluate pairwise correlation between the distributions of the two markers.

44.3.2 Immunohistochemistry and Quantitative Multiplex Immunofluorescence

Immunohistochemistry: The expression levels of five biomarkers (pEGFR, cox-2, cyclin D1, pStat3, and ki-67) were analyzed on the TMA sections by IHC. After TMA sections (4 μm) were deparaffinized, antigen was retrieved with 1 mM EDTA pH 7.5 (from a 100 mM stock). Slides were washed in TBS 0.05M pH 7.5 and 5% defatted powdered milk in TBS-Tween was applied as a blocking agent. For endogenous peroxidase inactivation, 3% hydrogenperoxidase and 0.1% sodium azide were applied. Slides were incubated overnight with the following previously characterized primary antibodies at the indicated dilution 22–34: anti-pEGFR (Tyr 1173) (1:50; Santa Cruz Biotechnolgy); anti-cyclin D1 (1:50, 1:100, 1:200, 1:600; Santa Cruz Biotechnology); anti-pStat3 (1:50, 1:100, 1:200, 1:600; Cell Signaling); anti-cox2 (1:400; Santa Cruz Biotechnology); anti-ki67 (1:100; Santa Cruz Biotechnology). The primary antibodies were then labeled with 50 μl of HRP conjugated secondary antibody (DakoCytomation, Envision+RSystem Labelled Polymer-HRP Anti-rabbit-HRP), washed, counterstained with hematoxylin, dehydrated, and mounted. Proper control was used with each batch. For each case, the staining intensities for the biomarkers were scored from 0 (no staining) to +1 (strong staining) independently by two pathologists. In cases where the interpretation was discordant, a final consensus was reached based on a discussion between the two pathologists. From all cells within a TMA core, the fraction of cells staining positive was assessed and multiplied by the staining intensity score to obtain the final score. For example, if only half of the epithelial cells in the TMA core show +1 staining, the final score would be 0.5. The final scores from three TMA cores generated from the same sample were then averaged to yield one final score per sample. The percentage of epithelial cells with nuclear staining of ki-67 was counted and recorded.

Quantitative Multiplex Immunofluorescence: Tumor cellularity before and after immunotherapy was assessed first by pathologic review and then confirmed with tumor cell count using qmIF. The FFPE tissue slides were stained with H&E for review and the representative tumor section slides were selected by the board-certified pathologists. The most representative area of the selected slides (one for biopsy and one for surgical tissue) was macrodissected for further multispectral imaging. The slides were stained with anti-human CD8

(a cytotoxic T-cell marker), FoxP3 (a regulatory T-cell marker), Tbet (a Th1-type helper T-cell marker), CD4 (a helper T-cell marker), and DAPI (a nuclear marker).

The digital image analysis, quantification, and classification of immunofluorescence-positive cells were overseen by a pathologist. Initially, images were imported, and marker-specific thresholds were established. The cells were quantified corresponding to each marker or combination thereof. This count was then standardized against the total area of the tissue region under examination, deriving cell-type density metrics and expressed as cells/mm^2.

44.3.3 Next-Generation Sequencing and Real-Time PCR

Signature miRNAs were discovered using deep sequencing analysis and validated by qRT-PCR. Targeted miRNA sequencing of the 100 OSCC internal subjects was performed with the HTG EdgeSeq system using the miRNA Whole Transcriptome Assay (HTG Molecular Diagnostics, Inc., Tucson, AZ). The Study Tracker randomized function was used to randomize the placement of samples and technical replicates to reduce potential intra-plating biases. Tumor samples were processed as singletons and human brain RNA controls in triplicate. Briefly, target capture was performed by adding nuclease protection probes to each sample to hybridize to the target mRNA. The library was prepared using HTG EdgeSeq PCR processing and quantified using HTG EdgeSeq KAPA Library Quantification for Illumina Sequencing. The automated HTG Library Calculator was used to ensure a sufficient concentration of samples for library pooling and also to determine the appropriate dilution for the library pool. The PhiX control adaptor-ligated library was spiked in the pooled library, which was heat-denatured and loaded into the well of the NextSeq sequencing cartridge. Sequencing was performed on the Illumina NextSeq platform. The sequencing data were imported into HTG EdgeSeq Parser software to align the FASTQ files obtained from the sequencer to miRBase v.21 sequences. Post-sequencing quality control (QC) was performed. Ninety-one of 100 samples passed HTG post-sequencing QC metrics.

From the deep sequencing analysis, promising prognostic miRNAs were selected and their expression levels were verified using the miScript II Rt kit (Qiagen). Briefly, 9 µL of isolated RNA was added to the cDNA master mix, composed of 5× miScript HiSpec Buffer, 10× miScript Nucleics Mix, miScript Reverse Transcriptase Mix, and water, to a total volume of 20 µL. The cDNA was incubated at 37 °C for 60 minutes, followed by 5 minutes incubation at 95 °C and then diluted 11 times. For amplification reactions, the miScript miRNA PCR Custom Array with a miScript SYBR Green PCR kit (Qiagen) was used in a 7300 qPCR system (Applied Biosystems, Foster City, CA), following the cycling conditions recommended by the supplier (15 minutes at 95 °C, followed by 40 cycles of 15 seconds at 94 °C, 30 seconds at 55 °C, and 30 seconds at 70 °C). The coefficient of variation was calculated and values < 5% were considered acceptable. Test samples were assayed in duplicate with the laboratory blinded to survival status and with 5% duplication after relabeling. Data was analyzed to determine the threshold cycle (Ct). The endogenous control RNU6–6p was used to normalize the relative expression of target miRNAs (ΔCt). The samples with undetermined Ct value for the control (RNU6–6p) were excluded from the analysis (nine samples from the internal test cohort, four from the internal validation cohort, and three from the external validation cohort). Those with an undetermined Ct for specific miRNAs were assigned a value of 39.99.

Association between categorical patient characteristics and binary patient survival outcomes, poor versus favorable, were analyzed using Fisher's exact test, Pearson's chi-squared test, and the Cochran–Armitage trend test. From deep sequencing analysis, the differential expression of miRNAs was assessed using the edgeR analysis pipeline, which assumes a negative binomial distribution of the read counts. We calculated the log2-fold change of normalized expression levels of miRNAs in patients with poor prognosis vs those with favorable prognosis. The p-value testing the differential expression levels of miRNAs between two prognostic groups was adjusted into a false discovery rate (FDR) following the Benjamin and Hochberg method. All miRNAs with an absolute log2-fold change > 1.1 and FDR ≤ 0.05 were selected. The differential expression of the top miRNAs was assessed by qRT-PCR, and the fold change of expression levels between two prognostic groups was evaluated using the $2^{-\Delta\Delta Ct}$ method.

In the internal test cohort, the univariate analysis was performed for all covariates— age, gender, and so on, as well as the expression levels of selected miRNAs. To develop the best prognostic model, we repeated the cross-validation procedure 1000 times by randomly segregating data into ~70% training set and ~30% testing set. We further considered variables that were included in the backward-stepwise selection Cox model on time-to-death with p-value ≤ 0.1 more than 10% of the times out of the 1000 cross-validations. Different Cox models with different sets of selected variables were constructed and the c-indexes were calculated. The model with the best performance (highest c-index) was selected as the final prognostic model. Using the multiple logistic regression model, we further constructed a receiver-operating characteristic (ROC) curve and calculated the area under the curve (AUC) for the internal test cohort using the same set of variables in the final prognostic model.

Using the final prognostic model derived from the internal test cohort, we generated the AUC for the internal and external validation cohorts. The final 5-plex prognostic marker panel was also utilized to generate a cancer-specific mortality risk score using the multivariate Cox regression analyses. We calculated the mortality risk score for every patient in the internal test and internal and external validation cohorts by summing the expression values of the selected miRNAs and covariates weighted by the regression coefficients obtained from the multivariate Cox regression analyses. Based on the individual mortality risk score, the patients were first stratified into higher vs lower mortality risk groups and then further stratified into high vs moderately high vs moderately low vs low risk groups. Kaplan–Meier curves were generated for the four mortality risk groups for the time-to-disease related death.

44.4 Results

44.4.1 Screening Markers in Exfoliative Buccal Cell Samples

Both biomarkers showed statistically significant increased levels of expression after the radiation exposure. The mean intensity for pChk2 before the radiation was 0.114 (SD = 0.035) and after the radiation was 0.139 (SD = 0.038). For γH2AX, the mean intensity before the radiation was 0.105 (SD = 0.033) and after the radiation was 0.125 (SD = 0.052). A paired *t*-test showed $p < 0.001$ for pChk2 and $p < 0.001$ for γH2AX (Figure 44.1 and 44.2).

44.4.2 Marker of Treatment Response in Cell-Based Tissue Microarray and Tissue Section

Immunohistochemical expression of phosphoactivated epidermal growth factor receptor (pEGFR), cyclooxygenase-2 (cox-2), and ki-67 were evaluated pre- and post-treatment (Figure 44.2). Although not statistically significant, the overall decrease in expression levels of pEGFR (27.5%), cox-2 (15.9%), and ki-67 positive cells (51.8%) were observed following EGCG treatment. Moreover, a detectable level of EGCG was found in saliva but not in plasma after the one-week treatment regime, demonstrating local availability of EGCG in oral mucosa without significant systemic absorption.

In the immunotherapy study, there was a nine-fold increase in activated helper T-cells contributing to anti-tumoral response. An increase in activated cytotoxic T-cells was also detected post-treatment (Figure 44.3).

44.4.3 Prognostic Marker Selection Using Next-Generation Sequencing and Real-Time PCR

The risk score from the 5-plex marker panel consisting of miRNAs-127–3p, 4736, 655–3p, TNM stage, and histologic grading stratified patients into four risk categories. Compared to the low-risk group, the highest-risk group had 23-fold increased mortality risk (hazard ratio 23, 95% confidence interval 13–42), with a median time-to-recurrence of six months and time-to-death of 11 months (vs > 60 months for each among low-risk patient; $p < 0.001$).

γH2AX pChk2

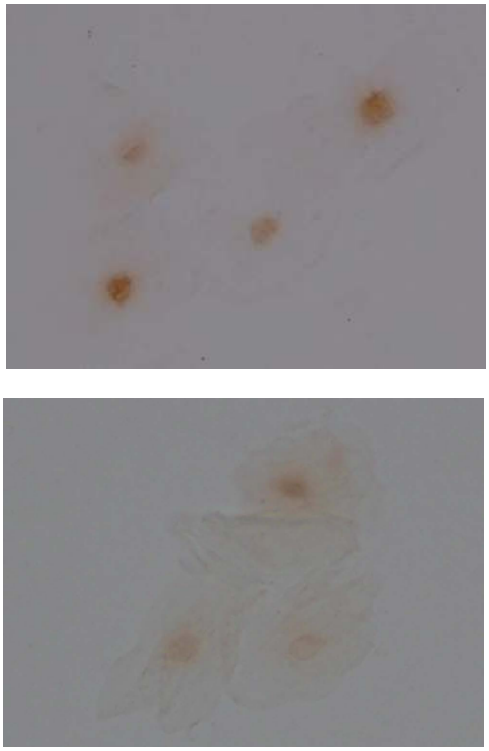
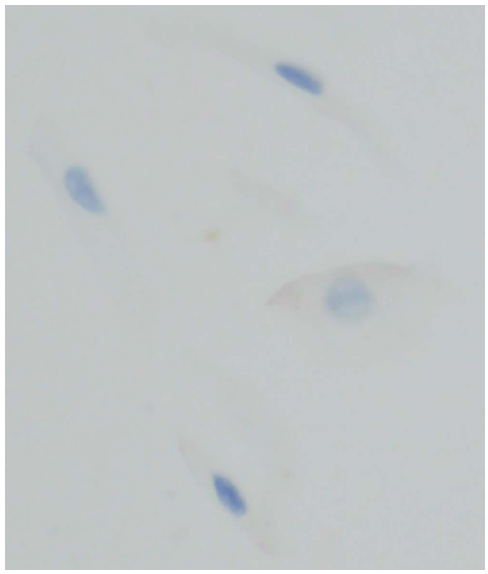


FIGURE 44.1 γH2AX and pChk2 positivity. Photomicrographs showing positive brown nuclear staining for gH2AX and pChk2 in the exfoliative oral epithelial cells.

Pretreatment



Post-treatment

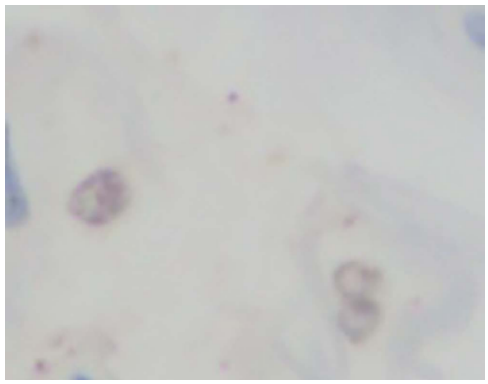


FIGURE 44.2 Photomicrograph of change in ki-67 marker expression in the exfoliated oral epithelial cells.

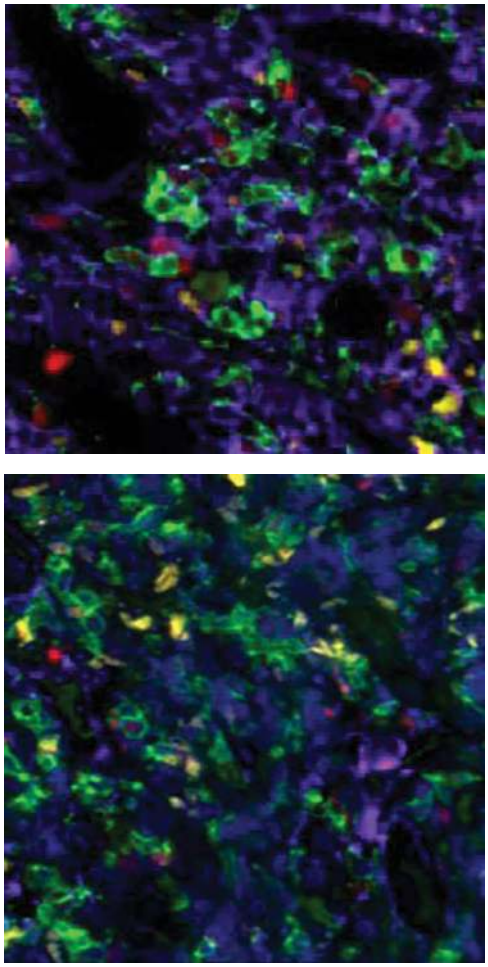


FIGURE 44.3 Immune markers expression in stroma CD4, CD8, Tbet, FoxP3 need to be CD4 (blue), CD8 (green), Tbet (red), FoxP3 (Yellow).

44.5 Discussion

We described the investigation of screening markers for exposure to radiation in the exfoliative buccal cell samples, which can be easily obtained using the swish-and-spit method. An adequate number of epithelial cells were obtained to assess two markers of DNA damage using immunohistochemistry. The mouthwash technique is a noninvasive means of obtaining exfoliative oral cells for radiation dosimetry studies. By rigorously rinsing the mouth, we were able to obtain analyzable oral epithelial cells, most likely from intermediate to superficial layers of the epithelium. Basal layer epithelial cells are not required for analysis since the ionizing radiation affects all layers of the epithelium. Small amounts of anucleated squames and keratotic materials were invariably included in each sample. Since we are measuring the expression of markers in the nucleus of oral epithelial cells, anucleated squames and parakeratins were not included in the intensity interpretation. Our results showed that the exfoliative oral cells obtained by the mouthwash technique serve as appropriate samples for biomarker analysis.

Following clinical trials, treatment marker assessment is crucial. If the trial therapy directly targets the tumor cells, it would be ideal to assess the molecular marker expression levels of tumor cells. In order to circumvent the invasive tissue biopsy procedures, exfoliative cells were collected and paraffin-embedded, and then a tissue microarray was constructed. It allowed for assessing multiple markers using IHC without the risk of depleting the paraffin-embedded cell samples. With immunotherapy study, immune cell profile changes were assessed using multiplex immunofluorescence (IF) assay on tissue slides. We were able to assess five marker expressions. We also identified cells with double and triple marker expression. Since inflammatory infiltration in the tumor microenvironment was an essential factor in determining treatment response, using the whole tissue section was beneficial instead of a small tissue segment selected for tissue microarray. Moreover, the ability to quantify cells using multiplex IF allows for rigorous and unbiased analysis, especially considering variation in IHC data evaluation and scoring.

Prognostic markers are useful in selecting patients at high risk for disease-specific death early on at the time of diagnosis, allowing more aggressive treatment course to increase survival. We studied an miRNA-based 5-plex marker panel optimized to differentiate between those who develop recurrence and die of cancer versus those who survive despite locoregional recurrence, in addition to discriminating between those who have cancer-specific death versus those who have five-year cancer-free survival. We discovered miRNA signatures using next-generation sequencing and validated the prognostic power of miRNAs using qRT-PCR. We then built a 5-plex prognostic model consisting of miRNAs and clinical covariates. The mortality risk score formula was driven from the final prognostic model to serve as a quantifiable risk assessment modality that can be readily applied in a clinical setting. To construct a robust prognostic marker panel, the prognostic model and the risk score formula driven from the internal test cohort (n = 191) were validated in an independent internal validation cohort (n = 101) as well as an external validation cohort (n = 259). The external validation cohort consisted of a heterogeneous population in the Northeastern and Midwestern US and Hawaii to ensure the generalizability of our findings across populations. The AUC of the ROC curve with the miRNA-based 5-plex marker panel was 0.83 ($p<.001$), 0.87 ($p<.001$), and 0.81 ($p<.001$) in internal test, internal validation, and external validation cohorts, respectively, demonstrating uniformly significant prognostic value. In a clinical setting, the prognostic risk score calculation can be performed as a part of the pathology work-up. The study results can then be utilized to guide decision-making in treatment selection and make risk-adjusted therapies possible for patients with early-stage oral cancer.

44.6 Future Trends

Careful selection of samples and study methods based on the study hypothesis is crucial in obtaining robust and reproducible data. The samples should be easily obtainable, and the

selected method should be able to handle a high volume of samples. The method that generates quantifiable data instead of descriptive data is preferred to enhance the rigor and reproducibility of the study.

REFERENCES

1. Sarhadi VK, Armengol G. Molecular biomarkers in cancer. *Biomolecules* 2022;12:1021.
2. Yoon AJ, Shen J, Wu HC, Angelopoulos C, Singer SR, Chen R, Santella RM. Expression of activated checkpoint kinase 2 and histone 2AX in exfoliative oral cells after exposure to ionizing radiation. *Radiat Res* 2009;171:771–775.
3. Yoon AJ, Shen J, Santella RM, Philipone EM, Wu HC, Eisig SB, Blitzer A, Close LG, Zegarelli DJ. Topical application of green tea polyphenol (-)-epigallocatechin-3-gallate (EGCG) for prevention of recurrent oral neoplastic lesions. *J Orofac Sci* 2012;4:43–50.
4. Satish T, Khan S, Levin M, Carvajal R, Yoon AJ. Treatment of recurrent mucosal melanoma of the oral cavity with topical imiquimod and pembrolizumab achieves complete histopathologic remission. *J Immunother Cancer* 2021;9:e001219.
5. Yoon AJ, Wang S, Kutler DI, Carvajal RD, Philipone E, Wang T, Peters SM, LaRoche D, Hernandez BY, McDowell BD, Stewart CR, Momen-Heravi F, Santella RM. MicroRNA-based risk scoring system to identify early-stage oral squamous cell carcinoma patients at high-risk for cancer-specific mortality. *Head Neck* 2020;42: 1699–1712.

45

Subharmonic-Aided Pressure Estimation for Evaluating Cancer

Hailee Mayer, Kibo Nam, Jason Shames, and Flemming Forsberg

45.1 Introduction

Neoadjuvant therapy is administrated to patients to downgrade and/or downstage cancer prior to starting the main treatment. In patients with non-metastatic breast cancer, neoadjuvant chemotherapy (NAC) is often given prior to surgical resection as it can offer benefits to the patient by shrinking the tumor and, in some cases, providing a complete pathologic response [1, 2]. In addition, NAC can also provide insight into a patient's response to chemotherapy, allowing for the distinction between treatment responders and non-responders. Moreover, this insight into a patient's response to treatment allows for modification if the patient is not responding positively to the current regimen [1, 2]. A patient's response to NAC can be indicated by changes in the interstitial fluid pressure (IFP) of the tumor [3]. In normal and healthy tissues, IFP is actively regulated through interactions between stromal cells and the extracellular matrix molecules. In cancerous tissue, IFP is generally about 10–30 mmHg higher than in normal tissue. There are a number of factors that contribute to increased IFP in tumor tissues, such as leaky and collapsed vessels, fibrosis, high cell density, a defective lymphatic system in the tumor, and contraction of the interstitial matrix [4–7]. High IFP in tumors has the potential to hinder and reduce drug delivery to the tumor, thus it has been suggested that using drugs that lower IFP can help to further improve a patient's response to NAC [8].Thus, there are several indicators that IFP may be useful in detecting and diagnosing cancer.

The most widely used method for IFP measurements is the invasive wick-in-needle micro puncture technique [4]. The wick-in-needle technique has been studied in cancers of the breast, cervix, head, neck, skin, lymph nodes, and others [6–9]. It requires custom-made needles with optimally sized side ports [6], and its invasive nature increases the risk to patients. Thus, a noninvasive alternative for evaluating the IFP of tumors would constitute an immense benefit to cancer detection and diagnosis.

Ultrasound contrast agents (UCAs) are comprised of gas-filled microbubbles ranging from 1 to 10 μm in diameter, with a lipid, protein, or polymer shell for stability. The gases used within UCAs have higher compressibility compared to the surrounding tissue, causing an impedance mismatch, leading to enhanced echogenicity from UCAs compared to the surrounding tissue. The enhanced backscattered ultrasound signals from UCAs can provide up to a 30 dB enhancement of ultrasound images [10], which aids in contrast-enhanced imaging of the vasculature that carries UCAs and results in better differentiation between vessels and surrounding tissue. The higher compressibility of the gas compared to surrounding tissue inspired Fairbank and Scully to propose the use of UCAs as noninvasive pressure monitors; the difference in compressibility causes any change in hydrostatic pressure to induce a change in the size of a microbubble [11, 12]. When the microbubbles undergo this pressure-induced change, the reflectivity and resonance frequency of the bubble will be affected [12, 13]. To date, pressure measurement studies with microbubbles include methods such as exploiting the changes in resonance frequency [11, 13], employing the disappearance time of the microbubbles [13, 14], and using the pulse-echo amplitude of a single bubble [15], but none of these methods have been translated clinically as their errors exceed what is clinically acceptable.

More recently, microbubble pressure estimation studies have focused on the nonlinear oscillations produced in microbubbles at higher acoustic pressures (typically above 100 kPa). The nonlinear oscillations of UCAs generate significant energy components in the received echoes, ranging from subharmonic to ultra-harmonic components [10]. These different components can be exploited for contrast-specific imaging modalities, such as subharmonic imaging (SHI), which transmits pulses at the fundamental frequency (f_0) and receives echoes only at the subharmonic ($f_0/2$). SHI is advantageous for contrast-enhanced imaging due to the significant subharmonic generation by the microbubbles compared to the limited, if any, subharmonic generation from the tissue [16]. The properties of subharmonic generation from UCAs also led to the development of subharmonic-aided pressure estimation, or SHAPE. SHAPE utilizes SHI and tracks variations in the subharmonic signal amplitude of the UCAs in response to changes in ambient pressure. SHAPE is most sensitive to changes in hydrostatic pressure when the acoustic power is optimized [17]. The generation of subharmonic signals from UCAs can be divided into three distinct stages: occurrence, growth, and saturation, which can be seen in Figure 45.1. In the occurrence phase, the subharmonic signal cannot be distinguished from the noise floor; during the growth phase, the subharmonic signal amplitude increases rapidly with increasing acoustic pressure; and in the saturation phase, the subharmonic signal has become saturated, and there are no longer observable changes in subharmonic signal from changes in acoustic pressure. On most clinical ultrasound machines, the acoustic pressure is represented by a parameter called the mechanical index (MI), which is calculated as the peak negative pressure divided by the square root of the center

DOI: 10.1201/9781003449942-50

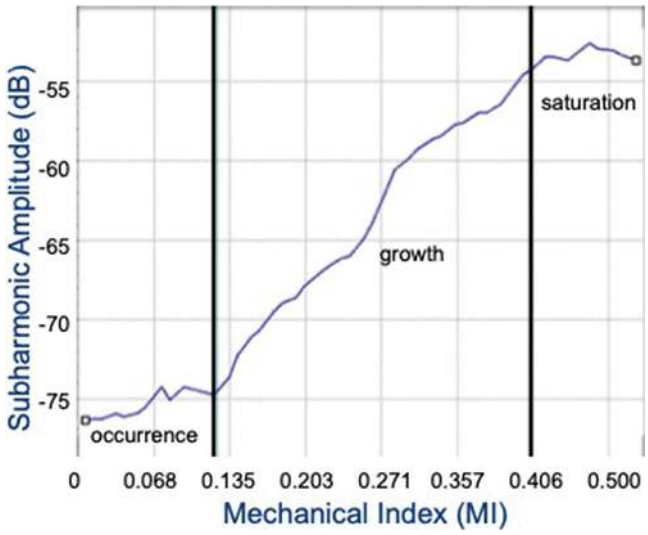


FIGURE 45.1 Generation of subharmonic signal from ultrasound contrast agents as a function of increasing mechanical index. The signal can be divided into three stages: occurrence, growth, and saturation.

frequency of the ultrasound pulse. The inflection point of the growth phase corresponds to the optimal acoustic power output. At the optimal acoustic pressure, there exists an inverse linear relationship for most commercially available UCAs between hydrostatic pressure and subharmonic amplitude. As the ambient pressure surrounding the UCAs increase, the amplitude of their subharmonic emissions will decrease. This inverse linear dependency can be used to noninvasively measure hydrostatic pressure. SHAPE has been studied in humans in evaluation of intra-cardiac pressures, portal hypertension, and in patients undergoing dialysis for chronic kidney disease [18–24].

Aside from tracking pressures for clinical diagnostic use, SHAPE has also been shown to be useful in estimating the interstitial fluid pressure of tumors of the breast and skin (i.e., melanoma) [25–28]. In this chapter, we focus on the use of SHAPE in studies estimating IFP. Using SHAPE to estimate IFP for cancer detection and therapy monitoring was first explored *in vitro* with two main objectives: establishing the optimal acoustic power levels for pressure estimation and determining SHAPE's utility at these optimal acoustic pressure levels by varying the hydrostatic pressure from 0 to 50 mmHg [25]. This pressure range is relevant to IFP as it can range from about −3 to 3 mmHg in normal tissue, and 10–30 mmHg higher than that in cancerous tumors [4]. After optimizing the frequency and acoustic output for SHAPE over a hydrostatic pressure range of 50 mmHg *in vitro*, an *in vivo* proof-of-concept study was carried out to explore the feasibility of SHAPE with those optimized acoustic parameters in swine with melanoma [25]. Following the feasibility study in swine, an *in vivo* study was completed in rats with breast cancer xenografts to compare the SHAPE estimated IFP to the invasive pressure monitor and determine the utility of SHAPE in monitoring changes in IFP after chemotherapy [26]. Once verified as an accurate estimate of IFP in preclinical studies, SHAPE was studied in 17 women scheduled to undergo NAC for breast cancer to determine whether three-dimensional SHAPE could help predict the response of breast cancer to NAC.

Most recently, this technique has been studied to characterize indeterminate breast lesions [28]. The SHAPE estimated pressure gradients between breast lesions and adjacent normal tissue were obtained in 21 indeterminate lesions (eight malignant and thirteen benign lesions) and compared to the scheduled ultrasound-guided needle biopsy results. The SHAPE gradient had an area-under-the-curve similar to the invasive method, helping to show the utility of 3D SHAPE for characterizing indeterminate breast lesions.

45.2 Materials

45.2.1 Ultrasound Contrast Agent

The UCA used in the studies is Definity (Lantheus Medical Imaging), which is approved for use in adult and pediatric contrast echocardiography in the United States (and for other applications across the world). Definity contains a perfluorocarbon gas (C_3F_8) encapsulated with a lipid shell. For human use, the recommended infusion dosage is 1.3 mL of prepared Definity microbubbles added to 50 mL of preservative free-saline, infused at a rate of 4.0–10.0 mL/min (titrated to effect). For indications that require a bolus injection, it is recommended to inject 10 μL/kg of activated Definity microbubbles within 30–60 seconds, followed by a 10 mL saline flush.

45.2.2 Animal Models

As a proof of concept, SHAPE was used to estimate melanoma tumor IFP in five Sinclair swine (weight: 9.5 ± 4.1 kg; Sinclair Bio Resources, Columbia, MO).

Following the proof of concept in swine, SHAPE was used to monitor IFP in rats with breast cancer xenografts.

45.2.3 Clinical Subjects

In the first-in-humans pilot study, 17 women with stage T1 or greater locally advanced breast cancer, who were scheduled to undergo neoadjuvant chemotherapy, were recruited for the study, and 11 were included in the final analysis.

In the prospective study, 25 patients scheduled for breast biopsy of a mass measuring at least 1 cm of the longest dimension and located within 3 cm of the skin surface were enrolled to evaluate tumor IFP using SHAPE compared to the pressure monitor reference.

45.3 Methods

45.3.1 *In Vitro* Setup and Data Analysis

SHAPE was initially explored *in vitro* by Halldorsdottir and colleagues in an acrylic water tank that could withstand pressures up to 200 mmHg to investigate SHAPE across physiological pressures, including the IFP levels encountered in breast tumors [17, 25]. The tank was lined with 25.4 mm of

Sorborthane to eliminate effects from standing waves, and 9.5 mm of gum rubber was used to couple the Sorborthane to the isotonic diluent used for the experiments. The reference hydrostatic pressure was monitored using a digital pressure gauge (OMEGA Engineering, Stamford, CT). A Sonix RP (Analogic Ultrasound; Peabody, MA, USA), operating in research mode, was used with a high-frequency linear array probe L14–5 (nominal bandwidth 5–14 MHz; actual bandwidth 3–12 MHz) [25, 26]. The Sonix RP's L14–5 probe was positioned at a 45-degree angle to a thin plastic acoustic window on the tank. The water tank was filled with 800 mL of isotonic diluent and immersed into a larger water bath also filled with isotonic diluent. Measurements were carried out at room temperature (25 °C). Definity was injected through an inlet on the tank in a dose of 0.2 mL/L to mimic clinical quantities. The solution was kept in suspension using a magnetic stirrer and 30 seconds of mixing was allowed before initiating measurements to ensure an even concentration of agent within the tank. Two transmission frequencies were used—6.7 MHz and 10 MHz, as they fell within the frequency band of the transducer and can be used for clinical breast imaging. Pulse inversion technique was implemented on the scanner for contrast-enhanced ultrasound to minimize scatter from tissue and increase sensitivity to bubble signals [10]. Measurements were taken at a fixed depth of 6 cm with a frame rate of 9 Hz with focus at 4.25 cm (positioned approximately 1.5 cm within the tank). The first objective of this study was to determine the optimal acoustic power setting for pressure measurements. Thus, the hydrostatic pressure was kept at 0 mmHg, and the acoustic power was increased with a 2 dB step size from −20 dB to 0 dB (full range of system) equivalent to 0.24 to 2.05 MPa peak to peak, measured with a 0.20 mm needle hydrophone. The subharmonic signals were plotted as a function of acoustic power, and the acoustic powers corresponding to the growth zone of the sigmoidal curve were selected for further SHAPE investigation with varying hydrostatic pressure [25]. Hydrostatic pressure was varied from 0 to 50 mmHg in steps of 10 mmHg, and pressure values were then compared using linear regression to the radiofrequency (RF) data acquired with the Sonix RP. A region of interest (ROI) of 4 × 4 mm was selected in each image, and then a Fourier transform of each vector within the ROI was obtained. The subharmonic amplitude was extracted from the spectra over a 1 MHz bandwidth around the subharmonic peak and averaged over the ROI and in frames corresponding to two seconds to eliminate noise. Linear regression analysis was used to determine the relationship between the hydrostatic pressure and the change in subharmonic amplitude.

45.3.2 IFP Estimation in Swine with Melanoma (Proof of Concept)

As an initial proof-of-concept study, SHAPE was tested on five Sinclair swine (weight: 9.5 ± 4.1 kg; Sinclair Bio Resources, Columbia, MO) with naturally occurring melanomas. The animals were sedated with an intramuscular injection of 0.04 mg/kg of Atropine (Med-Pharmex Inc., Pomona, CA, USA) and 5 mg/kg of Telazol (tiletamine/Zolazepam, Pfizer, New York, NY, USA). General anesthesia was maintained with 2%–4% of Isoflurane (Iso-thesia; Abbott Laboratories, Chicago, IL) through an endotracheal tube during the procedure. A heating pad was used to maintain body temperature, and after the procedure animals were euthanized with an intravenous injection of 0.25 mL/kg of Beuthanasia. All experiments were supervised by the Laboratory Animal Services Department and conducted in agreement with a protocol approved by Thomas Jefferson University's Institutional Animal Care and Use Committee.

For infusion, 3 mL of Definity was injected into a 50 mL saline bag and the Definity-saline mixture was administrated through an ear vein at a rate of 6.25 mL/min. A needle-based pressure monitor system (Stryker, Berkshire, UK) was used to obtain a reference standard to measure IFP within the tumors and 3–5 cm from the tumor periphery in the surrounding normal tissue. Pressure measurements were taken in triplicate both in normal tissue and in the melanoma, while simultaneously collecting ultrasound RF data. Two transmission frequencies, 6.7 MHz and 10 MHz, were used. The optimal acoustic power could not be determined *in vivo* due to time constraints, but acoustic power levels of −4 dB, −8 dB, and −12dB were used as they showed the most promise *in vitro*. An ROI was selected close to the tip of the pressure monitor needle, and then the subharmonic amplitude extracted from the RF data was compared to the pressure values acquired using the Stryker pressure monitoring system.

45.3.3 IFP Estimation in Rats with Breast Cancer

Following their proof-of-concept study in swine, Halldorsdottir and colleagues used SHAPE to monitor IFP in rats with breast cancer xenografts and, as before, compared the results to invasive pressure monitoring [26]. They hypothesized that SHAPE could be used to estimate the IFP in locally advanced breast cancers (LABC), thereby allowing the response to chemotherapy to be monitored. They developed a custom-programmed calibration for the SHAPE method on a Sonix RP scanner, using nude, athymic rats with breast cancer xenografts as the preclinical model. The obtained SHAPE curves from the calibration group were applied as an independent data set using a different group of rats with breast cancer xenografts undergoing paclitaxel chemotherapy. The modified Sonix RP scanner (Version 3.2.2 software, BK Ultrasound, Richmond, BC, Canada) was capable of real-time subharmonic amplitude detection for SHAPE [26]. A high-frequency linear array probe (L9–4/38) was used for scanning, as it contains a bandwidth suitable for breast imaging (4–9 MHz, center frequency: 5.5 MHz). The system was enabled with transmit frequencies of 8.0 MHz and 4.0 MHz for pressure estimation, and pulse inversion technique was used to minimize linear signal components. The scanner was enabled with a feature that allowed for optimization of the acoustic power, in large part based on a solution previously proposed by our group [29].

MDA-MB-231 breast cancer cells were cultured with 89% Dulbecco's minimal essential medium, 10% fetal bovine serum, and 1% penicillin-streptomycin [26]. The cells were implanted subcutaneously together with growth factor reduced Matrigel basement membrane matrix (BD Biosciences, San Jose, CA, USA). Eighty-nine athymic, nude, female,

six-week-old rats were randomly placed in either a calibration group or a treatment group (n = 25 and n = 64, respectively). MDA-MB-231 tumor cells were injected subcutaneously into the right mammary fat pad of the rats. Rats were inspected for tumor growth on days 7, 14, and 21. Definity was given to the rats in three bolus IV injections through a tail vein at a rate of 180 μL/kg, followed by a 0.2 mL saline flush. First, the optimal acoustic output was established on the Sonix RP, and then the IFP of the normal mammary fat pad, as well as the tumor, was measured with an invasive needle-based intra-compartmental Stryker pressure monitor, which provided the reference IFP. SHAPE data was acquired at the optimal acoustic output. The IFP measurements and SHAPE measurements were both carried out in triplicate and then averaged. The treatment group received a SHAPE exam before and after being given the chemotherapy agent paclitaxel (Mayne Pharma, Paramus, NJ, USA), to determine the ability of SHAPE to monitor changes in IFP. SHAPE and Stryker measurements were acquired before and 48 hours after administration of a single IV injection of 5 mg/kg of paclitaxel on day 21, 24, or 28. Tumor volumes were measured with calipers every other day to monitor and compare tumor growth. The methods described in the proof-of-concept study for extracting the subharmonic were followed [25] and a maximum intensity projection (MIP) was generated. Thresholding was performed based on previous work by our group [30] and resulted in three thresholds (100%, 115%, and 130%) used for analysis. The authors used the calibration group to produce calibration equations, and they were applied to the independent data set produced from the treatment group. Analysis of variance (ANOVA) was used to compare the three different thresholds. This analysis was then repeated on the treatment group (control/paclitaxel, pre-/post-treatment) using linear regression, ANOVA, and a paired t-test.

45.3.4 First In-Human Pilot Study

The first in-human study using SHAPE to evaluate IFP in breast tumors, by Nam and colleagues, was approved by the institutional review board at Thomas Jefferson University and was compliant with the Health Insurance Portability and Accountability Act. Seventeen women with stage T1 or greater LABC who were scheduled to undergo NAC were recruited for the study, and 11 were included in the final analysis. Ultrasound examinations were performed on subjects immediately before therapy, at completion of 10% (after one cycle), 60% (after seven or nine cycles), and 100% (after 12 or 16 cycles) of chemotherapy. These time points were selected as they represent a broad investigation across the course of therapy. The authors utilized a LOGIQ 9 scanner from GE HealthCare (Waukesha, WI, USA) with a 4D10L probe to obtain the SHAPE estimates of IFP. The scanner was modified to collect RF data from a 3D subharmonic mode. The system transmitted four cycle pulses at 5.8 MHz (receiving at 2.9 MHz) in subharmonic imaging mode. Like the previously discussed methods, a pulse inversion technique was used to minimize linear signals from tissue. A continuous infusion of 3 mL of Definity mixed with 50 mL of saline was given to the patient through the antecubital vein, and infusion rates varied from 4 to 10 mL/min [27]. The scanner was equipped with an acoustic output power optimization algorithm for SHAPE. Optimization of acoustic power was calibrated at the beginning of the baseline ultrasound examination for each patient. Following calibration, the infusion was halted, and once clearance of Definity bubbles was observed, three sets of RF data were collected for both 3D SHAPE and 3D SHI. The infusion was then restarted, and three sets of RF data were collected in 3D SHAPE and SHI at volume rates of 0.6–3.2 volumes per second. Figure 45.2 shows an example of the 3D ultrasound images of a malignant lesion (white arrows)

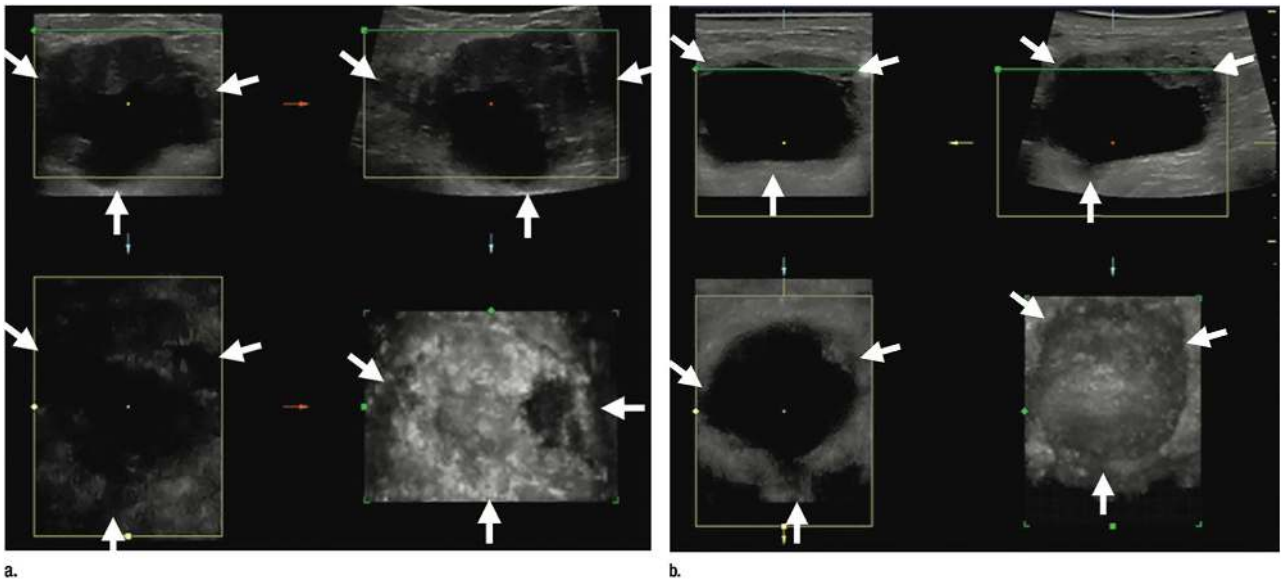


FIGURE 45.2 Three-dimensional ultrasound images (obtained by using 4D view) of breast cancer in 45-year-old woman. Images were acquired (a) before neoadjuvant therapy and (b) at completion of 10% of therapy. This viewer shows the three orthogonal planes of sagittal (top left), transverse (top right), and coronal (bottom left), as well as rendered volume (bottom right). Rendered volume is semitransparent representation of all volume pixels within the region of interest (ROI) boxes of top row. Tumor is located in center of each view (arrows).

without contrast agent (left) and with contrast agent (right). The analysis focused on the comparison of subharmonic amplitude gradient between the lesion and the surrounding tissues with and without Definity microbubbles present.

Tumor volume change during therapy was compared with imaging and pathologic response, as well as the resulting 3D SHI and SHAPE results at the same time points. Two ROIs were selected in each elevational plane for both 3D SHAPE and SHI. One of the ROIs was placed in the tumor and the other in the surrounding tissue [27]. The maximum magnitude from the resultant ROI frequency spectrum was extracted before and during infusion, and the ratio of the subharmonic signal during infusion to before infusion was calculated for each patient to account for possible attenuation-induced changes in the ROIs over the four time points. The difference in the obtained subharmonic signal ratios from the tumor and the surrounding area was calculated as a relative estimate of IFP in triplicate and averaged. This analysis was followed for all four time points for each patient. To analyze the 3D SHI data, the mean subharmonic signal intensity was used instead of the maximum magnitude from the frequency spectrum. The mean subharmonic signal intensity was calculated for each ROI and averaged over a volume and frames. The ratio of the mean subharmonic signal intensity during and before infusion was calculated, and the difference from the ROIs of the tumor and surrounding area was calculated in triplicate and averaged, like the 3D SHAPE data.

Clinically, the patient's response to NAC was evaluated pathologically by an experienced pathologist using a five-point grading scale developed locally. 1–5 corresponded to minimal therapeutic effect, moderate therapeutic effect with invasive tumor or vascular invasion, marked therapeutic effect with no or minimal invasive tumor, or complete therapeutic effect with no invasive disease, respectively. In addition, markers such as residual tumor measurement, tumor measurement change, lymph node measurement, and lymph node measurement change, as well as cellular information, were used to evaluate therapeutic effects. Imaging wise, 3D measurements were taken from baseline ultrasound images acquired after the completion of NAC. 3D measurements were considered more accurate for assessing the results from 3D data [27]. Pathologic response was assessed in two ways: patients with greater than 90% tumor volume reduction were classified as responders, and a locally developed five-point pathology grading scale. Patients with no therapeutic effect were grade 1, minimal therapeutic effect were grade 2, moderate therapeutic effect with invasive tumor or vascular invasion were grade 3, marked therapeutic effect with no or minimal invasive tumor were grade 4, and complete therapeutic effect with no invasive disease were grade 5.

45.3.5 Prospective Clinical Study

Twenty-five patients scheduled for breast biopsy of a mass measuring at least 1 cm of the longest dimension and located within 3 cm of the skin surface were enrolled. Tumor depth was limited to 3 cm to maintain ability of direct intra-compartmental pressure measurement using a pressure monitoring system equipped with a 6.3 cm needle (C2Dx, Schoolcraft, MI, USA). Prior to surgical biopsy, subjects underwent a contrast-enhanced ultrasound (CEUS) exam using a LOGIQ E9 scanner with a 3D linear RSP6–16 probe (GE HealthCare, Waukesha, WI, USA). Subjects received a continuous infusion of 3 mL of Definity mixed into a 50 mL saline bag through the antecubital vein [28]. The infusion rate was between 4 and 10 mL/min (titrated to effect). In this study, the acoustic power was optimized on an individual basis to achieve maximal sensitivity of SHAPE to estimate pressure. Once the optimal acoustic power was determined, the infusion was stopped for approximately five to eight minutes until the clearance of bubbles was seen. 3D conventional CEUS (second harmonic) and 3D SHAPE images were collected in triplicate. While operating under conventional CEUS mode, data was collected using a transmit pulse power of 29% and received gain of 48 dB for all examinations. The SHAPE exams were optimized individually and were, therefore, different from patient to patient. After collecting both conventional CEUS and 3D SHAPE images with no bubbles, the infusion was restarted, and both image types were again collected in triplicate. A paired data collection allowed for a direct comparison of the second and subharmonic signals from the breast lesions, with and without bubbles [28].

Conventional CEUS and SHAPE data volumes totaled 115–197 and 46–182 2D slices, respectively, depending on the field of view determined by the lesion size and depth. Data analysis was performed exclusively on the 2D slices that contained the lesion. The average second harmonic amplitude in the lesion volume was measured with conventional CEUS by subtracting the second harmonic signal amplitude in the lesion with no Definity bubbles from that with Definity bubbles present and averaging it over the lesion volume. The final average second harmonic amplitude of the lesion was determined by averaging the value of three volumes from three separate CEUS data sets. The pressure gradient between the lesion and adjacent tissue was measured via the gradient of average subharmonic amplitude obtained by subtracting the average subharmonic amplitude of the adjacent normal tissue from that of the lesion. ROI size of adjacent normal tissue was similar to the lesion ROI size, apart from cases where the lesion was too large to be contained within the field of view. The final average subharmonic amplitude estimated pressure gradient was calculated by averaging the results from the three SHAPE data sets [28].

45.4 Results

45.4.1 *In Vitro* Exploration

The *in vitro* exploration into optimal acoustic power levels resulted in four acoustic power levels covering the growth stage and its boundaries: −16, −12, −8, and −4 dB. These four acoustic power levels were used for further SHAPE exploration and correspond to 0.33, 1.06, 1.33, and 1.68 MPa and 0.24, 1.21, 1.52, and 1.78 MPa peak to peak for 6.7 MHz and 10 MHz, respectively. The acoustic power levels were normalized so that 1.0 corresponded to the maximum acoustic power level. An example of a B-mode image of one of the swine melanomas can be seen in Figure 45.3. An inverse linear relationship was seen in the water tank over a hydrostatic pressure change from

0 to 50 mmHg at both transmission frequencies and all acoustic power levels ($r^2 > 0.62$; $p < 0.05$). At a transmission frequency of 10 MHz, higher r^2 values were seen compared to 6.7 MHz transmission. Maximum sensitivity for SHAPE was at 10 MHz transmission and an acoustic power of −8 dB, which resulted in an 11.4 dB drop in subharmonic amplitude as hydrostatic pressure was increased from 0 to 50 mmHg ($r^2 = 0.95$; $p < 0.01$).

45.4.2 IFP Estimation in Swine with Melanoma (Proof of Concept)

In the proof-of-concept study involving five swine, the IFP was always below 11 mmHg in normal tissue, and in the melanomas, IFP was above 16 mmHg, while no significance was observed between tumor volume and IFP ($r^2 = 0.95$; $p < 0.01$). The full results of the study can be seen in Table 45.1 of the original article, but in summary, an inverse linear relationship was observed in all cases except for one, which did not have a statistically significant relationship in either the positive or negative direction (transmission = 10 MHz, acoustic power = −8 dB, $p = 0.054$). The 10 MHz transmission frequency resulted in stronger correlations between subharmonic amplitude and IFP for all animals, compared to 6.7 MHz. One swine was removed from the study due to technical difficulties, and because of time limitations, not all frequency/acoustic power settings were evaluated for each of the four swine. Three of the four swine had similar sensitivities for SHAPE (−0.19 ± 0.07 dB/mmHg), which was similar to the *in vitro* sensitivity of −0.22 dB/mmHg. The optimal acoustic parameters resulting from this study were 10 MHz transmission frequency and acoustic power of −4 or −8 dB.

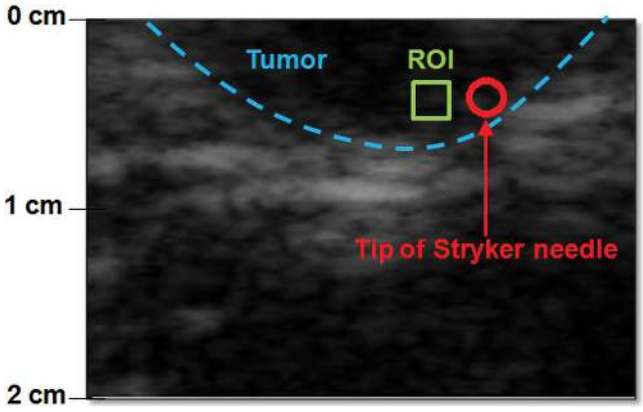


FIGURE 45.3 B-mode image of a melanoma in a swine. Reproduced from [24].

Source: Reprinted from Ultrasonics, Vol 54/edition 7, V.G. Halldorsdottir, J.K. Dave, J.R. Eisenbrey, P. Machado, H. Zhao, J.B. Liu, D.A. Merton, F. Forsberg, *Subharmonic aided pressure estimation for monitoring interstitial fluid pressure in tumours—In vitro and in vivo proof of concept, 1938–1944*, Copyright 2014, with permission from Elsevier.

45.4.3 IFP Estimation in Rats with Breast Cancer

45.4.3.1 Calibration Group

The calibration group included 25 of the 89 rats and were implanted with MDA-MB-231 cells. Sixteen (64%) exhibited tumor growth and 13 (52%) were successfully imaged. An inverse linear relationship was observed for all three thresholds, with the strongest correlation at the 115% threshold level, which can be seen in Figure 45.4 ($r = -0.64$ for 100%, $r = -0.69$ for 115%, and $r = -0.61$ for 130% ($p < 0.01$). The calibration equation resulting from the 115% threshold was:

$$\text{IFP} = \frac{\text{SHAPE} - 127.221}{-1.0624}$$

ANOVA showed that all three threshold groups differed significantly from each other ($p < 0.01$), and the differences

TABLE 45.1

A Summary of *In Vivo* Measurements Showing SHAPE Results Compared to the Pressure Monitor for Each Tumor Evaluated in the Swine. (Reproduced from [25].)

Tumor	Frequency [MHz]	Acoustic Output [dB]	Subharmonic Amplitude in Tissue [dB]	St. Dev [dB]	Subharmonic Amplitude in Tumor [dB]	St. dev [dB]	Change in Subharmonic Amplitude [dB]	Tissue IFP [mmHg]	Tumor IFP [mmHg]	Change in IFP [mmHg]	(r^2, p)	Slope [dB/mmHg]	Intercept [dB]
1	10	−4	66.52	4.08	55.27	1.90	−11.26	1	17	16	0.82, 0.01	−0.70	67.23
		−8	66.90	2.22	57.19	1.69	−9.71	1	17	16	0.90, < 0.01	−0.61	67.51
	6.7	−4	67.45	2.75	60.31	7.03	−7.14	2	17	15	0.40, 0.18	−0.48	68.40
		−8	67.92	0.86	60.67	8.24	−7.25	2	17	15	0.36, 0.20	−0.48	68.89
2	10	−4	67.87	2.17	60.52	0.67	−7.35	8	61	53	0.89, < 0.01	−0.14	68.98
		−8	70.18	0.90	56.90	2.01	−13.28	8	61	53	0.97, < 0.01	−0.25	72.19
4	10	−8	35.94	0.57	39.21	1.09	3.27	20	69	33	0.64, 0.05	0.07	34.61
		−12	36.13	0.93	35.53	4.19	−0.60	18	71	32	0.16, 0.51	−0.01	36.33
5	10	−8	44.56	1.13	37.84	1.27	−6.72	7	40	30	0.92, 0.03	−0.20	45.98
		−12	42.50	2.27	38.89	0.37	−3.61	8	40	0	0.63, 0.07	−0.11	43.40
	6.7	−8	49.36	0.87	46.48	0.42	−2.88	10	40	0	0.91, 0.02	−0.10	50.31
		−12	43.94	2.24	43.79	2.89	−0.15	10	40	0	0.10, 0.54	−0.01	43.99

Reprinted from Ultrasonics, Vol 54/edition 7, V.G. Halldorsdottir, J.K. Dave, J.R. Eisenbrey, P. Machado, H. Zhao, J.B. Liu, D.A. Merton, F. Forsberg, Subharmonic aided pressure estimation for monitoring interstitial fluid pressure in tumours—*In vitro* and *in vivo* proof of concept, 1938–1944, Copyright 2014, with permission from Elsevier.

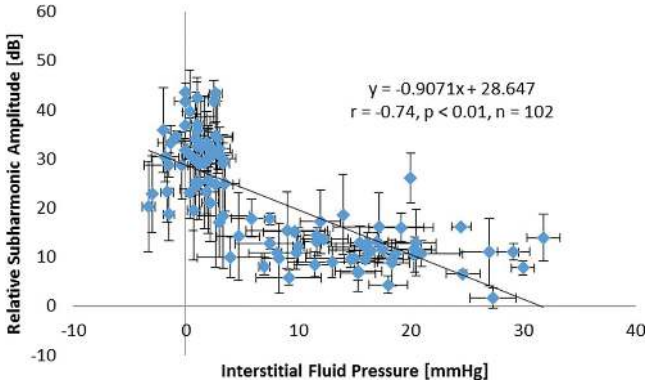


FIGURE 45.4 Example correlation curve at 115% threshold; the strongest observed inverse linear relationship out of the three threshold levels. Reproduced from [25].

Source: Reprinted from Ultrasound in Medicine and Biology, Vol 43/edition 7, V.G. Halldorsdottir, J.K. Dave, A. Marshall, A.I. Forsberg, T.B. Fox, J.R. Eisenbrey, P. Machado, Ji-Bin Liu, D.A. Merton, F. Forsberg, *Subharmonic-aided pressure estimation for monitoring interstitial fluid pressure in tumors: Calibration and treatment with paclitaxel in breast cancer xenografts, 1401–1410*, Copyright 2017, with permission from Elsevier.

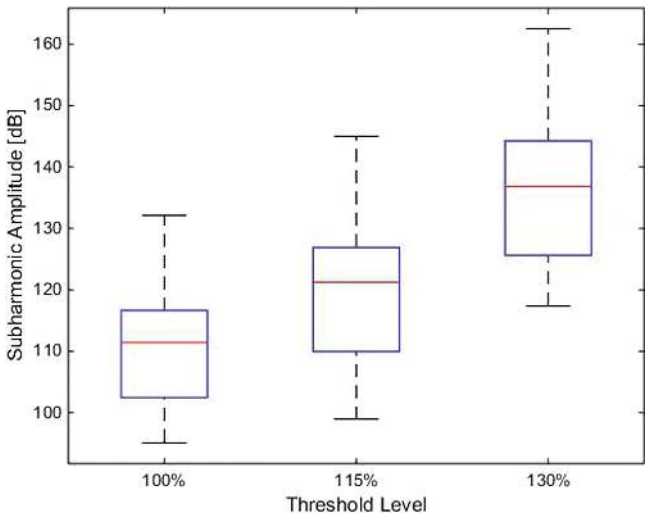


FIGURE 45.5 Boxplot results of the range of subharmonic amplitude (dB) seen at each threshold level (100%, 115%, and 130%). Reproduced from [25].

Source: Reprinted from Ultrasound in Medicine and Biology, Vol 43/edition 7, V.G. Halldorsdottir, J.K. Dave, A. Marshall, A.I. Forsberg, T.B. Fox, J.R. Eisenbrey, P. Machado, J-B. Liu, D.A. Merton, F. Forsberg, *Subharmonic-aided pressure estimation for monitoring interstitial fluid pressure in tumors: Calibration and treatment with paclitaxel in breast cancer xenografts, 1401–1410*, Copyright 2017, with permission from Elsevier.

between all groups were also significant ($p < 0.01$). The mean subharmonic amplitudes were 111.20 ± 10.59 dB for 100%, 119.70 ± 12.06 for 115%, and 136.51 ± 12.77 dB for 130%. The average difference between tumor IFP and normal tissue IFP was 13.26 ± 6.70 mmHg, and the average differences in subharmonic amplitude between tumor and tissue were −13.40 ± 5.25 dB for 100%, −14.45 ± 5.59 dB for 115%, and −15.47 ± 6.71 dB for 130%. The average difference between tumor IFP and normal tissue IFP was compared to the average difference between the subharmonic amplitude of the tumor and normal tissue in the 13 successfully imaged rats, using linear regression analysis, which showed no significant relationship between the two ($p > 0.35$).

45.4.3.2 Treatment Group

Sixty-four rats injected with MDA-MB-231 were evaluated in the treatment group. Fifty-nine percent of the rats grew tumors, and 26 (41%) were successfully imaged. Twelve of the rats with tumor growth were excluded for technical difficulties. Thus, six rats (two each on day 21, 24, or 28) did not receive paclitaxel and were designated as part of the control group. The calibration equations obtained from the calibration group were tested on the control group to determine efficacy in estimating the pressure differential between the tumor and surrounding tissue. The measured and calculated IFPs for 100% and 130% thresholds differed significantly ($p = 0.01$ for both groups), but in the 115% threshold group, a significant difference between measured and calculated IFP values was not observed ($p > 0.13$), further indicating the 115% threshold group being more suitable for SHAPE IFP estimation, compared to the 100% and 115% groups. Errors ranged from 7.80 to 8.55 mmHg and the strength of correlation ranged from 0.73 to 0.74, and the paired *t*-test showed that the measured and calculated groups did not differ significantly ($p > 0.14$).

In the treatment studies, an inverse linear relationship was observed for all three thresholds ($r = -0.72$ to -0.74, $n = 102$, $p < 0.01$). The highest correlation was at 100% ($r = -0.74$, $p < 0.01$). Following the results from the calibration study, all three thresholds differed significantly from each other ($p < 0.01$). A boxplot of the results for each threshold can be seen in Figure 45.5. The mean subharmonic amplitudes from the 100%, 115%, and 130% groups were 110 ± 11.7 dB, 122.0 ± 12.4 dB, and 135.5 ± 14.4 dB, respectively. No significant differences were observed in the slopes and intercepts of the calibration equations when compared with the slopes and intercepts of the treatment equations. Therefore, a universal equation could be derived for future pressure estimation. This idea was further exemplified by the 115% threshold group showing no significant differences between measured and calculated IFP values ($p > 0.14$). The average difference in IFP between tumor and tissue was 13.97 ± 6.72 mmHg, and the average difference between subharmonic amplitude in tumor and tissue were −16.49 ± 9.62 dB (100%), −18.76 ± 8.33 dB (115%), and −20.29 ± 9.97 dB (130%). No significant difference was observed when examining the difference in subharmonic amplitude between tumor and tissue and comparing it with the difference in IFP between tumor and tissue ($p > 0.15$), and no significant difference was seen between the treatment and control groups, regardless of threshold level ($p > 0.60$). Interestingly, there was no significant difference observed in IFP or SHAPE between pre- and post-paclitaxel treatment. In the end, there was no difference between IFP before and after treatment with paclitaxel. Therefore, there was no treatment response for SHAPE to evaluate.

45.4.4 Pilot Study

This HIPPA compliant pilot study recruited 17 women scheduled to undergo neoadjuvant therapy for breast cancer. Four patients left the study, and data from two patients were discarded due to technical error. Eight patients completed all four time points of the imaging protocol, and an additional three patients left the study after the imaging session performed at 10% completion of chemotherapy because of disease progression and were considered non-responders. At the completion of the study, six patients were classified as responders (> 90% tumor reduction), and five were partial responders or non-responders. The tumor volume change between the first two examinations (baseline and at 10% completion of chemotherapy) did not help differentiate responders from partial responders or non-responders based on imaging ($p = 0.28$) or pathologic assessment ($p = 0.27$).

When the 3D SHAPE and 3D subharmonic imaging results from all four examinations were compared with the imaging responses, only the results after 10% completion of chemotherapy showed significance (3.23 ± 1.41 dB vs. −0.88 ± 1.46 dB for SHAPE and 1.32 dB ± 0.73 vs. −0.82 ± 0.88 dB for subharmonic imaging; $p = 0.0011$ and 0.0017, respectively). The results from SHAPE and subharmonic imaging data were positive for responders (i.e., subharmonic signal increased more in the tumor than the surrounding area) and zero or negative for partial responders or non-responders. Based on pathologic assessment, 3D SHAPE and 3D subharmonic imaging did not show a significant difference between responders and partial responders or non-responders ($p = 0.13$ and 0.29, respectively).

Tumor volume change during therapy did not correlate with subharmonic signal changes. In three patients, the tumors increased, and in three, they decreased after completion of 10% of chemotherapy, although all the SHAPE and subharmonic imaging analyses showed an increase in subharmonic signal in the tumor compared to the surrounding tissue.

TABLE 45.2

Summary of Study Results

Subject Number	Direct Pressure Measurement (mmHg)		Difference in Mean Subharmonic Amplitude (a.u.)		Difference in Mean Second Harmonic Amplitude (a.u.)	Biopsy Result	Ki-67 Level (%)
	Tumor	Tissue	Tumor	Tissue	Tumor		
1	22.0	10.0	2.1	5.7	N/A	Malignant	36
2	13.0	3.0	−1.6	−11.3	N/A	Benign	N/A
4	38.0	8.7	1.2	0.4	N/A	Malignant	87
5	29.0	11.7	1.1	−4.2	N/A	Benign	N/A
6	26.3	8.3	0.2	−1.1	N/A	Malignant	28
8	12.7	11.7	−3.0	−0.9	N/A	Benign	N/A
9	5.3	13.3	4.8	−0.7	N/A	Benign	N/A
10	8.0	2.7	1.9	0.1	N/A	Benign	N/A
11	45.0	12.0	1.7	2.5	N/A	Malignant	72
12	35.3	11.0	0.0	−1.9	N/A	Benign	N/A
13	40.3	38.0	0.5	1.9	N/A	Benign	N/A
15	22.7	12.7	2.7	0.0	5.2	Benign	N/A
16	17.0	3.0	5.1	−1.2	16.7	Benign	N/A
17	35.7	16.0	1.4	1.3	14.9	Benign	N/A
19	25.0	14.7	1.0	−1.0	−0.2	Benign	N/A
20	15.3	5.3	9.9	7.4	18.6	Benign	N/A
21	26.0	11.0	4.7	2.6	30.8	Malignant	47
22	39.3	15.0	0.8	0.2	5.4	Malignant	54
23	34.0	10.0	0.0	−0.2	11.2	Malignant	30
24	24.0	12.0	2.1	−0.4	9.7	Benign	N/A
25	13.7	2.0	0.7	1.6	6.8	Malignant	81

Reproduced from [28]. a.u.: arbitrary unit; N/A: not available.

Reprinted from Ultrasound in Medicine and Biology, volume 50/edition 12, Kibo Nam, Mehnoosh Torkzaban, Jason P. Shames, Lydia Liao, Corinne E. Wessner, Priscilla Machado, Andrej Lyshchik, Flemming Forsberg, *Characterization of indeterminant breast lesions based on pressure estimates by noninvasive 3D contrast-enhanced ultrasound, 1854–1860, Copyright (2024), with permission from Elsevier.*

45.4.5 Prospective Clinical Study

Of the 25 subjects initially enrolled, 21 were included in the final analysis. All subjects were female, aged 21–69 (average age: 43 ± 15 years). Thirty-three percent of participants were White, 29% were Black or African American, and 9% were Hispanic. Clinical biopsy showed eight malignant (six were BI-RADS 4, and two were BI-RADS 5) and 13 benign lesions, all of which were classified as BI-RADS 4 [28]. Six of the malignant lesions were grade 3, and two were grade 2. Malignant tumor sizes were on average 4.3 ± 3.6 cm^3, while benign tumor sizes were on average 1.9 ± 2.2 cm^3. 3D SHAPE data was available for all 21 patients, but only ten had available 3D conventional CEUS data available (four malignant; six benign lesions) since 3D conventional CEUS was added to the study protocol after the start of the study. The mean second harmonic signal intensities did not differ significantly between malignant and benign lesions ($p > 0.57$), but the subharmonic signal gradients between the lesion and normal tissue by 3D SHAPE were higher in the benign lesions compared to the malignant, indicating lower pressure in the lesion than the surrounding tissue (2.86 ± 3.24 vs. −0.03 ± 1.72 a.u.; $p = 0.03$). The summary of these results can be found in Table 45.2 [28]. The area under the curve in detecting malignancy was 0.85, with 92% specificity and 50% sensitivity, and a weak inverse correlation existed between SHAPE gradients and directly measured pressure gradients ($r = -0.20$).

45.5 Discussion

From the preclinical exploration *in vitro* and *in vivo* in rats and swine to the clinical exploration in women with breast cancer, SHAPE has been shown to be useful in estimating the IFP of tumors using Definity. Definity was selected due to being commercially available in the United States, as it is approved by the Food and Drug Administration for use in echocardiography. Previous *in vitro* studies with lower-frequency (< 7 MHz) SHAPE in a water tank showed up to a 13 dB decrease in the subharmonic amplitude for Definity at 6.6 MHz transmission with hydrostatic pressure being varied from 0 to 186 mmHg

($r^2 = 0.98$, $p < 0.05$) [17]. Initially, an *in vitro* investigation was carried out to determine the optimal acoustic parameters for pressure estimation in the setting of IFP and showed that a transmission frequency of 10 MHz and an acoustic power of −8 dB was most optimal for SHAPE and resulted in an 11.4 dB drop in subharmonic amplitude as hydrostatic pressure was increased from 0 to 50 mmHg (sensitivity = 0.22 dB/mmHg; $r^2 = 0.95$; $p < 0.01$) [25]. These parameters were then used in an *in vivo* proof-of-concept study on five swine with melanoma. An inverse linear relationship was seen for all swine except for one (which did not have a significant relationship in either direction). The 10 MHz transmission showed stronger correlations between IFP and subharmonic amplitude for all animals compared to 6.7 MHz. The resulting maximum sensitivity in swine was −0.19 ± 0.07 dB/mmHg, which was comparable to what was seen *in vitro*. Like the *in vitro* study, the proof-of-concept study showed that using a transmission frequency of 10 MHz resulted in maximum sensitivity for SHAPE to estimate IFP. In this proof-of-concept study, the optimal acoustic power levels were −4 and −8 dB.

This concept was further studied in rats and included a calibration arm and treatment arm of the study [26]. The calibration arm helped determine the calibration equations that were applied to an independent data set from the treatment arm of the study. The calibration equations were then used to calculate IFP values, which were compared to the measured IFP values. Absolute errors ranged from 7.80 to 8.55 mmHg, and root mean square (RMS) errors ranged from 7.49 to 10.78 mmHg, which were in line with previous work by the same group. The treatment arm showed an inverse linear relationship between subharmonic amplitude and IFP for all three thresholds ($r = -0.72$ to -0.74, $p < 0.01$, $n = 104$), which was consistent with previously published results [12, 17, 20, 29, 31]. Like the calibration study, the treatment study showed that all threshold groups differed significantly from each other, and the slopes and intercepts for the calibration equations were very similar to the treatment study, suggesting a calibration curve may be used to estimate IFP, when utilizing the same scanner, acoustic parameters, contrast agent, and tumor type at a relatively uniform shallow depth. No statistically significant difference was seen in the subharmonic amplitude between the control and treatment groups due to the lack of treatment effect from the low dose of paclitaxel used in the study (5 mg/kg). This drug is more toxic to rats than humans, and the dose given was only 25% of what is normally used in humans, as that corresponds to the highest tolerable dose for rats [32]. This likely contributed to the lack of treatment effect seen in this study. The lack of treatment effect makes it impossible to draw any conclusions as to whether SHAPE is sensitive enough to measure changes in IFP for treatment monitoring based on this study.

In humans, it was first shown that 3D SHAPE and 3D subharmonic imaging could be used to investigate tumor IFP and vascularity as biomarkers to monitor neoadjuvant therapy response. The results from 3D SHAPE and 3D subharmonic imaging were closely related in this study; patients with greater tumor volume reduction showed higher subharmonic signal increases in the tumor than the surrounding area during the infusion of contrast agent after 10% completion of therapy. The increased subharmonic signal in SHAPE and subharmonic imaging data implies a reduction in IFP and increased vascularity, respectively. These results support the previously reported findings that decreased tumor IFP and increased vascularity can aid in the delivery of the chemotherapeutic agent in the early stages of treatment [4, 33]. Although 3D SHAPE and 3D subharmonic imaging matched the final tumor reduction rate at 10% completion, they did not correlate with the rate of tumor reduction throughout the entire duration of therapy. This was expected since tumor volume monitoring over time may not indicate the final outcome of the therapy regimen and is consistent with previous findings using IFP as an early response predictor [5]. Neither SHAPE nor SHI results were able to differentiate responders from partial responders or non-responders based on pathologic assessment, but both did correlate with imaging assessment and allowed for complete separation of responders and non-responders. This study did not consider whether factors such as age, cancer grade, treatment regimen, and breast cancer subtype impacted the prediction of pathologic response because of the limited sample size. Nam and colleagues attested that 3D SHAPE and 3D SHI combined with other predictors from MR imaging, CT, and PET, as well as patient demographics and cancer-specific factors, may improve the prediction of therapy response. The small sample size of this first-in-humans pilot study was the main limitation and did not allow for the authors to assess the correlation cancer type, grade, and hormone status with 3D SHAPE or 3D SHI. Similarly, the small sample size did not allow for the authors to determine if 3D SHAPE and 3D SHI can be used independently or in combination. However, this study showed the potential of 3D SHAPE and 3D SHI to predict the response of breast cancers to NAC as early as after 10% completion of therapy.

In the ongoing prospective study, the direct pressure gradient measurement showed that the IFP was always higher in the lesion compared to normal tissue (except for one case), and the IFP gradients between the malignant lesion and normal tissue were approximately 10 mmHg higher than those between the benign lesion and normal tissue. The subharmonic amplitude gradient was lower between the malignant lesion and normal tissue than between the benign lesion and normal tissue, implying IFP in the malignant lesions was higher than that in benign [28]. This study also compared IFP to tumor size, which showed a weak positive correlation ($r = 0.20$~0.30) between IFP and tumor size, which is in line with previous studies that also showed a positive correlation between IFP and tumor size [34–37].

45.6 Future Trends and Conclusions

The subharmonic signal gradient between a breast lesion and normal tissue, estimated by 3D SHAPE, shows potential in characterizing indeterminate breast lesions. Most recently, it has been shown that 3D SHAPE can help to estimate IFP, in patients with indeterminate breast lesions, to assess their response to neoadjuvant chemotherapy as early as after 10% completion of therapy. This was shown using the ultrasound contrast agent Definity. However, further exploration is needed

to optimize this technology for different ultrasound contrast agents and different ultrasound machines. Although a potential challenge, implementing 3D SHAPE and SHI with other commercially available agents and machines to evaluate IFP in breast tumors will help to maximize this technology's clinical translation and utility in evaluating IFP in breast tumors to help reduce the number of breast biopsies, thus improving the positive predictive value of diagnostic breast imaging.

WORKS CITED

1. V. Guarneri, A. Frassoldati, S. Giovannelli, F. Broghi and P. Conte, "Primary systemic therapy for operable breast cancer: A review of clinical trials and perspectives," *Cancer Lett.,* pp. 175–185, 2007.
2. D. Mauri, N. Pavlidis and J. Ioannidis, "Neoadjuvant versus adjuvant systemic treatment in breast cancer: A meta-analysis," *J. Natl. Cancer Inst.,* pp. 188–194, 2005.
3. K. Mcmasters and K. Hunt, "Neoadjuvant chemotherapy, locally advanced breast cancer, and quality of life," *J. Clin. Oncol.: Off. J. Am. Soc. Clin. Oncol.,* no. 17, pp. 441–444, 1999.
4. C.-H. Heldin, K. Rubin, K. Pietras and A. Ostman, "High interstitial fluid pressure—an obstacle in cancer therapy," *Nat. Rev. Cancer,* pp. 806–813, 2004.
5. S. Ferretti, P. Allegrini, M. Becquet and P. McSheehy, "Tumor interstital fluid pressure as an early response marker for anticancer therapeutics," *Nepolasia,* p. 874, 2009.
6. B. Curti, W. Urba, W. Alvord, J. Janik, J. Smith, K. Madara and D. Longo, "Interstital pressure of subcutaneous nodules in melanoma and lymphoma patients; changes during treatment," *Cancer Res.,* no. 53, pp. 2204–2207, 1993.
7. J. Less, M. Posner, Y. Boucher, D. Borochovits, N. Wolmark and R. Jain, "Interstitial hypertension in human breast and colorectal tumors," *Cancer Res.,* no. 52, pp. 6371–6374, 1992.
8. R. Jain, R. Tong and L. Munn, "Effect of vascular normalization by antiangiogenic therapy on interstital hypertension, peritumor edema, and lymphatic metastisis: Insights from a mathematical model," *Cancer Res.,* no. 67, p. 2729, 2007.
9. M. Milosevic, A. Fyles, D. Hedley, M. Pintilie, W. Levin, L. Manchul and R. Hill, "Interstitial fluid pressure predicts survival in patients with cervix cancer independent of clinical prognostic factors and tumor oxygen measurements," *Cancer Res.,* vol. 61, pp. 6400–6405, 2001.
10. B. Goldberg, J. Raichlen and F. Forsberg, *Ultrasound contrast agents: Basic principles and clinical applications*, 2nd ed. Inform. Healthc., pp. 15–24, CRC Press, 2001.
11. W. Fairbank and M. Scully, "A new noninvasive technique for cardiac pressure measurement: Resonant scattering of ultrasound from bubbles," *IEEE Trans. Bio-Med. Eng.,* pp. 107–110, 1977.
12. W. Shi, F. Forsberg, J. Raichlen, L. Needleman and B. Goldberg, "Pressure dependence of subharmonic signals from contrast microbubbles," *Ultrasound Med. Biol.,* pp. 275–283, 1999.
13. K. Ishihara, A. Kitabatake, J. Tanouchi, K. Fujii, M. Uematsu, Y. Yoshida, T. Kamada, T. Tamura, K. Chihara and K. Shirae, "New approach to noninvasice manometry based on pressure dependent resonant shift of elastic microcapsules in ultrasonic frequency characteristics," *Jpn. J. Appl. Phys. Suppl.,* pp. 125–127, 1988.
14. M. Postema, A. Bouakaz, N. de Jong and N. Bom, "Noninvasive microbubble-based pressure measurements: A simulation study," *Ultrason.,* pp. 759–762, 2004.
15. B. Hök, "A new approach to noninvasive manometry: Interaction between ultrasound and bubbles," *Med. Biol. Eng. Comput.,* pp. 35–39, 1981.
16. F. Forsberg, C. Piccoli, D. Merton, J. Palazzo and A. Hall, "Breast lesions: Imaging with contrast-enhanced subharmonic US- inital experience," *Radiol.,* pp. 718–726, 2007.
17. V. Halldorsdottir, J. Dave, L. Leodore, J. Eisenbrey, S. Park, A. Hall, K. Thomenius and F. Forsberg, "Subharmonic contrast microbubble signals for noninvasive pressure estimation under static and dynamic flow conditions," *Ultrason. Imag.,* vol. 33, no. 3, pp. 153–164, 2011.
18. I. Gupta, J. Eisenbrey, P. Machado, M. Stanczak, C. Wessner, C. Shaw, S. Gummadi, J. Fenkel, A. Tan, C. Miller, J. Parent, S. Schultz, M. Soulen, C. Sehgal, K. Wallace and F. Forsberg, "Diagnosing portal hypertesnion with noninvasice subharmonic pressure estimates from a US contrast agent," *Radiol.,* pp. 104–111, 2020.
19. J. Eisenbrey, J. Dave, V. Halldorsdottir, D. Merton, C. Miller, J. Gonzalez, P. Machado, S. Park, S. Dianis, C. Chalek, C. Kim, J. Baliff, K. Thomenius, D. Brown, V. Navarro and F. Forsberg, "Chronic liver disease: Noninvasive subharmonic aided pressure estimation of hepatic venous pressure gradient," *Radiol.,* vol. 268, no. 2, pp. 581–588, 2013.
20. J. Dave, S. Kulkami, P. Pangaonkar, M. Stanczak, M. McDonald, I. Cohen, P. Mehrotra, M. Savage, P. Walinsky, N. I. Ruggiero, D. Fishman, D. Ogilby, C. VanWhy, M. Lombardi and F. Forsberg, "Noninvasive intracardiac pressure measurements using subharmonic aided pressure estimation: Proof of concept in humans," *Ultrasound Med. Biol.,* vol. 43, no. 11, pp. 2718–2724, 2017.
21. C. Esposito, P. Machado, M. McDonald, M. Savage, D. Fischman, P. Mehrotra, I. Cohen, N. I. Ruggiero, P. Walinksy, A. Vishnevsky, K. Dickie, M. Davis, F. Forsberg and J. Dave, "Noninvasive evaluation of cardiac chamber pressures using subharmonic aided pressure estimation with definity microbubbles," *JACC Cardiovascular Imag.,* vol. 16, no. 2, pp. 224–235, 2023.
22. C. Esposito, P. Machado, M. McDonald, M. Savage, D. Fischman, P. Mehrotra, I. Cohen, N. I. Ruggiero, P. Walinksy, A. Vishnevsky, K. Dickie, M. Davis, F. Forsberg and J. Dave, "Evaluation of intra-cardiac pressures using subharmonic aided pressure estimation with Sonazoid microbubbles," *Radiol. Cardiothorac. Imag.,* vol. 6, no. 1, p. e230153, 2024. doi: 10.1148/ryct.230153. PMID: 38358329; PMCID: PMC10912883.
23. P. Machado, I. Gupta, J. Fenkel, S. Gummadi, M. Stanczak, C. Wessner, C. Shaw, S. Schultz, M. Soulen, K. Wallace, J. Eisenbrey and F. Forsberg, "Ultrasound pressure estimation for diagnosing portal hypertension in patients undergoing dialysis for chronic kidney disease," *J. Ultrasound Med.,* vol. 41, no. 9, pp. 2181–2189, 2022.
24. I. Gupta, J. Fenkel, J. Eisenbrey, P. Machado, M. Stanczak, C. Wessner, C. Shaw, C. Miller, M. Soulen, K. Wallace and F. Forsberg, "A noninvasive ultrasound based technique to

identify treatment responders in patients with portal hypertension," *Acad. Radiol.,* vol. 28, no. 1, pp. S128–S137, 2021.
25. V. Halldorsdottir, J. Dave, J. Eisenbrey, P. Machado, H. Zhao, J. Liu, D. Merton and F. Forsberg, "Subharmonic aided pressure estimation for monitoring interstitial fluid pressure in tumors- in vitro and in vivo proof of concept," *Ultrason.,* vol. 54, pp. 1938–1944, 2014.
26. V. Halldorsbottir, J. Dave, A. Marshall, A. Forsberg, T. Fox, J. Eisenbrey, P. Machado, J. Liu, D. Merton and F. Forsberg, "Subharmonic-aided pressure estimation for monitoring interstitial fluid pressure in tumors: Calibration and treatment with paclitaxel in breast cancer xenografts," *Ultrasound Med. Biol.,* vol. 43, no. 7, pp. 1401–1410, 2017.
27. K. Nam, J. Eisenbrey, M. Stanczak, A. Sridharan, A. Berger, T. Avery, J. Palazzo and F. Forsberg, "Monitoring neoadjuvant chemotherapy for breast cancer by using three-dimensional subharmonic aided pressure estimation and imaging with US contrast agents: Preliminary experience," *Radiol.,* vol. 285, no. 1, pp. 53–62, 2017.
28. K. Nam, M. Torkzaban, J. Shames, L. Liao, C. Wessner, P. Machado, A. Lyshchik and F. Forsberg, "Characterization of indeterminant breast lesions based on pressure estimates by noninvasive 3D contrast-enhanced ultrasound," *Ultrasound Med. Biol.,* vol. 50, no. 12, pp. 1854–1860, 2024. doi: 10.1016/j.ultrasmedbio.2024.08.010. Epub 2024 Sep 5. PMID: 39237425; PMCID: PMC11490378.
29. J. Dave, V. Halldorsdottir, J. Eisenbrey, D. Merton, J. Liu, P. Machado, H. Zhao, S. Park, S. Dianis, C. Chalek, T. K. E. D. Brown and F. Forsberg, "On the implementation of an automated acoustic output optimization algorithim for subharmonic aided pressure estimation," *Ultrason.,* vol. 53, pp. 880–888, 2013.
30. J. Eisenbrey, J. Dave, V. Halldorsdottir, D. Merton, P. Machado, J. Liu, C. Miller, J. Gonzalez, S. Park, S. Dianis, C. Chalek, K. Thomenius, D. Brown, V. Navarro and F. Forsberg, "Simultaneous grayscale and subharmonix ultrasound imaging on a modified commercial scanner," *Ultrason.,* vol. 51, pp. 890–897, 2011.
31. J. Dave, V. Halldorsdottir, J. Eisenbrey, J. Liu, F. Lin, J. Zhou, H. Wang, K. Thomenius and F. Forsberg, "In vivo subharmonic pressure estimation of portal hypertension in canines," *IEEE,* pp. 778–781, 2010.
32. S. Shord and J. Camp, "Intravenous administration of paclitaxel in Sprague-Dawley rats: What is a safe dose?," *Biopharm. Drug. Dispos.,* vol. 27, pp. 191–196, 2006.
33. W. Kuo, C. Chen, F. Hsieh, et al., "Vascularity change and tumor response to neoadjuvant chemotherapy for advanced breast cancer," *Ultrasound Med. Biol.,* vol. 34, no. 6, pp. 874–881, 2008.
34. M. Stohrer, Y. Boucher, M. Stangassinger and R. K. Jain, "Oncotic pressure in solid tumors is elevated," *Cancer Res.,* vol. 60, no. 15, pp. 4251–4255, 2000.
35. S. D. Nathanson and L. Nelson, "Interstital fluid pressure in breast cancer, benign breast conditions, and breast parenchyma," *Ann. Surg. Oncol.,* vol. 1, no. 4, pp. 333–338, 1994.
36. Y. Boucher, J. M. Kirkwood, D. Opacic, et al., "Interstital hypertension in superficial metastatic melanomas in humans," *Cancer Res.,* vol. 51, no. 24, pp. 6691–6694, 1991.
37. R. Gutmann, M. Leunig, J. Feyh, et al., "Interstitial hypertension in head and neck tumors in patients: Correlation with tumor size," *Cancer Res.,* vol. 52, no. 7, pp. 1993–1995, 1992.

46

Development and Implementation of a Method for Registering Prostate-Specific Antigen (PSA) Biosensor-Based Assay Results in a Personalized QR Code

Maria Daoutakou and Spyridon Kintzios

46.1 Introduction

Prostate cancer (PCa) is one of the most common cancers and is the second leading cause of cancer death among men in developed countries. Early detection is associated with better treatment options, leading to better outcomes for patients. Currently, the analysis of prostate-specific antigen (PSA) in the blood is the only clinical examination utilized for PCa screening. The PSA protein is a serine protease secreted by the epithelium of the prostate gland, mainly found in seminal plasma. PSA levels are typically low in healthy men and increase as a consequence of inflammation, hyperplasia, or cancer [1]. Due to the ganglion-like disposition of the prostate gland, PCa is often diagnosed at an advanced stage when no curative treatments are available. Although blood PSA levels greater than a threshold value in the range of 2.5 to 4 ng/mL are correlated with prostate diseases, this correlation is not sensible and results in significant false-positive and false-negative cases [2–4]. The false-positive cases lead to unnecessary invasive examinations such as biopsy, potentially causing side effects, while the false-negative cases may result in the progression of malign diseases [5, 6]. PSA testing has been routinely used during the last 40 years, resulting in a significant rise in the number of cases with low-grade and -volume tumors diagnosed clinically and treated by aggressive therapies. Complications of treatment (impotence and incontinence) might severely impact the life quality of patients. To reduce the number of unnecessary biopsies and re-biopsies, various diagnostic tests based on fPSA and proPSA fractions and their ratios with tPSA beta-microseminoprotein (BPSA), proPSA (p2PSA, pPSA), and %fPSA have been developed. They are good in differentiating PCa among men in a tPSA range of 2.5–10 ng/mL, not good for men in a tPSA range of 10–20 ng/mL, and perform poorly for men on finasteride therapy. Commercially available second-generation tests have been shown to discriminate cancer with 4–10 ng/mL results from normal prostates, but not with a 2.5–4 ng/mL tPSA range. In terms of clinical significance, fPSA and proPSA tests significantly improve specificity and can reduce the number of unnecessary biopsies by about 20%–40% [7–11].

The limitations of the PSA test have prompted the development of various adjunct testing modalities that can be used in combination, which would also promote the improvement of screening tools. Many efforts were dedicated to refining the existing PSA test and its clinical use, involving the standardization of the assay method used [12, 13]. The standard PSA assay is performed using chemiluminescent or fluorescent immunochemical IVD test systems. The measurement results are mostly recorded in printed format by an analyzer, which could be conveniently lost or accidentally discarded. In recent years, various electrochemical biosensing approaches have been reported employing a variety of biorecognition elements for the detection of PSA either in standard solutions or in biological samples. As an analytical concept, electrochemistry offers the advantage of rapid testing, relatively low cost and, quite often, portability. Among the very recent developments in this direction, electrochemical immunoassays stand out as the most favorable working principle [14, 15]. As a further advance in this field, a cell-based biosensor has been reported capable to complete semi-quantitative PSA assay within 5 min or less [16].

Irrespective of the assay method, PSA test results are mostly not readily available at the point of care and are recorded in plain text format, which makes them prone to misinterpretation. This could lead to incapability of obtaining a timely consultation with a medical professional even if the test result indicates a possible health risk.

A useful approach to enhance the link between the PSA assay results and medical consultation would be to have a unique identifier assigned to the PSA test result that could be subsequently used to fetch this test result from an external repository. However, the act of indexing and storing each PSA test result with the unique identifier is frequently an additional workload for healthcare professionals. Furthermore, a unique identifier alone would not be sufficient for a patient to find and fetch a particular test result, as it is impossible for the patient to remember the identifier.

A possible way to address the aforementioned issues would be to convert each PSA test result to a personalized machine-readable QR code. In essence, the QR code would encode the unique identifier of the PSA test result, which would subsequently allow its automatic retrieval. By storing this QR code on a medicine information booklet, the personalized QR code would be readily available for the patient immediately after receiving the PSA assay result. A personalized QR code would

DOI: 10.1201/9781003449942-51

help to link the PSA assay result and health information with a particular patient, thus minimizing the workload of the healthcare professional and preventing the loss of important health information.

Nevertheless, the conversion of the PSA assay results to personalized QR codes requires the implementation of both software and hardware. The software part includes the automatic indexing of the PSA assay results, converting the results to the QR code format, and storing this QR code image on the patient's personal health data repository. The hardware part is the necessary modification of the analyzer to allow for the printing of the personalized QR code.

In the present report we demonstrate the application of a bespoke smartphone app, coupled with a previously published bioelectric biosensor [16], capable of analyzing data regarding PSA cut-off values. The app is designed to encode the results from the biosensor device in a standard QR code, combined with personalized historical and current data.

46.2 Materials

46.2.1 Biosensor Setup

The biosensor system was based on the previous work by Mavrikou and colleagues involving, as biorecognition elements, membrane-engineered cells bearing anti-PSA antibodies on their surface. Briefly, 2.5×10^6 Vero cells in 40 μL PBS were incubated on ice with 400 μL of antibody (0.25 μg/mL) for 20 min. After incubation the mixture was transferred to an electric field of 1800 V/cm and two square electric pulses were applied. After modification the cells were stored at 37 °C, in 5% CO_2.

For the measurement of changes in the cell bioelectric potential, a custom-made single channel potentiostat (PalmSens, The Netherlands) was used as the readout device. The system recorded the electric signals from cells on the screen printed carbon working electrode (DRP-8X110, DropSens, Spain) and allowed a high speed of the assay (duration: 5 min).

46.2.2 Assay Process

Participants (n = 28) for this pilot study were recruited from a local community health center during their annual PSA test screening visit. Blood samples with PSA values measured by immunoradiometric assay (IRMA) technique [17] in the range from zero to 60 ng/mL were collected from an equal number of patients at Army Share Fund Hospital of Athens. Samples with PSA concentration below 4 ng/ml were considered negative (control). Prior to assay, samples were stored at −20 °C for one to three months. A drop of 45 μL of cell-containing solution (1.1×10^6 cells/mL) was deposited on the working electrode with the aid of a pipette, followed by 5 μL of sample. The response of the biosensor to the different samples was recorded as a time-series of bioelectric potential values for 5 min at a sampling rate of 2 Hz.

Results from each sample were expressed as a ratio to control (mean ± SD) according to the following equation:

Normalized Biosensor Response = Mean Final Value of Sample/Mean Final Value of Control

Results (normalized biosensor responses) equal or above the cut-off value of 4 ng/mL were considered positive.

46.3 Methods

46.3.1 Smartphone App Development

A user-friendly smartphone application was designed for implementing data input and generation of the assay-related QR code. Following end-user authentication and input of patient personal information, the user can activate the measurement button, thus triggering the application to communicate with the embedded biosensor device and upload the measurement results automatically. This straightforward, streamlined process is designed to minimize manual errors and to facilitate the interpretation of the assay results. Following the upload of the measurement data, a unique QR code is generated that encodes all the information, including the user's personal details and the test results.

The design and development of the registration system of PSA assessment with subsequent entry of results into the QR code was carried out using HTML, CSS, and JavaScript. Input validation was carried out using JavaScript at the stage of filling out the registration form and JavaScript file, without which registration is not possible.

The registration form consists of three blocks: personal data, input of the results of analysis for the content of PSA, and information about the appointment of additional research. The process of storing, processing, linking to results, and establishing a QR code takes place in two stages. After registration, the previous process occurs automatically, where the unique output number of the certificate is generated from the current date and time and additional numbers of sequence (if these numbers are occupied the sequence is increased) pre-distributed into the database. Validity checking occurs through a full match of the database tables to which the certificate number is associated. In the absence of a match, the system warns that this number is invalid. The second option of binding was established through a direct link to the database tables, and at decoding the question of personal data occurred (full name, identification number, or date of birth). The match goes through the same principle. Besides using only one input option, all options are equally valid. Both of these options were integrated into the QR code after the decoding options.

The concept of the assay from blood sampling to personalized QR code generation is depicted in Figure 46.1. The actual integrated device is shown in Figure 46.2.

46.4 Results

In the present study we demonstrated a system that allows rapid detecting prostate cancer risk and automatic registration of total prostate-specific antigen (tPSA) results in a personalized QR code. The implemented design was tested by examining the functionality and efficacy of the device via quality control testing. The registered test results could then be decoded to display the individual tPSA values. The automatic system was

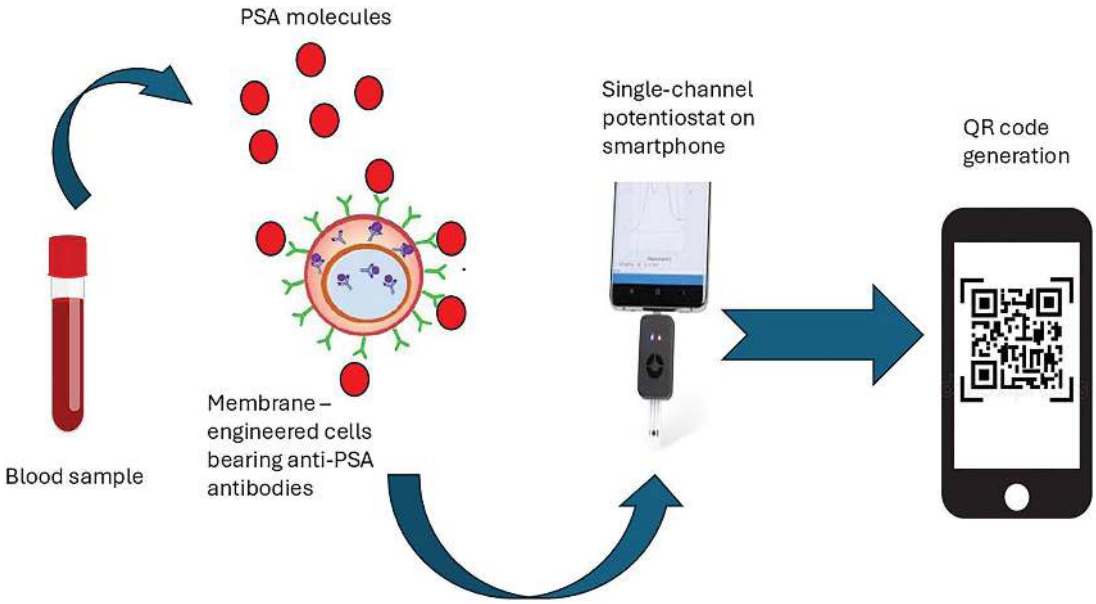


FIGURE 46.1 Schematic summary of the concept of the cell-based biosensor assay for PSA in blood and the generation of a QR code on a smartphone containing test result and patient historical data.

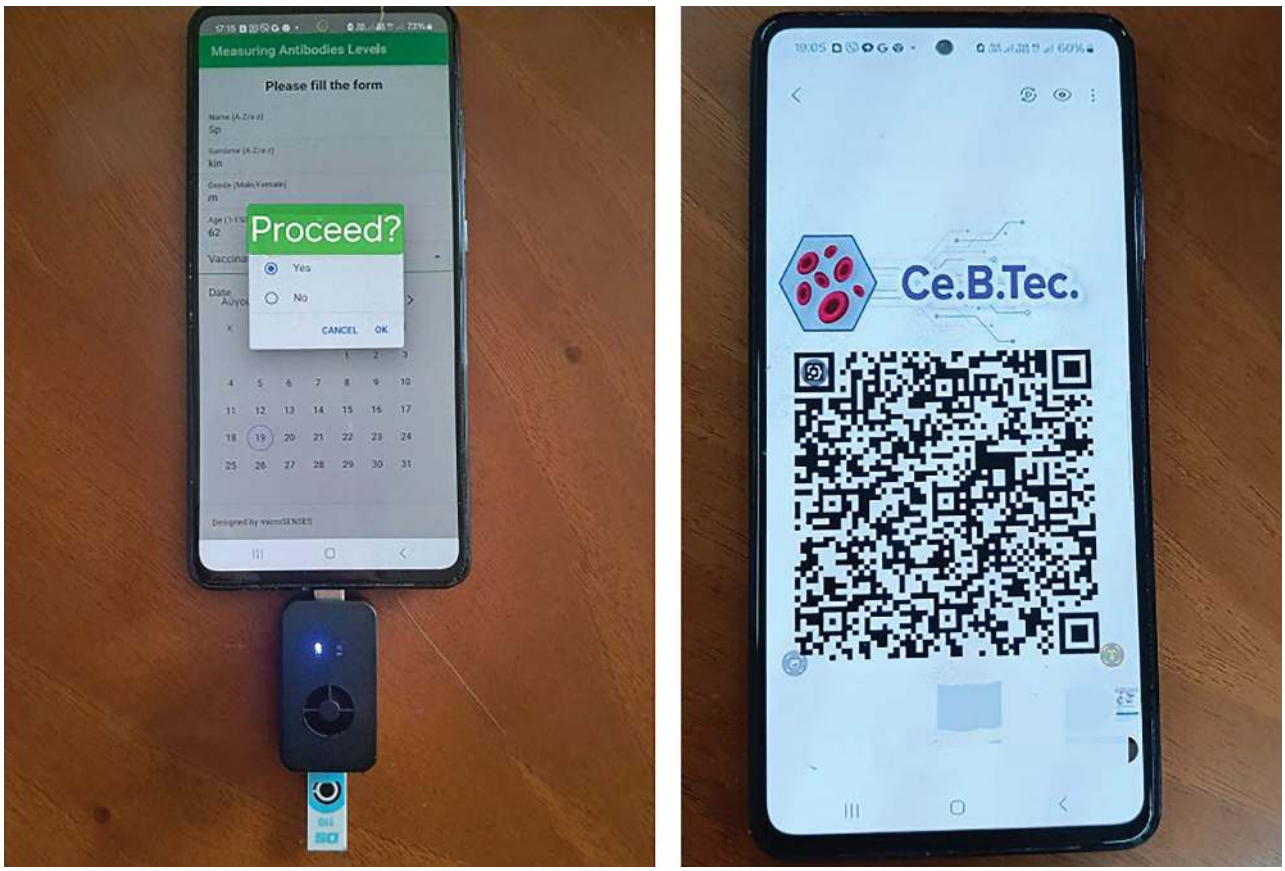


FIGURE 46.2 Representation of the user interface of the bespoke app developed as a modification of the approach by Apostolakis and colleagues [18]. Prior to the measurement of a blood sample with the aid of a portable bioelectric biosensor [16] (attached to the smartphone via a USB port), the user inputs personal and historical data (e.g., age, gender, etc.) (left). After the completion of the three-minute assay, a QR code is generated containing the test results (right).

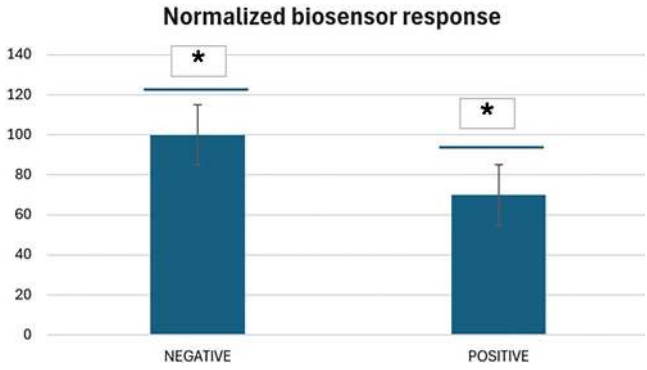


FIGURE 46.3 Normalized biosensor response (cumulative change of cell membrane potential) against 28 human serum samples classified in two different PSA concentration ranges (< 4 ng/mL and ≥ 4 ng/mL). Data are means ± SE of replications (n = 24) received on three different dates using different batches of membrane-engineered cells. * $p < 0.001$, significantly different from control samples.

adapted to store and present assay results in a personalized QR code, allowing patients to read the results through a mobile device and subsequently store and analyze them.

In agreement with our previously published report, we were able to successfully discriminate between control (negative) and positive blood samples at the cut-off value of ≥ 4 ng/mL PSA with the aid of the biosensor based on mammalian cells membrane-engineered with anti-PSA antibodies. Of the 28 patient blood samples tested, 19 scored positive by producing a normalized biosensor response on average 30% lower than control (Figure 46.3).

The integration of PSA analysis results with a QR code was implemented using a unique link to the analysis and decoding variants. The already-developed mechanism for certificate generation was used. A QR code with basic information about the registered PSA test is generated upon successful registration, allowing users to access their information through mobile devices. Registered data is securely stored in a MySQL database [19].

The methodology for acquiring, processing, transferring, and registering PSA assay results using a personalized QR code was demonstrated practically. A working model was created, consisting of software for processing the document with the assay results, an automatic generation of a personal QR code with the result, a website for registering QR codes, and a mobile application for displaying registered assay results.

46.5 Discussion

As shown in the present study, the integration of prostate-specific antigen results with QR code generation involved the use of a portable, POC diagnostic instrument designed to calculate PSA values and convert them into a QR code. This implementation is built on an embedded platform that combines a smartphone microcontroller (ESP32), a small display (ILI9341), and a Wi-Fi module integrated into the ESP32. Integrated systems based on the combination of these electronic components are capable of receiving the values of the results in digital format through the Internet, processing this information and immediately generating a QR code.

It is furthermore possible to integrate the demonstrated process with laboratory information systems (LISs) and a web application allowing the system to be used in any organization, accommodating laboratories, hospitals, clinics, and living houses with access to the Internet. A database storing patient PSA results could be developed alongside their unique QR codes. Data from different types of laboratories could be subjected to a unified format of communication with the web application, through which QR codes would be generated separately for any number of results and returned in a digital format. This opens new possibilities in the storage of results in easily accessible places and the protection against unauthorized access while retaining connection with the primary LISs.

Efficiency tests (analytical results not shown) demonstrated a very high degree of document processing accuracy (95% or higher) and a low probability of a personalized QR code result duplication (less than 0.03%). The system can adequately ensure the main goal of registering PSA biosensor-based assay results in a personalized QR code at the current stage of development and implementation.

The clinical significance of prostate-specific antigen for the early diagnosis of prostate cancer is unquestionable. However, the PSA test, just like any other medical test, can lead to false-positive and negative results, which in some cases can have dire consequences. Moreover, the PSA test is often performed elsewhere than the actual hospital, causing difficulties in making appointments with a specialist. The actual concentration of PSA in blood serum determines the risk of having prostate abnormalities, yet the practical potential of the PSA test is not fully utilized. In order to improve the quality of life of patients who have undergone a PSA test, the test result can be encoded in a personalized QR code. The QR code can then be used by health professionals to follow the patient's history [20–22]. This would make tracking the patient's history of the PSA test easy, which is important for quality prostate care and operations.

Today, most results from such procedures are provided by traditional methods, which consist of drawing several graphs and tables with numbers on them and printing them on paper sheets. This method has several disadvantages when being analyzed relative to the method proposed in this paper. Our approach aims to create a more versatile, user-friendly, and secure system for registering results from laboratory tests or health assessments.

Two limitations of the proposed approach should be acknowledged. First, requiring a permanent QR code for each patient necessitates a mechanism to link it to individual patients without violating data privacy laws. This challenge may hinder implementation at existing medical facilities where individual patients are represented only by a string of numbers or other forms of ID. In facilities where more information is known about each patient, such as an email, telephone number, and home address, assigning a QR code to each patient may be more feasible. Additionally, transitions to applying this technique to new patients could be gradual at some medical facilities.

Second, a permanent QR code could be problematic for patients suffering from conditions causing permanent changes

to their ID. These conditions may include mental illness where a patient can change their place of residence without reporting it to legal authorities. In legal systems where this change makes a patient lose their ID or change their ID number, the proposed method would no longer work. However, this limitation may seem negligible since patients suffering from such illness rarely have other coexisting conditions that would require prostate-specific antigen monitoring.

Nonetheless, two other solutions could potentially work around these limitations. One possibility is designing a QR code detector that would recognize the code itself, instead of registering it in a database linked to the patient's ID. Detected QR codes could be imported from photographs taken by lab technicians alongside PSA testing. On the one hand, such solutions could be neglected where scanning QR codes would be possible only in a small medical facility. On the other hand, in bigger medical facilities, such solution may dramatically reduce the number of medical errors made in analyzing urine samples and monitoring PSA assays results. In such systems, a given QR code could be permanently assigned to only one patient's urine sample, thus virtually excluding a possibility of analyzing results on a sample belonging to another patient. Additionally, the decision to implement such a solution depends on whether designing and implementing an automatic detector of QR codes in patient samples would be reasonably cheap and effective.

The other possibility to work around the proposed limitations could be making QR codes meta-data dependent on knowing both the QR code and the paper or tube where the QR code is printed. QR codes printed on paper sheets could include meta-data identifying a patient with a string of characters, while QR codes printed directly on tubes could include merely a random number of characters. In such a system, a database linking the first kind of QR code and their meta-data to a string of characters identifying the patient could be stored in the database. In this way, QRs printed on papers containing meta-data would need to be also scanned to access their meta-data. Nonetheless, scanning merely QR codes printed on tubes would remain ineffective as using their scanner would only have a bare chain of characters linking the QR code to a database with a meta-data of blank values.

In conclusion, the current work presented a novel method for registering biosensor-facilitated PSA assay results in a personalized QR code. This method integrates the results' encryption via a personalized identification number, its transformation into a QR code, and its assignment to the patient in the LIMS. An approach for the transformation of the PSA results into a QR code has been proposed, and the corresponding algorithm has been developed and implemented in the Java programming language. The developed software generates a QR code image file that encodes a string containing a personalized identification number and the PSA assay result. Consequently, this QR code may be printed out, attached to the testing sample or the patient's documentation, and used for remote access to the record containing the PSA testing result in the LIMS.

Despite showing promising results, several issues remain to be addressed. The proposed system requires reliable storage and secure transfer of the testing records between the LIMS and the equipment used for reading QR codes. Safety measures should be put in place to avoid situations where anyone could read the QR codes and access the records. An option for the sophisticated coding scheme of the strings embedded in the QR code may be implemented. In the current method, the PSA results are embedded as a sequence of decimal symbols. In a more sophisticated system, the PSA results may be encoded as numbers in floating-point format. Holding the string in such a format could accelerate the QR code reading process and support database integrity checkups. An extensive practical verification of the developed method would strengthen its applicability.

46.6 Future Trends

Innovative applications of artificial intelligence (AI) and machine learning may include the ability to harness data in order to predict health outcomes through tailoring recommendations to individuals and substantially reduce the time needed for diagnostic research. These technologies may also be used to develop "predictive health passports" that are generated through the use of complex profiles and patterns drawn from statistical data to inform health passport applications by patients and healthcare workers. The use of predictive AI for PSA is, in many ways, the natural evolution of the technology whose first applications have been seen in e-commerce management, high-frequency stock trading, and "clickbait" content discovery. For PSA, the concept may be to reduce fraudulent health passport applications by entrants from countries that require a visa or entry clearance through failure to meet the requirements or carrying the conditions of the stay that restrict a non-EEA national's access to the National Health System (NHS).

In the future, it is anticipated that standard verification and AI could further streamline the process of sharing health information. Digital solutions are anticipated to introduce user-friendly multilingual interfaces where an individual can store and share key health and medication information via an app, with real-time visa restrictions, health advisories, and vaccination and testing requirements populating based on the country of travel. Anticipated future trends in this area are driven by evolving consumer experience and user expectations, particularly concerning transparency and the increased use of AI and machine learning. In the area of digital health passports, these trends are expected to challenge how the digital tools are conceptualized and deployed. Revolutionary digital health passports or any tools reliant on the analysis of massive amounts of data to inform decision-makers have multiple unknown applications anticipated for the future. Partnerships between technology companies and providers will continue to be watched with interest, particularly around data privacy and freedoms. Data protection and privacy compliance continue to be treated as a key objective within the proposals. Once approved, standardization with international health passport standards would guide innovations in this context.

REFERENCES

1. Gandaglia G, Leni R, Bray F, Fleshner N, Freedland SJ, Kibel A, Stattin P, Van Poppel H, La Vecchia C. Epidemiology and Prevention of Prostate Cancer. *Eur Urol Oncol*. 2021:877–892.
2. Wang L, Lu B, He M, Wang Y, Wang Z, Du L. Prostate Cancer Incidence and Mortality: Global Status and Temporal Trends in 89 Countries From 2000 to 2019. *Front Public Health*. 2022;10:811044.
3. Ergun O, Capar E, Goger YE, Ergun AG. Can Expressed Prostatic Secretions Effect Prostate Biopsy Decision of Urologist? *Int Braz J Urol*. 2019;45(2):246–252.
4. Erdogan A, Polat S, Keskin E, Turan A. Is Prostate Volume Better Than PSA Density and Free/Total PSA Ratio in Predicting Prostate Cancer in Patients with PSA 2.5–10 ng/mL and 10.1–30 ng/mL? *Aging Male*. 2020;23(1):59–65.
5. Lundgren PO, Kjellman A, Norming U, Gustafsson O. Association between One-Time Prostate-Specific Antigen (PSA) Test with Free/Total PSA Ratio and Prostate Cancer Mortality: A 30-Year Prospective Cohort Study. *BJU Int*. 2021;128(4):490–496.
6. Gao XD, Miao Q, Zhang JL, Zhai JZ, Gui XM, Cai YH, Niu Q, Cai B. Clinical Application of Free/Total PSA Ratio in the Diagnosis of Prostate Cancer in Men Over 50 Years of Age with Total PSA Levels of 2.0–25.0 ng ml-1 in Western China. *Asian J Androl*. 2022;24(2):195–200.
7. Culp MB, Soerjomataram I, Efstathiou JA, Bray F, Jemal A. Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates. *Eur Urol*. 2020;77(1):38–52.
8. Bergengren O, Westerberg M, Holmberg L, Stattin P, Bill-Axelson A, Garmo H. Variation in Prostate-Specific Antigen Testing Rates and Prostate Cancer Treatments and Outcomes in a National 20-Year Cohort. *JAMA Netw Open*. 2021;4(5):e219444.
9. Oswald N, Lin T, Haaland B, Flynn M, Kawamoto K, Cooney KA, Lowrance W, Hanson HA, O'Neil B. Factors Associated with Appropriate and Low-Value PSA Testing. *Cancer Epidemiol*. 2020;66:101724.
10. Carlsson SV, Vickers AJ. Screening for Prostate Cancer. *Med Clin North Am*. 2020;104(6):1051–1062.
11. Verguet S, Feldhaus I, Jiang Kwete X, Aqil A, Atun R, Bishai D, Cecchini M, Guerra Junior AA, Habtemariam MK, Jbaily A, Karanfil O, Kruk ME, Haneuse S, Norheim OF, Smith PC, Tolla MT, Zewdu S, Bump J. Health System Modelling Research: Towards a Whole-Health-System Perspective for Identifying Good Value for Money Investments in Health System Strengthening. *BMJ Glob Health*. 2019;4(2):e001311.
12. Lee B, Mahmud I, Marchica J, Dereziński P, Qi F, Wang F, Joshi P, Valerio F, Rivera I, Patel V, Pavlovich CP, Garrett TJ, Schroth GP, Sun Y, Perera RJ. Integrated RNA and Metabolite Profiling of Urine Liquid Biopsies for Prostate Cancer Biomarker Discovery. *Sci Rep*. 2020;10(1):3716.
13. Lima AR, Pinto J, Amaro F, Bastos ML, Carvalho M, Guedes de Pinho P. Advances and Perspectives in Prostate Cancer Biomarker Discovery in the Last 5 Years through Tissue and Urine Metabolomics. *Metabolites*. 2021;11(3):181.
14. Dowlatshahi S, Abdekhodaie MJ. Electrochemical Prostate-Specific Antigen Biosensors Based on Electroconductive Nanomaterials and Polymers. *Clin Chim Acta*. 2021;516: 111–135.
15. Aidoo-Brown J, Moschou D, Estrela P. Multiplexed Prostate Cancer Companion Diagnostic Devices. *Sensors (Basel)*. 2021;21(15):5023.
16. Mavrikou S, Moschopoulou G, Zafeirakis A, Kalogeropoulou K, Giannakos G, Skevis A, Kintzios S. An Ultra-Rapid Biosensory Point-of-Care (POC) Assay for Prostate-Specific Antigen (PSA) Detection in Human Serum. *Sensors (Basel)*. 2018;18(11):3834.
17. Foroutan H, Najafi R, Babaei MH, Shafii M. Preparation of Prostate Specific Antigen Standards for Immunoradiometric Assay. *Iran J Radiat Res*. 2008;6:51–58.
18. Apostolakis A, Barmpakos D, Mavrikou S, Papaionannou GM, Tsekouras V, Hatziagapiou K, et al. System for Classifying Antibody Concentration Against Severe Acute Respiratory Syndrome Coronavirus 2 S1 Spike Antigen with Automatic Quick Response Generation for Integration with Health Passports. *Explor Digit Health Technol*. 2024;2:20–29.
19. Xia XQ, McClelland M, Wang Y. TabSQL: A MySQL Tool to Facilitate Mapping User Data to Public Databases. *BMC Bioinformatics*. 2010;11:342.
20. Jia W, Wu B, Shao Y, Cao X, Wang D. Diagnostic Performance of Prostate Cancer Antigen 3 and the Prostate Health Index in Detecting Overall and Clinically Significant Prostate Cancer in Men at First Biopsy: A Meta-Analysis. *Int J Urol*. 2021;28(3):315–325.
21. Robinson HS, Lee SS, Barocas DA, Tosoian JJ. Evaluation of Blood and Urine Based Biomarkers for Detection of Clinically-Significant Prostate Cancer. *Prostate Cancer Prostatic Dis*. 2025;28(1):45–55. doi: 10.1038/s41391-024-00840-0. Epub 2024 Jun 10. PMID: 38858447.
22. Guldvik IJ, Braadland PR, Sivanesan S, Ramberg H, Kristensen G, Tennstedt P, Røder A, Schlomm T, Berge V, Eri LM, Lilleby W, Mills IG, Taskén KA. Low Blood Levels of LRG1 Before Radical Prostatectomy Identify Patients with High Risk of Progression to Castration-Resistant Prostate Cancer. *Eur Urol Open Sci*. 2022;45:68–75.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Section 5

Organel and Cellular Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

47

Deep Learning-Based Colon Segmentation for Accurate Colorectal Polyps Detection

Samir Harb, Mohamed Yousuf, Ahmed Elsayed, Asem Ali, Salwa Elshazly, and Aly Farag

47.1 Introduction

Colorectal cancer (CRC) is a type of cancer characterized by an uncontrolled growth of abnormal cells in the colon or rectum, ranking as the third most common cancer in the United States [1]. The American Cancer Society (ACS) [2] recommends that screening for CRC should start at age 45. Early diagnosis and treatment of CRC has been shown to be most effective of treating this cancer with significant recovery rate (e.g., [3]).

There are various methods available for CRC screening, differing in sensitivity, specificity, safety, invasiveness, feasibility, and cost [4]. Optical colonoscopy is considered the gold standard method for CRC screening. On the other hand, CT colonography (CTC) is noninvasive and performed without sedation in less time and is quite safe (e.g., [4, 5]). It is also less expensive than colonoscopy, with its drawbacks being less significant compared to the risks associated with colonoscopy-related procedures and hospitalization [4, 6]. Despite concerns about radiation exposure, CTC offers a favorable benefit-to-risk ratio [4, 7].

The CTC pipeline is comprised of four distinct steps, as shown in Figure 47.1. First, image segmentation isolates the lumen from surrounding abdominal tissues. The second step involves the 3D construction of the colon. This enables extracting the centerline for visualization purposes and ensuring proper alignment of supine and prone CTC scans. The third step focuses on visualization, presenting the lumen on radiologists' stations with detailed 3D and corresponding 2D CT views. Finally, in the fourth and crucial step, an in-depth analysis is conducted. This involves polyp detection, classification, archiving of relevant data, and the meticulous preparation of a comprehensive patient record. Through these steps, the CTC framework enables screening for CRC and assists other functionalities such as training and image-guided interventions related to combating CRC. There exists a large literature on CTC (e.g., [8–10]).

Segmenting colons accurately from abdomen CT scans poses significant challenges, as depicted in Figure 47.2. First, colons exhibit highly variable and asymmetric topology [11], and their positions vary across different CT images [12]. Second, distinguishing colons from surrounding structures is complicated by the presence of Hounsfield intensity regions containing soft tissues, air regions resembling gas-filled organs like the small intestine, and high attenuation structures akin to bones. Lastly, patient preparation imperfections, such as residual stool and lesions, can lead to disjointed colon segments. These complexities may confound segmentation algorithms, impeding accurate results [12].

Biomedical image segmentation approaches in the literature could be grouped into two main categories: (1) classic segmentation approaches, which typically employ techniques such as Markov random field (MRF) models [8], edge detection, and region growing [10, 13, 14]; and (2) deep learning (DL) approaches; which casts the segmentation as a classification problem solved by convolutional neural networks (CNNs) identified by deep learning techniques (e.g., [15–18]). MRF-based methods exploit the connectivity of neighboring pixels/voxels that belong to the object of concern. DL approaches exploit the available data to learn complicated high-level characteristics that can be used for segmentation. Classical approaches tend to focus on low-level traits that may not be as helpful for segmentation. Although DL approaches have shown better performance in segmentation, in general, Wasserthal and colleagues [18] reported that the colon posed the most significant challenges, with a failure rate of ~35% of cases. This failure primarily resulted from difficulties in accurately segmenting small parts of the colon.

Most DL segmentation approaches depend on a meta-architecture, which consists of encoder and decoder. The encoder has a CNN-based backbone that generates multi-resolution feature maps that are passed to the decoder via skip connections. The decoders fuse these features through deconvolution blocks to generate a final feature map. The segmentation head uses this feature map to generate a predicted segmentation mask. Numerous segmentation approaches, for example UNet [19], UNet++ [20], Linknet [21], FPN [22], PSPNet [23], PAN [24], DeepLabV3 [25], and DeepLabV3+ [26], have been proposed with the encoder-decoder architecture demonstrating significant success in various segmentation tasks.

In the medical imaging literature, the decision between employing two-dimensional (2D) and three-dimensional (3D) deep learning models holds paramount importance. While 2D-based models adeptly handle individual CT images, offering computational efficiency and adaptability to irregular sampling, 3D models excel in capturing interslice context at the

DOI: 10.1201/9781003449942-53

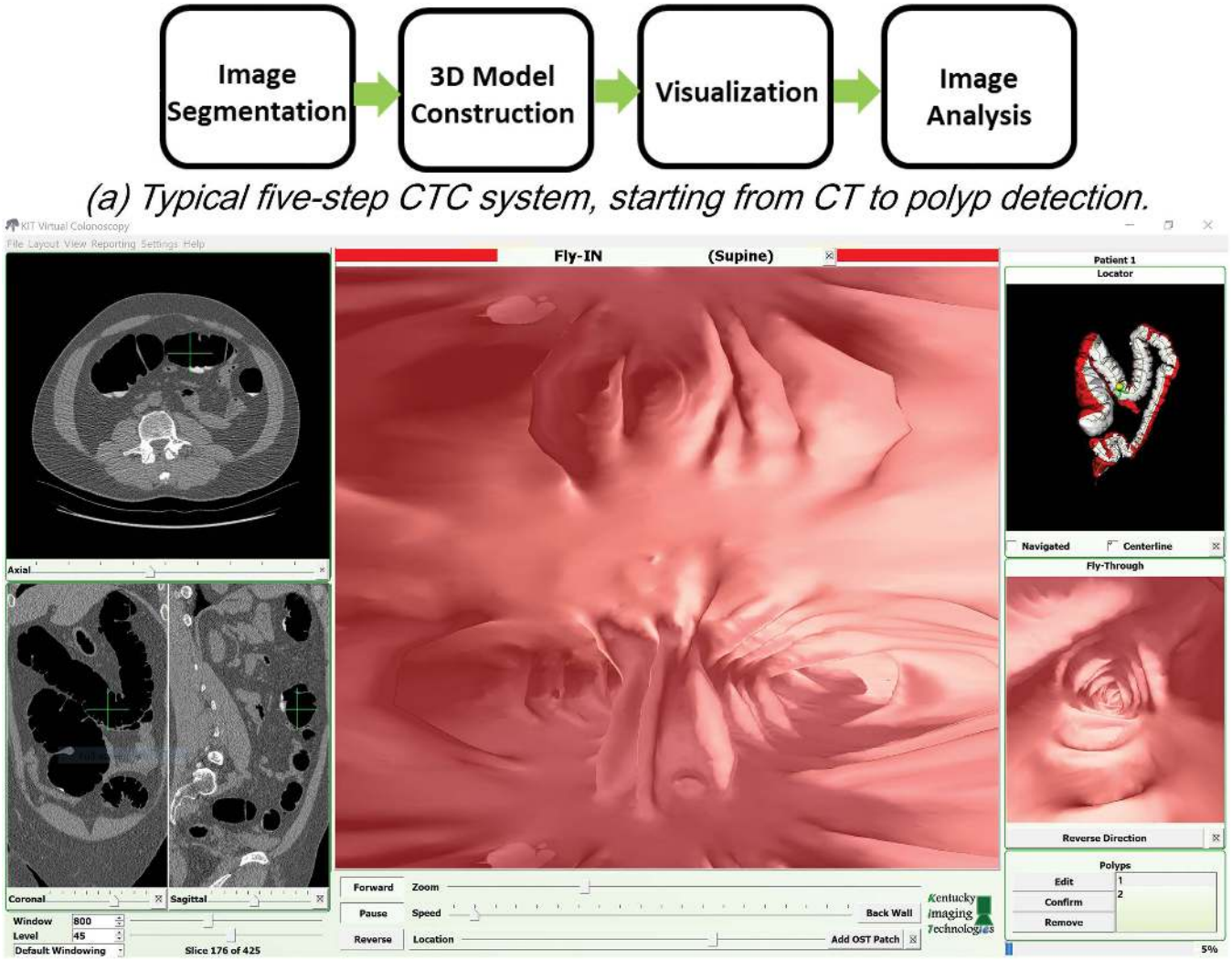


FIGURE 47.1 CTC platform: (a) typical pipeline; (b) CTC components: 2D CT slices, the 3D model reconstruction, and visualizations: Fly-In and Fly-Through.

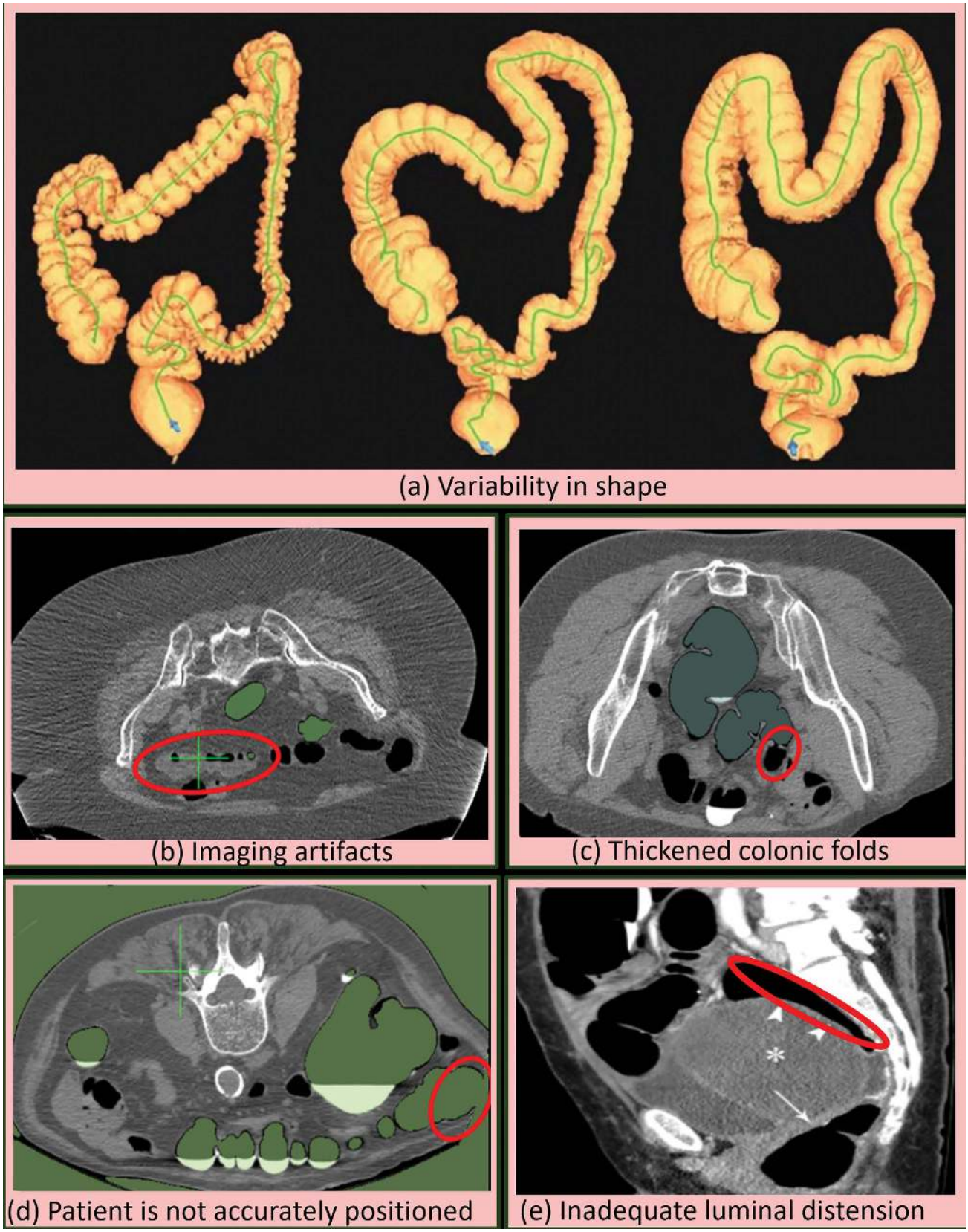


FIGURE 47.2 Examples of segmentation challenges due to inaccurate position, thickened colonic folds, and other artifacts.

cost of heightened computational demands due to an expanded parameter space. So, applying 3D networks to process CT datasets may entail downsampling, potentially compromising crucial information.

47.2 Materials

47.2.1 Dataset

We conducted our experiments on a private dataset provided by CTC experts from the University of Wisconsin. This dataset comprises scans of 49 patients in both supine and prone positions. Experts annotated the colon segments in these 98 CT scans. We normalized CT images with $\mu = 841.78$ and =1508.98 before being fed into the model for training or validation, where μ and σ are the mean and standard deviation, respectively, of the raw Hounsfield values over the whole dataset, which is available upon request. This research study was conducted retrospectively using human subject data made available in open access by ACRIN, Walter Reed Medical Center, and the University of Wisconsin [27]. The research protocol is governed by the University of Louisville IRB No. 07.0252.

47.2.2 Hardware

The used server hardware for the training and validation has the following specs: Processor: AMD EPYC 7702 P 64-Core Processor, RAM size:264 GB, used GPU: two Nvidia TITAN X with 24 GB of VRAM each.

For the CT device to generate the dataset, the manufacturer was GE Medical Systems, with Kilo Voltage Peak = 120, slice thickness 1.25, and X-ray tube current 180 mA.

47.2.3 Software and Packages

The Ubuntu 20.04.6 LTS is the main operating system that had been used, alongside with Python 3.9 as the main programming language. Further, Numpy, pandas, PyDicom, and OpenCV libraries have been used for preparing the CT images before processing.

For the CNN-based segmentation approach, to train all models, an AdamW optimizer [28] is used with an initial learning rate set to 10^{-4}, and a weight decay of 3^{-2}. Moreover, a cosine annealing learning rate [29] is used with T_{max} is set to the number of batches per epoch. All models are trained for ten epochs with batch size equal to eight, and image size 512×512. Moreover, we start training with a pretrained model from PyTorch Segmentation Models Library [30].

47.3 Methods

In this chapter, we present two quantitative algorithms for colon segmentation in CTC using MRF and DL.

47.3.1 MRF-Based Segmentation Approach

In a typical CTC scan, the lumen has two components: opacified fluid whose Hounsfield intensity is greater than 300 HU, and air, for which the characteristic peaks are almost at 1000 HU (e.g., [13]).

The problem is formulated as *maximum a posteriori* (MAP) estimation of a MRF, which involves finding the labeling that minimizes the following energy function $E(\mathrm{x})$ that combines both the spatial smoothness and data consistency to achieve an accurate segmentation of the colon in CT scans:

$$E(\mathbf{x}) = \sum_{\{i,j\}\in\mathcal{N}} V(x_i, x_j) + \sum_{i\in\mathcal{P}} D(x_i). \tag{47.1}$$

where $\mathcal{N}$ represents the set of neighboring pixel pairs (i, j), $V(x_i \mid x_j)$ is the potential function that penalizes label inconsistencies between neighboring pixels, and $D(x_i)$ is the data penalty term that measures how well the labeling x_i matches the observed image I_i. Minimizing energy function in Equation 47.1 using graph cut generates the optimal segmentation result.

First, we use expectation-maximization (EM) to calculate the empirical distributions of Hounsfield intensities in a DICOM volume $\mathbf{V}$, as shown in Figure 47.3b. This step is mandatory to identify the different classes, for example air, fat, fluid, and so on. Then, we identify colon regions using two thresholds, which are calculated using four empirical distributions of air, fat, muscle, and fluid. The threshold between air and fat and the other between muscle and fluid are used to generate O_{EM}. However, the initial segmentation results O_{EM} include noncolonic tissues, for example bones as shown in Figure 47.3c. To exclude these noncolonic components, we apply the region-growing (RG) algorithm, starting from the rectum, which can be easily identified as shown in Figure 47.3d. The RG algorithm is designed to generate an eroded colon O_{RG}. This guarantees that other organs, for example small intestine, are not merged with O_{RG}. Finally, to extract optimal segmentation, we exploit the pixel information in addition to its neighborhood to refine the initial segmentation O_{RG} through an optimization technique.

47.3.2 CNN-Based Segmentation Approach

Figure 47.4 illustrates a meta-architecture for a CNN-based classifier used for colon segmentation, consisting of an encoder and a decoder. The encoder has a CNN backbone that generates a multi-resolution feature map from a given input image or an image sequence. These feature maps are passed to the decoder via skip connections which use multiple upsampling blocks that fuse the multi-scale features and generate a final feature map. The final feature map is passed to the segmentation head to generate a segmentation mask that matches the input image resolution.

CNN-based segmentation deals with two main issues: the architecture and the encoder. We use architectures that are versatile with different backbones, and use the rule of thumb that a backbone with more parameters equates to a higher model capacity and, consequently, can handle more complex tasks. These two constraints effectively limit the number of architectures/backbones we should test.

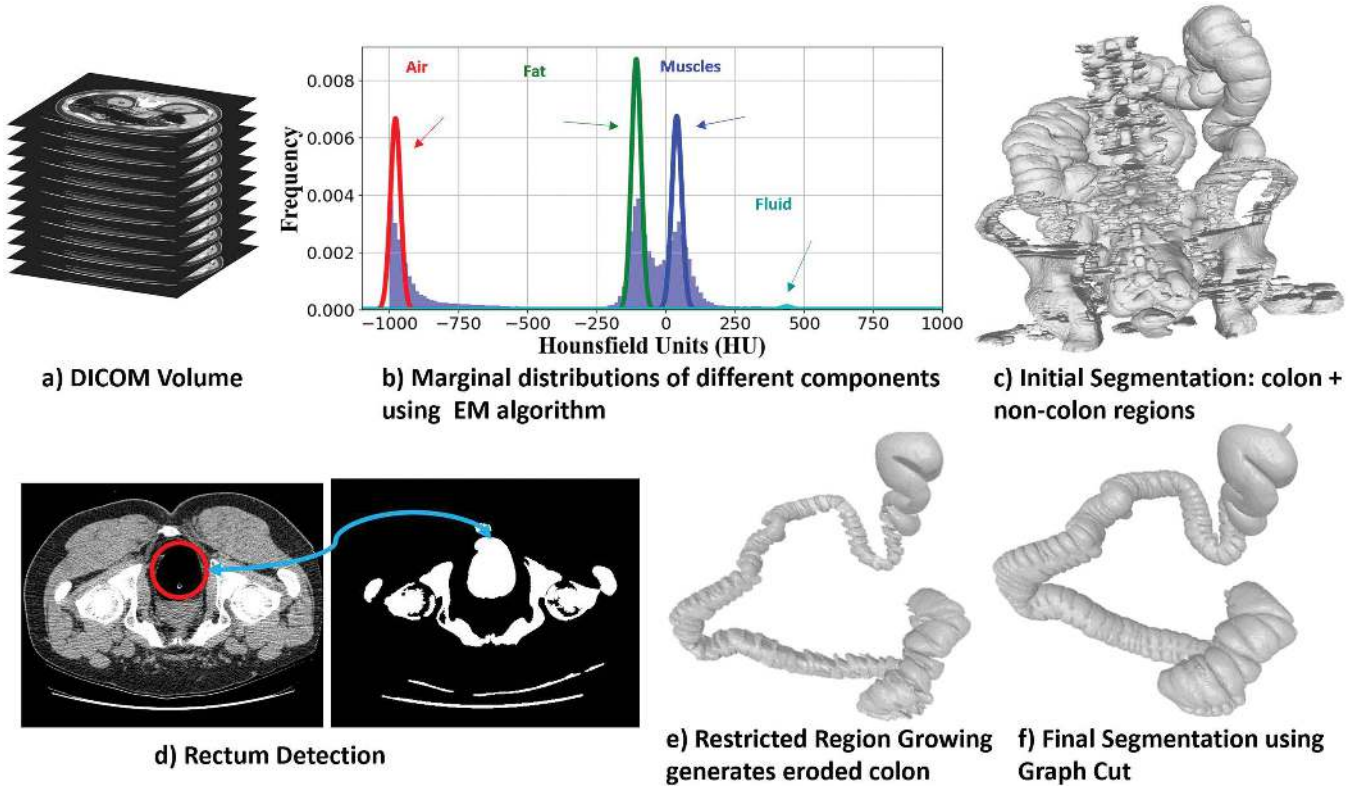


FIGURE 47.3 The classical approach for colon segmentation.

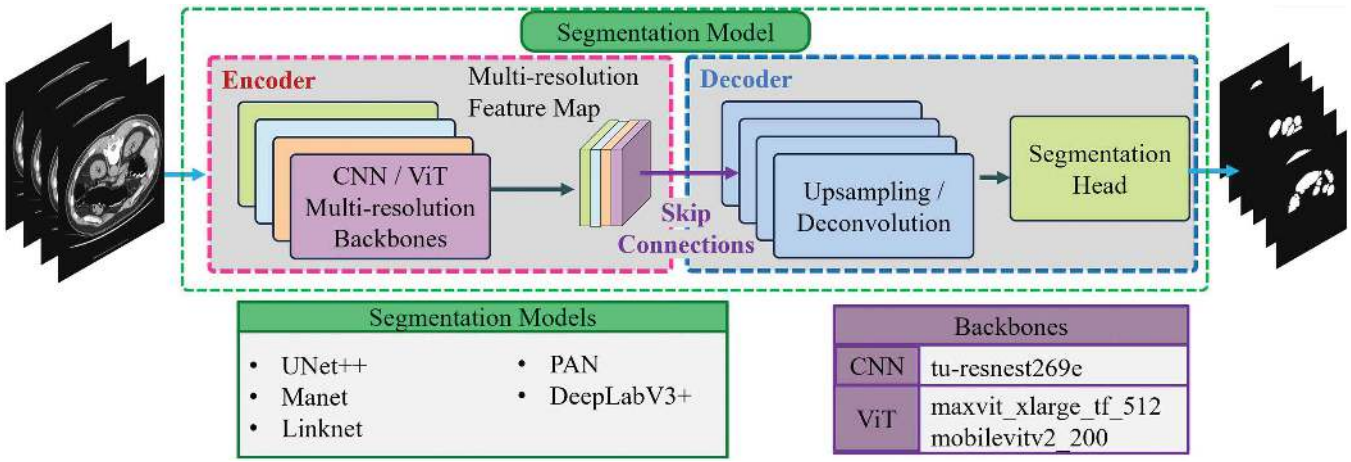


FIGURE 47.4 A meta-architecture consists of encoder and decoder. The encoder has a CNN-based backbone which generates multi-resolution feature maps that are passed to the decoder via skip connections. The decoders fuse these features through deconvolution blocks to generate a final feature map. The segmentation head uses this feature map to generate a predicted segmentation mask. During training, the custom dice uses the predicted and ground truth masks to update the network weights via back propagation.

While most segmentation problems treat the objects to be segmented as a whole, that is, all the pixels belonging to a certain class or object are treated as equally important, this should not be the case for colon segmentation. The reason is that pixels at the boundary of the colon are more important than pixels at its interior since almost all of a colon's interior is filled with air, which has a well-known Hounsfield value range, making it easy to segment by simple thresholding. Thus, to force the model to learn the intricate details of the colon boundary, we designed a per-pixel weighted dice loss that gives more weight to pixels near the boundaries of the colon (boundary weight). For each CT image pixel, a weight is assigned based on a weight map computed from each ground truth segmentation mask corresponding to this slice. The purpose of this modification is to give more importance to pixels near the boundary of the colon region in the loss calculation. This is often necessary in medical imaging tasks where precise boundary detection is crucial. The loss function is defined by:

$$DiL = 1 - \frac{2\sum_{i=1}^{N} p_i g_i}{\sum_{i=1}^{N} p_i^2 + \sum_{i=1}^{N} g_i^2}, \tag{47.3}$$

where $p_i \in [0,1]$ signifies the probability for the i -th pixel to be colon, and $g_i \in \{0,1\}$ represents the ground truth for the same pixel, and N represents the total number of pixels [31]. Further, $BL(x,b) = \frac{1}{N}\sum_{i=1}^{N} b_i \cdot x_i$ where represents the total number of pixels, b_i and x_i are the values of the boundary map and predicted segmentation mask at pixel i , respectively. The boundary map b is calculated by finding the minimum Euclidean distance from each pixel to any pixel on the boundary of the binary mask M . The distance transform assigns a higher value to pixels closer to the boundary, and a lower value to pixels in the interior of the binary mask. Where b_i is the value in the boundary map at position i , that is the value in the distance map at the same position:

$$DM(M)_i = \min_{r \in \text{Boundary}(M)} \| r - i \|. \tag{47.4}$$

47.4 Results

We conducted a series of experiments to rigorously evaluate the effectiveness of the colon segmentation approaches. Unfortunately, most of the previous colon segmentation work, for example [12, 17] evaluated their models on either a private dataset or a public dataset whose annotation is still private. Therefore, we will assess our model performance on our private dataset.

The goal of the first experiment is to evaluate our MRF-based segmentation approach. We compared the predicted segmentation mask x to the ground truth y using the intersection over the union metric $\frac{y \cap x}{y \cup x}$. The results show that the accuracy of the MRF-based approach is 90% for 70% of the scans, the remaining 30% of scans need to be manually refined by radiologists.

In the second experiment, we investigate which state of the art (SOTA) segmentation architecture and backbone are optimum for the CNN-based model. Among the encoder-decoder architectures proposed in the literature, UNet [19], UNet++ [20], Linknet [21], FPN [22], PSPNet [23], PAN [24], DeepLabV3 [25], and DeepLabV3+ [26] are considered the most successful ones. The brute force experimentation with all these models to pick the most accurate base model is not feasible. However, UNet++ and DeepLabV3+ are updated versions of UNet and DeepLabV3, respectively, and thus perform better. PSPNet and FPN perform poorly compared to all other models, PSPNet especially performs badly for detecting small objects and producing accurate pixel-level masks. So, we will not experiment with these specific two models. Taking this into account, we are only left with four models: UNet++, Linknet, PAN, and DeepLabV3+. There are hundreds of available backbones that can be used with these selected four segmentation models, and experimenting with all these variants is almost impossible. To tackle this problem, we will take the number of encoder parameters as the criteria for choosing backbones. This criterion is valid because as the number of parameters increases the model capacity increases, and this gives the model the ability to learn complex tasks such as colon segmentation. Conducting a series of experiments with different architectures and backbones, Table 47.1 shows that UNet++ with resnest269e backbone gives the best performance for the base model.

In the third experiment, we conducted a patient-independent seven-fold cross-validation for the base model with resnest269e encoder and UNet++ decoder. To enrich the training dataset, we incorporate deformable augmentations, which is an appropriate augmentation in case of soft organs like colon. The total size of the CTC-Dataset is 42,609 CT images. In each fold, data is divided into 42 patients/7 patients, on average 36,469/6140 images, for training/validation, respectively. Table 47.2 indicates the validation metrics for the base model compared to other models. We employed different well-known metrics [32] that quantify the pixel-wise agreement between the predicted and ground truth segmentation such as the dice coefficient, Jaccard coefficient/IOU, Hausdorff distance, sensitivity (SN), specificity (SP), and so on.

In addition, it is important to highlight that we conducted a comparative analysis by contrasting our proposed method with the nnUNet [33], which is the SOTA 3D-based method for medical image segmentation. Due to the computation burden, the 3D network was evaluated with low-resolution images. However, the 2D-based proposed model performs as

TABLE 47.1

Validation Dice and IOU Scores for Multiple Architectures with Different Backbones

Architecture/Backbone	Dice Score	IOU
PAN/resnest269e	97.1	95.46
DeepLabV3+/resnest269e	97.35	95.74
Linknet/resnest269e	97.36	95.96
UNet++/resnest269e	97.44	95.75

TABLE 47.2

Evaluation of Volume Segmentations for nnUNet, Base Model (UNet++/resnest269e), C-LSTM, and the Proposed Model Using the Metrics Reported in [32]

Model	DICE	JACRD	bAVD	KAPPA	ICCORR	MUTINF	PROB	SN	SP
nnUNet*	98.79	97.61	0.0133	98.73	98.79	25.78	5E-05	98.85	99.94
Base Model	97.44	95.75	0.3136	97.39	97.47	22.72	1E-04	98.01	99.89
C-LSTM	89.16	80.68	0.39	88.62	91.10	21.30	5E-04	96.38	99.08

* Due to computation burden, it was evaluated on half the resolution.

The dataset size for train/val: 29,430/7,786.

accurately as the 3D network without the need for down scaling the images and it has less computation complexity.

The primary challenge in traditional encoder-decoder networks lies in their inability to incorporate temporal information in a sequence of images such as colon CT scan. Therefore, we explore the fusion of C-LSTM with U-Net by replacing the convolutional layers in the encoder section with C-LSTM layers, as described in [34]. However, its performance is still lower than the base model, as shown in Table 47.2, because it has a larger number of parameters that should be optimized, and this hinders the network learning especially for long and high-resolution image sequences.

The last experiment explores the effect of accurate segmentation on preserving colorectal polyp information. We segment six scans, which have ten polyps, using different segmentation approaches. These scans are from the test set and are not used in the training process. For each segmented colon, we investigated the presence of the polyps. We then assigned a score if a polyp is not affected, 0 if a polyp is missed and 1/2 if a polyp is partially affected. Finally, sensitivity (i.e., percentage of polyps even partially exist from the total polyps) and accuracy (i.e., the scores average) are calculated as shown in Table 47.3. The results show that unlike classical approach, DL-based approaches can preserve the polyp's information even partially, that is, sensitivity = 100%, which helps radiologists in data annotation.

TABLE 47.3

Evaluation of Effectiveness of Colon Segmentation on Polyp Detection

Metric	MRF-based	Base Model
Sensitivity	90	100
Accuracy	55	75

47.5 Discussion

In this chapter, we described an MRF- and a DL-based architecture for colon segmentation. We described a meta-architecture that was trained with CTC-Dataset consisting of 98 prone and supine scans. Through rigorous experiments and evaluations, we demonstrated the superiority of the CNN-based architecture, achieving an impressive validation Dice score of 98.68%. Additionally, a seven-fold cross-validation reinforced the consistency and robustness of our approach across different patient groups. Evaluation of preserved polyp information in segmented colons shows promising results in colon segmentation, setting a strong foundation for future applications in clinical settings.

47.6 Future Trends

This chapter presented two quantitative algorithms for colon segmentation in CTC: an MRF-based approach and a CNN-based method. The MRF-based approach formulates the problem as MAP estimation, minimizing an energy function to achieve spatial smoothness and data consistency, and employs a region-growing algorithm to refine the segmentation and exclude non-colonic tissues. On the other hand, the CNN-based approach leverages an encoder-decoder architecture with multi-resolution feature maps and custom per-pixel weighted dice loss, emphasizing boundary pixels to improve segmentation accuracy, especially in regions critical to detecting polyps.

Improved accuracy of colon segmentation using CNN identified by deep learning suggests further development by integrating advanced techniques such as few-shot learning (FSL) to overcome the limitations of traditional CNNs, which require large, densely annotated datasets. Potential enhancements include incorporating 3D contextual information through sequential episodic training (SET), guided by a Markov random field–based model for colon anatomy. Pairing this approach with contrastive learning is expected to further improve segmentation accuracy and polyp detection. The MRF-based approach may be further examined with additional scans and pathological CTC cases with various types of polyps. Improving the accuracy of lumen segmentation will enhance the accuracy of the front end of CTC.

REFERENCES

1. R. L. Siegel, N. S. Wagle, A. Cercek, R. A. Smith and A. Jemal, "Colorectal cancer statistics, 2023," *CA: A Cancer Journal for Clinicians,* vol. 73, p. 233–254, 2023.
2. "The American Cancer Society," [Online]. Available: https://www.cancer.org/cancer/colon-rectal-cancer/detection-diagnosis-staging/acs-recommendations.html.
3. R. L. Siegel, A. N. Giaquinto and A. Jemal, "Cancer statistics, 2024," *CA: A Cancer Journal for Clinicians,* vol. 74, no. 1, pp. 12–49, 2024.
4. M. Jayasinghe, O. Prathiraja, D. Caldera, R. Jena, J. A. Coffie-Pierre, M. S. Silva, O. S. Siddiqui and J. A. Coffie-Pierre Jr, "Colon cancer screening methods: 2023 update," *Cureus,* vol. 15, 2023.

5. J. Sha, J. Chen, X. Lv, S. Liu, R. Chen and Z. Zhang, "Computed tomography colonography versus colonoscopy for detection of colorectal cancer: A diagnostic performance study," *BMC Medical Imaging,* vol. 20, p. 1–8, 2020.
6. S. Halligan, E. Dadswell, K. Wooldrage, J. Wardle, C. von Wagner, R. Lilford, G. L. Yao, S. Zhu and W. Atkin, "Computed tomographic colonography compared with colonoscopy or barium enema for diagnosis of colorectal cancer in older symptomatic patients: Two multicentre randomised trials with economic evaluation (the SIGGAR trials).," *Health Technology Assessment (Winchester, England),* vol. 19, p. 1–134, 2015.
7. D. J. Brenner and M. A. Georgsson, "Mass screening with CT colonography: Should the radiation exposure be of concern?," *Gastroenterology,* vol. 129, p. 328–337, 2005.
8. M. Yousuf, I. Alkabbany, A. Ali, S. Elshazley, A. Seow, G. Dryden and A. Farag, "An automatic colorectal polyps detection approach for Ct colonography," in *IEEE ICIP*, 2023.
9. M. Mohamad, A. Farag, A. M. Ali, S. Elshazly, A. A. Farag and M. Ghanoum, "Enhancing virtual colonoscopy with a new visualization measure," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, 2018.
10. D. Chen, R. Fahmi, A. A. Farag, R. L. Falk and G. W. Dryden, "Accurate and fast 3D colon segmentation in CT colonography," in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 2009.
11. Z. Ravindran, N. S. Das and others, "Automatic segmentation of colon using multilevel morphology and thesholding," in *2021 International Conference on Computer Communication and Informatics (ICCCI)*, 2021.
12. K. Gayathri Devi, R. Radhakrishnan and others, "Automatic segmentation of colon in 3D CT images and removal of opacified fluid using cascade feed forward neural network," *Computational and Mathematical Methods in Medicine,* vol. 2015, 2015.
13. J. J. Nappi, A. H. Dachman, P. MacEneaney and H. Yoshida, "Effect of knowledge-guided colon segmentation in automated detection of polyps in CT colonography," in *Medical Imaging 2002: Physiology and Function from Multidimensional Images*. Vol. 4683. SPIE, 2002.
14. D. K. Gayathri, R. Radhakrishnan and K. Rajamani, "Segmentation of colon and removal of opacified fluid for virtual colonoscopy," *Pattern Analysis and Applications,* vol. 21, p. 205–219, 2018.
15. G. S. K. D. Z. A. Malhotra and W. Enbeyle, "Deep neural networks for medical image segmentation," *Journal of Healthcare Engineering,* vol. 2022, 2022.
16. L. Guachi, R. Guachi, F. Bini and F. Marinozzi, "Automatic colorectal segmentation with convolutional neural network," *Computer-Aided Design and Applications,* vol. 16, p. 836–845, 2019.
17. A. Akilandeswari, D. Sungeetha, C. Joseph, K. Thaiyalnayaki, K. Baskaran, R. Jothi R., H. Al-Lohedan, D. M. Al-Dhayan, M. Karnan and K. Meansbo Hadish, "Automatic detection and segmentation of colorectal cancer with deep residual convolutional neural network," *Evidence-Based Complementary and Alternative Medicine,* vol. 2022, 2022.
18. J. Wasserthal, H. Breit, M. T. Meyer, M. Pradella, D. Hinck, A. W. Sauter, T. Heye, D. T. Boll, J. Cyriac, S. Yang, M. Bach and M. Segeroth, "Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images," *Radiology: AI,* vol. 5, 2023.
19. O. Ronneberger, P. Fischer and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, 2015.
20. Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Transactions on Medical Imaging,* vol. 39, p. 1856–1867, 2019.
21. A. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation," in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, p. 1–4, 2017.
22. T. Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan and S. Belongie, "Feature pyramid networks for object detection," in *IEEE CVPR*, 2017.
23. H. Zhao, J. Shi, X. Qi, X. Wang and J. Jia, "Pyramid scene parsing network," in *IEEE CVPR*, 2017.
24. H. Li, P. Xiong, J. An and L. Wang, "Pyramid attention network for semantic segmentation," *arXiv preprint arXiv:1805.10180,* 2018.
25. L. Chen, G. Papandreou, F. Schroff and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv preprint arXiv:1706.05587,* 2017.
26. L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
27. P. Wesp, S. Grosu, A. Graser, S. Maurus, C. Schulz, T. Knösel, M. P. Fabritius, B. Schachtner, B. M. Yeh, C. C. Cyran and others, "Deep learning in CT colonography: Differentiating premalignant from benign colorectal polyps," *European Radiology,* vol. 32, p. 4749–4759, 2022.
28. I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," *arXiv preprint arXiv:1711.05101,* 2017.
29. I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," *arXiv preprint arXiv:1608.03983,* 2016.
30. P. Iakubovskii, *Segmentation Models Pytorch,* GitHub, 2019.
31. R. Zhao, B. Qian, X. Zhang, Y. Li, R. Wei, Y. Liu and Y. Pan, "Rethinking dice loss for medical image segmentation," in *2020 IEEE International Conference on Data Mining (ICDM)*.
32. A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," *BMC Medical Imaging,* vol. 15, p. 1–28, 2015.
33. F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods,* vol. 18, p. 203–211, 2021.
34. A. Arbelle and T. R. Raviv, "Microscopy cell segmentation via convolutional LSTM networks," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, p. 1008–1012, 2019.

48

Emerging Blood-Based Tests for Colorectal Cancer Screening: Advances, Challenges, and Future Directions

Gloria D. Coronado, Jennifer L. Schneider, Cheryl A. Johnson, and Leslie Bienen

48.1 Introduction

Currently, adults can receive screening for colorectal cancer (CRC) using a variety of screening methods. Colonoscopy is considered the most invasive and is the most expensive and resource intensive, and stool testing (fecal immunochemical test [FIT] or FIT DNA testing) is the least invasive and lowest cost. Each type of test has advantages and disadvantages, and healthcare providers and patients must weigh tradeoffs when ordering and undergoing them. For example, colonoscopy requires an onerous bowel preparation, time off from work, and access to specialty care, can be prohibitively expensive, and can cause rare but serious complications (e.g., bowel perforation, bleeding). However, this method offers the benefit of immediate polyp removal if polyps are found and is considered the gold standard for detecting cancer and/or precancerous polyps. A recent pooled analysis from seven randomized clinical trials (663,319 participants having average colorectal cancer risk) reported a 20% reduction in CRC incidence (RR: 0.80, 95% CI: 0.77–0.83) and a 26% decrease in CRC mortality (RR: 0.74, 95% CI: 0.69–0.80) in the group receiving colonoscopy or sigmoidoscopy compared to standard care (no screening or screening with stool-based testing).[1]

FIT, on the other hand, can be done from home, has no serious risks, is low cost, and performs well at detecting blood or hemoglobin in stool and at finding cancers (the highest quality tests having 74–81% sensitivity and 93–94% specificity)[2]. However, sensitivity of FIT for detecting precancerous polyps is suboptimal (24% sensitivity for advanced adenomas for one-time testing), and a substantial number of people do not prefer stool testing, for various reasons. While sensitivity of FIT-DNA for detecting precancerous polyps is superior to FIT for one-time testing (42% sensitivity to detect advanced adenomas), the test is more costly than FIT. It is also notable that for test comparisons, one-time testing disadvantages FIT, which is recommended annually (versus FIT-DNA which is recommended every three years). Discomfort with handling fecal matter is a known deterrent to stool-testing options, and factors such as concern that stool tests are less effective than colonoscopies may contribute to some providers' preferences for more invasive methods of CRC screening. Additionally, adherence to CRC screening by any method is suboptimal, and rates of adherence are projected to fall even further as recent changes to guidelines expanded the age of eligibility and the aging population of the US pushes even more people into eligible ages for screening. In 2021, only 58% of eligible adults ages 45–75 in the United States were up-to-date with CRC screening recommendations, 22 percentage points below the target set by the National Colorectal Cancer Roundtable.[3, 4]

The aforementioned reasons, in addition to resource constraints around availability of specialty care for CRC screening, suggest that blood-based tests for CRC can offer multiple advantages that complement available CRC screening tests. Blood test availability and the fact that no specialty care is required to perform and analyze blood tests are enormous benefits in addressing low CRC screening rates due to resource constraints. Federally Qualified Health Center (FQHC) patients, for example, have lower rates of CRC screening completion (43% in 2022 data from the Uniform Data System[5]) than peers of similar ages in general populations, for multiple reasons. Scarcity of specialty care and difficulty of referring FQHC members to specialty clinics for colonoscopies is one factor depressing rates of follow-up for an abnormal noninvasive test result. Screening that starts with blood-based testing could lead to higher rates of CRC screening. Potential disadvantages of using blood-based screening tests for CRC include false negatives (leading to missed cancers or advanced adenomas), particularly in detecting adenomas, and higher cost than FIT or other stool tests.

Recently, Guardant Health (Redwood City, CA) released a next-generation sequencing assay to detect colorectal neoplasia-derived tumor signals in blood. The test, which has the brand name Shield, analyzes cell-free DNA in blood to detect genomic or epigenomic alterations associated with CRC. Shield is already being used as a screening test in South Korea, a country with high rates of CRC. Based on results from the ECLIPSE trial in the USA, which enrolled more than 20,000 patients, Guardant has applied for FDA approval for its blood test.[6, 7] Guardant reported that Shield "achieved 83% sensitivity for the detection of colorectal cancer in average-risk adults in the U.S., with a specificity of 90%." Guardant also reported that the test "demonstrated 13% sensitivity in detecting advanced adenomas." As far as we know, these results have not yet been validated in other studies. The test returns a binary result (normal or abnormal), and patients with an abnormal test result are advised to obtain a follow-up colonoscopy.

We (GC, JS) have conducted qualitative research to examine patient perspectives (n = 20; 15 with patients, five with providers) on the use of a blood test for CRC and found that the

DOI: 10.1201/9781003449942-54

majority of patients had positive feelings about their use and few hesitations.[8] In addition, interviewees noted positives such as convenience, lack of need for preparation, and that the procedure—a blood test—is something most patients are doing anyway at routine visits. However, literature on patients' and providers' perceptions of blood-based testing for CRC screening is scarce.

We also conducted a single health system randomized controlled trial to examine whether adherence to CRC screening recommendations in an organized FIT screening program could be improved by integrating a blood-based testing option opportunistically as part of standard health encounters.[9] We were interested in whether patients who had passively declined FIT testing (i.e., not returned routinely mailed FIT kits) would undergo a blood test to screen for CRC. In our study, about one-third of patients contacted by the research team who had not completed their FIT did complete a blood draw. This result was a marked improvement, by nearly 20 percentage points, over usual care in getting non-adherent but guideline-eligible patients to complete a CRC screening test.

48.2 Materials

Quantitative analyses were performed using SAS 9.4 (Cary, North Carolina).

48.3 Methods

48.3.1 Randomized Controlled Trial of Blood Test Adherence

48.3.1.1 Study Design and Participant Selection

This randomized controlled trial (ClinicalTrials.gov, NCT05937709) was conducted at Kaiser Permanente Northwest (KPNW), an integrated health system in Oregon and Southwest Washington. The trial was approved by the Institutional Review Board of KPNW. Participants were identified using electronic health record (EHR) data and included individuals aged 45–75 years who were eligible for CRC screening. Patients had received a FIT either in clinic or via mail but had not completed it in the prior three to nine months. Patients were excluded based on various criteria, including previous study opt-outs, interpreter needs, scheduled colonoscopies, cognitive impairment, severe mental illness, or lack of an email address in the EHR.

48.3.1.2 Randomization and Intervention

Participants were randomized (1:1) to receive an offer for a commercially available CRC blood test (Shield, by Guardant Health) or standard of care, which included reminders to complete FIT testing. Blood samples were processed in an accredited laboratory, and results were communicated to participants via an online portal and mailed letters. Those with abnormal results were referred for colonoscopy.

48.3.1.3 Study Outcomes and Statistical Analysis

The primary outcome was receipt of any CRC screening test within three months. Secondary outcomes included completion of follow-up colonoscopy among individuals with abnormal noninvasive test results. Logistic regression was used to compare screening completion between the blood draw and usual care conditions. Demographic characteristics were analyzed for associations with screening completion. Statistical analysis was performed using SAS 9.4, and results were considered significant if the two-sided P value was < 0.05. The study's sample size ($N = 2004$) provided sufficient power to detect meaningful differences in screening outcomes between intervention and usual care groups.

48.3.2 Patient and Provider Perceptions of Blood Test

48.3.2.1 Study Setting and Background

We recruited participants for qualitative interviews who were offered the blood-based CRC screening option during their scheduled appointments. Written informed consent was obtained from interested participants prior to the blood draw, with the possibility of follow-up interviews to gather insights about their screening experience.

48.3.2.2 Data Collection and Recruitment

Qualitative interviews were conducted with patients and providers using semi-structured interview guides developed by the research team. Patient interviews explored CRC screening history, motivations for accepting or declining the blood test, experiences with the screening process and results, and future actions. Provider interviews focused on awareness of blood-based screening options, reactions to the study's blood test, and recommendations for integrating blood-based testing into CRC screening protocols. Interviews were conducted by trained staff, audio-recorded, transcribed, and analyzed to explore barriers and facilitators of blood-based CRC screening acceptance.

48.3.2.3 Analysis

We used topical coding techniques to conduct a content analysis.[10–12] The qualitative team developed a code template based on interview guides and emergent themes, then extracted text relevant to coded topics and summarized it into reports, with key findings identified. The research team reviewed and refined these summaries to derive insights into patient and provider perspectives on blood-based CRC screening. The analysis process was iterative, involving regular team meetings for protocol consistency and transparency. The Consolidated Criteria for Reporting Qualitative Research (COREQ) was employed to aid in rigor and trustworthiness of interpretations.[13] The coding template and a sub-sample of illustrative quotes are provided in Table 48.2.

48.4 Results

48.4.1 Randomized Controlled Trial of Blood Test Adherence

In the blood-based intervention group, 548 out of 924 recruited patients (mean age = 60, 62% female, 80% non-Hispanic White) were successfully reached by phone, with 280 (51.1%) scheduling a blood-draw appointment and 206 (73.6%) attending as scheduled. Of those who attended, 204 (99%) consented to and completed the blood draw, yielding a high participation rate. Conversely, in the usual care group, 1013 initially eligible patients were reduced to 1003 following post-randomization exclusions. Table 48.1 gives the number of participants who adhered to screening in each study condition.

The intention-to-treat analysis showed that 30.5% of participants in the blood test group completed any CRC screening within three months, significantly higher than the 13% completion rate in the usual care group (Figure 48.1; OR: 2.94, 95% CI: 2.34, 3.70). Among those who completed the blood draw, 10.8% had abnormal test results and were referred for colonoscopy, with 50% completing it within six months. In comparison, 4.7% of those who completed FIT testing had abnormal results, with 70% completing colonoscopy within six months. Multivariable analysis showed that prior CRC screening significantly increased the odds of completing CRC screening, while age, sex, and race/ethnicity were not significant predictors of test completion. These findings underscore the effectiveness of the blood-based intervention in enhancing CRC screening rates within the study population.

TABLE 48.1

Colorectal Cancer Screening and Follow-Up Outcomes, by Study Condition within Three Months of Randomization

	Usual care	Intervention
Eligible participants	1003	1001
Recruited (intervention only)	–	924
Reached for recruitment phone call (% of n eligible)[a]	–	548 (59.3)
Scheduled for visit with research team (% of n reached)	–	280 (51.1)
Enrolled/consented for study (% of n scheduled)	–	204 (72.9)
Blood-based test completion	–	204[b]
FIT completion	115	99
Colonoscopy completion	15	12
Any colorectal cancer screening completion	130 (13.0)	305 (30.5)[b]

[a] 37 6 patients were not spoken with by phone.

[b] Includes 13 participants who completed the blood draw and whose samples were unable to be processed due to lab errors.

Citation: Coronado GD, Jenkins CL, Shuster E, and colleagues, Blood-based colorectal cancer screening in an integrated health system: a randomised trial of patient adherence. *Gut*. Mar 7 2024; 73(4):622–628. doi:10.1136/gutjnl-2023–330980

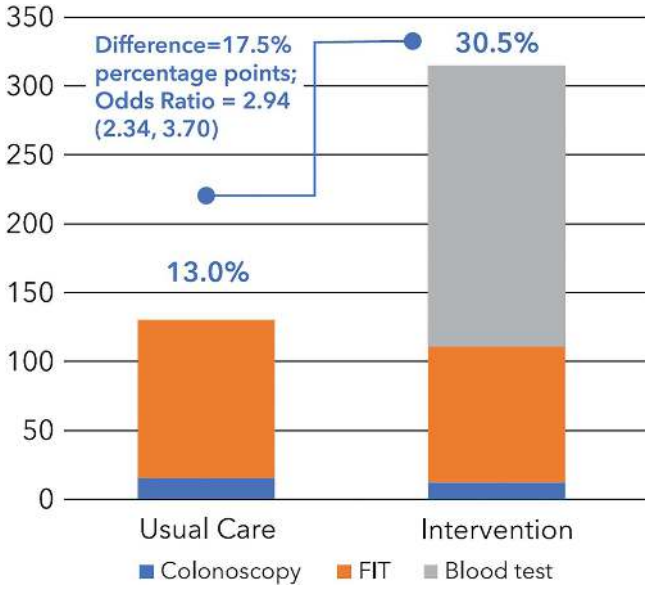


FIGURE 48.1 Patient adherence to blood-based testing for colorectal cancer, among adults 45–75 who were offered but did not complete a fecal test in the prior three to nine months, and had an upcoming in-person clinic visit (n = 2004; 1003 in usual care and 1001 in blood test offer).

48.4.2 Patient and Provider Perceptions of Blood Test

We conducted 20 interviews—15 with patients and five with healthcare providers. Among the patients, 11 completed the blood test, while four agreed to it but did not complete it in the study timeframe. The average age of those who completed the test was 58 years, were mostly female (73%), and had been members of the health system for about 21 years. In contrast, those who did not complete the test had an average age of 64 years, were mostly female (75%), and had been members for an average of 20 years. Provider interviews included three primary care providers, a gastroenterologist, and a medical assistant from the gastroenterology department.

Patient feedback showed that forgetfulness and procrastination were common reasons for not completing the FIT test in the past. Patients found the FIT process cumbersome and were often deterred by multiple steps involved, such as timing the collection with a bowel movement and mailing the sample. Patients perceived the blood test as simpler and more convenient, and they preferred being accountable to a scheduled blood draw rather than completing a FIT at home. The three-year screening interval was also appealing to patients, and most indicated the estimated out of pocket cost for the blood test (e.g., ~$25) acceptable. Logistical barriers, such as transportation issues or competing family priorities, were the primary reasons for not completing the blood test. Overall, patients generally expressed enthusiasm and relief at being offered the blood test over the FIT option and would prefer this option over FIT if given a choice in the future.

Providers expressed optimism about the blood test as an alternative to stool testing or colonoscopy, citing ease of implementation and patient compliance. Concerns centered on patient understanding of test results, accuracy in detecting

TABLE 48.2

Qualitative Interview Coding Template and Illustrative Quotes

Patient Experience and Feedback

Prior CRC screening and reasons for not completing a FIT in the past

"It seemed time consuming to do the [FIT] test. It was like I don't want to do that right now I'll do it later and I would just forget. And it was just not something that was simple—I am like I don't want to do that right now. Or the test would expire so I'd have to get another one. . . . So it wasn't like I was defiant. I have to be motivated sometimes to do something, so it wasn't motivating enough to do it." (completer, female, 10912)

"I had been delaying doing it [the FIT test] because of the gross factor." (completer, female, 10987)

Reasons and motivations to choose blood test option

"*I screamed hallelujah.* I mean I really did. That would be so awesome!" (non-completer, female, 10712)

"I just thought it'd be easier. I've had so many blood tests that I got very used to having to go up to [location], and so it was like nothing—I knew the parking arrangement and the best time to go and everything so it was very easy." (completer, male, 10531)

"I actually liked the idea because it was a specific time and place. I had to show up, I had to do the [blood] test. I couldn't just put it off forever, which is what I do with the FIT test. So that part of it was very positive for me." (completer, female, 10764)

"The blood draw would be a quick and easy way to get tested and I was eager to complete the screening when I had another appointment with my provider scheduled." (non-completer, male, 10794)

"I don't know how true it is, but I just felt like this way [blood test] they would be able to tell if there's anything in my system through the blood." (completer, female, 10518)

"I would rather have a blood test and do it real quick, so I don't have to do the icky parts [of FIT test]—the collecting of it, and taking the samples and putting it on the card and all that kind of stuff—you know and then mail it off." (completer, female 10943)

Questions and concerns

"The only question I had was how accurate it [blood test] was compared to the other test [FIT]. If I would've done the other test [FIT], would the results be the same?" (completer, female, 10912)

"What's the difference between the blood test and the FIT test, I just don't know that much about it." (non-completer, female, 10449)

Reasons for not completing blood test

"It was because of the winter storm, we were stuck . . . also I really wanted to be able to do it while I was going to get my blood draws anyway. And they said that wasn't possible. And that was my biggest reason for not doing it." (non-completer, female, 10991)

"I don't think I ever made it over there because it was on the same day as an appointment with my son. And that sort of overtook my day." (non-completer, male, 10449)

Reaction to results

I completely missed the big letter that said negative. And then I looked at it again and realized that that was up at the top of the page in very bold letters." (completer, female, 10764)

"I got [blood test results] via mail, all very clear . . . but the [FIT] still keeps popping up on my healthcare reminder that I need to do my little [FIT] test—so how to tie the two together?" (completer, female, 10954)

Future Action

"I would 100% do this blood test [if given a choice] . . . it was really convenient but even if it wasn't I would still make this choice even if I had to make a special trip to the lab to have blood taken." (completer, female, 10987).

"I feel like it was just as trustworthy as all the other tests that we take . . . I don't like the yearly FIT. I don't like how you have to get your stool . . . this way [blood test] it's clean. Fast. And I don't have to carry it around." (completer, female, 10518).

"I love that . . . I would much rather do it every 3 years if it was just a blood test, than every year." (completer, female, 10943)

"It's almost like my contact lenses at $500. . . . I would still pay for it [blood test]. I would figure out a way to do it. That is how important it is to me." (completer, female, 10518)

Provider Reaction and Feedback

Awareness and general reaction

"I've definitely heard of these tests and the potential availability for them . . . I think it will be a greater acceptance for doing blood-based tests over having to do FIT tests." (PCP, male, 105)

Advantages of blood test

"So it's just like all these steps [in doing a FIT]. . . . What I really like about this blood test is just the practical nature, once the patient buys in and says, 'yeah I'll do it', you just go to the lab and you get it done. " (PCP, female, 104)

"It's very easy as a clinician to say, 'okay we need to just check your routine labs' and people are very willing to do that at the time. . . . I have a suspicion that if a blood-based test becomes available it will be more popular and a higher percentage of people will be interested in doing it over the stool check." (PCP, male, 105)

"Definitely, stool is a very inconvenient thing for patients to do it. It's just not going to work. Blood test is something that would get a lot of compliance." (GI, male, 101)

Disadvantages of blood test

"I would say in primary care if I had to guess, that for 10–15% of patients cost, even if it seemed pretty negligible, was an issue." (PCP, female, 103)

"The only thing that would probably bring up more unintended consequences would be additional screening based on getting a positive [blood] test and having a negative colonoscopy. Because then somebody would think logically that 'well gee if I have some kind of DNA saying that there's a tumor in me, where is the tumor?' So that may entail additional scans or things like that." (PCP, male, 105)

(Continued)

TABLE 48.2 (*Continued*)

Qualitative Interview Coding Template and Illustrative Quotes

Patient Experience and Feedback
"I think we need to be able to have a blood test that [can] detect a fair amount of advanced adenomas and that we can provide our patients and society the confidence that we're truly catching it before it becomes a cancer" (male, 101)
"It [FIT] looks like it's anywhere between 60–80% for sensitivity so it looks like this [blood test] is a better test for sensitivity. That's actually reassuring." (PCP, female, 104)
Experience communicating results
"By and large we got most patients scheduled . . . once we kind of communicated their reason as to why." (MA Scheduler, female, 102)
"And then [lead GI provider] created a blurb of what the expectation was: 'If negative then no further follow up. If positive a screening colonoscopy is recommended' . . . I added [documentation]], so I feel like it was pretty well laid out for all the providers to know that this is being sent as a result of the positive Guardant test." (MA Scheduler, female, 102)
"I would say over 50% of patients pushed back on the seriousness or the need for it [follow up colonoscopy]. I think it's common. I don't think better or less [than abnormal FIT], I think it's about average." (MA Scheduler, female, 102)
"One of the biggest concerns is patients just don't seem to have understood the concern [if blood test abnormal]." (GI, male, 101)
Future actions
"If there was an order, someone wouldn't even necessarily have to make an extra trip into the lab." (PCP, female, 103)
"I think it would be a game changer for our CRC screening [rates] and HEDIS measures." (MA Scheduler, female, 102)
Patient and Provider Advice
"I think anybody would benefit from a conversation about accuracy. If it could be combined with other blood draws, that would be ideal. And having it available at all the locations." (completer, female, 10764)
"It would be nice to have from a primary care provider education standpoint and [for the] patient, a visual [showing] this is the efficacy of FIT testing for advanced adenoma versus colon cancer. This is the blood-test, and this is a screening colonoscopy." (PCP, female, 103)

precancerous polyps, and possible cost to patients with more limited financial means. Providers highlighted the need for educational resources and EHR tools to support both provider and patient education and tracking of CRC screening choice and completion. Overall, both patients and providers viewed the blood test as a promising option that could improve CRC screening rates due to its convenience and perceived effectiveness. Providers emphasized the importance of widespread education and seamless integration into existing healthcare workflows to maximize its impact.

48.5 Discussion

In recent years, many new technologies for detecting CRC and advanced adenomas have become available. This includes blood-based tests as well as enhanced stool-based tests (FIT-DNA testing and FIT-RNA testing). As these technologies come on the market, we will need research to assess how they can best be used to achieve the goal of detecting early forms of CRC or precancerous polyps, thereby reducing incidence and mortality from CRC. All screening programs using noninvasive CRC screening tests will also need to focus on ensuring that patients who have abnormal test results have access to and obtain follow-up colonoscopies.

Despite the enthusiasm among patients and providers for blood-based CRC testing options, it will be important to balance this enthusiasm with accurate information about the performance and relative merits of this option. Low sensitivity for the detection of advanced adenomas could mean that blood-based testing may not be able to prevent CRC (which can be accomplished with the removal of adenomas on colonoscopy), which accounts for 74% of the mortality reduction.[14] Patients and providers may opt for blood test options given the greater convenience. Patients who obtain blood testing may experience a false sense of security, believing, as noted in the qualitative interviews, that the blood test is more accurate than other available tests for CRC screening, which is not the case.

Future research is needed that compares the performance, uptake, cost, and cost-effectiveness of available CRC screening modalities. For example, abundant research has reported higher uptake of a single offer of FIT compared to colonoscopy and that this relative high uptake can lead to nearly equivalent detection of CRC compared to the offer of colonoscopy.[15–19] While high uptake of blood-based screening is anticipated, recent findings from microsimulation modeling show that FIT was more effective and less costly than blood-based screening even when blood-based screening uptake was 20 percentage points higher than for FIT.[20] Future research might corroborate findings from microsimulation models and empirical data.

A recent publication by Liang and colleagues demonstrated increases in CRC screening when patients in a Veterans' Administration health system were re-offered a FIT or colonoscopy, then offered a blood-based test (i.e., Septin).[21] The study (n = 359 participants) further showed that offering the blood-based test did not divert patients from other established forms of screening. These findings showed an overall improvement in CRC screening by 7.5 percentage points. Our study showed an even higher overall improvement in CRC screening rates (19.4 percentage points). Moreover, qualitative interviews we conducted showed that providers had a preference for offering new forms of screening to all eligible patients, and patients believed that the blood-testing technology was more accurate than stool-based testing. These perceptions raise the possibility that patients may complete blood-based testing, even though less invasive, higher performing, and less costly tests are available.

Future research might also compare the rates of participation in follow-up colonoscopy among individuals with abnormal non-colonoscopic screening test results. Our data show that only 11/22 (50%) of patients who obtained the blood test and had an abnormal result obtained a follow-up colonoscopy within six months. This compared to 7/10 (70%) among patients who completed a FIT.[9] However, the numbers were too small (n = 32 with abnormal test result) to make a statistical comparison. While previous research shows a higher follow-up colonoscopy completion rate for FIT-DNA versus FIT, a statistic that is often attributable to the belief on the part of patients or providers that an abnormal FIT result, but not a FIT-DNA result, can be explained by other GI-related conditions (i.e., hemorrhoids). These findings are consistent with national data showing that rates of follow-up colonoscopy are only 56%, and rates in FQHCs range from 18% to 57%.[22, 23]

Concerted efforts may be needed to spur follow-up colonoscopy completion so that proportions rise to the level recommended by the US Multi-Society Task Force (80%). This is critical because of the increased risk for CRC, for late-stage CRC detection, and for mortality from CRC for those who delay follow-up colonoscopy by three months or longer. Moreover, the benefits of noninvasive testing are nullified when no follow-up colonoscopy is obtained. As more costly initial screening tests enter the market, more healthcare waste can be expected unless efforts are undertaken to ensure that timely follow-up colonoscopy is obtained. Developing a reportable quality measure and aligning rates with healthcare incentives could be useful.[24] Patient navigation has been effective in some settings; however, there is limited evidence for its effectiveness for follow-up colonoscopy, specifically.[25] Efforts to expand the workforce to increase the number of providers who can perform colonoscopy deserve further exploration. Primary care providers may be able to communicate the importance of follow-up colonoscopy when they recommend initial screening. In addition, models that place a colonoscopy provider in primary care facilities can facilitate the transition from primary to specialty care for some patients.

48.6 Future Trends

Advancing population health demands a dual focus: advancing research on cancer detection and diagnosis, while also expanding access to existing tools within communities. The introduction of blood-based options holds potential to transform CRC screening practices in the United States. These options offer distinct advantages over traditional methods. However, careful attention is required to ensure that lower-performing, high-cost options do not replace more effective, cost-efficient options already available on the market.

A significant improvement in CRC screening adherence was observed among patients who opted for blood-based testing instead of FIT within organized screening programs. This shift represents a 17.5 percentage point increase over usual care. These findings underscore the potential for blood-based testing to enhance CRC screening participation. Nonetheless, it is crucial to explore the tradeoffs between improved adherence and potential reductions in accuracy for detecting advanced adenomas and subsequent health outcomes. Further research is needed to evaluate the overall impact and effectiveness of blood-based CRC screening in clinical settings.

REFERENCES

1. Han C, Wu F, Xu J. Effectiveness of Sigmoidoscopy or Colonoscopy Screening on Colorectal Cancer Incidence and Mortality: A Systematic Review and Meta-analysis of Randomized Controlled Trial. *Front Oncol*. 2024;14:1364923. doi:10.3389/fonc.2024.1364923
2. Lin JS, Perdue LA, Henrikson NB, Bean SI, Blasi PR. Screening for Colorectal Cancer: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA*. 2021;325(19):1978–1998. doi: 10.1001/jama.2021.4417.PMID: 34003220
3. National Colorectal Cancer Roundtable. *Reigniting Colorectal Cancer Screening As Communities Face and Respond To The COVID-19 Pandemic: A Playbook*. Accessed Sep 9, 2022. https://nccrt.org/resource/a-playbook-for-reigniting-colorectal-cancer-screening-as-communities-respond-to-the-covid-19-pandemic/
4. American Cancer Society. *Colorectal Cancer Facts & Figures 2023–2025*. 2023. https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/colorectal-cancer-facts-and-figures/colorectal-cancer-facts-and-figures-2023.pdf
5. Health Resources & Services Administration (HRSA). *2022 National Health Center Data*. Accessed Sep 23, 2023. https://bphc.hrsa.gov/uds/datacenter.aspx
6. Chung DC, Gray DM, 2nd, Greenson JK, et al. *Clinical Validation of a Cell-free DNA Blood-Based Test for Colorectal Cancer Screening in an Average Risk Population*. Presented at: Digestive Diseases Week Annual Meeting; 2023; Chicago, IL.
7. Chung DC, Gray DM, 2nd, Singh H, et al. A Cell-Free DNA Blood-Based Test for Colorectal Cancer Screening. *N Engl J Med*. Mar 14 2024;390(11):973–983. doi:10.1056/NEJMoa2304714
8. Schneider JLJCJCCGD. "I was Screaming Hallelujah": Patient and Provider Perceptions of Blood-Based Testing for Colorectal Cancer Screening. *PLoS One*. Nov 2023. doi:10.1371/journal.pone.0295685
9. Coronado GD, Jenkins CL, Shuster E, et al. Blood-Based Colorectal Cancer Screening in an Integrated Health System: A Randomised Trial of Patient Adherence. *Gut*. Mar 7 2024;73(4):622–628. doi:10.1136/gutjnl-2023-330980
10. Patton MQ. *Qualitative Research and Evaluation Methods*. Third Edition. Sage Publications, Inc.; 2002.
11. Köhler T, Smith A, Bhakoo V. Templates in Qualitative Research Methods: Origins, Limitations, and New Directions. *Organ Res Methods*. Apr 1 2022, 2021;25(2):183–210. doi:10.1177/10944281211060710
12. Morse JM. Critical Analysis of Strategies for Determining Rigor in Qualitative Inquiry. *Qual Health Res*. Sep 2015;25(9):1212–1222. doi:10.1177/1049732315588501
13. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. 2007;19(6):349–357.

14. Knudsen AB, Trentham-Dietz A, Kim JJ, et al. Estimated US Cancer Deaths Prevented With Increased Use of Lung, Colorectal, Breast, and Cervical Cancer Screening. *JAMA Netw Open*. Nov 1 2023;6(11):e2344698. doi:10.1001/jamanetworkopen.2023.44698
15. Gupta S, Halm EA, Rockey DC, et al. Comparative Effectiveness of Fecal Immunochemical Test Outreach, Colonoscopy Outreach, and Usual Care for Boosting Colorectal Cancer Screening Among the Underserved: A Randomized Clinical Trial. *JAMA Intern Med*. Oct 14 2013;173(18):1725–1732. doi:10.1001/jamainternmed.2013.9294
16. Inadomi JM, Vijan S, Janz NK, et al. Adherence to Colorectal Cancer Screening: A Randomized Clinical Trial of Competing Strategies. Randomized Controlled Trial Research Support, N.I.H., Extramural. *Archives of Internal Medicine*. Apr 9 2012;172(7):575–582. doi:10.1001/archinternmed.2012.332
17. Singal AG, Gupta S, Tiro JA, et al. Outreach Invitations for FIT and Colonoscopy Improve Colorectal Cancer Screening Rates: A Randomized Controlled Trial in a Safety-Net Health System. *Cancer*. Feb 1 2016;122(3):456–463. doi:10.1002/cncr.29770
18. Quintero E, Castells A, Bujanda L, et al. Colonoscopy versus Fecal Immunochemical Testing in Colorectal-Cancer Screening. *N Engl J Med*. Feb 23 2012;366(8):697–706. doi:10.1056/NEJMoa1108895
19. Pilonis ND, Bugajski M, Wieszczy P, et al. Participation in Competing Strategies for Colorectal Cancer Screening: A Randomized Health Services Study (PICCOLINO Study). *Gastroenterology*. Mar 2021;160(4):1097–1105. doi:10.1053/j.gastro.2020.11.049
20. van den Puttelaar R, Nascimento de Lima P, Knudsen AB, et al. Effectiveness and Cost-Effectiveness of Colorectal Cancer Screening With a Blood Test That Meets the Centers for Medicare & Medicaid Services Coverage Decision. *Gastroenterology*. Mar 18 2024. doi:10.1053/j.gastro.2024.02.012
21. Liang PS, Zaman A, Kaminsky A, et al. Blood Test Increases Colorectal Cancer Screening in Persons Who Declined Colonoscopy and Fecal Immunochemical Test: A Randomized Controlled Trial. *Clin Gastroenterol Hepatol*. Apr 8 2023. doi:10.1016/j.cgh.2023.03.036
22. Bharti B, May FFP, Nodora J, et al. Diagnostic Colonoscopy Completion After Abnormal Fecal Immunochemical Testing and Quality of Tests Used at 8 Federally Qualified Health Centers in Southern California: Opportunities for Improving Screening Outcomes. *Cancer*. Sep 3 2019. doi:10.1002/cncr.32440
23. Mohl JT, Ciemins EL, Miller-Wilson LA, Gillen A, Luo R, Colangelo F. Rates of Follow-up Colonoscopy After a Positive Stool-Based Screening Test Result for Colorectal Cancer Among Health Care Organizations in the US, 2017–2020. *JAMA Netw Open*. Jan 3 2023;6(1):e2251384. doi:10.1001/jamanetworkopen.2022.51384
24. Ciemins EL, Mohl JT, Moreno CA, Colangelo F, Smith RA, Barton M. Development of a Follow-Up Measure to Ensure Complete Screening for Colorectal Cancer. *JAMA Netw Open*. Mar 4 2024;7(3):e242693. doi:10.1001/jamanetworkopen.2024.2693
25. Selby K, Baumgartner C, Levin TR, et al. Interventions to Improve Follow-Up of Positive Results on Fecal Blood Tests: A Systematic Review. *Annals of Internal Medicine*. Oct 17 2017;167(8):565–575. doi:10.7326/M17-1361

49

Microfluidic Isolation and Molecular Characterization of Circulating Tumor Cells in Prostate Cancer

Rea Pittie, Victor Putaturo, Shyamala Maheswaran, Daniel A. Haber, Mehmet Toner, Avanish Mishra, and David T. Miyamoto

49.1 Introduction

Prostate cancer is the second most common cause of cancer-related death among men in the United States [1]. Although metastatic prostate cancer initially responds to androgen deprivation therapy (ADT), the emergence of castration-resistant prostate cancer (CRPC) is common, necessitating treatment with next-line therapies, including androgen receptor pathway inhibitors, taxane chemotherapy, radionuclides, and poly(ADP-ribose) inhibitors [2]. However, treatment responses are heterogeneous, and biomarkers to guide appropriate therapy selection and monitor treatment are needed. Genomic analysis of single tumor biopsy specimens may not be representative of intra-tumor heterogeneity [3]. Additionally, the most common site for prostate cancer metastasis is the bone, which is difficult to biopsy [4]. These challenges have led to efforts to develop minimally invasive liquid biopsies based on peripheral blood analyses [5].

Circulating tumor cells (CTCs) are cancer cells that disseminate from the primary tumor and intravasate into the bloodstream [6]. CTC burden is correlated with patient survival, and molecular analysis of CTCs can provide information useful for predicting treatment outcomes [7–9]. However, viable CTCs in circulation are extremely rare, posing a challenge to their study. Commonly used platforms for CTC isolation include the FDA-cleared Veridex CellSearch system [10], which uses ferrofluid-based immunomagnetic capture of cells expressing epithelial cell adhesion molecule (EpCAM) and, therefore, does not detect cells with low expression due to cell state changes such as epithelial-mesenchymal transition (EMT). To overcome this limitation, we developed the microfluidic CTC-iChip and ^{LP}CTC-iChip platforms [11], which enable label-free CTC enrichment through negative selection regardless of the expression of specific cell surface epitopes. In this chapter, we summarize microfluidic technologies and biological assays developed by our group for the molecular profiling of CTCs from prostate cancer patients.

49.1.1 CTC-iChip

The CTC-iChip (Figure 49.1A) is a microfluidic system that performs high-throughput depletion of normal blood cells from whole blood for enrichment of CTCs [11, 12]. Isolation of CTCs is achieved by labeling whole blood with streptavidin-coated magnetic beads functionalized with biotinylated antibodies that bind to white blood cells (WBCs). The fluidics of the CTC-iChip are controlled by a single pressure source at the blood and buffer inlets (Figure 49.1B), which can process approximately 20 mL of whole blood per hour [12]. The symmetrically parallelized architecture of the chip incorporates two distinct microfluidic modules. First, size-based sorting of nucleated cells such as red blood cells (RBCs) and platelets is achieved through non-equilibrium inertial separation array (NISA) devices, an inertial microfluidic technology consisting of arrays of islands (200 μm long × 50 μm wide × 52 μm tall) that exert wall-induced lift forces on flowing particles to create different migration paths for cells dependent on their size. WBCs and CTCs experience a bigger deflection from the wall than RBCs and platelets due to their larger size. The array is deterministically designed so that each island siphons a small portion of the flow (3.6%) nearest to the wall, where RBCs and platelets reside, while larger mononucleated cells remain in their original path [13–15]. Each island siphons a fixed portion of the flow until all the blood is redirected to the expanding waste row. After all the blood is siphoned, the rows in the array shift up such that islands are reintroduced into the waste row, and the stream of mononucleated cells is shifted one row closer to the product stream at the bottom of the microchannel. Ultimately, the array undergoes multiple shifts until WBCs and possible CTCs are found on the bottom row, while RBCs and platelets are kept in the topmost row and exit into the product and waste outlet of the module, respectively.

The second negative selection module consists of two magnetically activated cell sorting stages (MACS) working in series to remove labeled WBCs [12]. Cells are aligned by inertial focusing channels that use the balance of a shear gradient lift force, a wall interaction force, and a drag force caused by a secondary flow, also known as Dean flow [12]. The aligned cells then enter a 1 mm wide channel to undergo the first stage of magnetic sorting. A quadrupolar arrangement of rectangular neodymium–iron–boron magnets (N52 grade) creates a mean magnetic flux gradient of 200 T/m, deflecting cells labeled with a relatively high number of beads to the center of the channel towards the waste [12]. The remaining cells enter a second set of inertial focusing channels, followed by a second stage of magnetic sorting. Here, a higher magnetic gradient

DOI: 10.1201/9781003449942-55

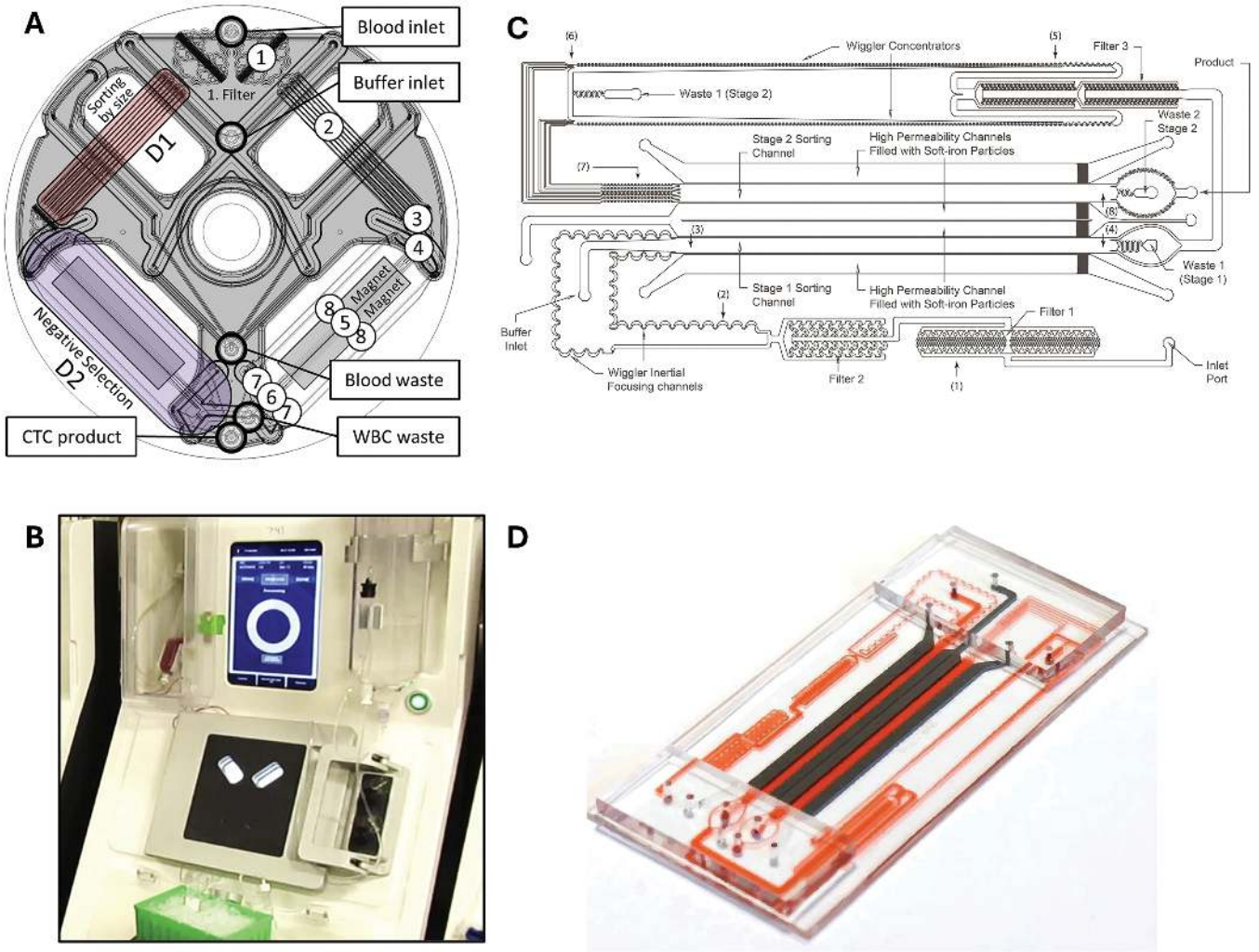


FIGURE 49.1 (A) Microfluidic CTC-iChip system, (B) isolator system used with CTC-iChip for whole blood processing, (C–D) A schematic diagram illustrating high-throughput magnetic sorter used in the ^{LP}CTC-iChip system and the microfabricated chip for processing leukapheresis products. This device exploits magnetic lenses to achieve 50-fold higher cell sorting throughput than the CTC-iChip system. (Figures reproduced from Edd et al. [11] and Mishra et al. [16].)

(425 T/m) results in the deflection of WBCs labeled with even a single bead into the waste while not disturbing any unlabeled CTCs traveling to the product outlet. Overall, the CTC-iChip enables the efficient removal of normal blood cells such as RBCs (99.95%), platelets (99.9%), and WBCs (99.98%), while also achieving a high yield of CTCs in the product [12].

49.1.2 LPCTC-iChip

Although the isolation of CTCs using the CTC-iChip enables their molecular characterization, the rarity of CTCs (medial 0 to 10 CTCs per 10 mL blood) has hindered the development of routine clinical applications [11]. Leukapheresis is a routine clinical procedure optimized for the isolation of rare cell populations in the blood [17]. This method draws blood from the antecubital vein in one arm, subjecting it to continuous centrifugal enrichment of mononuclear cells into a "leukopak" of approximately 100 mLs, and returning RBCs, platelets, and plasma back to the patient via the contralateral antecubital vein [17–19]. Given the similar density between tumor cells and mononuclear leukocytes (1050–1080 kg/m^3), standard leukapheresis parameters allow for the enrichment of CTCs [16]. A leukapheresis procedure can interrogate 3 L of blood per hour, and the total number of WBCs, CTCs, and platelets in the enriched leukopak is 50- to 100-fold higher than in whole blood, thus presenting an attractive approach to isolating large numbers of CTCs suitable for a variety of clinical applications [16]. Nevertheless, microfluidic technologies such as CellSearch are capable of processing only 5% of a leukopak [18]. Given the tremendous increase in volume and cell concentration in these samples, existing microfluidic technologies suffer from clogging and reduced throughput [20]. The ^{LP}CTC-iChip is an ultra-high throughput microfluidic platform developed to address these challenges and sort CTCs from leukopak products (Figure 49.1C and D) [16].

The ^{LP}CTC-iChip consists of a debulking chip and a magnetic lens-based high-throughput cell sorter arranged in series for the processing of leukopak products through negative selection [16]. The entire leukopak sample incubated with biotinylated antibodies against WBCs flows through the initial debulking chip, comprised of NISA devices to remove excess antibodies, plasma, RBCs, and platelets. The product is then incubated with 1 μm streptavidin-conjugated magnetic beads for labeling of WBCs, which at this stage are 50- to 100-fold more in number than in 10 mL blood tubes. The ^{LP}CTC-iChip addresses the challenge of sorting these cells through a unique magnetic setup. Two 1600 mm wide channels on each side of the sorting channel are filled with soft iron particles, a high magnetic permeability material. These ferromagnetic microchannels are magnetized by the macro magnetic field exerted by the quadrupolar arrangement of rectangular magnets around the magnetic sorter, creating a localized magnetic field that is maximal along the sidewalls of the sorting channel. The magnetic field decays rapidly towards the center of the sorting channel, resulting in a high magnetic field gradient in this direction as high as 15,400 T/m near the sidewalls. This arrangement yields a ~30-fold enhancement of the magnetic force and allows for effective magnetic sorting of the 3 to 6 billion WBCs in a leukopak [16]. The overall design of the magnetic sorter is divided into two stages. After the initial flow of the sample through 40 μm filters, two asymmetric serpentine channels use a balance of shear-induced lift force and Dean-flow-based drag force to focus cells near the channel walls, where the magnetic field gradient is the highest (Figure 49.1C) [16]. WBCs labeled with a higher number of beads (> 10) are deflected to the center of the channel towards the stage 1 waste. The remaining cells flow into stage 2 after traversing another 40 μm filter and two cell concentrators that aid with the positioning of cells near the sidewalls for a maximal deflection into the waste. At the stage 2 sorting channel, every labeled cell is deflected into the waste, while thousands of CTCs flow into a product that is concentrated 11-fold [16].

49.2 Materials

49.2.1 Cancer Cell Lines and Human Blood Samples

Various cell lines were used for optimization of microfluidic CTC isolation technologies and molecular assays, including LNCaP, VCaP, 22Rv1, and DU 145, obtained from American Type Culture Collection. Blood from healthy volunteers was obtained under an Institutional Review Board (IRB)-approved protocol or from Research Blood Components. For clinical studies, blood was obtained from prostate cancer patients under various IRB-approved clinical protocols.

49.2.2 Leukopaks from Apheresis

Leukapheresis was performed on consenting patients under an IRB-approved clinical protocol (2020P000251). Leukopaks were obtained by sampling the entire blood volume using the continuous mononuclear cell collection mode in the Spectra Optia System over a period of two hours.

49.2.3 Buffer

1% F68 Pluronic (Kolliphor P 188) was dissolved in phosphate buffered saline (PBS without $MgCl_2$ and $CaCl_2$) and filter-sterilized.

49.2.4 Antibodies

WBC depletion antibody cocktail consisted of biotinylated antibodies against CD45 (Thermo Fisher Scientific, clone HI30, IgG1, 0.25 μg/million cells), CD16 (BD Biosciences, clone 3G8, IgG1, 0.05 μg/million cells), CD66b (Novus Biologicals, clone 80H3, IgG1, 0.025 μg/million cells) in cold, sterile filtered 1X PBS and 0.1% BSA. Additional biotinylated antibodies against CD45RA (BioLegend, clone HI100, IgG2b, 0.5 μg/million cells) and CD3 (BioLegend, clone SK7, IgG1, 0.5 μg/million cells) were added for mass blood volume isolation.

49.2.5 Magnetic Beads

Dynabeads MyOne Streptavidin T1 superparamagnetic beads (1 μm; Invitrogen) were washed thrice with 0.01% TWEEN 20 in 1X PBS and 0.1% BSA in 1X PBS prior to use.

49.2.6 Chip Design and Microfabrication

The CTC-iChip was made of medical-grade cyclic olefin copolymer (COC) using variotherm injection molding by Stratec Biomedical AG (Germany). Fluidic resistances were precisely sculpted at the product and waste outlets of the channel to achieve the desired fluid fraction to be siphoned within the separation arrays. 16-array modules were included on one disk. The ^{LP}CTC-iChip magnetic sorter was fabricated using PDMS soft lithography technique. Microfeatures were achieved through photolithography by exposing the SU-8 layer to 365 nm UV radiation through a chrome mask, followed by treatment with a Baker BTS-220 SU-8 developer. The main PDMS layer was 1 mm thick, and 3 mm thick PDMS pads were bonded via plasma and used as fluidic inlet and outlet ports for press-fit tubing connections. After baking the device, soft magnetic iron powder (CAS#: 7439-89-6) was injected into sorting channels. The magnetic sorter was enclosed in a 6061-aluminum alloy manifold holding six permanent magnets in two parallel quadrupolar configurations.

49.3 Methods

49.3.1 Whole Blood Processing with CTC-iChip

Whole blood was incubated with antibody cocktail for 30 minutes. Washed beads were added to the blood and incubated for 20 minutes to magnetically label the WBCs. The blood was diluted 1X with buffer and processed through the CTC-iChip. The device uses a buffer and a pressure of 48.0 PSI to push blood through the microfluidic channels of the chip. RBCs and WBCs were deflected into waste bags, and the product was stored on ice.

49.3.2 Leukopak Processing with ^{LP}CTC-iChip

Leukopaks were aliquoted into 50 mL tubes and incubated with 5 mM ethylenediaminetetraacetic acid for 10 minutes at room temperature. WBC depletion antibody cocktail was added and incubated for 30 minutes at room temperature. The sample was processed through the debulking chip, after which DNAse I (100 U/mL) was added for digestion of neutrophil extracellular traps (NETs) that can induce clogging. The sample was incubated with 1 μm superparamagnetic beads at a concentration of 15 beads per WBC for 20 minutes at room temperature, and then processed through two MAGLENS sorters in parallel to obtain a final product enriched in CTCs.

49.3.3 CTC Immunofluorescence Imaging

CTC product was fixed for 10 minutes with 0.5% formaldehyde and centrifuged onto a microscope slide at 2000 rpm for 5 minutes using EZ Megafunnels and Shandon Cytospin 4 centrifuge (Thermo Fisher Scientific). Immunofluorescence staining was done with one of several fluorophore-labeled antibody cocktails, including antibodies against EpCAM, CK8, CK18, and CK19 to stain epithelial tumor cells, CD45, CD66b, and CD16 to stain WBCs, and DAPI to stain nuclei. Other antibodies against prostate cancer markers such as PSA and PSMA were used in separate channels as needed. Cells were imaged with a multispectral fluorescence microscope (Akoya Biosciences) at 20x or 40x magnification using appropriate filter sets. Analysis was performed with Indica Labs Halo digital image analysis software.

49.3.4 Droplet Digital PCR of Bulk CTC Mixture

Following RNA extraction of CTC pellets and complementary DNA (cDNA) synthesis, primer/probe mixes and Bio-Rad ddPCR Super-mix for Probes were added to cDNA. Droplets were generated using Bio-Rad's automated droplet generator. Each droplet underwent a 45-cycle PCR with 70 °C step-down in between the denaturation and annealing steps in the thermocycler. Droplets containing the transcript of interest were detected with the Bio-Rad QX 200 droplet reader system. For detailed methods, refer to Miyamoto and colleagues [21].

49.3.5 Single-Cell RNA-Sequencing and DNA Methylation Analyses

Following microfluidic enrichment of whole blood, CTCs were screened by microscopic visualization of cells stained with Alexa Fluor 488-conjugated antibodies against EpCAM, CDH11, and/or PSMA, and single cells were transferred into PCR tubes containing RNA protective lysis buffer using the Eppendorf TransferMan NK 2 micromanipulator. RNA was amplified and sequencing libraries were prepared for RNA-seq via the ABI SOLiD or Illumina platform [22]. Alternatively, paired single-cell DNA methylation and RNA-seq analysis was performed from the same single cells lysed in DNA/RNA lysis buffer [23]. The nucleus was magnetically separated from cytoplasm using Magnetic MyOne Carboxylic Acid Beads (Invitrogen Cat#65011). The supernatant (containing cytoplasmic RNA) was used for single-cell RNA-seq using the SMART-seq2 protocol, while the pellet (beads and intact nucleus) was resuspended in DNA methylation lysis buffer and used for whole genome bisulfite sequencing on a HiSeq X sequencer (Illumina). For nanopore sequencing, CTCs were subjected to high molecular weight DNA extraction (QIAGEN) and libraries were prepared using rapid barcoding kits (Nanopore#SQK-RBK004). Sequencing was on a Nanopore MinION device (Oxford Nanopore).

49.4 Results

49.4.1 CTC Isolation Yield

The CTC-iChip was rigorously tested for performance using 11 different cell lines spiked into up to 20 mL healthy donor blood, demonstrating a median recovery yield of spiked cells of 99.5%, with a median carryover of 445 WBCs per mL input blood [12]. The yield was consistent across a range of input cell concentrations (19 to 5000 cells/mL) and cell lines, including mesenchymal cells such as MDA-MB-231 [12]. The ^{LP}CTC-iChip was tested with three clinical leukapheresis samples (average volume 64.2 mL) spiked with 5000 cultured CTCs, demonstrating an 86.1% CTC recovery yield and removal of 99.97% of WBCs, > 99.99% of RBCs, and > 99.99% of platelets [16] (Figure 49.2). The excellent spiked cell yields and

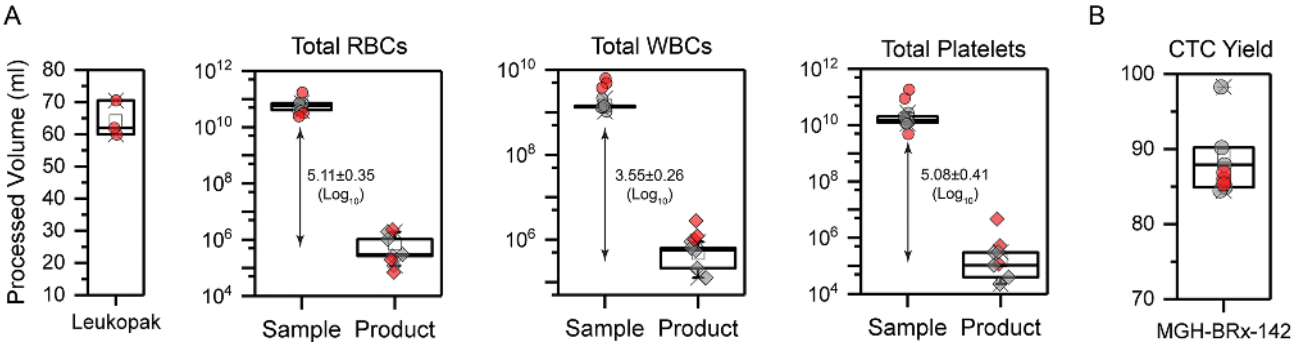


FIGURE 49.2 Performance of [LP]CTC-iChip. (A) Leukopak volumes and quantification of depletion of cellular components of blood using the [LP]CTC-iChip. (B) CTC isolation yield after processing of spiked cell samples through the [LP]CTC-iChip. (Figure reproduced from Mishra et al. [16].)

normal blood cell depletion performance of the CTC-iChip and [LP]CTC-iChip laid the groundwork for their subsequent testing in processing prostate cancer patient blood samples. In a recent pre-print manuscript [24], six cancer patient-derived leukapheresis products were processed, interrogating a mean blood volume of 5.83 liters. The authors demonstrated a median yield of 2799 CTCs per patient, which is 100-fold higher than typical 10 mL blood samples. Remarkably, a prostate cancer patient yielded 58,125 CTCs, including single cells and clusters of cells. Such a high number of CTCs can reveal phenotypic and molecular heterogeneity using multi-omics approaches.

49.4.2 Immunofluorescence Imaging of CTCs

Microscopy imaging of CTCs enriched from patient blood and immunostained with cancer-specific markers allowed discrimination from WBCs, and therefore enumeration of CTCs [25]. A greater number of CTCs correlated with higher prostate cancer stage and tumor burden, and a decrease in CTC numbers after ADT and other therapies correlated with therapeutic response [25]. Beyond enumeration, immunofluorescence to measure expression of specific proteins in CTCs provided insights into the biological state of CTCs, such as Ki-67 to indicate proliferation [25] and PSA/PSMA ratio to indicate androgen receptor (AR) signaling activity [26].

49.4.3 RNA-Based Digital CTC Signatures of Prostate Cancer

Immunofluorescence imaging is limited due to impurities in CTC populations and difficulties in scoring CTC images. To overcome this limitation, we established a high-throughput, cancer-specific RNA-based molecular signature for the sensitive detection of prostate CTCs using a droplet digital PCR (ddPCR) assay following microfluidic enrichment [21]. We identified a panel of genes found to have high prostate cancer expression but minimal expression in normal blood cells (*KLK3*, *KLK2*, *TMPRSS2*, *AGR2*, *FOLH1*, *HOXB13*, *FAT1*, *STEAP2*, *AR-V7*) and established a digital CTC score (CTC_M Score) for the sensitive RNA-based detection of CTCs without the need for microscopy [21]. We recently applied this ddPCR assay to analyze CTCs from patients with metastatic CRPC treated with the radionuclide Radium-223 [27] (Figure 49.3A). CTC RNA signatures (particularly *AR-V7*) were highly predictive of prognosis (Figure 49.3B). Thus RNA-based digital quantification of CTC signatures can predict treatment outcomes in prostate cancer.

49.4.4 Single-Cell RNA-Sequencing of Prostate CTCs

Although RNA profiling of CTC mixtures is useful for CTC detection and bulk expression analysis, it ignores the heterogeneity of individual tumor cells. We demonstrated that microfluidic isolation of CTCs yields viable cells with intact RNA that can be sequenced at the single-cell level [22, 23]. We performed individual cell micromanipulation and single-cell RNA-seq on 77 CTCs isolated from 13 prostate cancer patients using the CTC-iChip [22]. Ninety-two percent of CTCs expressed epithelial markers, 60% expressed stem cell markers, and 47% expressed cell proliferation markers. Analysis of AR mRNA splice variants showed heterogenous patterns among individual CTCs, with 43% expressing at least one AR variant and 18% with simultaneous expression of more than one AR variant. Unbiased analysis of activated pathways in CTCs revealed significant enrichment for non-canonical Wnt signaling in enzalutamide-resistant compared to enzalutamide-sensitive CTCs. Using leukapheresis products, scaled single-cell analytic technologies (paired single-cell RNA-seq and DNA-CNV scoring) can now be deployed within individual patients. For instance, >50,000 CTCs were isolated from a single patient with metastatic prostate cancer, with matched single-cell CNV demonstrating clonal patterns of aneuploidy and RNA analysis identifying two unsuspected divergent cancer cell subpopulations, one driven by FGFR signaling and one with inflammation-associated signatures [24]. Altogether, single-cell RNA-seq revealed heterogeneous signaling pathways in CTCs that may result in differential responses to therapy.

49.4.5 DNA Methylation Analysis in Prostate CTCs

DNA methylation profiling can provide complementary information to RNA-seq in characterizing biological states within

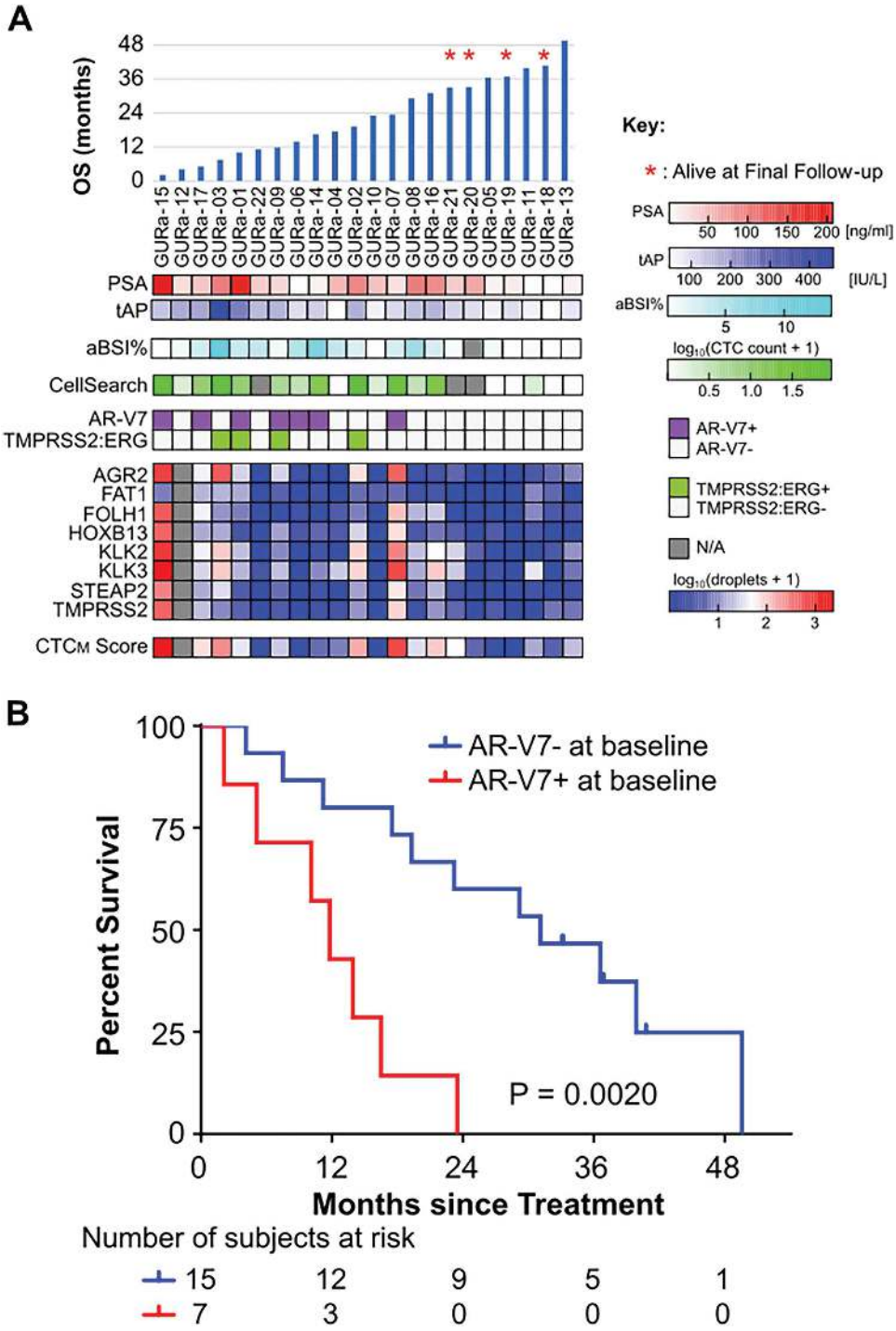


FIGURE 49.3 CTC biomarkers of Ra-223 response in mCRPC. (A) Bar graph shows overall survival (OS) in mCRPC patients after Ra-223. Heatmap shows level of signal from various assays in each patient, including CTC expression of genes measured by ddPCR. (B) Kaplan–Meier estimates of OS by baseline *AR-V7* status. (Figure reproduced from Saylor et al. [27].)

cells. We demonstrated the feasibility of simultaneous RNA-seq and whole genome DNA methylation sequencing from the same single CTCs [23] (Figure 49.4A). Using this method, we mapped 40 core partially methylated domains (PMD) that are uniform across different stages of prostate cancer including CTCs [23] (Figure 49.4B). These PMDs contained silenced genes enriched for immune genes, prominently including a cluster of genes involved in anti-tumor immune surveillance including *CD1* and *IFI16* genes. We inferred that DNA hypomethylation takes place during early prostate cancer tumorigenesis, given the presence of these PMDs even in low grade (Gleason score 6) cancers. Based on these results, we tested a new diagnostic strategy using long-read nanopore sequencing of core DNA hypomethylation domains as a molecular surrogate

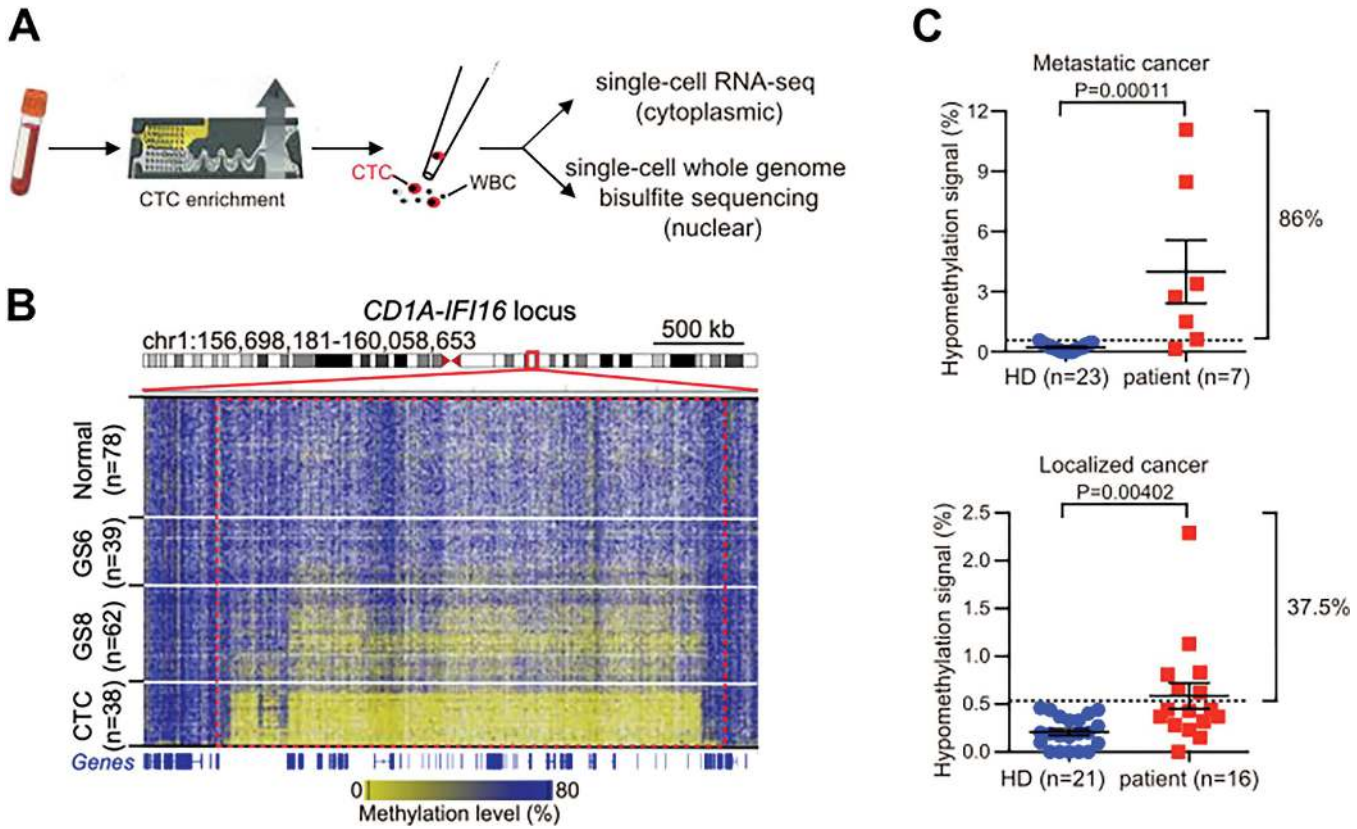


FIGURE 49.4 (A) Schematic of CTC microfluidic enrichment and paired single-cell RNA-seq and DNA methylation sequencing from individual CTCs. (B) IGV representation of single-cell DNA methylation data showing progressive demethylation at *CD1A-IFI16* locus from normal prostate, to Gleason score 6–8 prostate cancer, to CTCs. (C) Scatter plot showing DNA hypomethylation signal by nanopore sequencing in healthy donor (HD) vs. metastatic or localized prostate cancer blood samples. (Figure reproduced from Guo et al. [23].)

for CTCs and showed that this method can detect CTCs even in patients with early stage prostate cancer [23] (Figure 49.4C). Hence, DNA methylation analysis of CTCs provided insights into early tumorigenesis, leading to the development of a new potential assay for early detection of prostate cancer.

49.5 Discussion

The detection and analysis of rare CTCs in the blood has many potential clinical applications in prostate cancer, ranging from early detection, treatment monitoring, and detection of minimal residual disease to therapy selection in advanced disease [9]. The CTC-iChip microfluidic technology allows for the isolation of untagged, viable CTCs suitable for sophisticated molecular analyses, including immunofluorescence imaging, RNA expression analysis with ddPCR, and single-cell next-generation sequencing, which can be leveraged for the development of novel clinical assays. The ^{LP}CTC-iChip further increases the potential yield to many thousands of CTCs from leukopaks derived from sampling several liters of blood [16], potentially enabling clinical applications previously restricted by CTC numbers such as early detection of cancer. Both the CTC-iChip and ^{LP}CTC-iChip overcome limitations of prior CTC isolation technologies based on positive selection for specific cell surface epitopes such as EpCAM, which results in biased enrichment for epithelial populations [11]. Cells undergoing EMT or neuroendocrine transdifferentiation may not be captured by a positive selection approach but are effectively isolated using the CTC-iChip [11]. In addition, although other CTC isolation methods allow for enumeration of CTCs, complex molecular analyses are often limited due to the absence of viable cells. While enumeration is a useful as a prognostic biomarker, molecular analysis using the CTC-iChip unlocks the potential of CTCs as predictive biomarkers [28].

Despite their many strengths, the CTC-iChip and ^{LP}CTC-iChip platforms have several limitations. These technologies are not optimized for the capture of CTC clusters, which are associated with a 50 to100-fold increase in metastasis compared to single CTCs [29]. Clusters may split apart in circulation, such that 7.5 mL blood is estimated to contain just one to two clusters. These rare clusters may be further lost during microfluidic isolation when squeezing through the inertial separation array microposts, or by magnetophoresis if associated with magnetically labeled WBCs.

Our studies in prostate cancer using microfluidic CTC isolation technologies have yielded a suite of blood-based assays with a wide range of potential clinical uses. In the setting of metastatic prostate cancer, we have developed promising noninvasive CTC biomarkers that predict responses to AR pathway inhibitors and radionuclide therapies and warrant further validation in clinical trials [21, 27]. However, novel therapies continue to emerge from the pipeline each year, offering new treatment options while increasing the complexity of treatment selection. Recent therapeutic developments in prostate cancer include the FDA approval of Lutetium-177-PSMA-617 [30], and the ongoing evaluation of antibody-drug conjugates (ADC) that target various antigens expressed by prostate cancer cells, including PSMA, TROP-2, STEAP-1, DLL-3, B7-H3, and HER2 [31]. There is an urgent need for continued development of noninvasive biomarkers to precisely guide the selection of appropriate therapy for each individual patient. CTC molecular analyses may be ideally suited for this purpose, especially given the ability to measure and detect intercellular heterogeneity in CTC antigen expression in real time.

The blood-based early detection of prostate and other cancers is arguably an even greater technical challenge, although with tremendous potential upside in enabling early intervention when cancer is still localized and curable [32]. Our attempts to isolate CTCs from patients with localized prostate cancer using the CTC-iChip have yielded detectable CTCs in a minority of patients [21, 25], even when using advanced nanopore DNA methylation sequencing-based technology [23]. A potential solution would be to shift from assaying a tube of blood towards using the ^{LP}CTC-iChip to process leukopak samples for early cancer detection applications. Leukopaks can generate sufficient numbers of intact cancer cells to enable a genuine "cell-based liquid biopsy" with multi-analyte analyses applicable to a diverse array of cancers. This strategy would effectively enable the sampling of many liters of patient blood to detect the very rare CTCs shed by localized prostate cancers. Although there are logistical challenges associated with incorporating leukapheresis into cancer screening, it is a safe and routinely performed clinical procedure, and this strategy warrants further testing and consideration.

49.6 Future Trends

Future directions could include incorporating the use of a recently developed size-based cluster chip as the first step in the microfluidic CTC isolation process [33]. In addition, the enriched CTCs are often contaminated by WBCs and platelets, which can hinder immunofluorescence and next-generation sequencing analyses. Further purification steps using fluorescence-activated cell sorting or micromanipulation methods can achieve 100% purity and single cells, enabling single-cell sequencing. Another major area of development is the multi-omics analysis of purified CTCs. Efforts are underway to simplify paired-RNA and DNA sequencing of single CTCs, along with whole exome sequencing, in a cost-effective manner. Finally, the current ^{LP}CTC-iChip is comprised of fluidically unconnected devices [16]. Efforts are underway to increase system efficiency by creating a single integrated monolithic chip.

In summary, the CTC-iChip and ^{LP}CTC-iChip are highly effective platforms for the gentle and efficient microfluidic isolation of CTCs from the blood of prostate cancer patients. The negative selection strategy enables the enrichment of label-free CTCs that are viable and amenable to a wide range of detailed molecular characterization assays, including next-generation sequencing. These blood-based assays have the potential to serve as predictive biomarkers to guide the use of new therapies based on real-time molecular profiling of CTCs, although such assays require further validation prior to clinical use. A challenging but promising future direction is the development of ultra-sensitive CTC-based early detection assays that rely on the high-throughput processing of leukapheresis products that sample the entire blood volume of patients.

REFERENCES

1. R. L. Siegel, A. N. Giaquinto and A. Jemal, "Cancer statistics, 2024," *CA Cancer J Clin,* vol. 74, no. 1, pp. 12–49, Jan–Feb 2024, doi: 10.3322/caac.21820.
2. R. J. Rebello *et al*., "Prostate cancer," *Nat Rev Dis Primers,* vol. 7, no. 1, p. 9, Feb 4 2021, doi: 10.1038/s41572-020-00243-0.
3. M. C. Haffner *et al*., "Genomic and phenotypic heterogeneity in prostate cancer," *Nat Rev Urol,* vol. 18, no. 2, pp. 79–92, Feb 2021, doi: 10.1038/s41585-020-00400-w.
4. R. Mehra *et al*., "Characterization of bone metastases from rapid autopsies of prostate cancer patients," *Clin Cancer Res,* vol. 17, no. 12, pp. 3924–3932, Jun 15 2011, doi: 10.1158/1078-0432.CCR-10-3120.
5. K. Pantel and C. Alix-Panabieres, "Real-time liquid biopsy in cancer patients: Fact or fiction?," (in eng), *Cancer Res,* vol. 73, no. 21, pp. 6384–6388, Nov 1 2013, doi: 0008-5472.CAN-13-2030 [pii] 10.1158/0008-5472.CAN-13-2030.
6. K. Pantel and R. H. Brakenhoff, "Dissecting the metastatic cascade," (in eng), *Nat Rev Cancer,* vol. 4, no. 6, pp. 448–456, Jun 2004, doi: 10.1038/nrc1370 nrc1370 [pii].
7. D. C. Danila *et al*., "Circulating tumor cell number and prognosis in progressive castration-resistant prostate cancer," (in eng), *Clin Cancer Res,* vol. 13, no. 23, pp. 7053–7058, Dec 1 2007, doi: 13/23/7053 [pii] 10.1158/1078-0432.CCR-07-1506.
8. A. J. Armstrong *et al*., "Prospective multicenter study of circulating tumor cell AR-V7 and taxane versus hormonal treatment outcomes in metastatic castration-resistant prostate cancer," *JCO Precis Oncol,* vol. 4, 2020, doi: 10.1200/PO.20.00200.
9. W. L. Hwang, K. L. Hwang and D. T. Miyamoto, "The promise of circulating tumor cells for precision cancer therapy," *Biomark Med,* vol. 10, no. 12, pp. 1269–1285, Dec 2016, doi: 10.2217/bmm-2016-0192.
10. FDA Clearance Document for Veridex LLC. CellSearch(TM) Circulating Tumor Cell Kit. *Premarket notification—expanded indications for use—metastatic prostate cancer*. Feb 26 2008. Accessed June 4, 2024. [Online]. Available: https://www.accessdata.fda.gov/cdrh_docs/pdf7/k073338.pdf.
11. J. F. Edd *et al*., "Isolation of circulating tumor cells," *iScience,* vol. 25, no. 8, p. 104696, Aug 19 2022, doi: 10.1016/j.isci.2022.104696.
12. F. Fachin *et al*., "Monolithic chip for high-throughput blood cell depletion to sort rare circulating tumor cells," *Sci Rep,* vol. 7, no. 1, p. 10936, Sep 7 2017, doi: 10.1038/s41598-017-11119-x.
13. B. R. Mutlu *et al*., "Non-equilibrium inertial separation array for high-throughput, large-volume blood fractionation," *Sci Rep,* vol. 7, no. 1, p. 9915, Aug 30 2017, doi: 10.1038/s41598-017-10295-0.
14. D. Di Carlo, "Inertial microfluidics," (in eng), *Lab Chip,* vol. 9, no. 21, pp. 3038–3046, Nov 07 2009, doi: 10.1039/b912547g.
15. D. Di Carlo, J. F. Edd, K. J. Humphry, H. A. Stone and M. Toner, "Particle segregation and dynamics in confined flows," (in eng), *Phys Rev Lett,* vol. 102, no. 9, p. 094503, March 06 2009, doi: 10.1103/PhysRevLett.102.094503.
16. A. Mishra *et al*., "Ultrahigh-throughput magnetic sorting of large blood volumes for epitope-agnostic isolation of circulating tumor cells," *Proc Natl Acad Sci U S A,* vol. 117, no. 29, pp. 16839–16847, Jul 21 2020, doi: 10.1073/pnas.2006388117.
17. J. C. Fischer *et al*., "Diagnostic leukapheresis enables reliable detection of circulating tumor cells of nonmetastatic cancer patients," *Proc Natl Acad Sci U S A,* vol. 110, no. 41, pp. 16580–16585, Oct 8 2013, doi: 10.1073/pnas.1313594110.
18. M. B. Lambros *et al*., "Single-cell analyses of prostate cancer liquid biopsies acquired by apheresis," *Clin Cancer Res,* vol. 24, no. 22, pp. 5635–5644, Nov 15 2018, doi: 10.1158/1078-0432.CCR-18-0862.
19. K. C. Andree *et al*., "Toward a real liquid biopsy in metastatic breast and prostate cancer: Diagnostic LeukApheresis increases CTC yields in a European prospective multicenter study (CTCTrap)," (in eng), *Int J Cancer,* vol. 143, no. 10, pp. 2584–2591, Nov 15 2018, doi: 10.1002/ijc.31752.
20. K. C. Andree *et al*., "Toward a real liquid biopsy in metastatic breast and prostate cancer: Diagnostic LeukApheresis increases CTC yields in a European prospective multicenter study (CTCTrap)," *Int J Cancer,* vol. 143, no. 10, pp. 2584–2591, Nov 15 2018, doi: 10.1002/ijc.31752.
21. D. T. Miyamoto *et al*., "An RNA-based digital circulating tumor cell signature is predictive of drug response and early dissemination in prostate cancer," *Cancer Discov,* vol. 8, no. 3, pp. 288–303, Mar 2018, doi: 10.1158/2159-8290.CD-16-1406.
22. D. T. Miyamoto *et al*., "RNA-seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance," *Science,* vol. 349, no. 6254, pp. 1351–1356, Sep 18 2015, doi: 10.1126/science.aab0917.
23. H. Guo *et al*., "DNA hypomethylation silences anti-tumor immune genes in early prostate cancer and CTCs," *Cell,* vol. 186, no. 13, pp. 2765–2782.e28, Jun 22 2023, doi: 10.1016/j.cell.2023.05.028.
24. A. Mishra *et al*., "Tumor cell-based liquid biopsy using high-throughput microfluidic enrichment of entire leukapheresis product," *Nat Commun,* vol. 16, no. 32, 2025, doi: 10.1038/s41467-024-55140-x.
25. S. L. Stott *et al*., "Isolation and characterization of circulating tumor cells from patients with localized and metastatic prostate cancer," (in eng), *Sci Transl Med,* vol. 2, no. 25, p. 25ra23, Mar 31 2010, doi: 2/25/25ra23 [pii] 10.1126/scitranslmed.3000403.
26. D. T. Miyamoto *et al*., "Androgen receptor signaling in circulating tumor cells as a marker of hormonally responsive prostate cancer," (in eng), *Cancer Discov,* vol. 2, no. 11, pp. 995–1003, Nov 2012, doi: 2159-8290.CD-12-0222 [pii] 10.1158/2159-8290.CD-12-0222.
27. P. J. Saylor *et al*., "Circulating and imaging biomarkers of Radium-223 response in metastatic castration-resistant prostate cancer," *JCO Precis Oncol,* vol. 8, p. e2300230, Feb 2024, doi: 10.1200/PO.23.00230.
28. W. L. Hwang, H. M. Pleskow and D. T. Miyamoto, "Molecular analysis of circulating tumors cells: Biomarkers beyond enumeration," *Adv Drug Deliv Rev,* vol. 125, pp. 122–131, Feb 1 2018, doi: 10.1016/j.addr.2018.01.003.
29. N. Aceto *et al*., "Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis," *Cell,* vol. 158, no. 5, pp. 1110–1122, Aug 28 2014, doi: 10.1016/j.cell.2014.07.013.
30. O. Sartor *et al*., "Lutetium-177-PSMA-617 for metastatic castration-resistant prostate cancer," *N Engl J Med,* vol. 385, no. 12, pp. 1091–1103, Sep 16 2021, doi: 10.1056/NEJMoa2107322.

31. M. Sardinha, A. F. Palma Dos Reis, J. V. Barreira, M. Fontes Sousa, S. Pacey and R. Luz, "Antibody-drug conjugates in prostate cancer: A systematic review," *Cureus,* vol. 15, no. 2, p. e34490, Feb 2023, doi: 10.7759/cureus.34490.
32. D. S. Micalizzi, L. V. Sequist and D. A. Haber, "Deploying blood-based cancer screening," *Science,* vol. 383, no. 6681, pp. 368–370, Jan 26 2024, doi: 10.1126/science.adk1213.
33. J. F. Edd *et al.*, "Microfluidic concentration and separation of circulating tumor cell clusters from large blood volumes," *Lab Chip,* vol. 20, no. 3, pp. 558–567, Feb 7 2020, doi: 10.1039/c9lc01122f.

50

Chemical Imaging for Next-Generation Histopathology

Kianoush Falahkheirkhah, Kevin Yeh, and Rohit Bhargava

50.1 Introduction

Clinical and research activities related to cancer are often anchored in traditional histopathology, where a variety of stains are employed to assess morphology in processed biopsies or surgical specimens. Typically, thin sections of stained tissue are analyzed under a microscope by skilled pathologists to examine the tissue morphology. The advent of machine learning (ML), especially deep learning, has already begun to enhance this conventional diagnostic process through numerous digital pathology workflows.[1] However, many of these digital processes rely on the information content of traditional pathology. Recent advancements have introduced the concept of "digital chemical pathology,"[2] an approach brought about by the fusion of digital pathology with chemical imaging (CI) techniques, notably infrared (IR) imaging. Instead of simply recording morphology, IR imaging also records the global chemical composition of tissue at every pixel. Sensitive to the major classes of biomolecules[3]—proteins, lipids, carbohydrates, and nucleic acids—IR imaging may also permit more detailed molecular analyses of important biomolecules native to tissue, for example of collagen.[4–14] This integration can potentially enable a more detailed analysis of patient samples with new contrast, improving diagnostic accuracy and patient care. Innovations in IR imaging[15–29] not only aim to equip pathologists with cutting-edge tools, but also promise more time-efficient and less costly examinations of samples by obviating the need for reagents and eliminating steps in tissue processing. Most notably, IR CI seeks to be completely nonperturbative by using only low-intensity light. By maintaining the samples' original condition, this approach alleviates the need for traditional staining methods, thereby streamlining the diagnostic process. With IR imaging rapid advancements in recent years, this study focuses on highlighting its development and application to automate and refine the pathology process.

CI techniques based on IR imaging have garnered considerable recent attention due to the development of several types of new instruments that are specifically suited for pathology[30–33] with different contrast and sensing mechanisms, pixel sizes spanning from 10s of nanometers[34–44] to micrometers. While some of these techniques promise unprecedented resolution, imaging of thick samples, live tissue, or ultrathin sections, the needs in pathology are relatively narrow. For the bulk of tasks in routine pathology, which typically deals with thin sections as a matter of routine, imaging of micron-scale-thick samples with micron-scale resolution is usually sufficient. However, speed in imaging large sections (~cm × cm) is critical. Moreover, data quality needs to be consistent and high for ML methods to be effective and robust. Typically, data quality is measured by conventional microscopy measures such as the modulation transfer function for spatial characteristics and spectral signal-to-noise ratio (SNR) for chemical composition.[32, 45–48] One of the challenges is that the majority of label-free microscopy techniques rely on optical systems initially devised for light microscopy, which only offer compatibility up to near-infrared wavelengths. The pursuit of all-reflective optical systems for vibrational CI instruments has often resulted in a tradeoff, sacrificing performance to accommodate longer-wavelength light across a wider spectral range but at diminished intensities. Moreover, the spectral range of IR spectroscopy is ~20X larger than the bandwidth of optical microscopes and measurement of this large bandwidth is challenging. Fourier transform IR (FT-IR) spectroscopy has remained the predominant technique for label-free vibrational spectroscopy, thanks to its unmatched ability to capture hyperspectral data reliably, particularly in instances where preliminary sample information is lacking. Nonetheless, the inherent limitations of contemporary FT-IR microscopes become evident in scenarios demanding high-throughput analysis or when examining biological specimens with complex morphology for medical research necessitating advanced understanding behind the theory of image formation.[49] Utilizing prior knowledge about a sample, such as identifying cancerous markers in a tissue section, can significantly reduce the spectrum of required spectral features, thereby streamlining the analysis. This led to the idea of only measuring some, relevant frequencies rather than the whole spectrum.[50–54] The advent of quantum cascade laser (QCL) technology, which integrates these mid-IR sources into multiplexed assemblies, offers a solution forward. These devices, capable of high-intensity emission that is tunable across the mid-IR fingerprint range, herald the potential to transcend FT-IR limitations. With the development of first-generation wide-field QCL-based systems[45, 55–57] for discrete frequency infrared spectroscopy (DFIR), subsequent scanning-based systems,[21, 31, 32, 54] polarization sensitive systems,[58, 59] and photothermal-type systems,[25, 30, 60–64] the forthcoming years promise to further advance the capabilities of vibrational CI, paving the way for more precise, efficient diagnostic methodologies.

Alongside the progress in DFIR, significant strides have been made in ML, especially within areas like pathology. Deep learning (DL), a subfield of ML, employs artificial neural networks designed to emulate the complex processing patterns

DOI: 10.1201/9781003449942-56

of the human brain. By analyzing large datasets, DL learns to recognize patterns and autonomously generate solutions, marking a transformative impact across multiple domains.[65] Its application extends from enhancing vision-based tasks—such as image segmentation[66] and image synthetization[67]—to redefining the capabilities in language processing, including sentiment mining[68] and the development of large language models like ChatGPT. Within the realm of medical sciences, particularly in pathology, DL models contribute significantly by identifying histological components,[69] predicting disease prognosis[70] and patient survival rates.[71] The convergence of these technologies—QCL-based DFIR systems and ML, particularly DL, in pathology—illustrates a promising horizon for the field of digital chemical pathology. This manuscript aims to shed light on such innovative applications by detailing instances of IR-based virtual staining,[72, 73] segmentation,[74–77] and deepfake histology.[78] These examples are poised to redefine the traditional pathology workflow, offering a more streamlined, accurate, and insightful diagnostic process. By leveraging the synergy between advanced spectroscopy techniques and machine learning algorithms, we are entering a new era in pathology. This potential paradigm shift promises not only to enhance our understanding of diseases at the molecular level but also to significantly improve patient outcomes through faster, more precise diagnostics. As we delve into the specifics of these advancements, it becomes clear that the future of pathology lies in the integration of cutting-edge technology and AI, announcing a new era of digital chemical pathology that will transform how we conduct medical diagnostics and develop treatment strategies.

In this chapter, we provide a background for CI methods for histopathology and provide a perspective on future developments. We especially focus on providing a flavor of methods that incorporate the examination of three distinct datasets, each serving a unique purpose within the scope of our study: virtual staining, image segmentation, and the creation of deepfake histology models.

50.2 Materials

For the virtual staining task, we used a data set from a breast tissue microarray (TMA) (BR1003 from US Biomax Inc.), which is preserved through formalin-fixing and paraffin-embedding (FFPE). This array comprises an assembly of 101 tissue cores, each with a diameter of 1 mm and a thickness of 5 μm, derived from a cohort of 47 patients. Building upon data from a previous study,[31] this dataset was generated by imaging at 12 distinct bands within the mid-infrared fingerprint region. The images were captured at a spatial resolution of 2 μm, using a high-precision IR objective lens with a numerical aperture (NA) of 0.7. Turning our attention to image segmentation, we leveraged data from a prior study,[32] which involved colon surgical specimens from Carle Health in Urbana, Illinois. Briefly, these samples were carefully embedded in paraffin and sectioned into 100 consecutive slices, each 5 μm in thickness, and placed onto low-emissivity glass microscopy slides removed. The DFIR imaging was conducted with meticulous precision, integrating two coadditions per tile, utilizing a 10×0.4 NA objective, and adhering to a 2 μm pixel size. We selected spectral bands of high contrast, notably at the wavenumbers 1079, 1658, and 1765 cm^{-1}. Finally, for the development of the deepfake histology model, we engaged with a set of eight colon tissue microarrays. These arrays consist of 1 mm diameter samples that have been previously scrutinized using FT-IR spectroscopy in earlier studies.[79] In summary, the IR images were meticulously gathered using an Agilent Stingray imaging system configured to high magnification mode, achieving a pixel size of 1.1 μm and a spectral resolution of 4 cm^{-1}. Subsequently, we developed a segmentation model that identified ten critical histological components of the colon, including epithelium (mature), mucin, epithelium (proliferative), necrosis, reactive stroma, blood, inflammatory cells, non-reactive stroma, muscle, and loose stroma. These semantic maps, in conjunction with the corresponding H&E-stained images, were utilized to assemble the dataset required for the deepfake model. Further details are available in the preceding study.[78]

50.3 Method

For the virtual staining component, we employed a generative adversarial network (GAN) framework, analogous to that described in a previous study.[80] In brief, the generative model, which is based on a U-Net architecture,[81] and the discriminative model, which adopts a multiscale strategy, were both initialized with random weights. The optimization of the parameters for both the generator and the discriminator was performed using the Adam optimizer.[82] We started with an initial learning rate of 10^{-4}, which was decreased by a factor of 0.95 every 1000 iterations up to a total of 100,000 iterations. At each iteration, we applied random affine transformations to the images. The images, sized at 256×256 pixels, were processed in batches of four. Their sequence was randomized and reshuffled after each training epoch. Our training set consisted of 924 images, with an additional 85 images reserved for validation. For the image segmentation task, the methodology has been thoroughly outlined in the previous study.[32] To summarize, we utilized 300 patches of size 256×256 for training and another 100 for validation. A U-Net was employed to classify IR images into three distinct histological features: epithelium, goblet cells, and lamina propria. We incorporated batch normalization after each convolutional layer to hasten the training process and applied a softmax function at the output layer. The optimization was guided by the cross-entropy loss function. The network underwent training over 10,000 iterations, again using the Adam optimizer with a learning rate of 10^{-4}. Regarding the deepfake histology model, we adhered to the protocol set forth in a previously published study.[78] Patches of 512×512 pixels were extracted, resulting in a training dataset of 1632 image patches and a testing set of 500 patches. To enrich the dataset, each patch underwent augmentation through random rotations and reflections. The generator model was based on the SPADE architecture,[83] while the discriminator model utilized the pix2pixHD framework.[67] The Adam optimizer, with a starting learning rate of 0.0002, adjusted by a factor of 0.95 every 1000 iterations, was used throughout 200,000

iterations of training. All models were trained using NVIDIA V100 GPUs and operated on an Intel Xeon CPU E5–2680 v4 at 2.4 GHz. The implementations were carried out using PyTorch version 1.3, CUDA 10.1, and Python 3.7.1.

50.4 Results

IR imaging holds the promise of significantly transforming the pathology workflow, thereby enhancing both diagnostic processes and the overall patient experience within the healthcare sector. This innovative approach is particularly appealing due to its potential to streamline operations and improve outcomes, as detailed in prior research[33] and illustrated in Figure 50.1. The figure delineates the conventional histopathology process through Steps 1–5, contrasting it with the streamlined approach of digital chemical pathology, which spans from Step 6 to Steps 7a–c. The traditional workflow is characterized by extensive manual involvement, necessitating the sample handling, coordination of reagents, dependence on specialized personnel, and utilization of analytical equipment. Additionally, this conventional process is often hampered by the need for meticulous quality control and adherence to stringent procedural standards, which further compounds the complexity and time required for diagnosis. Conversely, digital chemical pathology simplifies the process into essentially two broad steps: IR imaging followed by ML analysis.[80] This simplification markedly reduces manual intervention, thereby saving significant effort. ML analysis, within this context, offers a broad spectrum of applications. One notable application is the utilization of ML models to convert IR images into stained images familiar to pathologists (virtual staining[72, 76, 84]). This technique, while still necessitating expert diagnosis, minimizes human interaction in the workflow. It preserves the original condition of the sample for further examination and enables virtual staining of a single sample into multiple types of stains. This capability not only saves material and processing costs but also reduces the variability associated with different technicians, sample batches, and specific staining protocols, thereby ensuring more consistent data. Another innovative application involves using ML to directly identify histological components of interest and automatically predict the presence of cancer or other diseases. Although this method still requires expert supervision, it significantly reduces the workload of pathologists and serves as an efficient screening tool until comprehensive results can be obtained from specialists. Furthermore, the adoption of color-coded images, rather than relying on the conventional, limited range of dyes and stains, eases the process of interpretation. It offers an invaluable opportunity to examine each cellular component either individually or within the spatial-chemical context of other components, paving the way for more nuanced and detailed analysis.

Building on these technological advancements, Figure 50.2 showcases an application of the described virtual staining process, demonstrating how it translates into practical visualizations for pathology purposes. We have developed a GAN-based framework capable of generating realistic H&E-stained data. This is achieved by harnessing 12 distinct bands from the DFIR fingerprint spectrum region. We assess the efficacy of our framework across two breast tissue pathologies, including normal tissue (Figure 50.2A) and cancerous tissue (Figure 50.2B). The virtual staining vividly brings out the changes in the number of cell layers from normal tissue to invasive tumor, showcasing the notable disappearance of the basement membrane as the disease progresses from a noninvasive to an invasive state. This evidence supports the potential of our digitally computed virtual stains, derived from label-free IR imaging of tissue samples, to serve as a reliable tool in identification of key tissue features and the resulting decision-making. These virtually stained images, produced quickly from IR images without the need for any chemical stains, not only facilitate immediate analysis but also ensure the preservation of the tissue sample for future examination or archival purposes.

For our second application, we focus on segmenting colon tissue into three fundamental histological categories: lamina propria, goblet cells, and epithelium (Figure 50.3). The lamina propria, a supportive connective tissue layer, is crucial for the immune defense within the mucosa, hosting a complex network of blood vessels and immune cells. Goblet cells, scattered among the epithelial cells, secrete mucus, playing a key role in protecting and lubricating the mucosal surfaces. The epithelium itself acts as a critical barrier, safeguarding against pathogens and facilitating absorption. This segmentation is achieved by utilizing only three specific IR wavelengths—1079, 1658, and 1765 cm^{-1}—combined with a sophisticated deep learning algorithm, allowing for the precise identification of these tissue types without the necessity for traditional stains or dyes. The deployment of our DFIR technology significantly speeds up this process. The efficiency and speed of our approach not only streamline the segmentation task but also render the application of this technology in clinical settings feasible. This advancement represents a leap forward in clinical diagnostics, enabling quick, stain-free analysis of tissues, thereby enhancing diagnostic workflows and patient care with its swift, noninvasive nature.

These two applications highlight how DFIR instruments and ML technologies can potentially improve the pathology workflow. Another critical element in ML model development is the creation of datasets. Often, research faces challenges in gathering enough examples to represent rare diseases, unusual manifestations of common illnesses, and complex cases, without assembling prohibitively large datasets. Such difficulties in dataset compilation can result in ML models that perform adequately for common scenarios but falter when confronted with less common, rare, or intricate situations, leading to underwhelming, imprecise, or equivocal outcomes.[85] Moreover, acquiring accurate ground truth annotations demands expertise from specialist pathologists, a resource that is not always easily available.[86] Additionally, variations in annotations are not rare, necessitating the pursuit of consensus among experts or the application of statistical methods to reconcile differences in the assessments by different pathologists, as well as inconsistencies in a single pathologist's evaluations. In response to these challenges, and building upon insights from prior research, we have pioneered a deepfake histology model capable of generating synthetic H&E images of colon tissue

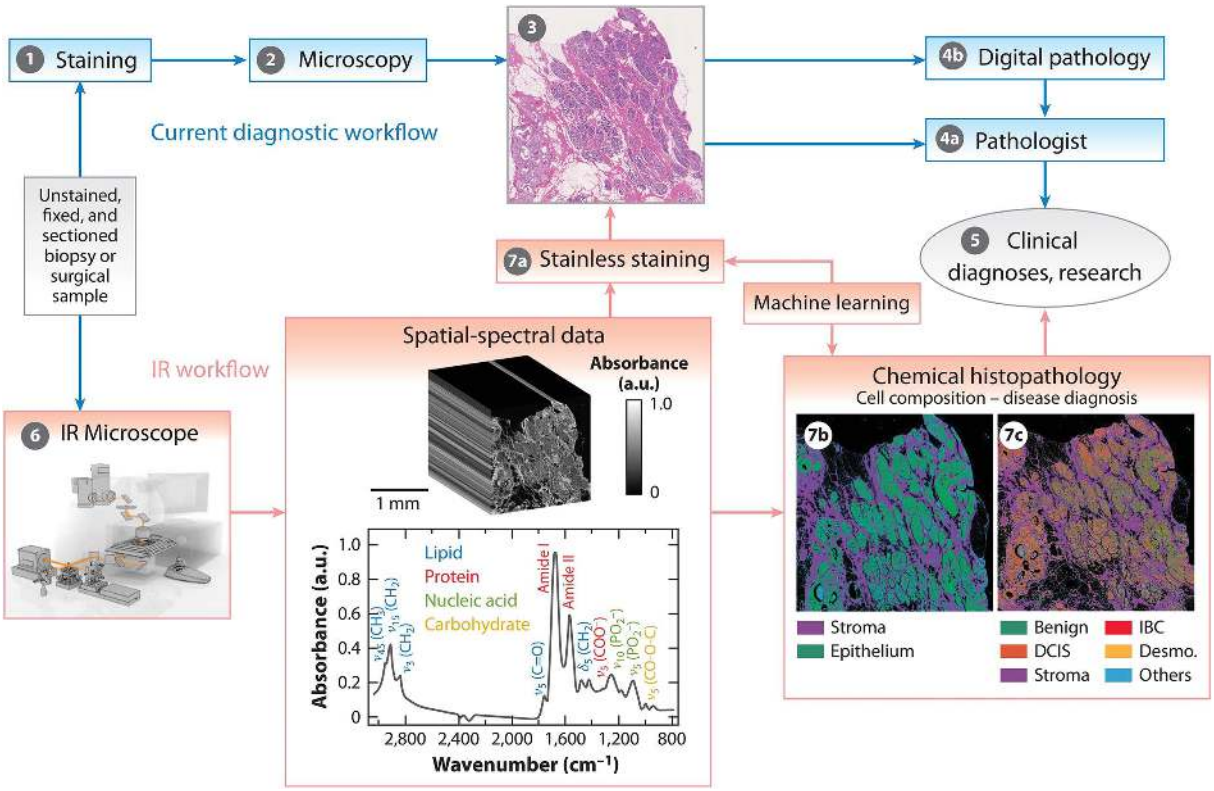


FIGURE 50.1 *Comparative schematic illustrating the conventional histopathology workflow alongside the proposed digital chemical pathology.* This figure is highlighting key procedural differences and potential improvements in diagnostic accuracy and efficiency. (The figure is reproduced with permission from a prior study.[33])

based on semantic maps that delineate ten critical histological components of the colon. This innovative model leverages class maps derived from the classification of IR. Given that IR imaging exhibits heightened sensitivity to chemical variations compared to traditional H&E staining, it facilitates a more detailed and accurate identification of histological units. These units, in turn, serve as precise inputs for our deepfake model, enhancing its ability to produce highly realistic and detailed synthetic images. This approach not only addresses the limitations posed by conventional staining methods but also opens new avenues for accurate and efficient dataset completion, bridging gaps in ML development with cutting-edge computational techniques. We showcase the application of this method in Figure 50.4, where we illustrate the process of generating an H&E image from a provided semantic map (Figure 50.4A). This illustrates our ability to create a lifelike H&E image (Figure 50.4B) from the initial semantic map. By modifying the semantic map (Figure 50.4C)—for instance, introducing blood cells or eliminating a gland among other adjustments—we can directly impact the resulting image, producing a modified synthesized H&E image (Figure 50.4D). These modifications showcase the adaptability and accuracy of our model in responding to alterations in input data, presenting a versatile instrument for histological examination and study.

50.5 Discussion

In this chapter, we have provided an overview and delved into two instances where IR imaging technology can impact pathology processes, along with showcasing an innovative application of deepfake technology in histology for dataset enhancement and generation. The first application, virtual staining, illustrates the capacity of deep learning techniques to apply virtual stains to unlabeled and unstained tissue samples. This process leverages DFIR imaging at various wavenumbers, offering detailed chemical insights into tissues within a practical timeframe. This advancement renders histological analyses more clinically viable and expedites the process compared to traditional FTIR imaging systems. Furthermore, virtual staining methods facilitate the quick and uniform staining of tissue specimens, opening up new pathways to streamline the conventional, labor-intensive histological staining workflows typical in routine pathology. Additionally, this technology preserves the tissue for subsequent analyses, such as immunohistochemical staining, enhancing tumor screening and treatment precision. The technology we propose also holds potential for application to fresh tissue samples and various staining markers, significantly broadening its utility. It's noteworthy that the virtual staining only requires actual stained images during the model's training phase; after which the model is capable of

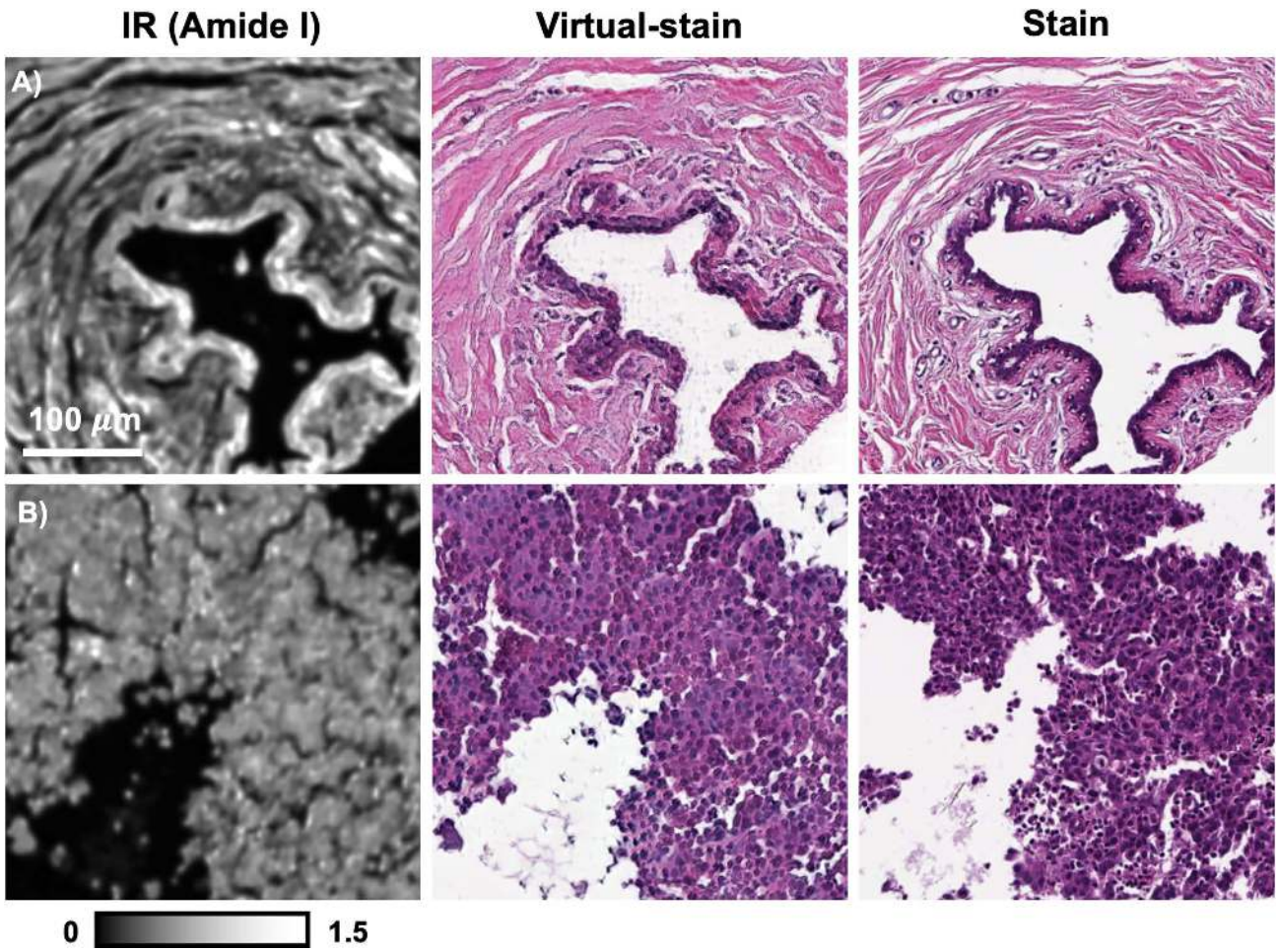


FIGURE 50.2 *Visual comparison of virtual staining*. Displayed are IR images, virtual-stained images, and traditional H&E-stained images for three distinct histological classes. Panel (A) represents the normal case, while (B) depicts the cancerous condition.

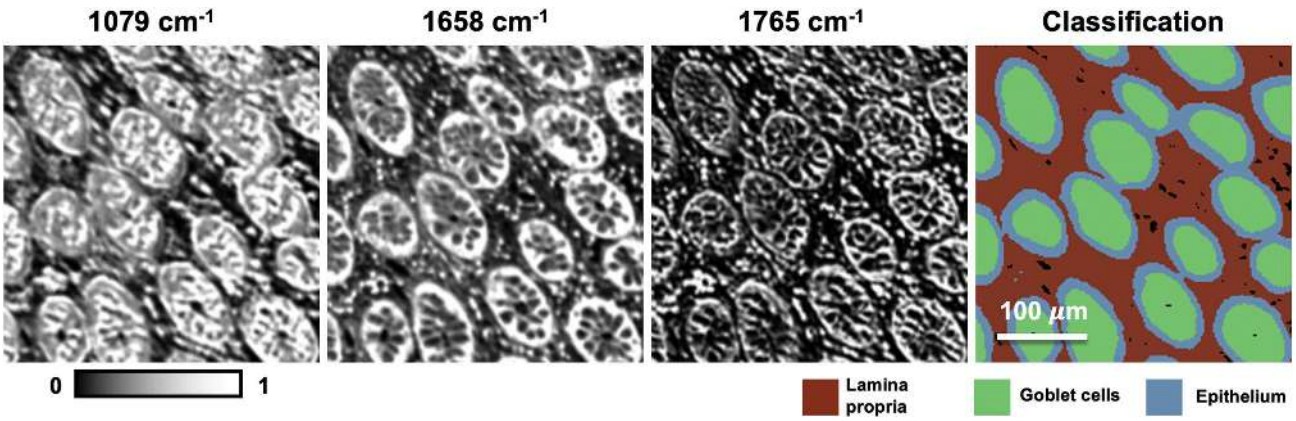


FIGURE 50.3 *Visual comparison of colon tissue segmentation*. Displayed from left to right, this figure demonstrates the utilization of three distinct IR bands—1079, 1658, and 1765 cm^{-1}, respectively—employed in the development of our segmentation model, followed by the resultant classification.

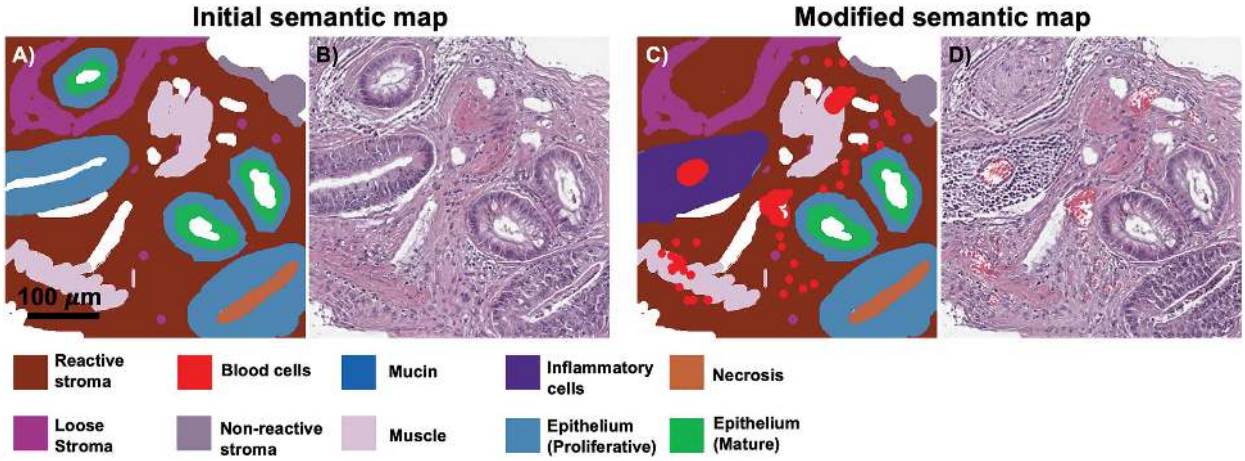


FIGURE 50.4 *Visual demonstration of deepfake model application on colon tissue.* (A) A showcase of the original semantic map, with (B) the synthesized H&E image based on this map. (C) Illustration of the semantic map after modifications, and (D) presentation of the altered deepfake H&E image resulting from these changes.

generating "stain-free" stained images. Although IR images typically exhibit lower resolution compared to H&E images, the virtual stain images benefit from an inherent "super-resolution" effect. This enhancement makes them comparable to traditional stain images, facilitated by the GAN formulation wherein the generator seeks to mimic the distribution of the H&E dataset to effectively deceive the discriminator. This approach can also integrate visible-light images of unstained tissues with IR data, compensating for the lack of spatial detail in IR images and enriching the histological analysis of virtually stained tissues by accentuating crucial structures that IR imaging alone might not fully resolve.

An additional crucial aspect of virtual staining to consider is the safety and reliability of the technique. This consideration is sparked by the necessity to evaluate the confidence in image-to-image translation technologies, especially in the context of virtual staining. To address this, uncertainty quantification of the generated images plays a pivotal role. Traditional methods that focus on pixel-level assessment may not suffice for virtual staining; instead, a more holistic approach is required to gauge this uncertainty at a perceptual level. Prior research[87] has paved the way by introducing perceptual metrics designed specifically to mimic human perception, which can be used for this type of uncertainty evaluation. These metrics allow for a nuanced analysis of the generated images, providing insights beyond mere pixel accuracy by considering the overall perceptual integrity of the image. This form of assessment proves invaluable for determining the virtual staining technique's safety, where one might anticipate lower uncertainty levels in regions less prone to errors and conversely higher uncertainty in areas where errors are more likely. Furthermore, this perceptual uncertainty assessment can serve as a cornerstone for ensuring the safety of the virtual staining process. By identifying areas of high uncertainty, researchers and clinicians can apply extra scrutiny or additional verification steps, thereby enhancing the reliability of the diagnoses based on virtually stained images. Such advancements not only contribute to the technical robustness of virtual staining but also bolster the confidence of pathologists and healthcare professionals in utilizing these digitally enhanced images for clinical diagnosis, ultimately aiming to achieve a balance between innovation and patient safety.

Besides virtual staining, in the second example, we explore the application of IR imaging and ML technologies for the segmentation of colon tissue into three crucial histological units. Traditionally, IR-based segmentation models primarily utilized pixel-wise information due to the redundancy of spectral bands. However, advancements in DFIR technology have facilitated faster imaging by narrowing down the number of bands imaged. This advancement necessitates approaches that incorporate both spatial and spectral data to offset the diminished spectral information. Echoing findings from previous research,[46] we demonstrate that leveraging a convolutional neural network (CNN) to analyze both spatial and spectral inputs enables the classifier to achieve segmentation closely matching the ground truth. This CNN-based classifier circumvents the traditional need for manually defined metrics common in IR imaging,[31] streamlining the feature engineering process for specific histological analyses. This development is at the forefront of current research, with emerging methodologies to identify and classify pixels for segmentation purposes.[88] Notably, similar to the virtual staining application, incorporating a visible-light image, which is relatively quick to acquire, can significantly enhance segmentation accuracy by providing detailed spatial context of the samples. Thus, a visible image can inform the segmentation model with critical details about the tissue's texture and structure, enhancing its precision. The integration of this technique with standard clinical imaging practices holds the potential to improve the diagnostic capabilities of chemical imaging systems, facilitating a smoother

transition and broader acceptance within clinical settings. Nonetheless, this innovative approach requires meticulous validation to confirm its reliability and robustness.

Deepfake technologies have shown great promise in augmenting datasets and adding complexity that is not easily accessible, as illustrated by our method that facilitates the generation and modification of tissue images with lifelike accuracy. Yet, the process of crafting artificial histological images is not without its hurdles and drawbacks. One primary obstacle includes the computational intensity required to produce images of high fidelity, especially for broad studies or applications demanding quick turnarounds. Nonetheless, advancements in processing power and algorithmic efficiency offer potential solutions, underscoring the importance of meticulously considering computational demands. Innovations in reducing model size, enhancing parallel processing capabilities, and optimizing for edge device deployment have the potential to broaden the application of deepfake technology within the realm of histology imaging. Additionally, the possibility of bias in the foundational data—arising from how samples are collected, prepared, or annotated—presents a challenge, as it may influence the synthetic images, skewing results or diminishing the model's adaptability to a wide array of situations. Addressing this issue requires the assembly of diverse and inclusive datasets alongside the implementation of strategies to counteract biases during model training, aiming to produce synthetic images that are not only of high quality but also widely applicable. Concerning the integration of deepfake technology for dataset enrichment and augmentation, data privacy emerges as a critical issue. The fabrication of deepfake images might reduce some privacy risks linked with using genuine patient specimens, yet the reliance on patient data for building the underlying model framework persists. Implementing stringent data protection and anonymization protocols for both the source and generated images is vital, ensuring privacy preservation without detracting from the datasets' value. This strategic approach to managing privacy concerns, alongside the technological enhancements and bias mitigation efforts, underscores the multifaceted considerations essential for harnessing deepfake technology effectively and ethically in medical research.

50.6 Future Trends

In this chapter, we have explored the integration of IR imaging with ML for pathology applications, alongside a novel utilization of deepfake technology for histological purposes. Our first case study demonstrated the capacity of IR imaging, when combined with ML algorithms, to perform virtual staining of histological samples. This method addresses several limitations inherent in traditional staining techniques, such as the requirement for multiple samples for various stains, staining inconsistency, and the dependence on skilled technicians. Furthermore, we developed an IR-based model for the segmentation of colon tissue into three critical histological categories. This approach showcased the ability of IR imaging to identify tissue classes swiftly and accurately without the reliance on conventional staining methods, thereby offering a more efficient diagnostic tool. Last, we introduced an advanced deepfake histology model aimed at enriching datasets and facilitating dataset completion for enhanced ML model training. This model stands to revolutionize data preparation for ML tasks in pathology by providing a versatile tool for generating and modifying histological images. Collectively, these examples highlight the significant potential of employing IR imaging in conjunction with ML and deepfake technologies to improve clinical practices and research methodologies. The advancements discussed here promise not only to elevate diagnostic precision and efficiency but also to drive innovation in medical research, signifying a significant advancement in the integration of technology within healthcare. While IR imaging delivers detailed insights by thoroughly measuring chemical compositions, it is essential to harness this wealth of information for the identification of novel biomarkers, the optimization of personalized treatment strategies, and the improvement of predictive analytics in the medical field.

Acknowledgments

Work reported in this chapter was supported by the National Institutes of Health via grant numbers R01GM142172, R01EB009745, R01CA260830, R21CA263147, and P41EB031772 and by the National Science Foundation via grant numbers 2243257 and 2229986.

REFERENCES

1. Serag, A., Ion-Margineanu, A., Qureshi, H., McMillan, R., Saint Martin, M. J., Diamond, J. et al. Translational AI and deep learning in diagnostic pathology. *Front Med (Lausanne)* 6, 185 (2019).
2. Falahkheirkhah, K. *Digital chemical pathology* (University of Illinois Urbana Champaign, 2023).
3. Baker, M. J., Trevisan, J., Bassan, P., Bhargava, R., Butler, H. J., Dorling, K. M. et al. Using Fourier transform IR spectroscopy to analyze biological materials. *Nat Protoc* 9, 1771–1791 (2014).
4. Doyle, B. B., Bendit, E. G. & Blout, E. R. Infrared spectroscopy of collagen and collagen-like polypeptides. *Biopolymers* 14, 937–957 (1975).
5. Belbachir, K., Noreen, R., Gouspillou, G. & Petibois, C. Collagen types analysis and differentiation by FTIR spectroscopy. *Anal Bioanal Chem* 395, 829–837 (2009).
6. Petibois, C., Gouspillou, G., Wehbe, K., Delage, J.-P. & Déléris, G. Analysis of type I and IV collagens by FT-IR spectroscopy and imaging for a molecular investigation of skeletal muscle connective tissue. *Anal Bioanal Chem* 386, 1961–1966 (2006).
7. Cheheltani, R., McGoverin, C. M., Rao, J., Vorp, D. A., Kiani, M. F. & Pleshko, N. Fourier transform infrared spectroscopy to quantify collagen and elastin in an in vitro model of extracellular matrix degradation in aorta. *Analyst* 139, 3039–3047 (2014).
8. Paschalis, E. P., Verdelis, K., Doty, S. B., Boskey, A. L., Mendelsohn, R. & Yamauchi, M. Spectroscopic characterization of collagen cross-links in bone. *J Bone Miner Res* 16, 1821–1828 (2001).

9. Stani, C., Vaccari, L., Mitri, E. & Birarda, G. FTIR investigation of the secondary structure of type I collagen: New insight into the amide III band. *Spectrochim Acta A Mol Biomol Spectrosc* 229, 118006 (2020).
10. West, P. A., Torzilli, P. A., Chen, C., Lin, P. & Camacho, N. P. Fourier transform infrared imaging spectroscopy analysis of collagenase-induced cartilage degradation. *J Biomed Opt* 10, 014015 (2005).
11. Bakir, G., Girouard, B. E., Wiens, R., Mastel, S., Dillon, E., Kansiz, M. et al. Orientation matters: Polarization dependent IR spectroscopy of collagen from intact tendon down to the single fibril level. *Molecules* 25, 4295 (2020).
12. Kamińska, A. & Sionkowska, A. Effect of UV radiation on the infrared spectra of collagen. *Polym Degrad Stab* 51, 19–26 (1996).
13. Lazarev, Y. A., Grishkovsky, B. A. & Khromova, T. B. Amide I band of IR spectrum and structure of collagen and related polypeptides. *Biopolymers* 24, 1449–1478 (1985).
14. Jakobsen, R. J., Brown, L. L., Hutson, T. B., Fink, D. J. & Veis, A. Intermolecular interactions in collagen self-assembly as revealed by Fourier transform infrared spectroscopy. *Science (1979)* 220, 1288–1290 (1983).
15. Lewis, E. N., Levin, I. W., Fraser, J. C., Arens, J. F., Kidder, L. H. & Colarusso, P. Infrared spectroscopic imaging: From planetary to cellular systems. *Appl Spectrosc* 52:3, 106A–120A (1998).
16. Bhargava, R. Infrared spectroscopic imaging: The next generation. *Appl Spectrosc* 66, 1091–1120 (2012).
17. Pilling, M. & Gardner, P. Fundamental developments in infrared spectroscopic imaging for biomedical applications. *Chem Soc Rev* 45, 1935–1957 (2016).
18. Ewing, A. V. & Kazarian, S. G. Infrared spectroscopy and spectroscopic imaging in forensic science. *Analyst* 142, 257–272 (2017).
19. Wrobel, T. P. & Bhargava, R. Infrared spectroscopic imaging advances as an analytical technology for biomedical sciences. *Anal Chem* 90, 1444–1463 (2018).
20. Steiner, G. & Koch, E. Trends in Fourier transform infrared spectroscopic imaging. *Anal Bioanal Chem* 394, 671–678 (2009).
21. Yeh, K., Lee, D. & Bhargava, R. Multicolor discrete frequency infrared spectroscopic imaging. *Anal Chem* 91, 2177–2185 (2019).
22. Kazarian, S. G. Perspectives on infrared spectroscopic imaging from cancer diagnostics to process analysis. *Spectrochim Acta A Mol Biomol Spectrosc* 251, 119413 (2021).
23. Kazarian, S. G. & Chan, K. L. A. Micro-and macro-attenuated total reflection Fourier transform infrared spectroscopic imaging. *Appl Spectrosc* 64, 135A–152A (2010).
24. Lux, L., Phal, Y., Hsieh, P.-H. & Bhargava, R. On the limit of detection in infrared spectroscopic imaging. *Appl Spectrosc* 76, 105–117 (2022).
25. Bai, Y., Zhang, D., Lan, L., Huang, Y., Maize, K., Shakouri, A. et al. Ultrafast chemical imaging by widefield photothermal sensing of infrared absorption. *Sci Adv* 5, eaav7127 (2019).
26. Wang, H., Lee, D., Cao, Y., Bi, X., Du, J., Miao, K. et al. Bond-selective fluorescence imaging with single-molecule sensitivity. *Nat Photonics* 17, 846–855 (2023).
27. Ishigane, G., Toda, K., Tamamitsu, M., Shimada, H., Badarla, V. R. & Ideguchi, T. Label-free mid-infrared photothermal live-cell imaging beyond video rate. *Light Sci Appl* 12, 174 (2023).
28. Xia, Q., Yin, J., Guo, Z. & Cheng, J.-X. Mid-infrared photothermal microscopy: Principle, instrumentation, and applications. *J Phys Chem B* 126, 8597–8613 (2022).
29. Astratov, V. N., Sahel, Y. Ben, Eldar, Y. C., Huang, L., Ozcan, A., Zheludev, N. et al. Roadmap on label-free superresolution imaging. *Laser Photon Rev* 17, (2023).
30. Schnell, M., Mittal, S., Falahkheirkhah, K., Mittal, A., Yeh, K., Kenkel, S. et al. All-digital histopathology by infrared-optical hybrid microscopy. *Proc Natl Acad Sci* 117, 201912400 (2020).
31. Mittal, S., Yeh, K., Suzanne Leslie, L., Kenkel, S., Kajdacsy-Balla, A. & Bhargava, R. Simultaneous cancer and tumor microenvironment subtyping using confocal infrared microscopy for all-digital molecular histopathology. *Proc Natl Acad Sci U S A* 115, 651–660 (2018).
32. Yeh, K., Sharma, I., Falahkheirkhah, K., Confer, M. P., Orr, A. C., Liu, Y. T. et al. Infrared spectroscopic laser scanning confocal microscopy for whole-slide chemical imaging. *Nat Commun* 14:1, 1–12 (2023).
33. Bhargava, R. Digital histopathology by infrared spectroscopic imaging. *Annu Rev Anal Chem* 16, 205–230 (2023).
34. Dazzi, A. & Prater, C. B. AFM-IR: Technology and applications in nanoscale infrared spectroscopy and chemical imaging. *Chem Rev* 117, 5146–5173 (2017).
35. Schwartz, J. J., Jakob, D. S. & Centrone, A. A guide to nanoscale IR spectroscopy: Resonance enhanced transduction in contact and tapping mode AFM-IR. *Chem Soc Rev* 51, 5248–5267 (2022).
36. Mathurin, J., Deniset-Besseau, A., Bazin, D., Dartois, E., Wagner, M. & Dazzi, A. Photothermal AFM-IR spectroscopy and imaging: Status, challenges, and trends. *J Appl Phys* 131, (2022).
37. Dazzi, A., Prazeres, R., Glotin, F. & Ortega, J. M. Analysis of nano-chemical mapping performed by an AFM-based ("AFMIR") acousto-optic technique. *Ultramicroscopy* 107, 1194–1200 (2007).
38. Cinque, G., Kelley, C. S., Frogley, M. D., Filik, J., Wehbe, K., Fitzpatrick, A. et al. World first for diamond in synchrotron-based IR photothermal nanospectroscopy. *Synchrotron Radiat News* 29, 37–39 (2016).
39. dos Santos, A. C. V. D., Hondl, N., Ramos-Garcia, V., Kuligowski, J., Lendl, B. & Ramer, G. AFM-IR for nanoscale chemical characterization in life sciences: Recent developments and future directions. *ACS Meas Sci Au* 3, 301–314 (2023).
40. Hinrichs, K. & Shaykhutdinov, T. Polarization-dependent atomic force microscopy–infrared spectroscopy (AFM-IR): Infrared nanopolarimetric analysis of structure and anisotropy of thin films and surfaces. *Appl Spectrosc* 72, 817–832 (2018).
41. Nguyen-Tri, P., Ghassemi, P., Carriere, P., Nanda, S., Assadi, A. A. & Nguyen, D. D. Recent applications of advanced atomic force microscopy in polymer science: A review. *Polymers (Basel)* 12, 1142 (2020).
42. Kenkel, S., Mittal, S. & Bhargava, R. Closed-loop atomic force microscopy-infrared spectroscopic imaging for nanoscale molecular characterization. *Nat Commun* 11, 3225 (2020).

43. Kenkel, S., Mittal, A., Mittal, S. & Bhargava, R. Probe--sample interaction-independent atomic force microscopy--infrared spectroscopy: Toward robust nanoscale compositional mapping. *Anal Chem* 90, 8845–8855 (2018).
44. Kenkel, S., Gryka, M., Chen, L., Confer, M. P., Rao, A., Robinson, S. et al. Chemical imaging of cellular ultrastructure by null-deflection infrared spectroscopic measurements. *Proc Natl Acad Sci* 119, (2022).
45. Yeh, K., Kenkel, S., Liu, J. N. & Bhargava, R. Fast infrared chemical imaging with a quantum cascade laser. *Anal Chem* 87, 485–493 (2015).
46. Falahkheirkhah, K., Yeh, K., Mittal, S., Pfister, L. & Bhargava, R. Deep learning-based protocols to enhance infrared imaging systems. *Chemom Intell Lab Syst* 217, 104390 (2021).
47. Phal, Y., Yeh, K. & Bhargava, R. Design considerations for discrete frequency infrared microscopy systems. *Appl Spectrosc* 75:9, 1067–1092 (2021).
48. Falahkheirkhah, K., Yeh, K., Mittal, S., Pfister, L. & Bhargava, R. A deep learning framework for morphologic detail beyond the diffraction limit in infrared spectroscopic imaging. *arXiv preprint arXiv:1911.04410* (2019).
49. Davis, B. J., Carney, P. S. & Bhargava, R. Theory of mid-infrared absorption microspectroscopy: II. Heterogeneous samples. *Anal Chem* 82, 3487–3499 (2010).
50. Kodali, A. K., Schulmerich, M., Ip, J., Yen, G., Cunningham, B. T. & Bhargava, R. Narrowband midinfrared reflectance filters using guided mode resonance. *Anal Chem* 82, 5697–5706 (2010).
51. Liu, J.-N., Schulmerich, M. V., Bhargava, R. & Cunningham, B. T. Sculpting narrowband Fano resonances inherent in the large-area mid-infrared photonic crystal microresonators for spectroscopic imaging. *Opt Express* 22, 18142 (2014).
52. Yeh, K. & Bhargava, R. Discrete frequency infrared imaging using quantum cascade lasers for biological tissue analysis. In *Biomedical Vibrational Spectroscopy 2016: Advances in Research and Industry*, SPIE Digital Library, vol. 9704 (2016).
53. Kole, M. R. *Discrete frequency chemical imaging with stimulated Raman scattering microscopy* (University of Illinois Urbana Champaign, 2017).
54. Tiwari, S., Raman, J., Reddy, V., Ghetler, A., Tella, R. P., Han, Y. et al. Towards translation of discrete frequency infrared spectroscopic imaging for digital histopathology of clinical biopsy samples. *Anal Chem* 88, 10183–10190 (2016).
55. Kole, M. R., Reddy, R. K., Schulmerich, M. V, Gelber, M. K. & Bhargava, R. Discrete frequency infrared microspectroscopy and imaging with a tunable quantum cascade laser. *Anal Chem* 84, 10366–10372 (2012).
56. Kröger, N., Egl, A., Engel, M., Gretz, N., Haase, K., Herpich, I. et al. Quantum cascade laser–based hyperspectral imaging of biological tissue. *J Biomed Opt* 19, 111607 (2014).
57. Bassan, P., Weida, M. J., Rowlette, J. & Gardner, P. Large scale infrared imaging of tissue micro arrays (TMAs) using a tunable quantum cascade laser (QCL) based microscope. *Analyst* 139, 3856–3859 (2014).
58. Mukherjee, P., Ghosh, A., Spegazzini, N., Lamborn, M. J., Monwar, M. M., Deslauriers, P. J. et al. Relating post-yield mechanical behavior in polyethylenes to spatially varying molecular deformation using infrared spectroscopic imaging: Homopolymers. *Macromolecules* 51, 3836–3844 (2018).
59. Phal, Y., Yeh, K. & Bhargava, R. Concurrent vibrational circular dichroism measurements with infrared spectroscopic imaging. *Anal Chem* 93, 1294–1303 (2021).
60. Li, Z., Aleshire, K., Kuno, M. & Hartland, G. V. Super-resolution far-field infrared imaging by photothermal heterodyne imaging. *J Phys Chem B* 121, 8838–8846 (2017).
61. Yin, J., Lan, L., Zhang, Y., Ni, H., Tan, Y., Zhang, M. et al. Nanosecond-resolution photothermal dynamic imaging via MHZ digitization and match filtering. *Nat Commun* 12:1, 1–11 (2021).
62. Zhang, D., Li, C., Zhang, C., Slipchenko, M. N., Eakins, G. & Cheng, J. X. Depth-resolved mid-infrared photothermal imaging of living cells and organisms with submicrometer spatial resolution. *Sci Adv* 2, (2016).
63. Xia, Q., Yin, J., Guo, Z. & Cheng, J. X. Mid-infrared photothermal microscopy: Principle, instrumentation, and applications. *J Phys Chem B* 126, 8597–8613 (2022).
64. Yin, J., Zhang, M., Tan, Y., Guo, Z., He, H., Lan, L. et al. Video-rate mid-infrared photothermal imaging by single-pulse photothermal detection per pixel. *Sci Adv* 9, (2023).
65. Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* 521, 436–444 (2015).
66. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Xplore, 3431–3440 (2015).
67. Wang, T.-C., Liu, M.-Y., Zhu, J.-Y., Tao, A., Kautz, J. & Catanzaro, B. High-resolution image synthesis and semantic manipulation with conditional gans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Xplore, 8798–8807 (2018).
68. Zhang, L., Wang, S. & Liu, B. Deep learning for sentiment analysis: A survey. *Wiley Interdiscip Rev Data Min Knowl Discov* 8, e1253 (2018).
69. Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D. et al. Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med* 24:10, 1559–1567 (2018).
70. Jaroensri, R., Wulczyn, E., Hegde, N., Brown, T., Flament-Auvigne, I., Tan, F. et al. Deep learning models for histologic grading of breast cancer and association with disease prognosis. *npj Breast Cancer* 8:1, 1–12 (2022).
71. Echle, A., Rindtorff, N. T., Brinker, T. J., Luedde, T., Pearson, A. T. & Kather, J. N. Deep learning in cancer pathology: A new generation of clinical biomarkers. *Br J Cancer* 124, 686–696 (2021).
72. Mayerich, D., Walsh, M. J., Kadjacsy-Balla, A., Ray, P. S., Hewitt, S. M. & Bhargava, R. Stain-less staining for computed histopathology. *Tech (Singap World Sci)* 3, 27–31 (2015).
73. Zimmermann, E., Mukherjee, S. S., Falahkheirkhah, K., Gryka, M. C., Kajdacsy-Balla, A., Hasan, W. et al. Detection and quantification of myocardial fibrosis using stain-free infrared spectroscopic imaging. *Arch Pathol Lab Med* (2021).
74. Fernandez, D. C., Bhargava, R., Hewitt, S. M. & Levin, I. W. Infrared spectroscopic imaging for histopathologic recognition. *Nat Biotechnol* 23, 469–474 (2005).

75. Udelhoven, T., Naumann, D. & Schmitt, J. Development of a hierarchical classification system with artificial neural networks and FT-IR spectra for the identification of bacteria. *Appl Spectrosc* 54, 1471–1479 (2000).
76. Berisha, S., Lotfollahi, M., Jahanipour, J., Gurcan, I., Walsh, M., Bhargava, R. et al. Deep learning for FTIR histology: Leveraging spatial and spectral features with convolutional neural networks. *Analyst* 144, 1642–1653 (2019).
77. Tiwari, S., Falahkheirkhah, K., Cheng, G. & Bhargava, R. Colon cancer grading using infrared spectroscopic imaging-based deep learning. *Appl Spectrosc* (2022). doi:10.1177/00037028221076170.
78. Falahkheirkhah, K., Tiwari, S., Yeh, K., Gupta, S., Herrera-Hernandez, L., McCarthy, M. R. et al. Deepfake histologic images for enhancing digital pathology. *Laboratory Investigation* 103, 100006 (2023).
79. Tiwari, S., Kajdacsy-Balla, A., Whiteley, J., Cheng, G., Hewitt, S. M. & Bhargava, R. Inform: Infrared-based organizational measurements of tumor and its microenvironment to predict patient survival. *Sci Adv* 7, eabb8292 (2021).
80. Falahkheirkhah, K., Mukherjee, S. S., Gupta, S., Herrera-Hernandez, L., McCarthy, M. R., Jimenez, R. E. et al. Accelerating cancer histopathology workflows with chemical imaging and machine learning. *Cancer Res Commun* 3, 1875–1887 (2023).
81. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer Nature, 234–241 (2015).
82. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. In *International Conference for Learning Representations (ICLR)* (2015).
83. Park, T., Liu, M.-Y., Wang, T.-C. & Zhu, J.-Y. Semantic image synthesis with spatially-adaptive normalization. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2337–2346 (2019).
84. Lotfollahi, M., Berisha, S., Daeinejad, D. & Mayerich, D. Digital staining of high-definition Fourier transform infrared (FT-IR) images using deep learning. *Appl Spectrosc* 73, 556–564 (2019).
85. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M. et al. A survey on deep learning in medical image analysis. *Med Image Anal* 42, 60–88 (2017).
86. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. & Aerts, H. J. W. L. Artificial intelligence in radiology. *Nat Rev Cancer* 18:8, 500–510 (2018).
87. Zhang, R., Isola, P., Efros, A. A., Shechtman, E. & Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Xplore, 586–595 (2018).
88. Yang, J., Xu, J., Zhang, X., Wu, C., Lin, T. & Ying, Y. Deep learning for vibrational spectral analysis: Recent progress and a practical guide. *Anal Chim Acta* 1081, 6–17 (2019).

51

GALAD Score and Extracellular Vesicle-Based Assays for Early Detection of Hepatocellular Carcinoma

Yi-Te Lee, Yazhen Zhu, Sungyong You, Vatche G. Agopian, Hsian-Rong Tseng, and Ju Dong Yang

51.1 Introduction

Hepatocellular carcinoma (HCC), which accounts for more than 80% of primary liver cancers,[1, 2] is one of the leading causes of cancer-related death in the United States (US).[3] Despite recent advances in diagnosis and treatments, the survival of patients with HCC is still suboptimal and the mortality rate is increasing in more than half of the states.[4] Current clinical practice guidelines suggest semiannual ultrasound with serum alpha-fetoprotein (AFP) for HCC surveillance among patients with liver cirrhosis or high-risk chronic hepatitis B,[5, 6] aiming to detect HCC at an early stage that can be potentially cured. However, the sensitivity of ultrasound with serum AFP is only around 63%.[7] Recent studies have shown that less than 30% of patients receive routine HCC surveillance.[8] As such, development of accurate and convenient biomarkers is required to complement current surveillance tools and improve patients' outcomes.[9]

Several serum-based biomarkers, such as lens culinaris agglutinin-reactive AFP (AFP-L3) and des-gamma-carboxy prothrombin (DCP) have been investigated and validated for early detection of HCC.[9] However, these biomarkers still suffered from limited sensitivity or specificity.[9] To overcome this, the GALAD score was established, which is calculated based on gender, age, AFP-L3%, AFP, and DCP.[10] The GALAD score was subsequently validated in European and Asian populations and demonstrated great potential for distinguishing early-stage HCC from at-risk patients with liver cirrhosis or chronic hepatitis B.[9] Recently, our team validated its performance for the first time in the US cohorts and developed a novel risk score model combining the GALAD score and liver ultrasound results.[11]

Extracellular vesicles (EVs) are lipid bilayer-enclosed particles naturally released by both tumor and normal cells, with tumor cells releasing more EVs than normal cells.[4, 12] Studies have shown the molecular cargoes in tumor-derived EVs, such as RNA, proteins, metabolites, and lipids, participate in cell-to-cell communications, tumorigenesis, and cancer progression. The cargoes of tumor-derived EVs are reflective of the tumor of origin, and tumor-derived EVs are present in circulation at relatively early stages of disease.[13] Therefore, targeting and profiling tumor-derived EVs is regarded as a promising liquid biopsy approach for early-stage cancer detection.[4, 9, 12] However, the commonly used methods for isolating EVs, including ultracentrifugation, filtration, and precipitation, could not efficiently distinguish tumor-derived EVs from those released by normal cells or platelets.[4] This limitation may hinder their utility as cancer biomarker.

Over the past decade, our joint research team has been dedicated to developing *in vitro* diagnostic (IVD) devices for enriching rare circulating tumor cells and tumor-derived EVs from patients' blood.[14–17] We have shown our click chemistry-mediated platforms, EV Click Chips and Click Beads,[14, 17] can efficiently enrich HCC-derived EVs. With downstream molecular analyses and clinical validation in phase 2 biomarker studies, HCC EV digital scoring assay (EV Click Chip + reverse-transcription droplet digital PCR [RT-ddPCR]) and HCC EV surface protein assay (Click Beads + real-time immuno-PCR) demonstrated great promise in early detection of HCC.[14, 17]

In this chapter, we will discuss the performances of GALAD score and our EV-based assays for early detection of HCC.

51.2 Materials

51.2.1 Measurement of GALAD Score

GALAD score was calculated based on gender, age, AFP-L3%, AFP, and DCP (Figure 51.1a) using the formula described in the Methods section. The demographic information was collected with the patients provided the consent forms. AFP, AFP-L3%, and DCP were measured as part of their standard clinical care.

51.2.2 Antibodies for Enrichment of HCC EVs

Goat anti-human EpCAM (R&D Systems), rabbit anti-human ASGPR1 (LifeSpan BioSciences, Inc.), sheep anti-human GPC3 (R&D Systems, Inc.), and goat anti-human CD147 (R&D Systems, Inc.) were used to target HCC-associated surface proteins and enrich HCC EVs. These four antibodies were incubated with trans-cyclooctene (TCO)-PEG_4-N-hydroxylsuccinimide (NHS) ester (0.5 mM, Click Chemistry Tools Bioconjugate Technology Company) in phosphate buffered saline (PBS) at the mole ratio of 1:4, respectively, at room temperature for 30 min. The individual TCO-conjugated antibody was prepared just before use.

DOI: 10.1201/9781003449942-57

51.3 HCC EV Digital Scoring Assay

The HCC EV digital scoring assay (Figure 51.1b) consists of (i) EV Click Chips for enriching HCC EVs and (ii) RT-ddPCR for quantifying a panel of mRNA markers in the enriched HCC EVs.

51.3.1 EV Click Chip

An EV Click Chip is composed of two functional components: (i) tetrazine (Tz)-grafted Si nanowire substrate (SiNWS): a patterned SiNWS covalently functionalized with disulfide bonds that link to terminal Tz motifs, and (ii) an overlaid polydimethylsiloxane (PDMS) chaotic mixer, housed in a custom-designed microfluidic chip holder.

The fabrication of Tz-grafted SiNWS began with introducing 10–15 μm densely packed Si nanowires (diameter=100–200 nm) onto SiNWS for facilitating click chemistry-mediated HCC EV enrichment. The incorporation of disulfide bonds and terminal Tz motifs onto SiNWS was carried out via a three-step chemical modification procedure. (i) Silanization: the SiNWS were first immersed in a freshly prepared piranha solution (H_2SO_4/H_2O_2, 2:1 (v/v), Sigma–Aldrich) for 1 h, followed by rinsing with deionization (DI) water and ethanol successively, three times. After drying under nitrogen flow, the resultant SiNWS were sealed in a vacuum desiccator for treatment with (3-mercaptopropyl) trimethoxysilane vapor (211.4 mg,200 μL, Sigma–Aldrich) for 45 min to introduce thiol groups onto the SiNWS. (ii) Incorporation of disulfide bond: OPSS-PEG-NH_2 (0.30 mg, 3.8 mM, Nanocs Inc.) was incubated with freshly prepared HS-SiNWS in dimethyl sulfoxide (DMSO,200 μL) solution for 2 h to introduce disulfide linkers with terminal amine groups. Then the amine-terminated SiNWS (H_2N-SiNWS) were rinsed with ethanol three times. (iii) To graft Tz motifs, the H_2N-SiNWS was incubated with Tz-sulfo-NHS ester (0.32 mg, 3.8 mM, Click Chemistry Tools Bioconjugate Technology Company) in PBS (200 μL, pH = 8.5) for 1 h. The resulting Tz-grafted SiNWS were rinsed with DI water three times. After drying under nitrogen flow, the Tz-grafted SiNWS were stored at −20 °C.

The overlaid PDMS chaotic mixer was prepared by thermally curing PDMS pre-polymer (Sylgard184) on a Si-based replicate mold, where the herringbone patterns were fabricated by photolithography, followed by inductively coupled plasma etching.

51.3.2 RT-ddPCR

The HCC EVs enriched by EV Click Chips were lysed by QIAzol Lysis Reagent. RNA was extracted using a miRNeasy Micro Kit (Qiagen, USA). Then the complementary DNA (cDNA) was synthesized using a Thermo Scientific Maxima H Minus Reverse Transcriptase Kit. ddPCR experiments were performed on a QX200 system (Bio-Rad Laboratories, Inc.). All primers and probes were purchased from Thermo Fisher Scientific. Data were analyzed using the QuantaSoft™ software to quantify the corresponding copy numbers of gene transcripts detected in each assay.

51.4 HCC EV Surface Protein Assay

The HCC EV surface protein assay (Figure 51.1c) consists of (i) Click Beads for enriching HCC EVs and (ii) real-time immuno-PCR for quantifying subpopulations of HCC EVs.

51.4.1 Click Bead

Click Beads were produced by a three-step chemical modification procedure. (i) The silica microbeads with natural hydroxyl (silica microbeads [MBs], Bangs Laboratories, Inc.) were first treated with nitric acid (2 mol/L, Sigma-Aldrich) for 30 min, followed by washing with DI water and ethanol successively. (ii) The resultant silica MBs were then resuspended in (3-aminopropyl) triethoxysilane/ethanol solution (4%, v/v, Sigma-Aldrich) for 45 min to introduce terminal amine groups. Then the amine-terminated silica microbeads (NH_2-modified MBs) were washed with ethanol three times. (iii) To graft methyltetrazine (mTz) motifs, the NH_2-modified MBs was incubated with methyltetrazine-sulfo- NHS Ester (3.8 mM, Click Chemistry Tools) in PBS (200μL, PH=8.5) for 1 h. The resulting mTz-grafted MBs were washed with DI water three times. After drying under nitrogen flow, the mTz-grafted MBs were stored at 4 °C and retrieved just before use. Four trans-cyclooctene (TCO)-conjugated HCC-associated antibodies (anti-EpCAM, anti-ASGPR1, anti-CD147, and anti-GPC3) were prepared as described earlier.

51.4.2 Real-Time Immuno-PCR

To conduct real-time immuno-PCR, we designed and conjugated DNA barcodes onto the antibodies for targeting CD63 and CD9, which are the common EV surface markers. The DNA barcodes were purchased from IDT and the sequences were provided in our manuscript.[17] N_3-DNA was first produced by incubating DNA barcodes with azido-PEG_4-NHS (Click Chemistry Tools) in the molar ratio of 1:20 at room temperature for 4 h. Surplus azido-PEG_4-NHS was removed using Amicon Ultra−0.5 Centrifugal Filter Unit (10K, Millipore). Next, we prepared the diarylcyclooctyne (DBCO)-antibodies (anti-CD63 and anti-CD9). A buffer exchange was performed using 40K ZebaTM spin desalting columns (Thermo Scientific) to 50 mM borate buffered saline (BBS, pH 8.4). Antibodies were incubated for 45 min with NHS-PEG_5-DBCO (Click Chemistry Tools) in the molar ratio of 1:10 at room temperature. Surplus NHS-PEG_5-DBCO was removed using Amicon Ultra−0.5 Centrifugal Filter Unit (50K, Millipore). Finally, strain-promoted azide-alkyne cycloaddition (SPAAC) was used for functionalization of DBCO-antibodies with N_3-DNA for overnight at 4 °C to conjugate the DNA-antibodies. Conjugation of antibody and DNA was performed with a molar ratio of 1:2, 1:4, and 1:10 in PBS to determine functionalization efficiency. Surplus N_3-DNA was removed using Amicon Ultra−0.5 Centrifugal Filter Unit (50 kDa, Millipore) for two times. DNA-antibodies were stored in a glycerin/PBS (1:1) solution at −20 °C and retrieved just before use.

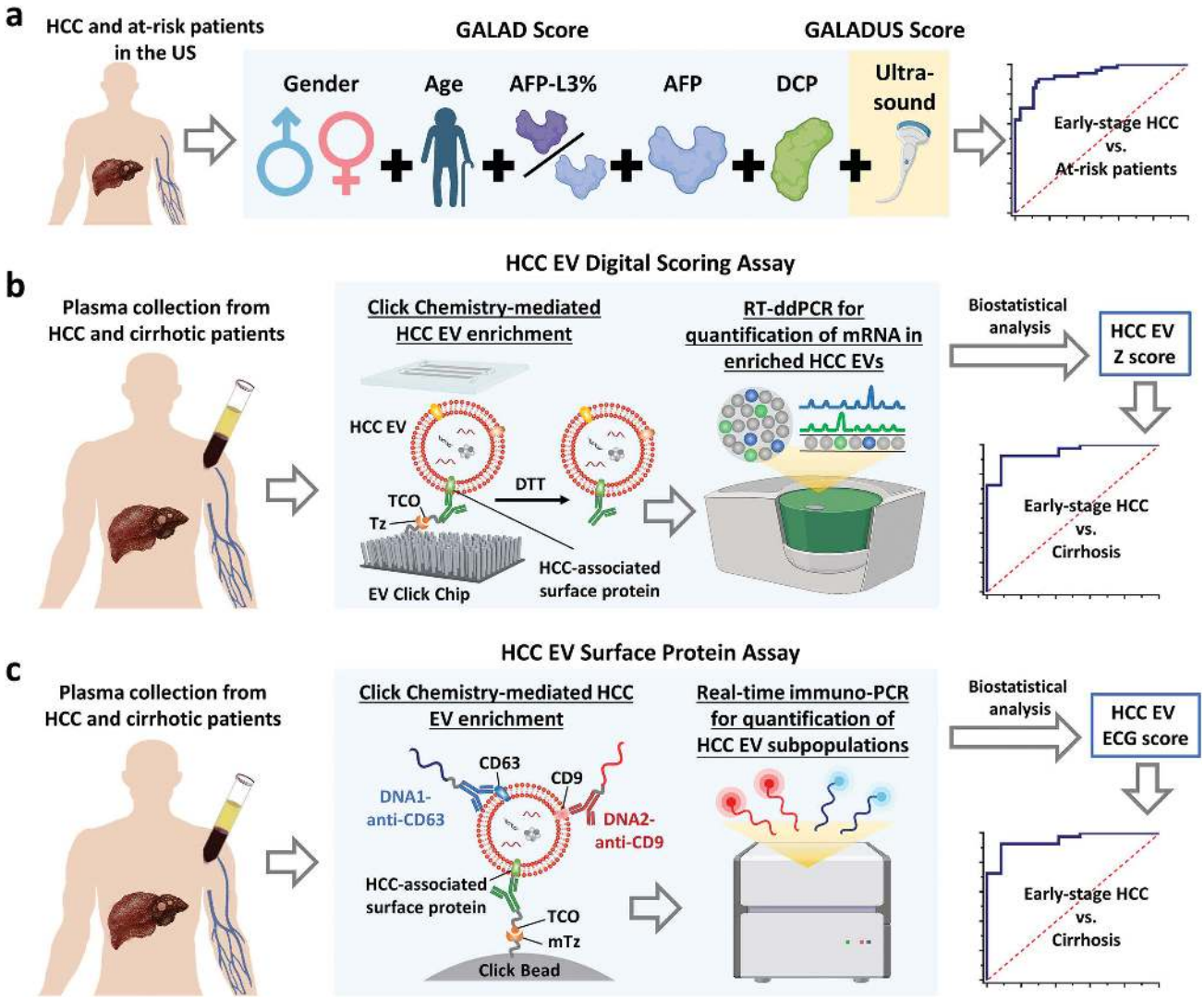


FIGURE 51.1 *The study and assays included in this chapter.* (a) The performance of GALAD score, calculated based on gender, age, AFP-L3%, alpha-fetoprotein (AFP), and des-gamma carboxyprothrombin (DCP), along with GALADUS score (GALAD score + liver ultrasound), for detecting early-stage hepatocellular carcinoma (HCC) are evaluated in the United States (US) populations. (b) HCC EV digital scoring assay is comprised of EV Click Chip and reverse-transcription droplet digital PCR (RT-ddPCR). In conjunction with the use of antibody cocktail targeting HCC-associated surface proteins, EV Click Chip enables the click chemistry-mediated capture of HCC extracellular vesicles. A pair of highly reactive click chemistry motifs, i.e., tetrazine (Tz) and trans-cyclooctene (TCO), are grafted onto EV Click Chip and HCC EVs, respectively, resulting in the immobilization of the HCC EVs. Subsequently, the exposure to 1,4-dithiothreitol (DTT) leads to the cleavage of the embedded disulfide bonds to release the immobilized HCC EVs. The enriched HCC EVs can then be subjected to RT-ddPCR to obtain HCC-specific genes signatures and calculate HCC EV Z scores, which are used for detecting early-stage HCC from cirrhosis. (c) HCC EV surface protein assay is comprised of Click Bead and real-time immuno-PCR. Patient plasma samples are first incubated with TCO-conjugated HCC-associated antibodies, followed by a second incubation with DNA-conjugated EV-specific antibodies: DNA1-anti-CD63 and DNA2-anti-CD9. Click Beads, which are grafted with methyltetrazine (mTz), are then added to the incubated plasma samples for click chemistry-mediated capture of HCC EV subpopulations. On-Bead real-time immuno-PCR is applied to quantify each subpopulation of HCC EVs. The resulting HCC EV surface protein signatures are analyzed to generate the HCC EV ECG score (derived from EpCAM$^+$ CD63$^+$ HCC EVs, CD147$^+$ CD63$^+$ HCC EVs, and GPC3$^+$ CD63$^+$ HCC EVs) for detecting early-stage HCC from cirrhosis.

51.5 Preparation of Artificial Plasma Samples

HepG2 and Hep3B cells were cultured in 18 Nunc EasYDish dishes (145cm^2, Thermo Fisher Scientific) for 72 h. Then the culture medium was switched to serum-free culture medium (Thermo Fisher Scientific) to starve the cells for 24–48 h. The serum-free culture medium incubated with cells was finally collected for EV isolation. After first centrifugation at 300 × *g* (4 °C) for 10 min to remove cells and cell debris, the supernatant was collected and transferred to new tubes and centrifuged at 2800 × *g* (4 °C) for 10 min to further eliminate the remaining cellular debris and large particles. The supernatant was carefully transferred to Ultra-Clear Tubes (38.5 mL, Beckman Coulter, Inc., USA), followed by ultracentrifugation using an

Optima L-100 XP Ultracentrifuge (Beckman Coulter, Inc., USA) at 100,000×*g* (4 °C) for 70 min. The EV pellets at the bottom of the tubes were carefully collected and resuspended in 200 µL fresh PBS. For the preparation of artificial plasma samples, each 10 µL aliquot of EV pellets was spiked into 90 µL healthy donors' plasma or cirrhotic patients' plasma.

51.6 Collection of Patient Plasma Samples

Peripheral venous blood samples were collected from patients or healthy donors with written informed consent from each patient or healthy donor according to the institutional review board (IRB) protocols. A 10.0-mL ethylenediaminetetraacetic acid vacutainer tube (BD Medical, Fisher Cat. #BD 366643-1) was used for blood sample collection. The whole blood was centrifuged at 530 g for 10 min, and the supernatant was retrieved and centrifuged at 4600 g for 10 min. The final supernatant was collected as plasma samples and stored in −80 °C refrigerators. All the plasma samples were retrieved just before being subjected to HCC EV digital scoring assay or HCC EV surface protein assay.

51.7 Methods

51.7.1 Calculation of GALAD Score

GALAD score was calculated based on the following formula[10]: GALAD score = –10.08 + 1.67 × [Gender (1 for male, 0 for female)] + 0.09 × [Age] + 0.04 × [AFP-L3%] + 2.34 × log[AFP] + 1.33 × log[DCP].

51.7.2 Phase 2 Biomarker Studies for Validation of GALAD Score in the US Populations

In our study,[11] GALAD score was validated in two US cohorts: the Mayo test cohort and the multicenter Early Detection Research Network (EDRN) cohort. The Mayo test cohort consisted of 60 patients with early-stage HCC (by Barcelona Clinic Liver Cancer [BCLC] stage 0-A HCC) and 180 patients with liver cirrhosis or chronic hepatitis B, who were identified using the Mayo Clinic's institutional clinical database. The multicenter EDRN cohort included the data from the National Cancer Institute EDRN phase 2 biomarker case–control study,[18] which included 233 early-stage HCC patients and 412 cirrhosis controls. Control patients were confirmed to not have HCC by either negative contrast-enhanced multiphasic CT, MRI, or liver biopsy at the time of GALAD score measurement, or at least six months of follow-up after GALAD score measurement.

51.8 HCC EV Digital Scoring Assay (EV Click Chip + RT-ddPCR) for Early Detection of HCC

HCC EV Digital scoring assay is implemented through a two-step workflow (Figure 51.1b).

51.8.1 Step 1: Enrichment of HCC EVs by EV Click Chip

A total of 500 µL of plasma for each patient was incubated with the antibody cocktail, which comprised of TCO-anti-EpCAM (50 ng), TCO-anti-ASGPR1 (50 ng), and TCO-anti-CD 147 (25 ng), for 30 min at room temperature. After incubation, the sample was injected into EV Click Chip microfluidic devices at 1.0 mL/h. To release the captured HCC EVs, 100 µL 1,4-dithiothreitol (DTT) solution (50 mM) was injected into the EV Click Chips at 1.0 mL/h and the released EVs were collected in 1.5 mL RNase-free Eppendorf tubes for subsequent RNA extraction.

51.8.2 Step 2: Absolute Quantification of 10 HCC-Specific mRNA Markers by RT-ddPCR

The HCC EVs enriched from EV Click Chips were lysed by 700 µL QIAzol Lysis Reagent. RNA was extracted using a miRNeasy Micro Kit (Qiagen, USA) according to the manufacturer's instructions. Then the complementary DNA (cDNA) was synthesized using a Thermo Scientific Maxima H Minus Reverse Transcriptase Kit according to the manufacturer's instructions. Ten µL of total cDNA (12 µL) was divided into five tubes to detect the ten HCC-specific genes with two fluorescence filters in each tube. DdPCR experiments were performed on a QX200 system (Bio-Rad Laboratories, Inc.) and analyzed by QuantaSoft™ software.

51.9 HCC EV Surface Protein Assay (Click Bead + Real-Time Immuno-PCR) for Early Detection of HCC

HCC EV surface protein assay is implemented through a three-step workflow (Figure 51.1c). An internal control, fetal bovine serum (FBS), was used instead of plasma samples to measure the background binding of the DNA1-anti-CD63 and DNA2-anti-CD9 to Click Beads.

51.9.1 Step 1: Sequentially Labeling Each Subpopulation of HCC EVs in Plasma

A total of 400 µL of plasma for each patient were evenly aliquoted into four Eppendorf tubes. In the first incubation, these four 100µL plasma aliquots were incubated with the respective TCO-conjugated HCC-associated antibodies: TCO-anti-EpCAM (50 ng), TCO-anti-CD147 (25 ng), TCO-anti-GPC3 (50 ng), and TCO-anti-ASGPR1 (50 ng), respectively. In the second incubation, these aliquots were cocktailed with DNA-conjugated EV-specific antibodies: DNA1-anti-CD63 and DNA2-anti-CD9, at room temperature for 30 min. No TCO-conjugated HCC-associated antibody was added into the internal control, which was only incubated with DNA1-anti-CD63 and DNA2-anti-CD9.

51.9.2 Step 2: Click Chemistry-Mediated Capture of Subpopulations of HCC EVs onto Click Beads

Click Beads[19] were blocked with protein free buffer and then 5% bovine serum albumin (BSA) solution before use. The plasma samples labeled with TCO-antibodies and DNA barcodes were then incubated with the Click Beads (0.1 mg) for 30 min for isolation of each subpopulation of EVs. Click Beads with isolated EVs were collected by centrifuge at 10,000 g for 2 min. All plasma samples were subjected to Click Beads underwent only one freeze-thaw cycle. The internal control was processed with the same protocol.

51.9.3 Step 3: On-Bead Duplex Real-Time Immuno-PCR for Quantifying Each Subpopulation of HCC EVs

Click Beads with isolated EVs were resuspended in 500 μL PBS with 0.2% BSA and then collected by centrifuge at 10,000g for 2 min. After three-time washing step with PBS with 0.2% BSA, Click Beads were then washed by PBS twice to wash off unbounded DNA1-anti-CD63 and DNA2-anti-CD9. After five-time washing steps, 90 μL double-distilled water was added to the sediments of each well and mixed thoroughly. Nine μL of the mixed suspension was loaded into the real-time PCR system (CFX Connect™ Real-Time PCR Detection System, Bio-Rad). Raw data of real-time PCR runs were evaluated using the corresponding software system (Bio-Rad CFX Manager 3.1, Bio-Rad). The relative signal was calculated using the following equation:

$$\text{Relative signal} = 2^{-(cq(sample) - cq(\text{internal control}))}$$

51.9.4 Phase 2 Biomarker Studies for Clinical Validation of HCC EV Digital Scoring Assay and HCC EV Surface Protein Assay

After optimization of the HCC EV digital scoring assay and HCC EV surface protein assay, we performed phase 2 (case-control) biomarker studies, respectively, to validate the performance of these assays for early detection of HCC. To avoid overemphasizing the assays' performance, we reported the assays' performance for early HCC detection by restricting the case group to patients with treatment naïve early-stage HCC (defined by BCLC stage 0-A HCC) and the control group to patients with liver cirrhosis, who are required to receive semi-annual ultrasound surveillance for HCC according to the clinical practice guideline.

51.10 Statistical Analysis

The area under the receiver operator characteristic (AUROC) was used to assess the performance of GALAD score, HCC EV digital scoring assay, and HCC EV surface protein assay for detection of early-stage HCC. Youden's index was used to find the optimal cutoffs for each score/assay. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at the optimal cutoff were reported.

For the study evaluating the performance of GALAD score in the context of standard surveillance test with liver ultrasound, a GALADUS score was developed through combining GALAD and liver ultrasound using multivariable logistic regression model that included all variables in the GALAD score and the binary result of liver ultrasound.

For the study evaluating the performance of HCC EV digital scoring assay, an HCC EV Z score was calculated by incorporating the expression of the ten mRNA markers measured by HCC EV digital scoring assay using weighted Z-score method.[20]

For the study evaluating the performance of HCC EV surface protein assay, an HCC EV ECG score was calculated by combining the readouts of three EV populations associated with HCC (see Results section) using a multivariable logistic regression model.

51.11 Results

51.11.1 GALAD Score for Early Detection of HCC in the US Populations

In the Mayo test cohort, the AUROC of GALAD score for detecting early-stage HCC (BCLC stage 0-A) from the controls was high at 0.92 (95% confidence interval [CI] = 0.88–0.96; best cutoff = 1.18, sensitivity = 92%, specificity = 79%, Figure 51.2a). In addition, the AUROC of the GALAD score was higher than that of ultrasound for detecting early-stage (BCLC stage 0-A, 0.92 vs. 0.82, $P < 0.01$) and very early-stage HCC (BCLC stage 0, 0.91 vs. 0.78, p<0.001). The performance of the GALAD score remains excellent when limiting the early-stage HCC and control cases to those with liver cirrhosis (AUROC = 0.91, 95% CI = 0.86–0.95) or after frequency matching of the Child-Pugh score between HCC cases and controls (AUROC = 0.89, 95% CI = 0.83–0.94). When applying the best cutoff reported in the previous study at −0.63, GALAD score's performance was still excellent with an AUROC of 0.89 (95% CI = 0.82–0.94). The sensitivity and specificity were 82% and 86%, respectively, for early-stage HCC detection, and 73% and 86%, respectively, for very early-stage HCC detection.

In the multicenter EDRN cohort, the AUROC of GALAD score for early-stage HCC detection was 0.88 (95% CI = 0.85–91). At a cutoff of −0.17, the GALAD score had a sensitivity of 76% and a specificity of 86%. The AUROC of GALAD score for detecting very early-stage HCC was 0.86 (95% CI = 0.76–96). Importantly, the AUROC of GALAD score remained excellent in all the subgroups stratified by common HCC etiologies (Figure 51.2b). When the cutoff was set at –0.63, the best cutoff reported in the previous study, the GALAD score had a sensitivity of 79% and a specificity of 79% for HCC detection.

Lastly, we evaluated if combining GALAD score with liver ultrasound could further enhance the performance in detecting early-stage HCC compared to either test alone in the Mayo test

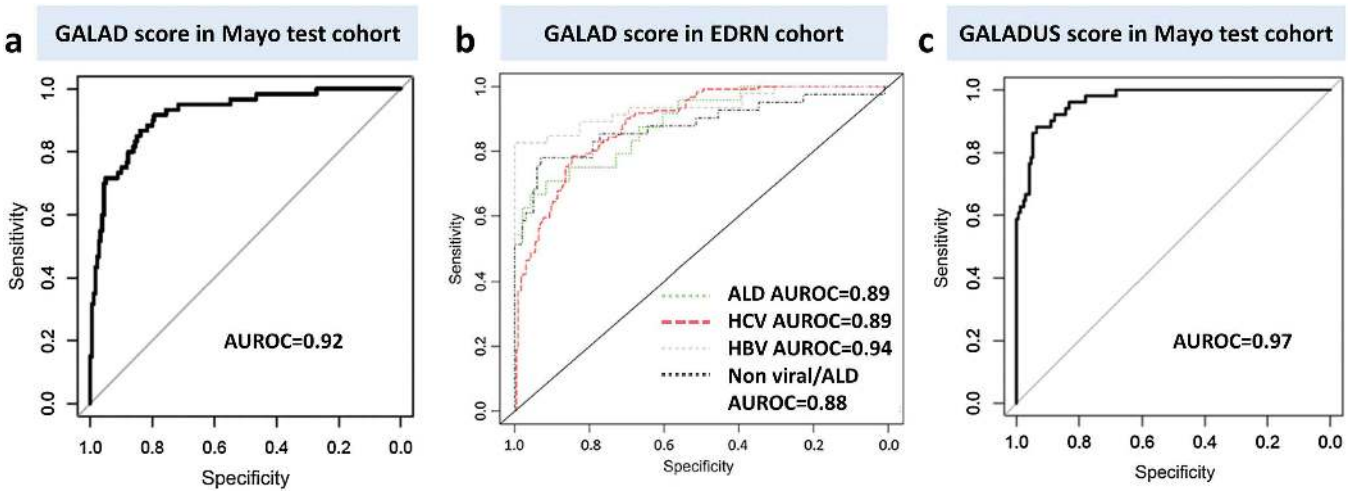


FIGURE 51.2 The performance of GALAD score for detecting early-stage HCC in (a) the Mayo test cohort and (b) the subgroups in the multicenter EDRN cohort. (c) The performance of GALADUS score (GALAD score + liver ultrasound) in the Mayo test cohort.

ALD, alcoholic liver disease; AUROC, area under the receiver operating characteristic curve; EDRN, Early Detection Research Network; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus.

cohort. The GALADUS score was calculated as: GALADUS score = –12.79 + 1.74 × [Gender (1 for male, 0 for female)] + 0.09 × [Age] + 0.04 × [AFP-L3%] + 2.44 × log[AFP] + 1.39 × log[DCP] + 3.56 × [Ultrasound result (1 for positive ultrasound, 0 for negative ultrasound)]. The AUROC of GALADUS score was high at 0.97 (95% CI = 0.95–0.99; best cutoff = 0.5, sensitivity = 88%, specificity = 94%; Figure 51.2c), which was significantly higher than GALAD score (AUROC = 0.92) and ultrasound (AUROC = 0.82).

51.11.2 HCC EV Digital Scoring Assay (EV Click Chip + RT-ddPCR) for Early Detection of HCC

EV Click Chip was first validated using the artificial plasma samples prepared by spiking HCC cell-derived EVs into plasma from healthy donors or patients with liver cirrhosis. We first tested the recovery yield of the EVs using individual TCO-conjugated antibodies (TCO-anti-EpCAM, TCO-anti-CD147, TCO-anti-GPC3, and TCO-anti-ASGPR1) and the combination of different TCO-conjugated antibodies (antibody cocktail). We found that the combination of TCO-anti-EpCAM, TCO-anti-CD147, and TCO-anti-ASGPR1 resulted in the highest recovery yield, which outperformed any single antibodies or other combinations. After determining the optimal sample flow rate 1.0 mL/h for EV Click Chips, we further compared the recovery yield and purity of HCC EVs between EV Click Chips and ultracentrifugation, the commonly used method for collection of EVs, using the prepared artificial plasma samples. The recovery yield of EV Click Chip was 82.7 ± 1.34% and the purity of EVs was 90.2 ± 6.2%, which were significantly higher than those obtained from ultracentrifugation. In terms of reproducibility of EV Click Chip, the percent coefficient of variation (%CV) was 1.12%–12.65% for the intra-assay variability and 3.88% for inter-assay variability, suggesting its stable performance in enriching HCC EVs.

After optimizing EV Click Chips and confirming the reproducibility, we proposed an HCC EV digital scoring assay by coupling EV Click Chips and RT-ddPCR for quantification of ten well-validated HCC-specific mRNA markers (i.e., *AFP*, *GPC3*, *ALB*, *APOH*, *FABP1*, *FGB*, *FGG*, *AHSG*, *RBP4*, and *TF*) for early detection of HCC (Figure 51.1b). For this, we collected 158 plasma samples from five cohorts, including (i) patients with treatment naïve HCC ($n = 46$), (ii) patients with liver cirrhosis ($n = 26$), (iii) patients with chronic hepatitis B or C without liver cirrhosis ($n = 25$), (iv) healthy donors ($n = 23$), and (v) patients with primary malignancies other than HCC, with or without liver metastases ($n = 38$). As depicted in the heatmaps (Figure 51.3a), higher signals were observed in patients with HCC, compared with those from the noncancer cohorts (i.e., cirrhosis, hepatitis, and healthy donors) and patients with other cancers. Furthermore, signals differed between early-stage (BCLC stage 0-A) and advanced-stage (BCLC stage B-C) HCC patients, highlighting the potential of this assay for reflecting the tumor burden.

Next, we used the weighted Z-score method to incorporate the expression of the ten mRNA markers measured by HCC EV digital scoring assay into a score named HCC EV Z score. As shown in the box plot (Figure 51.3b), the HCC EV Z score of patients with HCC (both early-stage and advanced-stage HCC) was significantly higher ($P<0.001$) than those of individuals with cirrhosis, chronic hepatitis without cirrhosis, healthy donors, or other cancers. Finally, we examined the potential of HCC EV Z score to distinguish early-stage HCC (BCLC stage 0-A) from at-risk liver cirrhosis. The AUROC was 0.93 (95% CI = 0.86–1.00; sensitivity = 94%, specificity = 89%, Figure 51.3c), which significantly outperformed the most widely used serum AFP, the only HCC biomarker being validated through the five phases of biomarker development (AUROC = 0.69, 95% CI = 0.55–0.83; sensitivity = 66.7%, specificity = 72.0%) for detection of early-stage HCC.

51.11.3 HCC EV Surface Protein Assay (Click Bead + Real-Time Immuno-PCR) for Early Detection of HCC

Similar to HCC EV digital scoring assay, the reproducibility of HCC EV surface protein assay was excellent after optimization. The optimized HCC EV surface protein assay was utilized to obtain the HCC EV surface protein signatures of the UCLA cohort consisting of 45 patients with early-stage HCC and 61 patients with liver cirrhosis. As mentioned in the Methods section, four HCC-associated antibodies (i.e., anti-EpCAM, anti-CD147, anti-GPC3, and anti-ASGPR1) were applied to enrich the EVs and two conjugated DNA-antibodies (i.e., DNA1-anti-CD63 and DNA2-anti-CD9) targeting EV markers were utilized to indicate the presence of enriched HCC EVs. As a result, we obtained the readouts of eight subpopulations of HCC EVs (i.e., $EpCAM^+$ $CD63^+$ HCC EVs, $CD147^+$ $CD63^+$ HCC EVs, $GPC3^+$ $CD63^+$ HCC EVs, $ASGPR1^+$ $CD63^+$ HCC EVs, $EpCAM^+$ $CD9^+$ HCC EVs, $CD147^+$ $CD9^+$ HCC EVs, $GPC3^+$ $CD9^+$ HCC EVs, and $ASGPR1^+$ $CD9^+$ HCC EVs). The patients with early-stage HCC had significantly higher signals of these eight EV subpopulations than those with liver cirrhosis ($P<0.005$ in all the EV subpopulations), indicating these EV subpopulations could be potential biomarkers for HCC.

Next, we selected the top three HCC EV subpopulations strongly associated with HCC (i.e., $EpCAM^+$ $CD63^+$ HCC EVs, $CD147^+$ $CD63^+$ HCC EVs, and $GPC3^+$ $CD63^+$ HCC EVs) to establish a model named HCC EV ECG score for early detection of HCC using logistic regression. HCC EV ECG score is defined as:

$$\text{HCC EV ECG score} = -9.54338 + 0.13544 \times [EpCAM^+ CD63^+ \text{ HCC EVs}] + 0.35729 \times [CD147^+ CD63^+ \text{ HCC EVs}] + 0.37513 \times [GPC3^+ CD63^+ \text{ HCC EVs}]$$

HCC EV ECG score showed excellent accuracy for distinguishing patients with early-stage HCC from cirrhotic controls in the UCLA training cohort with an AUROC of 0.95 (95% CI = 0.90–0.99; Figure 51.4a). At the optimal cutoff of −0.40, the sensitivity and specificity of HCC EV ECG score for early-stage HCC detection were 91% (95% CI = 79%–96%) and 90% (95% CI = 80%–95%), respectively, with the PPV of 87% (95% CI = 76%–94%) and NPV of 93% (95% CI = 84%–97%). We also demonstrated that HCC EV ECG scores from patients with early-stage HCC were significantly higher than those from patients with BCLC stage B-C HCC ($n = 6$),

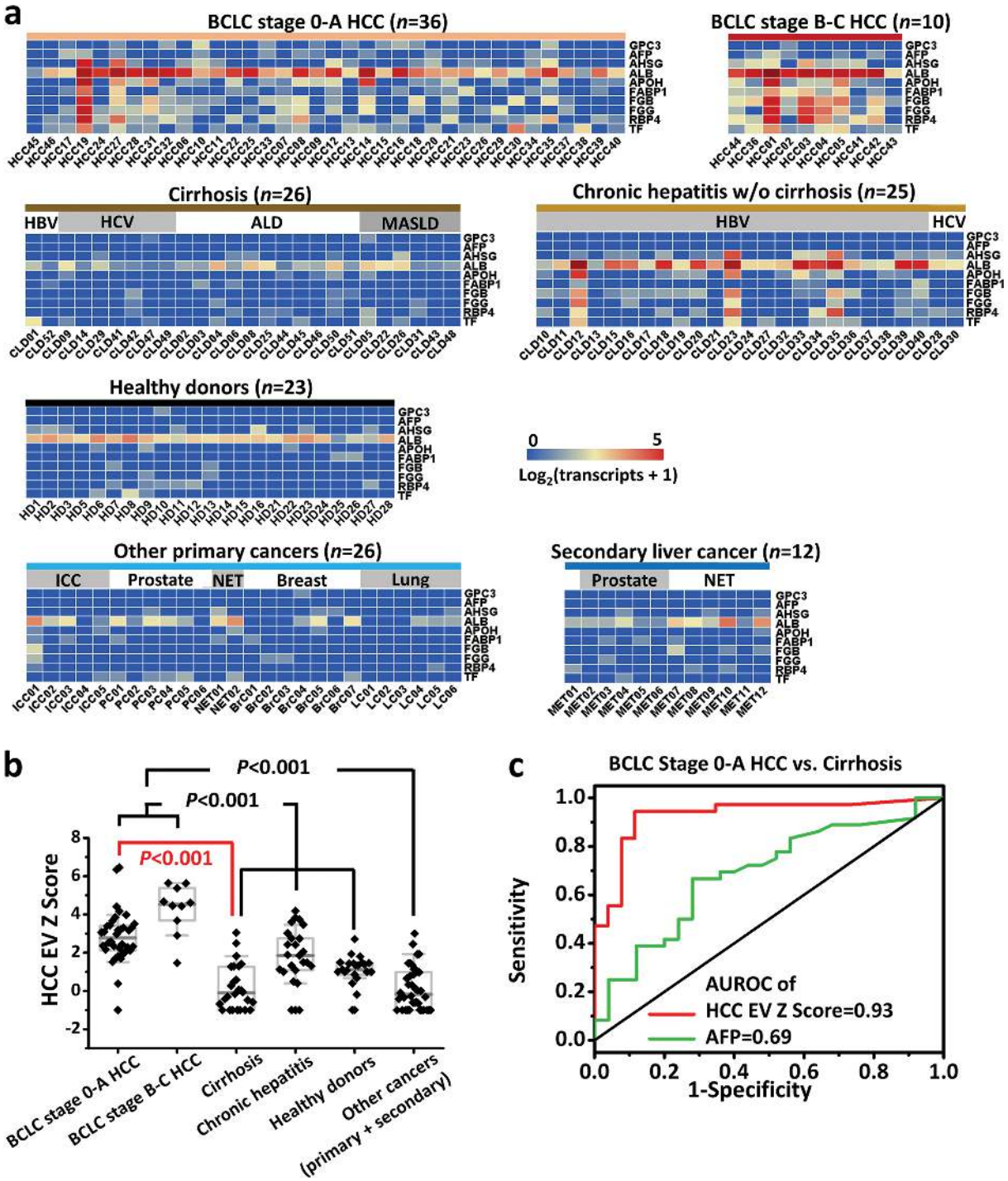


FIGURE 51.3 HCC EV digital scoring assay for enrichment of HCC EVs and quantification of 10 HCC-specific mRNA markers in the enriched HCC EVs. (a) Heatmaps depicting relative signal intensities of the ten HCC-specific genes across different patient cohorts. (Upper row) Patients with treatment naïve HCC are grouped according to BCLC staging system from early (BCLC stage 0-A, n = 36) to intermediate and advanced stages (BCLC stage B-C, n = 10). (Second row) Patients with liver cirrhosis (n = 26) and chronic hepatitis without cirrhosis (n = 25). (Third row) Healthy donors (n = 23). (Last row) Patients with primary cancers other than HCC (n = 26) and secondary liver cancers (n = 12). Primary copy numbers are $\log_2$-transformed and scaled to the highest value for each gene. (b) Box plots showing that the HCC EV Z scores, which were calculated based on the expression of the ten mRNA markers, are capable of distinguishing HCC (n = 46) from at-risk chronic liver diseases (n = 51) covering both liver cirrhosis (n = 26) and chronic hepatitis (n = 25), healthy donors (n = 23), as well as other cancers (other primary cancers + secondary liver cancers, n = 38). (c) The performance of HCC EV Z score and serum AFP for differentiating early-stage HCC (BCLC stage 0-A, n = 36) from at-risk liver cirrhosis (n = 26).

AFP, alpha-fetoprotein; ALD, alcoholic liver disease; AUROC, area under the receiver operating characteristic curve; BCLC, Barcelona Clinic Liver Cancer; EDRN, Early Detection Research Network; EV, extracellular vesicle; HBV, hepatitis B virus; HCC, hepatocellular carcinoma; HCV, hepatitis C virus; ICC, intrahepatic cholangiocarcinoma; MASLD, metabolic dysfunction–associated steatotic liver disease; NET, midgut neuroendocrine tumor.

breast cancer, prostate cancer, or thyroid cancer (P<0.001, respectively), confirming the score was specific to HCC.

To further validate HCC EV ECG score for detecting early-stage HCC, a total of 35 patients with early-stage HCC and 37 at-risk patients with liver cirrhosis recruited at Cedars-Sinai Medical Center (CSMC) were utilized as an independent validation cohort. Similarly, the readouts of the three EV subpopulations included in HCC EV ECG score (i.e., $EpCAM^+$ $CD63^+$ HCC EVs, $CD147^+$ $CD63^+$ HCC EVs, and $GPC3^+$ $CD63^+$ HCC EVs) were significantly higher in patients with HCC than patients with liver cirrhosis (P<0.001). In this validation cohort, the AUROC of HCC EV ECG score for detecting early-stage HCC remained excellent as 0.93 (95% CI = 0.87–0.99; Figure 51.4b). Using the same cutoff value derived from the UCLA training cohort (−0.40), HCC EV ECG score had great accuracy of detecting early-stage HCC with sensitivity of 91% (95% CI = 77%–98%), specificity of 81% (95% CI = 65%–92%), PPV of 82% (95% CI = 70%–90%) and NPV of 91% (95% CI = 76%–93%).

We also compared the performance of HCC EV ECG score with serum AFP for detecting early-stage HCC in all the participants in this study (UCLA+CSMC cohorts, n = 172). As demonstrated in Figure 51.4c, HCC EV ECG score outperformed serum AFP in distinguishing early-stage HCC from cirrhosis (AUROC = 0.94 [95% CI = 0.90–0.97] vs. 0.79 [95% CI = 0.72–0.86], P<0.001). Lastly, we showed the performance of HCC EV ECG score remained excellent in the subpopulations stratified by etiology (viral vs. non-viral). The AUROC was 0.95 (95% CI = 0.90–1.00; Figure 51.4d) for patients with viral hepatitis, 0.97 (95% CI = 0.91–1.00; Figure 51.4e) for patients with metabolic dysfunction–associated steatotic liver disease (MASLD), and 0.95 (95% CI = 0.90–1.00; Figure 51.4f) for patients with alcoholic liver disease (ALD).

51.12 Discussion

Here, we reported the performance of GALAD score, a serum biomarker-based statistical model, for the detection of early-stage HCC in at-risk patients in two large US cohorts. Our joint research team also developed two EV-based assays for quantifying mRNAs and surface proteins of HCC EVs, respectively, and evaluated their performance in detecting early-stage HCC from cirrhosis in phase 2 biomarker studies. Overall, these biomarkers hold great promise for improving the early detection of HCC.

GALAD score was first developed using the United Kingdom cohorts for detection of HCC in patients with chronic liver diseases.[10] Its performance remained stable and excellent among different races/ethnicities and HCC etiologies in the following validation studies, including ours, confirming the GALAD score's ability to distinguish early-stage HCC from at-risk controls in the US population.[11] Recently, our team conducted a meta-analysis and showed that the AUROC of the GALAD score for detecting early-stage HCC was 0.83 and 0.73, respectively, in phase 2 and phase 3 biomarker studies.[9] Although promising, the results suggested that its performance may be overestimated in phase 2 case-control studies. Currently, there are only two phase 3 biomarker studies conducted in the US.[21,22] More phase 3 or 4 studies in other geographic regions are thus required before its clinical application.

For our EV-based assays (HCC EV digital scoring assay and HCC EV surface protein assay), the most unique feature is the utilization of the click chemistry-mediated EV enrichment. Compared to the conventional immunoaffinity-based EV capture approaches, which are driven by the dynamic binding between a pair of antigens (on EVs) and antibodies (on the substrates), click chemistry-mediated methods have several advantages. First, it provides a rapid and stable reaction. The inverse-electron-demand Diels–Alder cycloaddition[23] between Tz/mTz and TCO motifs (a rate constant[24] of $10^4/M^{-1}s^{-1}$) enables balanced chemical properties without the presence of a catalyst. In addition, the ligation between Tz-grafted chips/mTz-grafted beads and TCO-grafted EVs is specific, irreversible, and insensitive to biomolecules, water, and oxygen, leading to immobilization of the EVs with improved capture efficiency and reduced nonspecific trapping of particles in the background. Our studies have demonstrated superior performance in HCC EV enrichment using click chemistry-mediated methods than immunoaffinity-based approaches.

Considering the heterogeneity of HCC cells and their derived EVs, we adopted different strategies to target the HCC EVs. In the first study using HCC EV digital scoring assay, we applied an antibody cocktail, including TCO-anti-EpCAM, TCO-anti-CD147, and TCO-anti-ASGPR1, which aimed to cover most of the HCC EVs. On the other hand, our second study using HCC EV surface protein assay sought to dissect the different populations of HCC EVs and investigate their association with HCC. Each approach provides its own strengths. For example, the antibody cocktail captures a larger amount of HCC EVs, enabling downstream molecular analysis that could potentially provide further biological insights. Moreover, the subpopulations of HCC EVs can serve as biomarkers on their own. Profiling mRNA or other markers in these EV subpopulations, although challenging due to their scarcity, may lead to the discovery of more specific HCC biomarkers and warrant future investigation.

In addition to the innovation of our EV enrichment technologies, adherence to the International Liver Cancer Association (ILCA) White Paper on HCC biomarker development[25] allows us to rigorously evaluate the performance of our assays for detection of early-stage HCC. In these phase 2 case-control biomarker studies, we restricted the patients with HCC to BCLC stage 0-A and the controls to those with liver cirrhosis, which reduces the risk of overestimating the assays' performance and reflects the actual clinical scenario using the biomarkers. Given the dynamic changes in HCC etiologies from chronic viral hepatitis to MASLD, a rapidly increasing disease worldwide, it is noteworthy that the accuracy of HCC EV surface protein assay remained stable across different underlying liver diseases including those with MASLD.

There are some limitations in our EV-based studies. First, the sample sizes of both studies are limited to the two academic medical centers. Larger phase 2 biomarker studies are required for further validation of the results. Second, given the retrospective design of a phase 2 biomarker study, HCC patients were detected by the current surveillance tool, liver ultrasound with/without serum AFP, and diagnosed before

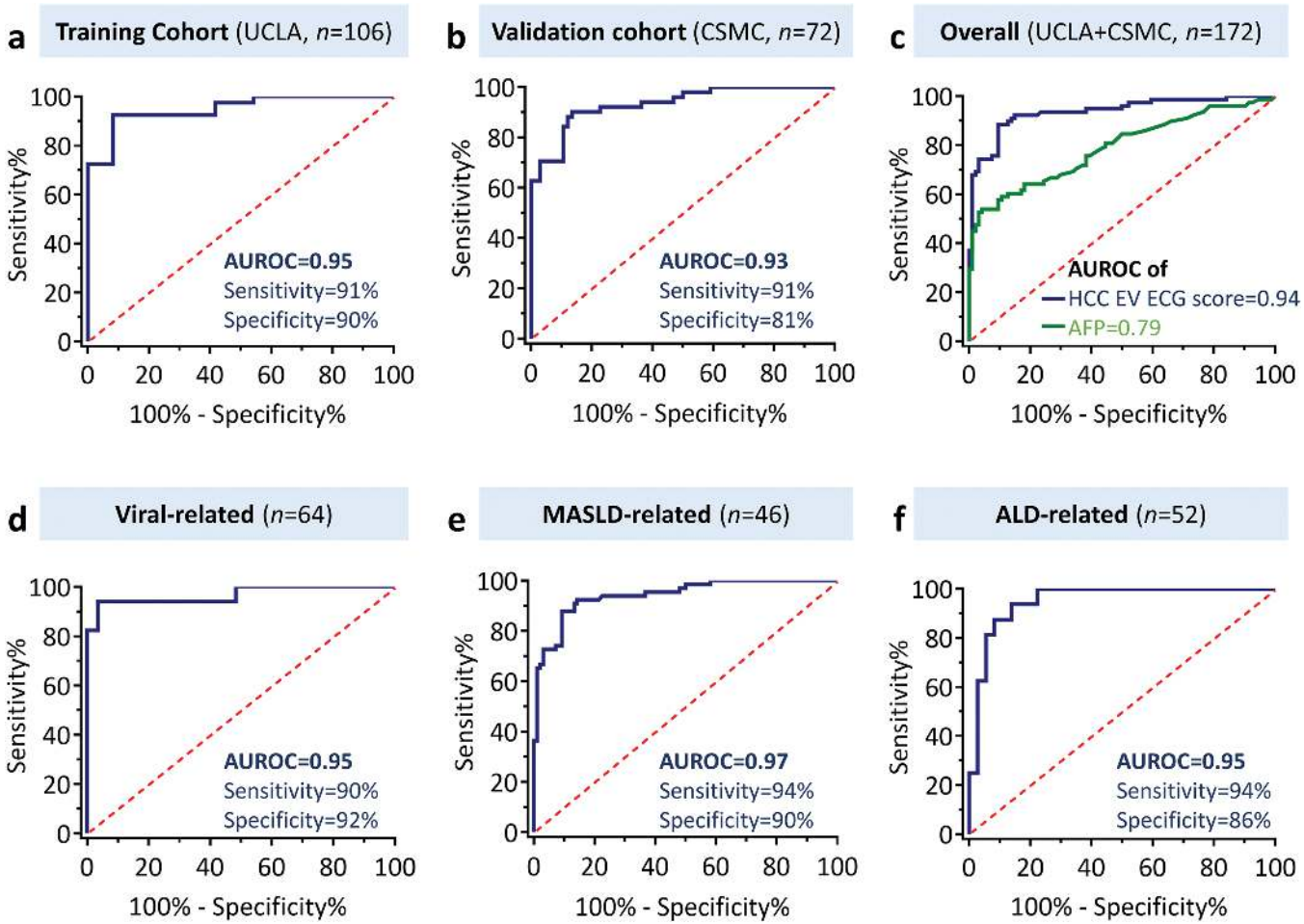


FIGURE 51.4 The performance of HCC EV surface protein assay and the resultant HCC EV ECG score for detecting early-stage HCC from liver cirrhosis in (a) the training UCLA cohort; (b) the validation Cedars-Sinai Medical Center (CSMC) cohort; (c) the combined UCLA + CSMC cohort with comparison to serum AFP; (d) the subgroup of patients with chronic viral hepatitis; (e) the subgroup of patients with metabolic dysfunction–associated steatotic liver disease (MASLD); and (f) the subgroup of patients with alcoholic liver disease (ALD).

AFP, alpha-fetoprotein; ALD, alcoholic liver disease; AUROC, area under the receiver operating characteristic curve; CSMC, Cedars-Sinai Medical Center; HCC, hepatocellular carcinoma; MASLD, metabolic dysfunction–associated steatotic liver disease.

enrollment. As such, it is not feasible to directly compare the performance of the assays with liver ultrasound with/without serum AFP for detecting early-stage HCC. While we reported the comparison between serum AFP and our assays, performance of standard surveillance test (AFP or ultrasound) might have been overestimated. Finally, despite adherence to the ILCA guideline, there could be some potential biases in a phase 2 case-control study. Subsequent phase 3 and 4 biomarker studies will address this concern and confirm the ability of GALAD, HCC EV digital scoring assay, and HCC EV surface protein assay for detecting early-stage HCC.

In conclusion, GALAD score and our EV-based assays hold great promise for improving the identification of HCC at a curable stage and enhancing patients' long-term outcomes. Further validation in a larger phase 2 multicenter biomarker study and subsequent phase 3–4 studies are essential to confirm their utility in clinical settings.

REFERENCES

1. Yang JD, Hainaut P, Gores GJ, et al. A Global View of Hepatocellular Carcinoma: Trends, Risk, Prevention and Management. *Nat Rev Gastroenterol Hepatol* 2019; 16:589–604.
2. Lee YT, Singal AG, Luu M, et al. Increasing Incidence of Intrahepatic Cholangiocarcinoma Relative to Hepatocellular Carcinoma in the United States. *Gastro Hep Adv* 2022;1: 121–124.
3. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71:209–249.
4. Lee YT, Wang JJ, Luu M, et al. State-Level HCC Incidence and Association with Obesity and Physical Activity in the United States. *Hepatology* 2021;74:1384–1394.
5. European Association for the Study of the Liver. Electronic Address EEE, European Association for the Study of the L. EASL Clinical Practice Guidelines: Management of Hepatocellular Carcinoma. *J Hepatol* 2018;69:182–236.
6. Singal AG, Llovet JM, Yarchoan M, et al. AASLD Practice Guidance on Prevention, Diagnosis, and Treatment of Hepatocellular Carcinoma. *Hepatology* 2023;78:1922–1965.
7. Tzartzeva K, Obi J, Rich NE, et al. Surveillance Imaging and Alpha Fetoprotein for Early Detection of Hepatocellular Carcinoma in Patients with Cirrhosis: A Meta-Analysis. *Gastroenterology* 2018;154:1706–1718.e1.
8. Wolf E, Rich NE, Marrero JA, et al. Use of Hepatocellular Carcinoma Surveillance in Patients with Cirrhosis: A Systematic Review and Meta-Analysis. *Hepatology* 2021;73:713–725.
9. Lee YT, Fujiwara N, Yang JD, et al. Risk Stratification and Early Detection Biomarkers for Precision HCC Screening. *Hepatology* 2023;78:319–362.
10. Johnson PJ, Pirrie SJ, Cox TF, et al. The Detection of Hepatocellular Carcinoma Using a Prospectively Developed and Validated Model Based on Serological Biomarkers. *Cancer Epidemiol Biomarkers Prev* 2014;23:144–153.
11. Yang JD, Addissie BD, Mara KC, et al. GALAD Score for Hepatocellular Carcinoma Detection in Comparison with Liver Ultrasound and Proposal of GALADUS Score. *Cancer Epidemiol Biomarkers Prev* 2019;28:531–538.
12. van Niel G, D'Angelo G, Raposo G. Shedding Light on the Cell Biology of Extracellular Vesicles. *Nat Rev Mol Cell Biol* 2018;19:213–228.
13. van Niel G, D'Angelo G, Raposo G. Shedding Light on the Cell Biology of Extracellular Vesicles. *Nat Rev Mol Cell Biol* 2018;19:213–228.
14. Sun N, Lee YT, Zhang RY, et al. Purification of HCC-Specific Extracellular Vesicles on Nanosubstrates for Early HCC Detection by Digital Scoring. *Nat Commun* 2020;11:4489.
15. Lee YT, Sun N, Kim M, et al. Circulating Tumor Cell-Based Messenger RNA Scoring System for Prognostication of Hepatocellular Carcinoma: Translating Tissue-Based Messenger RNA Profiling into a Noninvasive Setting. *Liver Transpl* 2022;28:200–214.
16. Wang JJ, Sun N, Lee YT, et al. Prostate Cancer Extracellular Vesicle Digital Scoring Assay—a Rapid Noninvasive Approach for Quantification of Disease-Relevant mRNAs. *Nano Today* 2023;48.
17. Sun N, Zhang C, Lee YT, et al. HCC EV ECG Score: An Extracellular Vesicle-Based Protein Assay for Detection of Early-Stage Hepatocellular Carcinoma. *Hepatology* 2023;77:774–788.
18. Marrero JA, Feng Z, Wang Y, et al. Alpha-Fetoprotein, Des-Gamma Carboxyprothrombin, and Lectin-Bound Alpha-Fetoprotein in Early Hepatocellular Carcinoma. *Gastroenterology* 2009;137:110–118.
19. Sun N, Tran BV, Peng Z, et al. Coupling Lipid Labeling and Click Chemistry Enables Isolation of Extracellular Vesicles for Noninvasive Detection of Oncogenic Gene Alterations. *Adv Sci* 2022; In press.
20. Levine DM, Haynor DR, Castle JC, et al. Pathway and Gene-Set Activation Measurement from mRNA Expression Data: The Tissue Distribution of Human Pathways. *Genome Biol* 2006;7:R93.
21. Tayob N, Kanwal F, Alsarraj A, et al. The Performance of AFP, AFP-3, DCP as Biomarkers for Detection of Hepatocellular Carcinoma (HCC): A Phase 3 Biomarker Study in the United States. *Clin Gastroenterol Hepatol* 2023;21:415–423.e4.
22. Singal AG, Tayob N, Mehta A, et al. GALAD Demonstrates High Sensitivity for HCC Surveillance in a Cohort of Patients with Cirrhosis. *Hepatology* 2022;75:541–549.
23. Blackman ML, Royzen M, Fox JM. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels-Alder Reactivity. *J Am Chem Soc* 2008;130:13518–13519.
24. Karver MR, Weissleder R, Hilderbrand SA. Synthesis and Evaluation of a Series of 1,2,4,5-Tetrazines for Bioorthogonal Conjugation. *Bioconjug Chem* 2011;22:2263–2270.
25. Singal AG, Hoshida Y, Pinato DJ, et al. International Liver Cancer Association (ILCA) White Paper on Biomarker Development for Hepatocellular Carcinoma. *Gastroenterology* 2021;160:2572–2584.

52

Using Epidemiologic Research Methods to Adapt a Human Papillomavirus Assay for Cervical Cancer Screening in Low- and Middle-Income Countries

Louise Kuhn, Rakiya Saidu, and Lynette Denny

52.1 Introduction

52.1.1 Statement of the Problem

Cervical cancer remains a major threat to the health and well-being of women globally.[1] Largely a result of the failure of organized cervical cancer screening programs, there are gross disparities in incidence and mortality of cervical cancer in low- and middle-income countries (LMIC) compared to north American and Europe.[1] These disparities are exacerbated in sub-Saharan Africa due to the high prevalence of human immunodeficiency virus (HIV) among women in these regions.[2] Women living with HIV are at high risk of developing cervical cancer and its precursor lesions.[2]

Since the licensed vaccines against human papillomavirus (HPV), the etiologic agent of cervical cancer, are effective only for primary prevention of HPV infection (which is why a young age prior to usual initiation of sexual activity is chosen as the target age for the vaccine), secondary prevention through screening remains the only means of reducing the burden of cervical cancer for current generations of women.[1] Moreover, incomplete coverage of vaccination of the target populations within critical age groups, as well as the need for screening for cervical cancer precursor lesions associated with high-risk HPV types not included in the current vaccines, make continued screening programs essential for the foreseeable future.

Historically, cervical cytology (Pap test) has been the standard screening test in higher-income countries and, in those countries with a robust healthcare infrastructure that can support this resource-intensive screening approach, the cumulative incidence of cervical cancer has been significantly reduced.[3] However, organized screening with cytology has not been successfully implemented in any LMIC. The World Health Organization (WHO) recommends against cytology as a screening modality for countries without pre-existing, well-functioning, cytology-based programs.[4] Failure of organized cervical cancer screening in LMICs is largely due to the substantial attrition that occurs over the multi-step process required by the traditional model of screening with cytology. In addition, lack of resources, including insufficient colposcopy, cytology, and pathology services, introduces delays and bottlenecks that result in inefficient and ineffective programs.[5–7]

In sum, there remains a critical need to design effective cervical cancer screening programs suitable for implementation in LMICs, including those countries that have a high prevalence of HIV among women. Screening is essential to ameliorate the substantial morbidity and mortality associated with this common cancer in women.

52.1.2 Screen-and-Treat as a Potential Solution

Since the early 1990s, several investigators have evaluated alternative cervical cancer screening modalities to overcome the inefficiencies and bottlenecks of the traditional multi-step approach.[8–10] In the most efficient version of this model, screening and treating are done at a single visit.[11] The diagnostic step involving colposcopy and histological diagnosis is circumvented. In this way, screen-and-treat avoids the attrition associated with the traditional approach, making it more effective than the traditional approach, which is generally poorly implemented. Moreover, the screen-and-treat approach is not reliant on colposcopy and pathology services that are often in short supply or unavailable. However, without these steps in place, overtreatment is more common with screen-and-treat than with the traditional approach (Figure 52.1).

Visual inspection with acetic acid (VIA) is the screening test that has been most widely used in screen-and-treat programs to date.[12–14] VIA is an attractive screening test to use for a variety of reasons. This test requires visualizing the cervix after application of acetic acid, and generally nurses have been trained in undertaking VIA. VIA is a simple and inexpensive test and provides a result immediately, but quality control, training, and supervision remain major hurdles.[14, 15] Despite many advantages, VIA has several disadvantages. Notably, VIA has both low sensitivity and low specificity for detecting true cervical cancer precursor lesions in need of treatment. Low sensitivity means that cervical lesions that should have been treated are missed, leading to only marginally effective screening programs. Low specificity leads to overtreatment.

Since HPV is a necessary, albeit insufficient, cause of cervical cancer, HPV testing is the logical screening test for cervical cancer screening programs. Once HPV tests became commercially available, HPV testing has been universally integrated into cervical cancer screening programs in high-resource

DOI: 10.1201/9781003449942-58

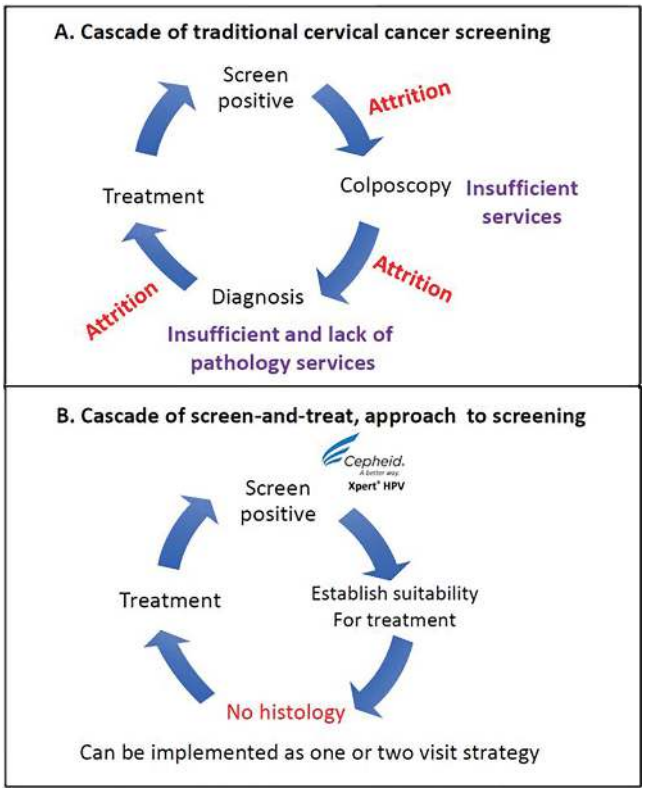


FIGURE 52.1 (A) Cascade of the traditional approach to cervical cancer screening showing the weak links, including attrition due to multistep process and bottlenecks due to lack of capacity in colposcopy and pathology services, that render this approach inefficient and ineffective. (B) Cascade of the alternative approach to cervical cancer screening, namely screen-and-treat, that simplifies the screening process, reducing attrition and reliance on colposcopy and pathology services, to render this approach more efficient and effective in low- and middle-income countries.

settings across the globe.[16] When incorporated into well-functioning screening programs in high-resource settings, previously reliant on cytology as the primary screening test, the high sensitivity of HPV tests increases the effectiveness of programs overall and allows for longer screening intervals.[16]

When HPV testing has been evaluated as the screening test within the screen-and-treat approach, it has consistently been demonstrated to lead to the best efficacy in terms of reduction in cervical cancer and its precursor lesions. In a head-to-head comparison of screen-and-treat using VIA and HPV testing we, and others, have found VIA-based screen-and-treat to be less effective than HPV-based approaches.[8–10, 17, 18] This is because HPV testing has very high sensitivity to detect cervical cancer precursor lesions. While specificity of HPV testing is worse than cytology, VIA and HPV testing have a similar profile of specificity introducing a similar extent of overtreatment when used in the screen-and-treat approach. Thus WHO recommends VIA-based screen-and-treat only when HPV-based alternatives are not available.[4]

Until recently, most HPV tests could not easily be run at the point of care. When HPV testing is incorporated into screen-and-treat programs in low-resource settings, point-of-care testing is desirable to allow for a single-visit approach to screening. Two-visit screen-and-treat with HPV testing is also cost-effective since this approach still minimizes attrition over that in the traditional model; nevertheless, a single-visit approach is better.[11]

52.1.3 Adapting Xpert HPV for Use in Screen-and-Treat

With all screening and diagnostic tests, there is an inevitable tradeoff between sensitivity and specificity. Specificity can be improved if reductions in sensitivity can be tolerated. Assays to detect HPV DNA have high sensitivity, hence their utility for primary screening in high-resource settings.[16] In the LMICs and for the screen-and-treat approach, HPV tests need to be run at the point of care in order to optimize screening participant retention. In addition, since a positive result flags a screening participant for immediate treatment, concerns have been raised about the relatively low specificity of HPV testing, particularly in high-risk groups such as women living with HIV.[2, 19]

Our goal was to investigate whether changes could be made to how an existing and validated HPV assay was utilized under real-world conditions to optimize HPV testing for cervical cancer screening within the screen-and-treat model. We were interested in the application of HPV testing in a resource-constrained context like South Africa that also has a high prevalence of HIV infection among women.

52.1.4 Why Xpert HPV

The first commercially available HPV DNA test was Hybrid Capture 1 (Digene). The update on this test was Hybrid Capture II (Qiagen), and this assay became the benchmark against which other commercially available tests were evaluated.[20–22] These tests were designed to maximize sensitivity since their primary use was envisioned in high-resource settings where not missing disease is the primary objective.[20–22] These tests were also not designed to be used at the point for care relying on already established centralized laboratories. For our study, we chose to investigate how to adapt the test developed by Cepheid, namely Xpert HPV™. Xpert HPV is a WHO pre-qualified HPV DNA test and has CE-marking. This test can easily be run as a point-of-care test, is validated, and has been successfully used in several programs.[23–28]

Xpert HPV is a real-time PCR assay that detects 14 high-risk HPV types in five fluorescent channels: (1) HPV 16, (2) HPV 18, 45, (3) HPV 31, 33, 35, 52, 58, (4) HPV 51, 59, and (5) HPV 39, 56, 66, 68. Output from the assay includes a cycle threshold (Ct) value from each of these five channels. Ct values have been determined by the manufacturers to define a positive result for each channel. Additionally, a sample adequacy control detects whether the sample contains human DNA in a sixth channel.

There are several advantages to this assay. Most pertinent to its use in cervical cancer screening programs in LMICs, Xpert HPV is very simple to run and can be operated by a

worker without laboratory training. A competitor assay, *care*HPV, was also developed with the intention to be contextually appropriate for use in low-resource settings.[29] However, technical demands of the assay require skilled technicians and advanced laboratory environments to run the assay properly.[30] *care*HPV also requires batching. Thus, a number of samples has to be collected before the test can be run efficiently, introducing additional logistical constraints into its use for screening.[30] In contrast, Xpert HPV does not require batching as each sample can be tested individually in the GeneXpert module system.[23–28] One of the important limitations of the Hybrid Capture 2 HPV assay, the guideline-recommended HPV assay for reference in validation studies,[21] is that the assay is not intended for use at the point of care. Testing at the point of care is important as it allows screening and treatment to be done at a single visit, significantly reducing costs and opportunities for attrition.

52.1.5 Ways in Which We Adapted Xpert HPV

52.1.5.1 Rationale for Type Restriction

The first way in which we adapted the assay to make it more suitable for screen-and-treat was by type restriction. There are 14 HPV types that through the accumulation of research over the years have been defined as "high risk" for development of cervical cancer and the types that should be covered by HPV assays for screening. However, there is not a consistent strength of association between each of these types individually and cervical cancer, with some of the types having stronger, and others weaker, associations with cancer and its precursor lesions.[31] The high-risk types also differ in their frequency in the population, hence making each individual type's attributable contribution quite variable.

The Xpert HPV assay provides grouped typing. The assay is designed to detect the 14 high-risk types in five independent channels: (1) HPV 16, (2) HPV 18 and/or 45, (3) HPV 31, 33, 35, 52, and/or 58, (4) HPV 51 and/or 59, and (5) HPV 39, 56, 66, and/or 68. The assay classifies each channel as positive or negative for the indicated group of HPV types based on cut-offs developed in the design of the assay. The first three channels include the HPV types that have the strongest associations with cervical cancer, including HPV 16, 35, and 45. HPV 18 is of significant interest as it is associated with adenocarcinoma, often overlooked by other screening modalities. We hypothesized that screen-positive could be defined purely on the basis of the first three channels and that channels 4 and 5 could be effectively ignored. Reclassifying the output of the assay in this would lead, we hypothesized, to improvements in specificity with only minor loss of sensitivity.

52.1.5.2 Rationale for Ct Cut-Offs

The second way in which we adapted the assay to make it more suitable for screen-and-treat was by making the cut-offs to define screen-positive more stringent. Strong associations between semi-quantitative ways of assessing HPV "load" and the likelihood of cervical cancer precursor lesions are consistently shown despite the technical challenges in accurate quantitation of the amount of HPV present. Xpert HPV is a PCR assay, and part of standard output is the cycle threshold of each HPV channel. In a PCR assay, the cycle threshold is a rough approximation of the reverse of the signal strength: the higher the Ct value the weaker the signal strength; the lower the Ct value the stronger the signal strength. The assay has been calibrated to define positivity on each channel at certain Ct cut-offs. For the optimal utility of this test for the screen-and-treat setting, we hypothesized that we should make the Ct cutoffs more stringent, that is, require a stronger signal strength (lower Ct value) before we would define the sample as positive to improve specificity without a major loss in sensitivity.

52.1.5.3 Operational Issues

We are particularly interested in the operational issues involved in running the assay at the point of care. The feature we found most compelling about the methodology was its ease of use. The cartridge design of the assay is such that an operator needs no technical laboratory skills. The instruments also needed to be robust enough to withstand the environmental conditions typically found in clinical sites, including wide variations in temperature, dust, rough handling, and surges of power. From a clinical point of view, reliability in terms of not returning invalid results and being consistently available to run tests when patient's samples were collected are also important.

Xpert HPV requires 60 minutes of run-time in the instrument. We therefore had to structure the clinical encounter to take this time into account. We also had to build in systems to ensure accurate labeling of the sample and adequate coordination between the workers involved in collecting the sample, running the assay, and reporting the results back to the patient. Each of these steps required quality control and can introduce delays if not efficiently designed. The instruments are designed to run assays simultaneously and each test can be introduced independently at any time. We used a four-module instrument able to run four assays simultaneously. We found this configuration suitable to the size and flow of our clinical site since it was rare to complete more than four clinical examinations in an hour. Each of these operational steps, and potential stumbling blocks, were monitored during our study.

52.2 Materials

To collect the cervical samples the following materials were needed:

- Gynecological bed or examination couch.
- Stool or chair for provider.
- Disposable examination gloves for provider.
- Bi-valve vaginal speculum.
- Lubricant to lubricate speculum, if available.
- Extended tip plastic spatula.
- Endocervical cytobrush.

To undertake the HPV testing the following materials were needed:

- Disposable examination gloves.
- Thin prep vials containing PreservCyt solution (Hologic, Bedford, MA).
- Xpert HPV™ cartridges (Cepheid, Sunnyvale, CA).
- Transfer pipettes.
- GeneXpert Instrument (Cepheid, Sunnyvale, CA).
- Barcode scanner.
- Computer.
- Printer.

52.3 Methods

52.3.1 Design of Clinical Study in Cape Town, South Africa

To address the question of how to adapt the Xpert HPV assay, we designed a clinical study that we conducted in Cape Town, South Africa. The study enrolled 1160 women who were aged 30–65 years. Approximately equal numbers of HIV-negative and HIV-positive women were recruited by design, since analyses stratified by HIV status were planned. Women were recruited at two sites in Cape Town to represent two distinct populations: (1) a screening population and (2) a referral population. At a primary care facility (Khayelitsha Site B Day Hospital, Western Cape Department of Health facility) we recruited women through community outreach aiming to be representative of healthy women seeking routine screening. This population was necessary to describe the prevalence of HPV and the prevalence of cervical cancer precursor lesions in the population. At a referral colposcopy clinic (Groote Schuur Hospital [GSH], tertiary hospital of the University of Cape Town), we recruited women referred to colposcopy because of abnormal cytology results obtained as part of existing screening services in the region. This population was necessary to yield a sufficient number of women with cervical cancer precursor lesions to estimate sensitivity accurately. The stratified design was chosen as it was the most efficient with the limited resources available to describe the key parameters of interest. Other inclusion criteria were (1) no history of any anogenital cancer, (2) no history of treatment for cervical dysplasia, (3) no hysterectomy, and (4) not pregnant. All women provided written informed consent. The study protocol was approved by the Institutional Review Boards of Columbia University and the University of Cape Town.

The study was designed to ensure that gold standard endpoint could be ascertained in all women to avoid verification bias. All women underwent colposcopy with cervical biopsies of any abnormal-appearing areas of the cervix. If a woman had no visible lesions on the cervix, an endocervical curettage (ECC) was performed. A loop electrosurgical excision procedure (LEEP) was performed if clinically indicated. The LEEP procedure is a treatment procedure but also provides a histological sample. For women recruited at the primary care facility, further follow-up with histological sampling ensured that disease was not missed. If any lesion was seen, it was biopsied or LEEP was performed. If no lesion was seen, a random biopsy was obtained from the 12 o'clock position. After initial pathology review in Cape Town for the clinical management of study participants, all histology slides were shipped to Columbia University and re-reviewed. Results discordant with the initial reads were reviewed by a third gynecologic pathologist as the tie-breaker for the final endpoint determination. Accurate determination of endpoint is essential, as error in the measurement of the gold standard diagnosis also introduces bias. We defined the study endpoint to be cervical intraepithelial neoplasia grade 2 or grade 3 or cancer (CIN2+) based on adjudicated pathology.

In the context of this pelvic examination, two cervical samples were obtained using an extended tip plastic spatula with endocervical cytobrush (Medscand, Berlin, Germany) (Box 52.1). The samples were placed in separate vials filled with 20 ml of PreservCyt solution (Hologic, Bedford, MA). The vial was gently inverted several times to ensure the liquid was adequately mixed, and then 1 ml of the liquid was pipetted out of the 20 ml vial using a pipette included with each kit and placed into the cartridge, which was then placed into the machine and the assay initiated on a linked computer (Box 52.2, Figure 52.2). A four-module GeneXpert instrument was installed at the primary care site in Khayelitsha, and a

BOX 52.1 MATERIALS AND METHODS FOR CERVICAL SAMPLE COLLECTION

Materials in examination room

- Gynecological bed or examination couch.
- Stool or chair for provider.
- Disposable examination gloves for provider.
- Bi-valve vaginal speculum.
- Lubricant to lubricate speculum, if available.
- Extended tip plastic spatula
- Endocervical cytobrush

Method of sample collection

a. Ensure privacy of the examination room.
b. Label the thin prep vials with patient details
c. Place the patient on the examination couch with the legs comfortably on the stirrups
d. Insert the speculum so that the cervix comes into view.
e. Insert the long arm of the plastic spatula into the cervical os, applying it firmly on the cervix and rotate 360 degrees 3–5 times, remove and place in the thin prep vial and shake the spatula about 10 times to transfer the cells from the spatula into the PreservCyt solution.
f. Repeat the last step using the cytobrush.
g. Close the PreservCyt vial tightly and send off sample for testing.

BOX 52.2 MATERIALS AND METHODS FOR XPERT HPV TESTING

Materials in the testing room

i. Disposable examination gloves
ii. Thin prep vials containing PreservCyt solution (Hologic, Bedford, MA).
iii. Xpert HPV cartridges
iv. Transfer pipettes
v. GeneXpert platform (comes with a barcode scanner and a printer)

Methods of to undertake Testing

1. Obtain one Xpert HPV cartridge and one transfer pipette.
2. Enter all patient information on the Xpert software on the computer

3. Open the Xpert cartridge lid

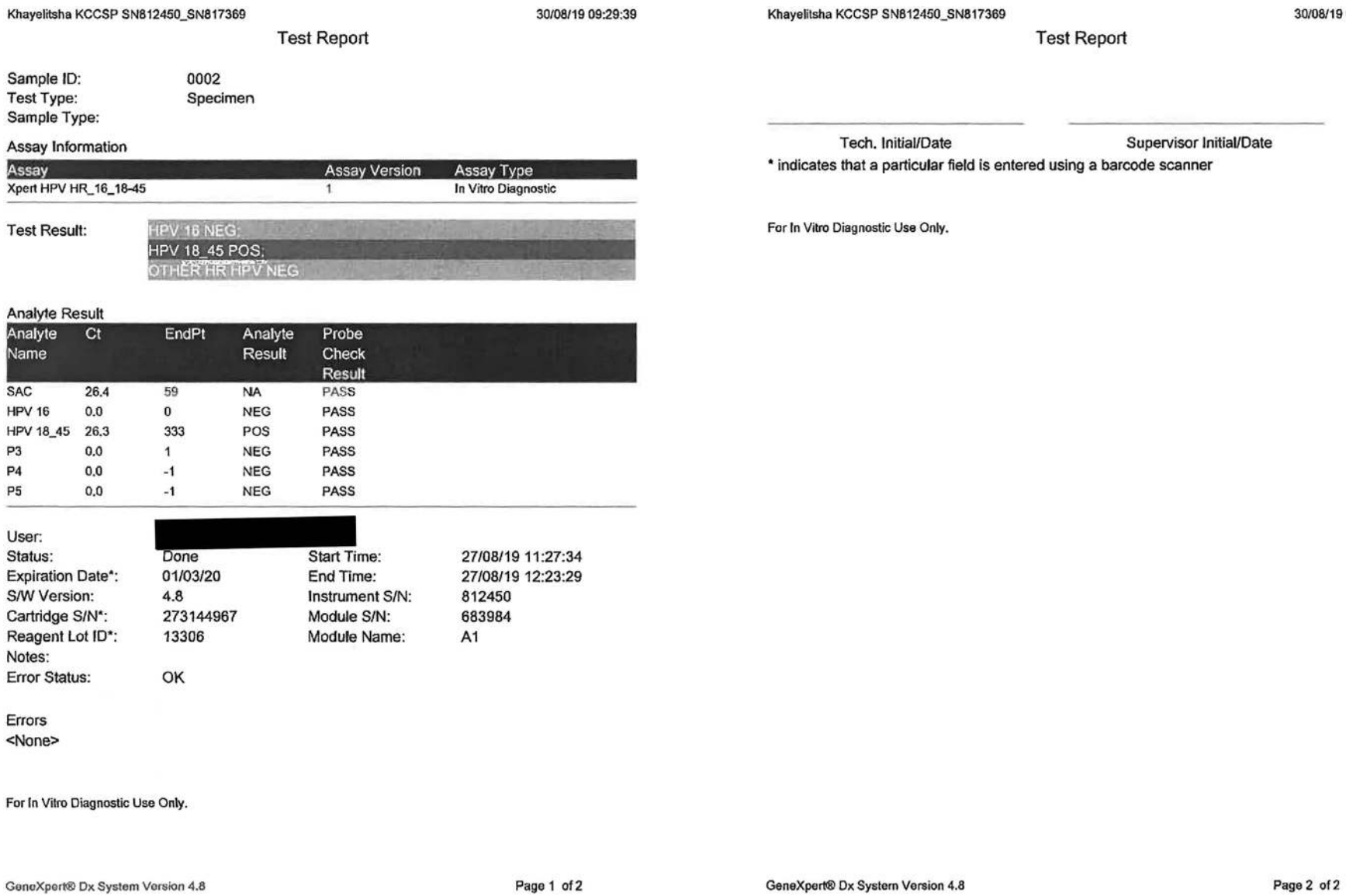
4. Gently invert the sample vial 8–10 times to mix the sample well

5. Fill the transfer pipette with 1ml of the sample

6. Empty the contents into the sample chamber

7. Close the Xpert cartridge kit
8. Scan the Xpert HPV cartridge barcode
9. Insert the prepared cartridge into a module in the GeneXpert machine and start the assay.

10. At the end of one hour, the result of the completed assay prints out automatically.



Khayelitsha KCCSP SN812450_SN817369 30/08/19 09:29:39

Test Report

Sample ID: 0002
Test Type: Specimen
Sample Type:

Assay Information

Assay	Assay Version	Assay Type
Xpert HPV HR_16_18-45	1	In Vitro Diagnostic

Test Result: HPV 16 NEG;
HPV 18_45 POS;
OTHER HR HPV NEG

Analyte Result

Analyte Name	Ct	EndPt	Analyte Result	Probe Check Result
SAC	26.4	59	NA	PASS
HPV 16	0.0	0	NEG	PASS
HPV 18_45	26.3	333	POS	PASS
P3	0.0	1	NEG	PASS
P4	0.0	-1	NEG	PASS
P5	0.0	-1	NEG	PASS

User: [redacted]
Status: Done
Expiration Date*: 01/03/20
S/W Version: 4.8
Cartridge S/N*: 273144967
Reagent Lot ID*: 13306
Notes:
Error Status: OK

Start Time: 27/08/19 11:27:34
End Time: 27/08/19 12:23:29
Instrument S/N: 812450
Module S/N: 683984
Module Name: A1

Errors
<None>

For In Vitro Diagnostic Use Only.

GeneXpert® Dx System Version 4.8 Page 1 of 2

Khayelitsha KCCSP SN812450_SN817369 30/08/19 09:29:39

Test Report

Tech. Initial/Date Supervisor Initial/Date

* indicates that a particular field is entered using a barcode scanner

For In Vitro Diagnostic Use Only.

GeneXpert® Dx System Version 4.8 Page 2 of 2

FIGURE 52.2 Printout from the Xpert HPV.

community health worker, with no prior laboratory experience, was trained in running the test. Samples collected from women recruited at the primary care site were run as soon as possible after collection and results reported to participants. Samples collected from women recruited at the colposcopy clinic were transported to the site the next day and run on the same machine.

52.3.2 Statistical Methods

The focus of our analysis was to obtain the optimal balance between sensitivity and specificity to make the assay most appropriate for screen-and-treat. For calculation of sensitivity, we used the cases of CIN2+ that had been identified in both the screening and referral populations. We had made this decision when we designed the study because achieving a large enough sample of cases of CIN2+ only from a screening population would have been too expensive. For calculations of specificity, which included in the denominator the absence of CIN2+ (i.e., CIN 1 and within normal limits on adjudicated pathology diagnosis), only women recruited from the primary care site were included. For calculations of the prevalence of HPV (i.e., screen-positive rate), positive predictive value (PPV), and negative predictive value (NPV), only women recruited from the primary care site were included. Confidence intervals (CI) of 95% around the screen-positive rate, sensitivity, specificity, PPV, and NPV were calculated using the binomial method. All statistical analyses were conducted in SAS v9.4 (Cary, NC).

To investigate the potential benefits of type restriction, we conducted multivariable logistic regression analysis with CIN2+ as the outcome and each of the five HPV channels from Xpert HPV as the predictors. We selected the channels that were statistically significant ($p < 0.05$) in these multivariable models. We then compared the screen-positive rate, sensitivity, specificity, PPV, and NPV of the test "as is"—defining the test as positive if any one of the five HPV channels was positive, vs. "type-restricted"—defining the test as positive only if the statistically significant channels were positive.

To investigate the potential benefits of changing the cut-offs on the test, we conducted multivariable logistic regression analysis with CIN2+ as the outcome and the reverse of the cycle threshold (rCt) from each of the five HPV channels from Xpert HPV as the predictors. Cycle threshold values are reported out for each of the five channels in the assay and are an approximation of the signal strength. If the Ct is low, there is more HPV. If the Ct is high, there is less HPV. Cut-off Ct values have been determined by the manufacturers to define a positive result for each channel. To correct the directionality, in other words to make sure that high values indicated more HPV and low values indicated less HPV, we defined rCt equal to 45 minus the observed Ct value if the sample was considered to be positive by the manufacturer (Ct value greater than 0 but below the manufacturer's Ct cut-off). For example, if the observed Ct value was 32, then the rCt value was 45 − 32 = 13. We defined rCt values equal to 0 if the sample was considered to be negative by the manufacturer.

For positive results, let reverse Ct (rCt) = 45 − Ct.

For negative results, let reverse Ct (rCt) = 0.

The final logistic regression models included rCt values from channels that, when included together in multivariable models, were significantly associated with disease ($p < 0.05$). From these final models were calculated receiver operating characteristic (ROC) curves. ROC curves plot sensitivity (y-axis) against 1-specificity (x-axis) at all observed values of the rCt-based cut-offs. The area under the curve (AUC) is a useful parameter that defines the strength of the association, with 1 being perfect sensitivity and specificity and all values less than 1 and greater than 0 indicating less than perfect sensitivity and specificity. By design, we had ensured that the maximum sensitivity that could be achieved was when the manufacturer-defined cut-offs were used. We were interested in determining the effect on specificity of shifting the Ct cut-offs on the assay to make it more difficult to detect CIN2+. We therefore looked up the calculated specificity on the ROC, when the sensitivity declined from the maximum observed (Figure 52.3). In our case, this was sensitivity values below 85%. We looked up specificity on the ROC curve when sensitivity was 85%, 80%, 75%, 70%, and 65%. This was intended as the primary output of our study. This range allowed us to consider the potential tradeoffs between sensitivity that we would need to attain a particular level of specificity. Specificity is important to reduce overtreatment. Some loss in sensitivity can be tolerated, particularly in populations who will undergo repeat screening.

Next we utilized the sensitivity/specificity combinations we had generated from the ROC curve to calculate positive predictive value (PPV), negative predictive value (NPV) and the screen-positive rate (SPR). For these calculations we used the observed prevalence (P_i) of CIN2+ in the screening population. The following algebraic formulae indicate how these parameters were calculated, where i refers to either HIV-negative or HIV-positive groups:

$$PPV_i = ([P_i * \text{sensitivity}_i)/([1-P_i] * [1-\text{specificity}_i] + [P_i * \text{sensitivity}_i]);$$

$$NPV_i = ([1-P_i] * \text{specificity}_i)/(P_i * [1-\text{sensitivity}_i] + [1-P_i] * \text{specificity}_i);$$

$$SPR_i = (P_i * \text{sensitivity}_i) + (1-P_i) * (1-\text{specificity}_i).$$

Since we calculated multivariable logistic regression models utilizing the rCt values from all the channels on the assay that were significantly associated with CIN2+, the actual Ct cut-off on a single channel is not defined if more than one channel is positive. In cases where samples are positive on only one channel, it is possible, for illustrative purposes, to calculate what the actual Ct value cut-off is for a given sensitivity/specificity combination. In a logistic regression model, the outcome Y (which presence or absence of CIN2+) = natural log (ln) of p/(1-p) where p is the probability level at a given sensitivity. From output generated as part of the calculation of the ROC curve one can look up "p" at the given sensitivity/specificity combination. We can then use this information in the following algebraic formulae to calculate the cut-off (K):

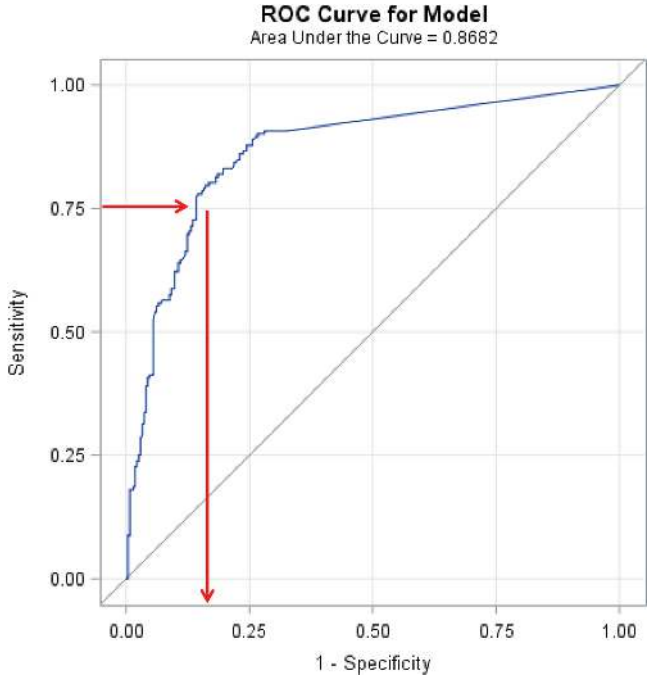


FIGURE 52.3 Example of a receiver operating characteristic (ROC) curve in which sensitivity is plotted on the y-axis and 1 minus specificity on the x-axis. The red arrows illustrate how we inferred specificity at sensitivity equal to 75%. A line is drawn from 75% on the y-axis and then specificity can inferred by dropping a point down to the x-axis.

Cut-off (K) = natural log (ln) of p/(1 − p)

From the logistic regression model:

Y = parameter estimate for HPV16 channel (Beta16) * rCt16 + parameter estimate for HPV1845 channel (Beta18)*rCt18 + parameter estimate for HPV31+ channel (Beta31)*rCt31 + intercept (Int).

For example, to calculate the cut-off (K) for HPV 16 (assuming no multiple types), terms Beta18 and Beta31 in the logistic regression model equal 0 and the inequality simplifies to:

$$\text{Beta16}*\text{rCt16} + \text{Int} < K$$
$$\text{rCt} < (K - \text{Int})/\text{Beta16}$$
$$\text{Ct16} < 45 - (K - \text{Int})/\text{Beta16}$$

52.4 Results

Utilizing only the approach restricting to the three channels detecting HPV types 16, 18, 45, 31, 33, 35, 52, 58 (manufacturer's cut-offs) led to some improvements in specificity with very little compromise in sensitivity.[32] Using genotype restriction only, sensitivity to detect CIN2+ was only modestly reduced to 87.1% (95% CI: 81.2–93.0) in women without HIV and 90.7% (95% CI: 86.4–95.0) in women living with HIV. By comparison, using the standard cut-off and detecting all 14 high-risk types in all five channels, the sensitivity to detect CIN2+ was 88.7% (95% CI: 83.1–94.3) in women without HIV and 93.6% (95% CI: 83.1–94.3) in women living with HIV. In contrast, despite these marginal reductions in sensitivity, when the definition of screen-positive was restricted to the three channels, specificities improved to 89.7% (95% CI: 86.5–92.8) in women without HIV and 67.5% (95% CI: 62.0–73.1) for women living with HIV. In the absence of this restriction, specificity was to 86.9% (95% CI: 83.4–90.4) in women without HIV and 59.9% (95% CI: 54.1–65.7) in women without and women with HIV, respectively, for all channel positives.[32]

However, combining this genotype restriction approach with shifts in the cut-offs on these three channels to define screen-positive (the method described earlier) produced the best results. Using the logistic regression model approach, only the rCt values for the first three channels were significantly predictive of CIN2+. Based on these models, we allowed sensitivity to range between 65% and 85%. We then looked up the corresponding specificity from the ROC curves from the multivariable logistic regression models. From this we could calculate that if sensitivity for CIN2+ is allowed to decline to 80%, applying more stringent cut-offs to the three channels detecting HPV types 16, 18, 45, 31, 33, 35, 52, 58 leads to specificity of 94.1% in women without HIV. At this cut-off, the PPV is 43.1% and NPV 98.8%, and 9.8% of the screening population would be defined as screen-positive. For women living with HIV, more stringent Ct values allowing sensitivity for CIN2+ to decline to 80% leads to specificity of 83.2%, PPV of 49.4%, NPV of 95.3%, and screen-positive rate of 27.5%.[33]

The full range of performance parameters related to shifting the cut-offs as we allow sensitivity to shift between 65%

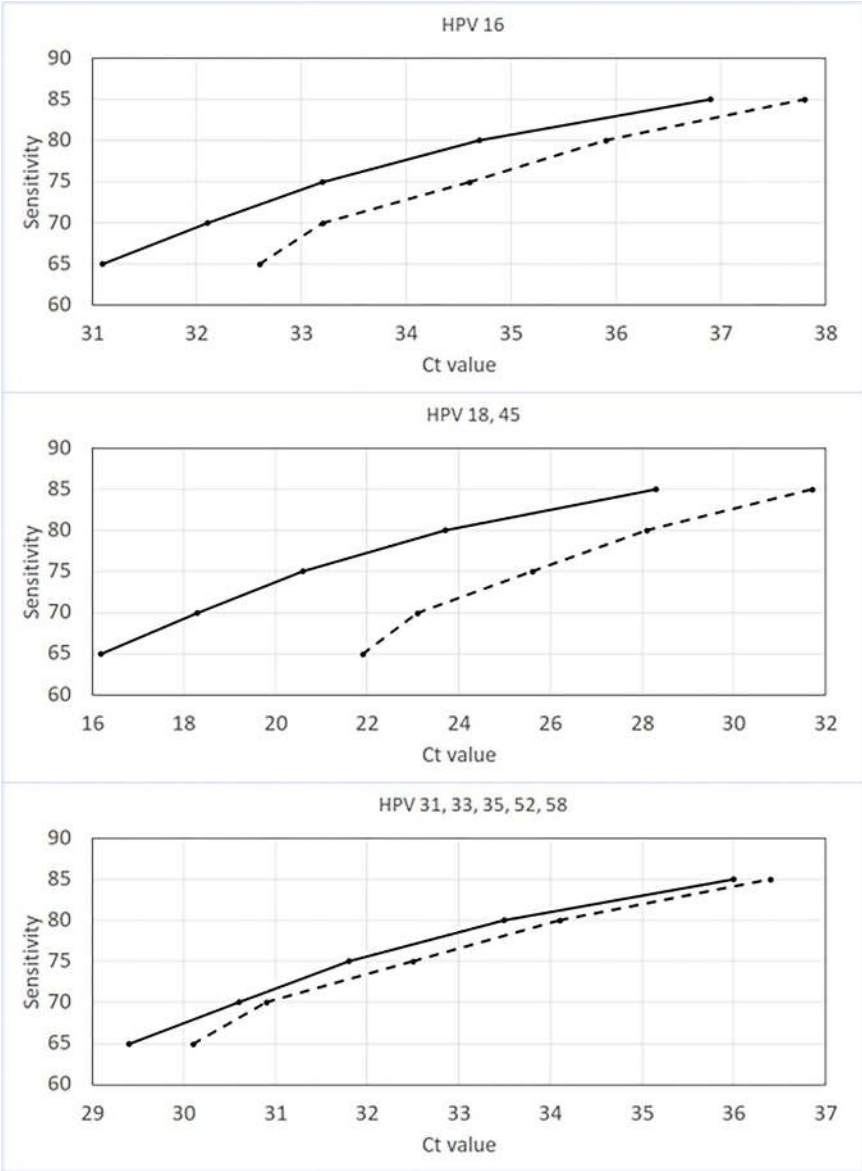


FIGURE 52.4 Cycle threshold cut-off points for the three channels of the Xpert HPV assay significantly associated with histopathologically confirmed cervical intraepithelial neoplasia grade 2 or worse (CIN2+). Illustration of cycle threshold cut-off points to detect CIN2+ at given sensitivity values ranging between 65% and 85% assuming that women are positive on only one of the three channels that detect (1) HPV 16, (2) HPV 18 or 45, and (3) HPV 31, 33, 35, 52 or 58.

and 85% is displayed in Table 52.1. In Figure 52.4 are the estimated Ct cut-offs of the three channels to achieve sensitivities between 65% and 85% calculated from the logistic regression models assuming no multiple types.[33]

52.5 Discussion

In typical examples of assay validation, the sensitivity/specificity balance is fixed by the manufacturer. In the case of HPV testing for screening for cervical cancer, the consensus balance of sensitivity/specificity favors sensitivity, as the assay is most widely used in high-resource settings where missing disease carries greatest salience. In low-resource settings, and in the screen-and-treat model, we argue that specificity should be favored to allow scarce resources to be targeted to the most high-risk women and to minimize overtreatment. Our approach allows for an available and approved assay to be adapted in this way without any new engineering. Xpert HPV is already designed to provide grouped HPV typing as well as the channel-specific Ct value as part of standard output. This information can be exported into a simple calculator to define screen-positive in different ways to suit the program. This allows decision-making about performance characteristics of the screening test to be shifted to program implementers.

Inserts

TABLE 52.1

Effect of Varying the Cycle Threshold Cut-Offs from the Three Channels on HPV Xpert Detecting HPV 16, 18, 45, 31, 33, 35, 52, and 58 by Specified Sensitivity Value.

	Sensitivity	Specificity	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)	Screen-Positive (95% CI)
CIN2+					
HIV-negative					
	65	96.4	50.0 (42.3–57.7)	98.0 (96.5–99.5)	6.9 (4.5–9.2)
	70	96.4	51.9 (33.3–59.5)	98.3 (97.0–99.7)	7.1 (4.7–9.5)
	75	95.3	47.2 (40.2–54.2)	98.6 (97.4–99.8)	8.4 (5.7–11.1)
	80	94.1	43.1 (36.7–49.5)	98.8 (97.6–100)	9.8 (6.9–12.8)
	85	91.3	35.4 (30.0–40.8)	99.1 (98.1–100)	12.7 (9.3–16.1)
HIV-positive					
	65	88.3	53.3 (46.6–60.1)	92.5 (89.3–95.7)	20.7 (16.5–25.0)
	70	87.2	52.9 (46.4–59.4)	93.4 (90.4–96.4)	22.5 (18.0–27.0)
	75	85.8	52.0 (45.8–58.2)	94.4 (91.5–97.3)	24.5 (19.8–29.2)
	80	83.2	49.4 (43.5–55.3)	95.3 (92.6–98.0)	27.5 (22.5–32.5)
	85	77.0	43.1 (37.8–48.4)	96.2 (93.7–98.7)	33.5 (28.1–39.0)
CIN3+					
HIV-negative					
	65	96.7	39.6 (31.8–47.4)	98.8 (97.7–99.9)	5.3 (3.1–7.4)
	70	95.9	36.2 (29.7–43.2)	99.0 (98.0–100)	6.2 (3.8–8.5)
	75	95.1	33.7 (27.3–40.1)	99.2 (98.3–100)	7.2 (4.6–9.6)
	80	93.4	29.0 (23.4–34.6)	99.4 (98.8–100)	8.9 (6.1–11.8)
	85	92.3	28.4 (23.1–33.7)	99.5 (98.8–100)	9.6 (7.1–13.2)
HIV-positive					
	65	88.4	34.2 (27.7–40.8)	96.5 (94.3–98.7)	16.1 (12.2–20.1)
	70	87.1	33.7 (27.5–40.0)	97.0 (95.0–99.0)	17.8 (13.6–21.9)
	75	86.1	33.5 (27.5–39.5)	97.4 (95.5–99.3)	19.1 (14.8–23.4)
	80	84.8	32.8 (27.1–38.5)	97.9 (96.2–99.6)	20.7 (16.2–25.2)
	85	74.8	24.0 (19.7–28.3)	98.3 (96.6–99.9)	30.3 (25.1–35.5)

Since we calculated multivariable logistic regression models utilizing the rCt values from all the channels on the assay that were significantly associated with CIN2+, the actual Ct cut-off on a single channel is not defined if more than one channel is positive. In the case of multi-channel positivity, the Ct cut-off on any one channel to define positive or negative status is dependent on the Ct values in the other channels. Thus in order to determine whether a sample is "screen-positive," a calculator has to be added that inputs the Ct values generated from all the relevant channels on the assay as designed. Ideally, this calculator could be automated into the assay software. We conducted a follow-on demonstration study of this approach.[34] In the demonstration project we did not have resources to fully automate this process. Instead, we used a simple calculator external to the assay software requiring manual entry of the Ct values.

In selecting optimal cut-offs on the assay, program implementers could weigh resource availability and community preferences. For example, if there are insufficient clinicians to treat the large numbers of women who would require treatment if sensitivity is set at 80%, cut-offs could be shifted to a lower sensitivity resulting in a smaller percentage of women in need of treatment. Preferences of the clinical community could also be taken into account. For example, for settings where overtreatment is less tolerated, cut-offs that ensure high PPV could be chosen. In contrast, where there are few alternative treatment options and women are likely to be infrequently screened, higher sensitivity could be selected and more overtreatment tolerated.

It will be important to validate this approach with Xpert HPV in other populations. It is also important to evaluate whether the approach will be as effective using other commercially available HPV assays that are configured to provide full or partial genotyping and that provide a signal strength indicator such as a cycle threshold value.

Acknowledgments

Research was supported by the National Cancer Institute (UH2/3 CA 189908).

REFERENCES

1. Denny L, de Sanjose S, Mutebi M, et al. Interventions to close the divide for women with breast and cervical cancer between low-income and middle-income countries and high-income countries. *Lancet (London, England)* 2017;389(10071):861–870. (In eng). DOI: 10.1016/s0140-6736(16)31795-0.
2. Heard I. Prevention of cervical cancer in women with HIV. *Current Opinion in HIV and AIDS* 2009;4(1):68–73. (In eng). DOI: 10.1097/COH.0b013e328319bcbe.
3. Laara E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries: Association with organised screening programmes. *Lancet (London, England)* 1987;1(8544):1247–1249. (In eng).
4. World Health Organization. *Comprehensive Cervical Cancer Control: A Guide to Essential Practice*. 2nd Edition. 2014. ISBN 978 92 4 154895 3. https://www.who.int/publications/i/item/9789241548953.
5. Denny L, Quinn M, Sankaranarayanan R. Chapter 8: Screening for cervical cancer in developing countries. *Vaccine* 2006;24(Suppl 3):71–77. (In eng). DOI: 10.1016/j.vaccine.2006.05.121.
6. Goldhaber-Fiebert JD, Denny LE, De Souza M, Wright TC, Jr., Kuhn L, Goldie SJ. The costs of reducing loss to follow-up in South African cervical cancer screening. *Cost Effectiveness and Resource Allocation: C/E* 2005;3:11. (In eng). DOI: 10.1186/1478-7547-3-11.
7. Denny L, De Sousa M, Kuhn L, Pollack A, Wright TC, Jr. Cervical cancer prevention—a paradigm shift? *Gynecologic Oncology* 2005;99(3 Suppl 1):S12. (In eng).
8. Denny L, Kuhn L, De Souza M, Pollack AE, Dupree W, Wright TC, Jr. Screen-and-treat approaches for cervical cancer prevention in low-resource settings: A randomized controlled trial. *JAMA* 2005;294(17):2173–2181. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16264158.
9. Denny L, Kuhn L, Hu CC, Tsai WY, Wright TC, Jr. Human papillomavirus-based cervical cancer prevention: Long-term results of a randomized screening trial. *Journal of the National Cancer Institute* 2010;102(20):1557–1567. (In eng). DOI: djq342 [pii] 10.1093/jnci/djq342.
10. Sankaranarayanan R, Nene BM, Shastri SS, et al. HPV screening for cervical cancer in rural India. *The New England Journal of Medicine* 2009;360(14):1385–1394. (In eng). DOI: 10.1056/NEJMoa0808516.
11. Goldie SJ, Kuhn L, Denny L, Pollack A, Wright TC. Policy analysis of cervical cancer screening strategies in low-resource settings: Clinical benefits and cost-effectiveness. *JAMA* 2001;285(24):3107–3115. (In eng). DOI: 10.1001/jama.285.24.3107.
12. Sankaranarayanan R, Nessa A, Esmy PO, Dangou JM. Visual inspection methods for cervical cancer prevention. *Best Practice & Research Clinical Obstetrics & Gynaecology* 2012;26(2):221–232. (In eng). DOI: 10.1016/j.bpobgyn.2011.08.003.
13. Wright TC, Jr., Denny L, Kuhn L, Goldie S. Use of visual screening methods for cervical cancer screening. *Obstetrics and Gynecology Clinics of North America* 2002;29(4):701–734. (In eng).
14. Mwanahamuntu MH, Sahasrabuddhe VV, Blevins M, et al. Utilization of cervical cancer screening services and trends in screening positivity rates in a 'screen-and-treat' program integrated with HIV/AIDS care in Zambia. *PLoS One* 2013;8(9):e74607. (In eng). DOI: 10.1371/journal.pone.0074607.
15. Parham GP, Mwanahamuntu MH, Kapambwe S, et al. Population-level scale-up of cervical cancer prevention services in a low-resource setting: Development, implementation, and evaluation of the cervical cancer prevention program in Zambia. *PLoS One* 2015;10(4):e0122169. (In eng). DOI: 10.1371/journal.pone.0122169.
16. Koliopoulos G, Nyaga VN, Santesso N, et al. Cytology versus HPV testing for cervical cancer screening in the general population. *The Cochrane Database of Systematic Reviews* 2017;8:Cd008587. (In eng). DOI: 10.1002/14651858.CD008587.pub2.
17. Gelband H, Sankaranarayanan R, Gauvreau CL, et al. Costs, affordability, and feasibility of an essential package of cancer control interventions in low-income and middle-income countries: Key messages from disease control priorities, 3rd edition. *Lancet (London, England)* 2016;387(10033):2133–2144. (In eng). DOI: 10.1016/s0140-6736(15)00755-2.
18. Chibwesha CJ, Frett B, Katundu K, et al. Clinical performance validation of 4 point-of-care cervical cancer screening tests in HIV-infected women in Zambia. *Journal of Lower Genital Tract Disease* 2016;20(3):218–223. (In eng). DOI: 10.1097/lgt.0000000000000206.
19. Mapanga W, Girdler-Brown B, Feresu SA, Chipato T, Singh E. Prevention of cervical cancer in HIV-seropositive women from developing countries through cervical cancer screening: A systematic review. *Systematic Reviews* 2018;7(1):198. (In eng). DOI: 10.1186/s13643-018-0874-7.
20. Arbyn M, Snijders PJ, Meijer CJ, et al. Which high-risk HPV assays fulfil criteria for use in primary cervical cancer screening? *Clinical Microbiology and Infection: The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases* 2015;21(9):817–826. (In eng). DOI: 10.1016/j.cmi.2015.04.015.
21. Meijer CJ, Berkhof H, Heideman DA, Hesselink AT, Snijders PJ. Validation of high-risk HPV tests for primary cervical screening. *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology* 2009;46(Suppl 3):S1–S4. (In eng). DOI: 10.1016/s1386-6532(09)00540-x.
22. Meijer CJ, Berkhof J, Castle PE, et al. Guidelines for human papillomavirus DNA test requirements for primary cervical cancer screening in women 30 years and older. *International Journal of Cancer* 2009;124(3):516–520. (In eng). DOI: 10.1002/ijc.24010.
23. Akbari A, Vanden Broeck D, Benoy I, Padalko E, Bogers J, Arbyn M. Validation of intra- and inter-laboratory reproducibility of the Xpert HPV assay according to the international guidelines for cervical cancer screening. *Virology Journal* 2018;15(1):166. (In eng). DOI: 10.1186/s12985-018-1076-6.
24. Cubie HA, Morton D, Kawonga E, et al. HPV prevalence in women attending cervical screening in rural Malawi using the cartridge-based Xpert((R)) HPV assay. *Journal*

of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology 2017;87:1–4. (In eng). DOI: 10.1016/j.jcv.2016.11.014.

25. Mbulawa ZZA, Wilkin TJ, Goeieman B, et al. Xpert human papillomavirus test is a promising cervical cancer screening test for HIV-seropositive women. *Papillomavirus Research (Amsterdam, Netherlands)* 2016;2:56–60. (In eng). DOI: 10.1016/j.pvr.2016.02.004.
26. Toliman PJ, Kaldor JM, Badman SG, et al. Performance of clinical screening algorithms comprising point-of-care HPV-DNA testing using self-collected vaginal specimens, and visual inspection of the cervix with acetic acid, for the detection of underlying high-grade squamous intraepithelial lesions in Papua New Guinea. *Papillomavirus Research (Amsterdam, Netherlands)* 2018;6:70–76. (In eng). DOI: 10.1016/j.pvr.2018.10.009.
27. Rabaan AA, Alfaraj SA, Alkhalifah MA. Comparison of the Cepheid Xpert HPV test and the HC2 high-risk HPV DNA test for detection of high-risk HPV infection in cervical smear samples in SurePath preservative fluid. *Journal of Medical Microbiology* 2018 (In eng). DOI: 10.1099/jmm.0.000723.
28. Einstein MH, Smith KM, Davis TE, et al. Clinical evaluation of the cartridge-based GeneXpert human papillomavirus assay in women referred for colposcopy. *Journal of Clinical Microbiology* 2014;52(6):2089–2095. (In eng). DOI: 10.1128/jcm.00176-14.
29. Qiao YL, Sellors JW, Eder PS, et al. A new HPV-DNA test for cervical-cancer screening in developing regions: A cross-sectional study of clinical accuracy in rural China. *The Lancet Oncology* 2008;9(10):929–936. (In eng) http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18805733.
30. Gage JC, Ajenifuja KO, Wentzensen N, et al. Effectiveness of a simple rapid human papillomavirus DNA test in rural Nigeria. *International Journal of Cancer* 2012;131(12):2903–2909. (In eng). DOI: 10.1002/ijc.27563.
31. Schiffman M, Doorbar J, Wentzensen N, et al. Carcinogenic human papillomavirus infection. *Nature Reviews Disease Primers* 2016;2:16086. (In eng). DOI: 10.1038/nrdp.2016.86.
32. Johnson LG, Saidu R, Mbulawa Z, et al. Selecting human papillomavirus genotypes to optimize the performance of screening tests among South African women. *Cancer Medicine* 2020;9(18):6813–6824. (In eng). DOI: 10.1002/cam4.3329.
33. Kuhn L, Saidu R, Boa R, et al. Clinical evaluation of modifications to a human papillomavirus assay to optimise its utility for cervical cancer screening in low-resource settings: A diagnostic accuracy study. *Lancet Glob Health* 2020;8(2):e296-e304. (In eng). DOI: 10.1016/s2214-109x(19)30527-3.
34. Denny L, Saidu R, Boa R, et al. Point-of-care testing with Xpert HPV for single-visit, screen-and-treat for cervical cancer prevention: A demonstration study in South Africa. *Scientific Reports* 2023;13(1):16182. (In eng). DOI: 10.1038/s41598-023-43467-2.

53

Organoid Models for Early Detection and Diagnosis of Colorectal Cancer

Jinghui Liang, Janvhi Suresh Machhar, and Chengyu Liang

53.1 Introduction

Colorectal cancer (CRC) stands as the third most frequently diagnosed and the second most fatal cancer globally [1, 2]. It is projected that global CRC incidence will be doubled by 2035, with a pronounced increase anticipated in less developed countries [3]. On the other hand, mortality rates for CRC in developed countries have decreased steadily in the past two decades. This reduction is largely due to advances in CRC treatment but also CRC screening, which allows for early detection and treatment of precancerous lesions such as adenomatous polyps [4]. While general population-based screening for CRC predominantly rely on noninvasive methods such as stool-based or blood tests, alongside radiological imaging, invasive procedures like flexible sigmoidoscopy or colonoscopy are also employed (Figure 53.1a) [4, 5]. Despite colonoscopy being considered as the gold standard for CRC detection, concerns persist regarding its efficacy. Studies have revealed that 5%–8% of all CRC cases are diagnosed in patients who had undergone colonoscopies three to five years prior [6–9]. These instances of cancer detection subsequent to previous colonoscopies are termed interval colorectal cancer (iCRC), signifying missed lesions during initial screening, rapid tumor growth, or the emergence of new lesions between screenings (Figure 53.1a) [10]. In fact, a large retrospective cohort study revealed that one in 13 CRC cases could be classified as early or missed CRCs, detected after an initial colonoscopy [7]. Despite notable strides in screening technologies and strategies, the effectiveness and reliability of CRC screening programs remain suboptimal. This highlights the pressing need for improved methods to facilitate early detection and early diagnosis of CRC.

CRC develops through stepwise accumulation of genetic and epigenetic alterations that disrupts the balance between stem cell and differentiation program, eventually leading to the transition from normal to hyperproliferative epithelium [11, 12]. Current cancer imaging technologies can only detect well established tumors consisting of approximately 10^9 cells [13]. Early neoplastic lesions, which give origins to malignant tumors, cannot be easily distinguished. Addressing this limitation is essential for developing more predictive and translational research methodologies. Recent advances in organoid technologies make it possible to identify early-stage changes associated with aberrant stem cell-like activity in a controlled environment. Organoids are three-dimensional organ-like structures that can be generated from induced pluripotent stem cells (iPSCs) or cells sourced directly from patient biopsies (patient-derived organoids, PDOs), representing a diverse array of cell types including normal stem/progenitor cells, differentiated cells, and cancerous cells [14, 15]. Furthermore, organoid cultures provide greater accessibility for manipulation and detailed biological investigations compared to traditional tissue culture models. Therefore, organoid technology has expanded extensively to a diverse array of applications including but not limiting to disease modeling, genetic profiling, drug discovery, personalized medicine, and cell therapy (Figure 53.1b) [16–19]. That said, colon organoids, representing colonic crypts with lifelike microanatomy *in vitro*, stand as promising resources for evaluating and dissecting intestinal stem cell (ISC) behavior and CRC (Figure 53.1c). However, despite current progress in using colon organoids for drug discovery, their potential in facilitating early diagnosis of CRC has been less explored.

In this chapter, we focus on organoid changes induced by a cancer-promoting frameshift (FS) mutation in *UVRAG*, which was previously identified in our study [20] as a target gene of microsatellite instability (MSI), a mutator phenotype in CRC caused by the deficiency of DNA mismatch repair (MMR) [21]. *UVRAG* FS mutation increased the expressivity of MMR deficiency in tumorigenesis and was detected in 33% CRC with MSI [20]. Moreover, the mouse model with inducible expression of $UVRAG^{FS}$ demonstrated age-related hyperplastic changes and spontaneous cancers [22]. To facilitate early detection and study of oncogenic $UVRAG^{FS}$-induced changes, we presented data from $UVRAG^{FS}$ murine colon organoids developed from $UVRAG^{FS}$ mouse model as well as $UVRAG^{FS}$ human colon organoids derived from asymptomatic, early-stage MSI patients. Characterization of organoid data revealed early-occurring event of aberrant stem cell and differentiation activity following $UVRAG^{FS}$ pathway in both contexts. Thus, the organoid tools that could rapidly model early-stage neoplastic alterations represent promising resources for CRC detection/diagnosis, overcoming the limitations of existing conventional approaches.

DOI: 10.1201/9781003449942-59

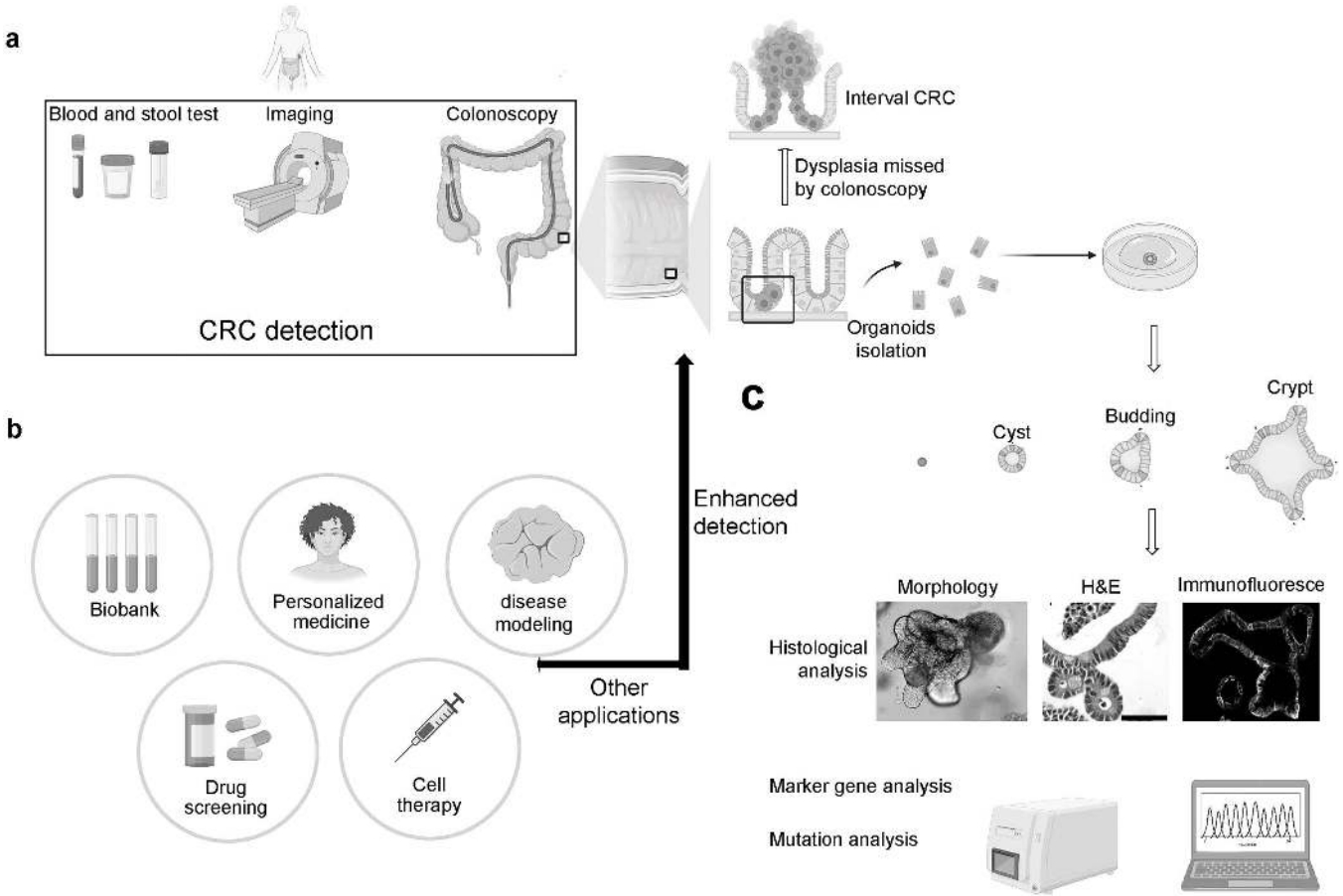


FIGURE 53.1 Organoid models can be served for detection and diagnosis in colorectal cancer. (a) Despite present cancer screening methods, the occurrence of missed CRC underscores the necessity for improved detection and diagnostic strategies. (b) Organoids have been widely used in studying tumor biology, identifying biomarkers, encompassing drug discovery, personalized medicine, and cell therapy. (c) Colon organoids, derived from stem cells isolated within crypts, exhibit complex budding structures resembling *in vivo* colonic epithelium. Histological analyses, including morphology assessments, hematoxylin and eosin (H&E) staining, and immunofluorescence, in conjunction with functional analyses such as marker gene expression and mutation analysis, are recruited to reveal the multidimensional features of organoids.

53.2 Materials

53.2.1 Biological Materials

Biopsy samples of intestinal tissue were collected from anonymized patients who had undergone colonoscopy. All patients gave informed consent for the storage and analysis of their samples. Patient materials were collected from the Hospital of the University of Pennsylvania with written informed consent under the protocol approved by the University of Pennsylvania Institutional Review Board. Mouse colon crypts were isolated from doxycycline-treated *UVRAG*^{FS} transgenic mice and littermate control mice as previously described [22]. All animal experiments were approved by the Institutional Animal Care and Use Committee (IACUC) of the Wistar Institute and performed in accordance with IACUC guidelines.

53.2.2 Biological Reagents

ROCK (Rho kinase) inhibitor Y-27632 dihydrochloride (Cat. No. M1817) was obtained from Abmole Bioscience. TrypLE Express Enzyme (Cat. No. 12604013) and TRIzol reagent (Cat. No. 15596018) were obtained from Thermo Fisher Scientific. iScript™ cDNA Synthesis Kit was obtained from BIO-RAD (Cat. No. 1708840) and Perfecta® SYBR® Green SuperMix Reaction Mixes (Cat. No. 66170261) were from Quantabio.

53.2.3 Organoid Culture

Advanced DMEM/F12 (Cat. No. 12634010) was obtained from Thermo Fisher Scientific. Mouse organoids were cultured in IntestiCult™ Organoid Growth Medium (mouse) (STEMCELL™ technologies, 06005) supplemented with 100 U/ml penicillin-streptomycin. Human organoids were cultured in IntestiCult™ Organoid Growth Medium (human) (STEMCELL™ technologies, 06010) supplemented with 100 U/ml penicillin-streptomycin. Organoids all grow at Matrigel (Corning, Cat. No. 356237) in a humidified incubator at 37 °C and 5% CO_2.

53.3 Methods

53.3.1 Sample Preparation for Organoid Culture

Murine colon tissues or human endoscopic biopsies were collected in ice-cold DMEM/F12 medium supplemented with 10 μM ROCK inhibitor Y–27632. Non-epithelial components were meticulously removed using surgical scissors or scalpels and forceps under a dissection microscope. The tissue underwent digestion in 25 ml of 5 mM EDTA for 30 minutes at 4 °C on a shaker. The digested tissue suspension was then filtered through a 70 μm cell strainer and centrifuged at 290 *g* for three minutes at 4 °C. The supernatant was carefully aspirated, and the pellet was resuspended in a mixture of standard organoid growth medium and Matrigel. This mixture was plated onto 48-well plates and allowed to solidify for 10–20 minutes in an incubator. Once solidified, 500 μl of organoid growth medium was added to each well, and the culture plate was incubated at 37 °C and 5% CO_2. The medium was changed every two to three days by carefully aspirating the old medium and replacing it with fresh, pre-warmed medium. It typically takes 9–12 days for human organoids and 6–9 days for mouse organoids to form complex, multi-lobed structures.

53.3.2 Organoid Passaging

Once the primary organoid cultures were established from tissue species, organoids were isolated, passaged, and replated into 48-well plates to facilitate further experimentation. For organoid passaging, TrypLE was added and incubated for a minimum of five minutes until organoids dissociated into single cells. The dissociated organoids were centrifuged at 290 *g* for three minutes at 4 °C. Upon completion of digestion, a wash with 10 ml of DMEM/F12 medium ensued, followed by centrifugation at 290 *g* for three minutes at 4 °C. The supernatant was aspirated, and the organoid pellet was resuspended in culture medium. The cell number was counted and approximately 100 organoids were resuspended in Matrigel and plated into 48-well plates as small droplets on the bottom of pre-warmed culture plates. After plating, the plates were inverted and incubated for 10–20 minutes until the Matrigel dome solidified. Pre-warmed medium was meticulously pipetted into the wells after solidification. The plates were then transferred to a humidified incubator at 37 °C with 5% CO_2 for continued culture.

53.3.3 Organoid Hematoxylin and Eosin (H&E) and Immunofluorescence (IF)

For H&E staining, colon organoids were fixed with 4% paraformaldehyde (PFA) for 48 h followed by standard H&E staining for morphology assessment. For IF analysis, organoids were isolated from Matrigel using recovery solution (Corning, Cat. No. 354253) and fixed with 4% PFA at room temperature (RT) for 20 min, followed by permeabilization with 0.5% Triton X-100 for 30 min. Samples were blocked using the blocking buffer (10% goat serum in PBS) for 2 h at RT and incubated overnight at 4 °C in primary antibodies against Muc2 (1:100, Thermo Fisher Scientific; PA5–103083) and Krt20 (1:200, Cell Signaling Technology; 13063). After gentle wash with PBS for three times, organoids were incubated in Alexa-568-conjuated secondary antibody for 1 h at RT. DNA was stained with DAPI (molecular probes). Organoids were mounted with the fluorescence mounting medium (Abcam, ab104135). The images of organoids were taken by confocal microscopy with a Leica Stellaris 8 confocal microscope.

53.3.4 Organoid Image Processing and Analyses

Brightfield images of colon organoids for morphological characterization were acquired using Echo revolve microscope (Echo, San Diego, CA). ImageJ/Fiji (version 1.51w, National Institutes of Health, Bethesda, MD) was used for image processing. The number of crypt domains—clearly defined budding protrusions—per organoid was counted manually in a blind manner. The branch length was measured by the straight line and oval selection tool of ImageJ/Fiji [23].

TABLE 53.1

Primer names and sequences for qPCR

Primers	Sequences
mouse *Gapdh* forward	5'-GAAGGCCGGGGCCCACTTGA-3'
mouse *Gapdh* reverse	5'-CTGGGTGGCAGTGATGGCATGG-3'
mouse *Axin2* forward	5'-GCGACGCACTGACCGACGAT-3'
mouse *Axin2* reverse	5'-GCAGGCGGTGGGTTCTCGGA-3'
mouse *Lgr5* forward	5'-GACAATGCTCTCACAGAC-3'
mouse *Lgr5* reverse	5'-GGAGTGGATTCTATTATTATGG-3'
mouse *Krt20* forward	5'-AGTTTTCACCGAAGTCTGAGTTC-3'
mouse *Krt20* reverse	5'-GTAGCTCATTACGGCTTTGGAG-3'
mouse *Muc2* forward	5'-AGGGCTCGGAACTCCAGAAA-3'
mouse *Muc2* reverse	5'-CCAGGGAATCGGTAGACATCG-3'
human *GAPDH* forward	5'-GGAGCGAGATCCCTCCAAAAT-3'
human *GAPDH* reverse	5'-GGCTGTTGTCATACTTCTCATGG-3'
human *AXIN2* forward	5'-CAACACCAGGCGGAACGAA-3'
human *AXIN2* reverse	5'-GCCCAATAAGGAGTGTAAGGACT-3'
human *LGR5* forward	5'-CTCCCAGGTCTGGTGTGTTG-3'
human *LGR5* reverse	5'-GAGGTCTAGGTAGGAGGTGAAG-3'
human *KRT20* forward	5'-GGACGACACCCAGCGTTTAT-3'
human *KRT20* reverse	5'-CGCTCCCATAGTTCACCGTG-3'
human *MUC2* forward	5'-GAGGGCAGAACCCGAAACC-3'
human *MUC2* reverse	5'-GGCGAAGTTGTAGTCGCAGAG-3'

53.3.5 RT-qPCR Analysis

RNA isolation from organoids was performed using TRIzol and complementary DNA (cDNA) was synthesized using iScript Reverse Transcription Supermix. Quantitative PCR (qPCR) were conducted using SYBR® Green Master SuperMix. Relative fold change in gene expression was determined by the $2^{-\Delta\Delta Ct}$ method. All $^{\Delta\Delta}Ct$ values were normalized to the housekeeping gene *Gapdh*. Table 53.1 shows the primers used for qPCR.

53.4 Results

53.4.1 Establishment of Murine Colon Organoid Culture

One major advantage of organoid model over traditional cell culture model is the recreation of the architecture and biology of their source organs *in vitro* through a stem cell-derived 3D culture system [24]. Under appropriate niche environment, isolated single ISCs self-organize to initially form cystic structures with a single central lumen. The cyst subsequently buds outward to form crypt-like structures through a process of proliferation, migration, and differentiation (Figure 53.1c) [25]. To demonstrate this property, we studied organoid formation using epithelium isolated from the mouse colon. To emulate the laminin-enriched crypt microenvironment, colonic crypts consisting of ISCs were plated into biologically derived matrix Matrigel supplemented with appropriate niche factors. This environment supports the survival, proliferation, and formation of organoids by ISCs [24]. Figure 53.2a illustrates the progressive changes of colon organoids, forming hollow cystic structures during days 1–3 after culture, and progressively exhibiting budding structures by day 6. Notably, during the differentiation process, a subset of cells undergoes apoptosis, contributing to structural degradation and disintegration of organoid assembly over time [26]. Consistent with this, after nine days' culture, we observed an increased presence of granular differentiated cells, which was accompanied by a darker appearance of the organoids, thicker organoid walls, and accumulation of debris in organoid lumen, suggesting organoid deterioration (Figure 53.2a). Hence, in our experience, organoid maturation, which harbors the well-differentiated cell types, is usually achieved around five to seven days' culturing. This outcome underscores the normal growth pattern exhibited by these organoids [27]. It is important to critically evaluate whether cultured organoids recapitulate the key aspects of crypt structures. For instance, when organoids were initially seeded at the base of matrix in direct contact with culture plate, they adopt a polygonal morphology and lacked the 3D structure and budding formations (Figure 53.2b). This deviation underscores the importance of optimal conditions and operation to sustain the proliferative and differentiative capacities of organoids for reliable investigations.

53.4.2 Functional Characterization of Murine Colon Organoids

To determine whether cultured organoids retain key functions and gene expression profile of colon, we analyzed the organoids by H&E staining and immunofluorescence. H&E staining of whole-mount organoid revealed a characteristic organization of bud formation following a three-day culture period and the discernible crypt-like structures within six days (Figure 53.2c). Consistent with this result, differentiation markers such as keratin 20 (KRT20) for epithelial cells and mucin2 (MUC2) for goblet cells were strongly expressed in the cytoplasm and/or cell membrane of organoids (Figure 53.2d), suggesting an established balance of self-renewal and multilineage differentiation of ISCs as previously reported [28]. Mucin secretion was also detected in the lumen of organoids after six-day culturing (Figure 53.2d). Furthermore, quantitative RT-PCR was performed for LGR5, a marker for ISCs, and KRT20 and MUC2, markers for differentiation. We observed a three- to four-fold increase of MUC2 and KRT20 expression on day 6 of organoid culture as compared to that of day 3, concomitant with a significant reduction of LGR5 (Figure 53.2e). These results underscore the rapid cell differentiation dynamics reminiscent of colon epithelial development, validating the utility of organoid as a model system for studying these processes [24, 25].

53.4.3 *UVRAG*^{FS} Murine Colon Organoids Recapitulate Early-Stage Tumor Cell Phenotype *In Vivo*

Due to the high prevalence of *UVRAG*^{FS} mutation in MSI patients and its cancer-promoting property *in vivo* [20, 22], we herein selected the *UVRAG*^{FS} mutant as the target for our organoid model study. Oncogenic *UVRAG*^{FS} is associated with poor prognosis and treatment resistance [20, 29]. CRC cell line experiments revealed that *UVRAG*^{FS} acts as a

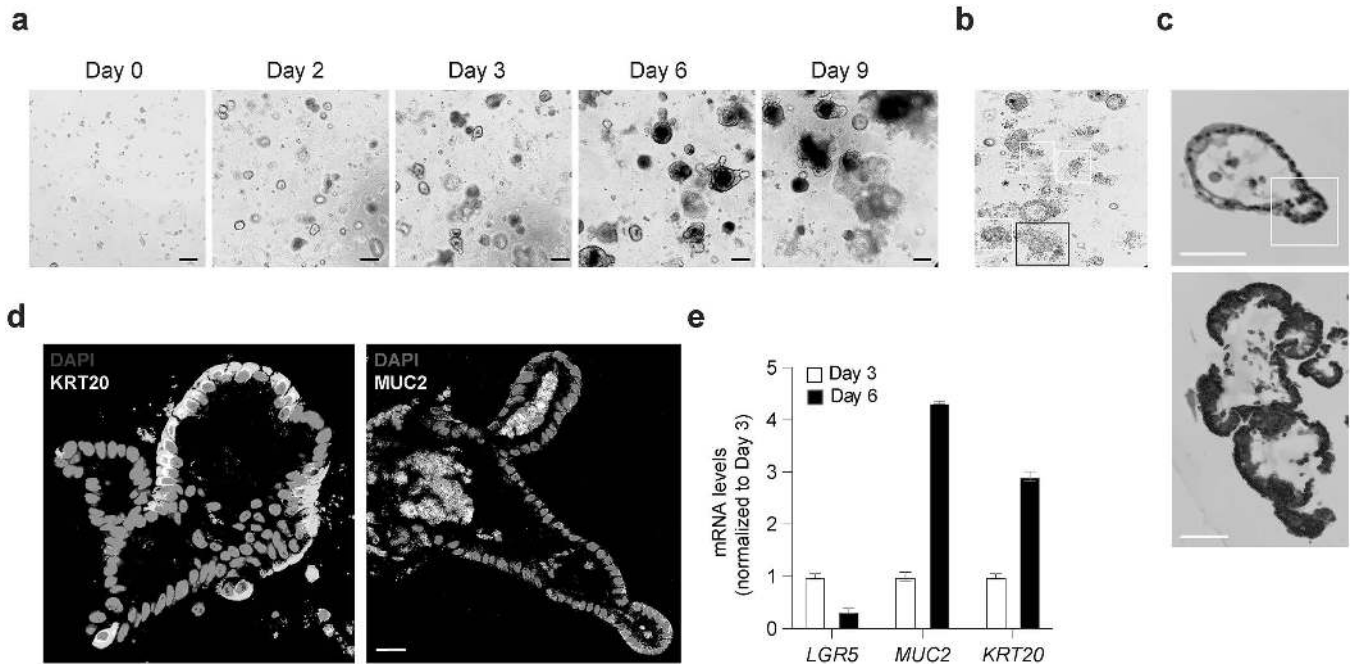


FIGURE 53.2 Morphological and functional analysis of murine colon organoids. (a) Representative phase-contrast images of time course culture of murine colon organoids. (b) Representative image of the murine colon organoid culture showing organoid death (white boxes) and suboptimal seeding (black box). (c) H&E staining of organoids that show bud formation (white box) after three-days' culture (top) and that form crypt structure after six days' culture (bottom). (d) Representative IF images for the differentiation markers *KRT20* (white, left image) and *MUC2* (white, right image), and DAPI (gray) for nuclei in colon organoids cultured for six days. (e) Real-time quantitative RT-PCR analysis for intestinal stem cell marker *LGR5* and differentiation markers *MUC2* and *KRT20* in organoids cultured for three days and six days. Data represent mean ± SD. Scale bars, 100 μm (a and c) and 10 μm (d).

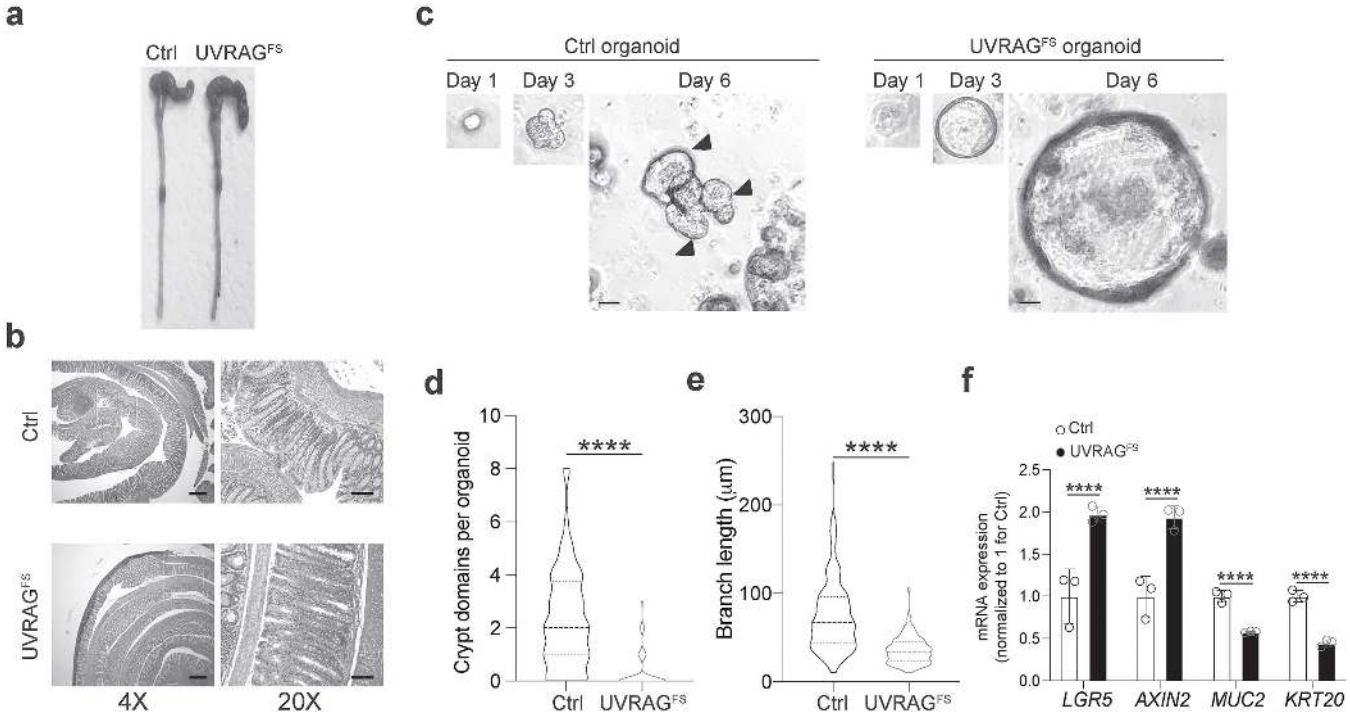


FIGURE 53.3 Structural and functional alterations in colon organoids derived from *UVRAG*^{FS} mice model. (a) Representative macroscopic images of colon from six-week-old *UVRAG*^{FS} mice and control (Ctrl) mice [22]. (b) H&E sections of the colons from six-week-old mice in (a). Scale bars, 300 μm (left) and 100 μm (right). (c) Representative brightfield images of time course culture of colon organoids from six-week-old *UVRAG*^{FS} mice and littermate control mice. Black arrowheads highlight crypt structures formed in the control (but not *UVRAG*^{FS}) organoid. Scale bars, 20 μm. (d and e) Quantification of crypt domains (d) and branch length (e) in organoid culture in (c). (f) Real-time quantitative RT-PCR analysis for intestinal stem cell markers (*LGR5* and *AXIN2*) and differentiation markers (*MUC2* and *KRT20*) in colon organoids in (c) cultured for six days. Mean ± SD is plotted (d–f). ****, $p < 0.0001$; two-tailed student's *t*-test.

dominant negative mutant of wild-type (WT) UVRAG and disrupts most if not all UVRAG-dependent tumor suppressing pathways [20, 29]. Furthermore, expression of oncogenic *UVRAG^{FS}* is sufficient to initiate spontaneous tumor growth in genetically engineered mouse model (GEMM) as shown in our previous work [22]. Despite the significant oncogenic effect of *UVRAG^{FS}*, hyperplastic changes of colon in *UVRAG^{FS}* GEMM could only be detected in aged mice between 8 and 12 months old [22]. To determine whether organoid culture could model the early-stage neoplastic transformation that follows *UVRAG^{FS}* induction, we isolated colon tissues from young (six-week-old) *UVRAG^{FS}* mice and control mice and developed *in vitro* organoid system as described previously (Figure 53.2). Notably, no distinguishable difference was observed in *UVRAG^{FS}* colon length and crypt structure compared to that of control (Figure 53.3a and b). However, upon seeding into organoid culture, wild-type organoids progressively formed crypt-like projections following a six-day culture as expected, whereas organoids from *UVRAG^{FS}* mice colon remained as giant "spheroids" and failed to form crypt-like structures (Figure 53.3c). To further validate this result, approximately 250 organoids showcasing crypt-like domains were captured using bright-field microscopy at day 6 and measured using ImageJ by Fiji. *UVRAG^{FS}* organoids showed strikingly decreased budding efficiency and crypt domain formation (Figure 53.3d and e), suggestive of impaired differentiation. Compared to controls, *UVRAG^{FS}* organoids displayed elevated mRNA expression of stem cell genes (*Lgr5* and *Axin2*) but decreased expression of differentiation markers (*Muc2* and *Krt20*) (Figure 53.3f). These observations suggest that expression of UVRAG^{FS} is sufficient to impair intestinal differentiation and align with characteristics typical of tumor-like organoids, consistent with previously published data [27, 30]. Thus, mouse colon organoid systems recapitulate core components of early-stage of crypt homeostasis and tumor-like phenotype, providing early sign and rapid modeling in early-stage cancer detection.

53.4.4 Patient-Derived *UVRAG^{FS}* Colon Organoids Forecast Disease Progression

To elucidate the clinical relevance of organoids as a tool for early detection of CRC, organoids were developed from patient-derived biopsy samples and cultured *ex vivo*. It is imperative to ascertain that patient-derived colon organoids (PDCOs) retain the genetic characteristics of the patient biopsies from which they were originally derived. To validate *UVRAG^{FS}* mutation being preserved in our organoid culture, PDCOs were established from *UVRAG^{WT}* and *UVRAG^{FS}* MSI patients and subjected to sequencing analysis after organoid passaging. As shown in Figure 53.4a, a heterozygous frameshift deletion that engages a single nucleotide A within the *UVRAG* A10-coding repeat was detected in *UVRAG^{FS}* PDCO, consistent with the genomic profile of the patient tissue. We have previously demonstrated that *UVRAG^{FS}* predisposes MSI cancers to increased genetic instability and cancer progression due to its significant roles in activating the epithelial-mesenchymal transition (EMT) program and promoting cell invasion [20]. To interrogate the *UVRAG^{FS}* genotype-tumor phenotype correlation in the *in vitro* organoid system, we established the organoid culture from an MMR-deficient *UVRAG^{WT}* patient with no tumor lesion detected during colonoscopy, and induced *UVRAG^{FS}* expression through lentivirus-mediated gene delivery as previously described [31]. As expected, control PDCO exhibited typical growth dynamics, manifesting crypt-like structures by day 4 with subsequently escalated crypt budding, resulting in the emergence of new multicellular bundling and branched structures by day 8 (Figure 53.4b). On the contrary, *UVRAG^{FS}* PDCO exhibited persistent cystic-like structures that failed to undergo budding or branching (Figure 53.4b), indicative of differentiation blocks, as seen with our mouse organoid data. Further imaging analysis revealed a reduction in crypt-like domains and diminished branch length in *UVRAG^{FS}* PDCO compared to controls (Figure 53.4c and d). In accord, induction of *UVRAG^{FS}* promoted the stem cell marker expression but suppressed cell lineage marker expression, indicating that organoid cells lose differentiation early after induction of oncogenic *UVRAG^{FS}* (Figure 53.4e). Although it is generally thought that loss of differentiation often occurs in late-stage tumors [32], our organoid data suggest that imbalance of stem cell-like behavior and differentiation process could be an early event that primes cancer initiation and progression. Although in-depth analysis of organoid plasticity and heterogeneity is certainly needed, our study highlights PDOs as reliable tools in forecasting oncogenic progression in an accelerated manner *in vitro*.

53.5 Discussion

When cancer is detected in an early stage, survival can be substantially improved [33]. While CRC is among the top common cancer types in the United States, it is also one of the most commonly misdiagnosed ones, partially because the early symptoms of CRC are easily confused with other digestive disorders such as inflammatory bowel disease, or because many people with early-stage CRC are asymptomatic [34]. Despite established screening programs of at-risk populations, current detection and diagnostic strategies of CRC vary in their sensitivity, specificity, and accuracy. The noninvasive stool-based tests, such as fecal immunochemical testing (FIT) for microscopic blood or cancer-associated DNA markers, though safe and easy to perform, do not reliably detect early-stage CRC and have high false-positive and false-negative rates [35]. The noninvasive liquid biopsy-based tests, including measuring circulating tumor DNA (ctDNA), circulating tumor cells (CTCs), and exosomes, in the blood hold promise for early diagnosis and recurrence detection, but are presently limited by their typical small sample size, instability, and/or lack of validated biomarkers [36, 37]. As such, colonoscopy, which allows visual examination of colonic surface and direct removal of precursor lesions, still remains the most efficient and reliable modality for CRC diagnosis and prevention [38]. However, the missed rate of colonoscopy is high even among experienced endoscopists, which was up to 26.3% for polyps, 22.4% for adenomas, and 11% for advanced adenomas

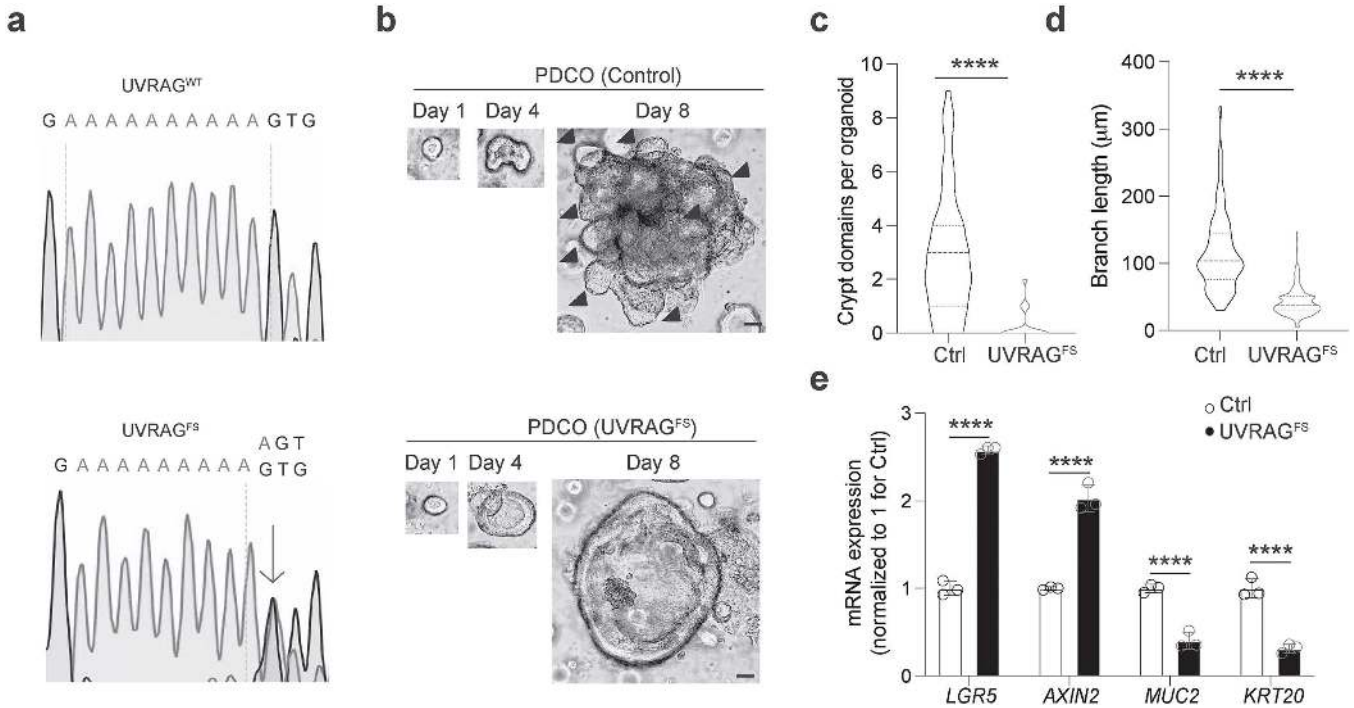


FIGURE 53.4 Patient-derived colon organoids (PDCOs) expressing *UVRAG*^{FS} show morphological changes and blocked cellular differentiation. (a) Representative PCR-based genotyping for *UVRAG* of outgrowing PDCO lines from *UVRAG* wild-type (WT) and *UVRAG* frameshift (FS) mutant patients. Arrow (right image) indicates the heterozygous FS deletion of one nucleotide A in the *UVRAG* A_{10}-coding repeat in *UVRAG*^{FS} PDO. (b) Representative brightfield images of time course culture of the PDO expressing vector control and *UVRAG*^{FS}. Black arrowheads highlight crypt structures formed in the control organoids after culturing for eight days. Scale bars, 20 μm. (c and d) Quantification of crypt domains (c) and branch length (d) in the organoid culture in (b). (e) Real-time quantitative RT-PCR analysis for intestinal stem cell markers (*LGR5* and *AXIN2*) and differentiation markers (*MUC2* and *KRT20*) in the PDOs of indicated genotypes eight days after culture. Mean ± SD is plotted (c–e). ****, $p < 0.0001$ versus control; two-tailed student's *t*-test.

according to a recent study of large cohorts of patients [39]. Notably, oncogenic changes initiated by cancer-promoting mutations may occur silently decades before a neoplastic lesion can be detected by endoscopic examination or advanced imaging technologies [33]. Tissue biopsy and genetic testing for CRC-related risk mutations, though generally more sensitive, cannot predict the onset and progression of the disease. Therefore, more reliable and predictive systems need to be developed to complement existing approaches in facilitating early detection, diagnosis, and treatment of CRC.

Recent advances in the development of stem cell-based 3D organoid system provide a superior representation of human tissues in a controlled environment *in vitro* [23, 40]. Unlike traditional 2D culture and patient-derived xenograft models, organoids derived from patient samples have been demonstrated to maintain the genomic, phenotypic, and functional changes of the parental tissues and have robust scalability [23, 40]. Given the notion that long-lived ISCs are major recipients of mutational insults that drive tumor initiation [41], ISCs-derived mini-gut organoids, which can be reliably passaged for expansion, serve as an extraordinary tool for probing early-stage CRC development in an accelerated manner, potentially revolutionizing diagnostic approaches. Indeed, a recent study by Dost and colleagues utilized human organoids to model early-stage lung adenocarcinoma (LUAD) and revealed early transcriptional changes as a result of oncogenic driver gene (*KRAS*) expression [42]. In this study, we explored the utility of colon organoids in the diagnosis of early-stage neoplastic changes that take months to see in mice and probably even longer in patients. We directly compared colon organoids derived from $UVRAG^{FS}$ mice with those from control mice at an early-stage time point. We observed $UVRAG^{FS}$ organoids exhibiting a spheroidal morphology devoid of the typical crypt-like architecture observed in normal organoids. Transcriptional profiling of stem cell and differentiation markers revealed loss of differentiation occurring early in day 6 $UVRAG^{FS}$ organoids culture. But in $UVRAG^{FS}$ mice, hyperplastic changes of colons could only be detected by hispathological and fluorescent imaging six to eight months after oncogenic $UVRAG^{FS}$ induction [22]. Thus, we speculate that the colon organoids can recapitulate the early stage of differentiation block and premalignant changes induced by $UVRAG^{FS}$ preceding CRC initiation. Although the disrupted balance between stem cell and differentiation programs is a defining property of CRC [12, 43], it is often thought to be a feature of established or advanced CRC. However, evidence suggests that process of aberrant developmental reprogramming of crypt-restricted ISCs can occur early and quickly, even before the cancer initiation [44]. This possibility is also supported by our data from $UVRAG^{FS}$ organoids, showing severely impaired differentiation in organoid culture from six-week-old $UVRAG^{FS}$ mice. Furthermore, our PDO studies showed that the mutational status of *UVRAG* was retained during organoid culture and passaging, consistent with prior studies demonstrating a high degree of concordance of mutations and copy number variations between organoids and parental tissues [45]. In humans, $UVRAG^{FS}$ mutation is correlated with a higher CRC tumor stage and worse survival [20]. In mouse models, CRC tumors with gain of $UVRAG^{FS}$ functions were highly metastatic and dedifferentiated [20]. Consistently, organoids derived from normal colonic tissue from MSI patient failed to generate differentiated structures upon $UVRAG^{FS}$ expression and showed high LGR5 and AXIN2 expression and low KRT20 and MUC2 expression. These results emphasize a $UVRAG^{FS}$-dependent oncogenic program that correlates with those observed in our GEMM and human datasets. Because longitudinal expansion of organoids enables the detection of low-frequency subclonal mutations with potential oncogenic implications, the utility of PDO system holds promise in augmenting clinical strategies for detecting precancerous lesions, particularly in the context of iCRC, where conventional diagnostic modalities may fall short in detecting subtle pathological changes despite normal colonoscopy findings.

In addition to early cancer detection, PDOs from CRC patients are amendable to high-throughput sequencing to analyze the mutational landscape [46]. Due to the enrichment of stem cell population, sequencing of PDOs enables detection of subclonal mutations at low frequency that were not readily detectable in primary samples, facilitating the search for targets for CRC [46]. Furthermore, PDOs can be used to screen anti-cancer drugs to identify the most effective therapy. To date, different drug screening tests on PDOs of CRC have been reported, which show high drug-genotype correlations and recapitulate patient-specific drug sensitivities [46–48]. These approaches also enable the identification of hidden vulnerabilities exposed upon targeted therapy, which could be used to combat therapy resistance. For instance, a recent study by Mertens and colleagues [49] utilizes an imaging-based screening method for PDOs of CRC and discriminated the microtubule-targeting agents, such as taxanes and vinca-alkaloids, as a single class of compounds to be most effective when combined with EGFR/MEK inhibitors. These studies demonstrate the feasibility of organoid platforms in assessing treatment response of individual patients and their potential to be implemented in personalized anti-cancer care.

Despite numerous advantages of organoid systems, challenge remains in developing and optimizing technologies to reliably harvest and analyze these mini-organs *in vitro* to mostly replicate the complexity and functionality of native tissues. Premature passaging of organoids may compromise the optimal pace of expansion, while delaying the passaging process may culminate in diminished cell viability and suboptimal expansion rates. Thus, meticulous consideration of the ideal passaging timing and condition is essential to uphold the integrity and reproductivity of organoid cultures [26]. With further technological development, particularly those enabling the creation of more complex multicellular structures, such as assembloids, to integrate multiple tissue types [50], we expect that organoid systems will be successfully applied to early cancer detection and precision medicine and thereby improve outcomes for patients with CRC and other cancer.

In conclusion, this study underscores the predictive value and transformative potential of organoid models in capturing the early stage of neoplastic changes and forecasting oncogenic progression. By combining morphological and molecular analyses, organoids offer valuable insights into the complex dynamics underlying CRC development and hold the promise

for personalized medicine approaches tailored to individual patient needs.

53.6 Future Trends

Future trends may involve advanced technologies like single-cell sequencing and CRISPR-based genome editing to elucidate cancer pathogenesis and identify therapeutic targets. The integration of organoids with microfluidics and artificial intelligence is anticipated to significantly enhance drug screening efficiency and precision oncology initiatives [51, 52]. Continued refinement and application of organoid models across cancer types are poised to revolutionize cancer research and clinical practice.

REFERENCES

1. Gopal, S. and Sharpless, N.E. (2021) Cancer as a global health priority. *JAMA*. doi: 10.1001/jama.2021.12778. Epub ahead of print. PMID: 34357387.
2. Dagnaw, G. G. and Dejene, H. (2025) Colorectal cancer in Ethiopia: Epidemiological trends, diagnostic and laboratory capacities, and challenges. *Sem Oncol* 52 (1), 19–26. PMID: 39537473.
3. Papamichael, D. et al. (2015) Treatment of colorectal cancer in older patients: International society of geriatric oncology (SIOG) consensus recommendations 2013. *Ann Oncol* 26 (3), 463–476.
4. Siegel, R.L. et al. (2019) Cancer statistics, 2019. *CA Cancer J Clin* 69 (1), 7–34.
5. Williams, J.R. and Ayscue, J.M. (2023) Colorectal cancer screening: A review of current screening options, timing, and guidelines. *J Surg Oncol* 127 (8), 1236–1246.
6. Arain, M.A. et al. (2010) CIMP status of interval colon cancers: Another piece to the puzzle. *Am J Gastroenterol* 105 (5), 1189–1195.
7. Singh, H. et al. (2010) Rate and predictors of early/missed colorectal cancers after colonoscopy in Manitoba: A population-based study. *Am J Gastroenterol* 105 (12), 2588–2596.
8. Gorski, T.F. et al. (1999) Colorectal cancer after surveillance colonoscopy: False-negative examination or fast growth? *Dis Colon Rectum* 42 (7), 877–880.
9. Cooper, G.S. et al. (2012) Prevalence and predictors of interval colorectal cancers in medicare beneficiaries. *Cancer* 118 (12), 3044–3052.
10. Sanduleanu, S. et al. (2015) Definition and taxonomy of interval colorectal cancers: A proposal for standardising nomenclature. *Gut* 64 (8), 1257–1267.
11. Nguyen, L.H. et al. (2020) Pathways of colorectal carcinogenesis. *Gastroenterology* 158 (2), 291–302.
12. Hanahan, D. (2022) Hallmarks of cancer: New dimensions. *Cancer Discov* 12 (1), 31–46.
13. Frangioni, J.V. (2008) New technologies for human cancer imaging. *J Clin Oncol* 26 (24), 4012–4021.
14. Crespo, M. et al. (2017) Colonic organoids derived from human induced pluripotent stem cells for modeling colorectal cancer and drug testing. *Nat Med* 23 (7), 878–884.
15. Huch, M. and Koo, B.K. (2015) Modeling mouse and human development using organoid cultures. *Development* 142 (18), 3113–3125.
16. Al Shihabi, A. et al. (2022) Personalized chordoma organoids for drug discovery studies. *Sci Adv* 8 (7), eabl3674.
17. Kim, M. et al. (2019) Patient-derived lung cancer organoids as in vitro cancer models for therapeutic screening. *Nat Commun* 10 (1), 3991.
18. Clevers, H. (2016) Modeling development and disease with organoids. *Cell* 165 (7), 1586–1597.
19. Fan, Y.Y. et al. (2019) Murine colonic organoid culture system and downstream assay applications. *Methods Mol Biol* 1576, 171–181.
20. He, S. et al. (2015) Truncating mutation in the autophagy gene UVRAG confers oncogenic properties and chemosensitivity in colorectal cancers. *Nat Commun* 6, 7839.
21. Sinicrope, F.A. and Sargent, D.J. (2012) Molecular pathways: Microsatellite instability in colorectal cancer: Prognostic, predictive, and therapeutic implications. *Clin Cancer Res* 18 (6), 1506–1512.
22. Quach, C. et al. (2019) A truncating mutation in the autophagy gene UVRAG drives inflammation and tumorigenesis in mice. *Nat Commun* 10 (1), 5681.
23. Onozato, D. et al. (2021) Generation of budding-like intestinal organoids from human induced pluripotent stem cells. *J Pharm Sci* 110 (7), 2637–2650.
24. Sato, T. et al. (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459 (7244), 262–265.
25. Date, S. and Sato, T. (2015) Mini-gut organoids: Reconstitution of the stem cell niche. *Annu Rev Cell Dev Biol* 31, 269–289.
26. Pleguezuelos-Manzano, C. et al. (2020) Establishment and culture of human intestinal organoids derived from adult stem cells. *Curr Protoc Immunol* 130 (1), e106.
27. Naser, A.N. et al. (2022) Three-dimensional culture of murine colonic crypts to study intestinal stem cell function ex vivo. *J Vis Exp* (188). doi: 10.3791/64534.
28. Fujii, M. et al. (2018) Human intestinal organoids maintain self-renewal capacity and cellular diversity in Niche-inspired culture condition. *Cell Stem Cell* 23 (6), 787–793. e6.
29. He, S. and Liang, C. (2015) Frameshift mutation of UVRAG: Switching a tumor suppressor to an oncogene in colorectal cancer. *Autophagy* 11 (10), 1939–1940.
30. Yip, H.Y.K. et al. (2018) Colon organoid formation and cryptogenesis are stimulated by growth factors secreted from myofibroblasts. *PLoS One* 13 (6), e0199412.
31. Maru, Y. et al. (2016) Lentivirus-based stable gene delivery into intestinal organoids. *Methods Mol Biol* 1422, 13–21.
32. Nieto, M.A. (2013) Epithelial plasticity: A common theme in embryonic and cancer cells. *Science* 342 (6159), 1234850.
33. Crosby, D. et al. (2022) Early detection of cancer. *Science* 375 (6586), eaay9040.
34. Lamprell, K. et al. (2023) People with early-onset colorectal cancer describe primary care barriers to timely diagnosis: A mixed-methods study of web-based patient reports in the United Kingdom, Australia and New Zealand. *BMC Prim Care* 24 (1), 12.
35. Carethers, J.M. (2023) Stool-based screening tests for colorectal cancer. *JAMA* 329 (10), 839–840.
36. Dasari, A. et al. (2020) ctDNA applications and integration in colorectal cancer: An NCI colon and rectal-anal task forces whitepaper. *Nat Rev Clin Oncol* 17 (12), 757–770.

37. Mauri, G. et al. (2022) Liquid biopsies to monitor and direct cancer treatment in colorectal cancer. *Br J Cancer* 127 (3), 394–407.
38. Zauber, A.G. et al. (2012) Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* 366 (8), 687–696.
39. Jiang, W. et al. (2023) Risk factors related to polyp miss rate of short-term repeated colonoscopy. *Dig Dis Sci* 68 (5), 2040–2049.
40. Lau, H.C.H. et al. (2020) Organoid models of gastrointestinal cancers in basic and translational research. *Nat Rev Gastroenterol Hepatol* 17 (4), 203–222.
41. Barker, N. et al. (2009) Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* 457 (7229), 608–611.
42. Dost, A.F.M. et al. (2020) Organoids model transcriptional Hallmarks of oncogenic KRAS activation in lung epithelial progenitor cells. *Cell Stem Cell* 27 (4), 663–678.e8.
43. de The, H. (2018) Differentiation therapy revisited. *Nat Rev Cancer* 18 (2), 117–127.
44. Bala, P. et al. (2023) Aberrant cell state plasticity mediated by developmental reprogramming precedes colorectal cancer initiation. *Sci Adv* 9 (13), eadf0927.
45. Lee, S.H. et al. (2018) Tumor evolution and drug response in patient-derived organoid models of bladder cancer. *Cell* 173 (2), 515–528.e17.
46. van de Wetering, M. et al. (2015) Prospective derivation of a living organoid biobank of colorectal cancer patients. *Cell* 161 (4), 933–945.
47. Vlachogiannis, G. et al. (2018) Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science* 359 (6378), 920–926.
48. Ganesh, K. et al. (2019) A rectal cancer organoid platform to study individual responses to chemoradiation. *Nat Med* 25 (10), 1607–1614.
49. Mertens, S. et al. (2023) Drug-repurposing screen on patient-derived organoids identifies therapy-induced vulnerability in KRAS-mutant colon cancer. *Cell Rep* 42 (4), 112324.
50. Andersen, J. et al. (2020) Generation of functional human 3D cortico-motor assembloids. *Cell* 183 (7), 1913–1929.e26.
51. Zuchowska, A. et al. (2024) Review: 3D cell models for organ-on-a-chip applications. *Anal Chim Acta* 1301, 342413.
52. Maramraju, S. et al. (2024) AI-organoid integrated systems for biomedical studies and applications. *Bioeng Transl Med* 9 (2), e10641.

Section 6

Computaional – Digital Technologies

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

54

Experimental Methods for Establishing and Maintaining Canine Bladder Cancer Organoids

Mohamed Elbadawy, Christopher Zdyrski, Hannah Wickham, Corey F. Saba, Savantha Thenuwara, Alison P. Mosichuk, Dipak Kumar Sahoo, Jonathan P. Mochel, and Karin Allenspach

54.1 Introduction

Rodent models, particularly mice, have extensively been used to study human diseases due to their cost-effectiveness and ease of access to genetically engineered animals. Adult stem cell (ASC)-derived organoids are three-dimensional cellular aggregates grown in an extracellular matrix (ECM) and stem cell supporting media. These organoids have the ability to accurately recapitulate the epithelial architecture of the primary tissue and can self-renew, self-organize, and differentiate into all epithelial cell types found in the parent tissue. Recent advances in organoid technology have opened avenues for the development of novel and fundamental research such as disease modeling, drug testing, and future transplantation attempts.[1–3] ASC-derived organoids have been developed from multiple epithelial tissues and species, including mouse models, alongside ASC-derived organoids from healthy and diseased human tissues. However, the use of murine models in biomedical research does not always allow for an accurate prediction of the efficacy of compounds in a preclinical setting.[4] This has been attributed to the fact that mouse models of human diseases often fail to accurately mimic the underlying human diseases. For chronic conditions, such as cancer, diabetes mellitus, Alzheimer's disease, inflammatory bowel disease, and cancer, it can be difficult to accurately model human disease in rodents due to their short lifespan and different anatomy and physiology, as well as vastly different diet and intestinal microbiome.[5–7]

The use of more diverse models, including dogs, has recently gained momentum. Dogs develop spontaneous diseases that are analogous to human diseases, which are typically more representative of chronic disease states than rodent models of diseases. One factor seen as advantageous for the biopharmaceutical industry is the fact that dogs have a relatively large body size, allowing for improved comparative pharmacological and dose-finding studies of compounds. Additionally, the longer lifespan of dogs compared to rodents, and their anatomy and physiology more closely resembling humans, are added benefits when using dogs. On the other hand, using the *in vivo* dog model, either by assessing safety and toxicity in healthy dogs (on purpose-bred research dogs) or testing the efficacy of compounds in dogs with analogous diseases as humans (in canine clinical trials to gain preclinical efficacy for human trials), is expensive and ethically questionable if using healthy dogs for research. Hence, using ASC-derived organoids from epithelial tissues of both healthy dogs and dogs that develop spontaneous diseases closely resembling those in humans could help bridge this divide.[8] Recently, there has been progress in developing larger bioarchives of canine organoid cell lines, which can be made available to researchers and eventually diminish the use of live dogs in biopharmaceutical research[9–11] and an Integrated Canine Data Commons are now available at National Cancer Institiute (https://canine-commons.cancer.gov/#/).

To this end, cancer has recently gained recognition as one of the diseases that can benefit from using dogs as a natural disease model. Dog and human cancers share numerous important characteristics, including a similar histological appearance, incidence ratio, clinical behavior such as the development of metastatic disease, immune response patterns, and genetic and histological heterogeneity within tumors and among patients.[12, 13] In dogs, bladder cancer (BC) comprises about 1.5%–2% of all naturally occurring cancers, a rate similar to that documented in humans.[14, 15] In both species, BC occurs mainly at older ages and with similar clinical symptoms including hematuria, which is the most common. Also, a great resemblance was noted in the microscopic picture and prostate involvement ratio between canine and human BC.[14] Similar histologic variants of BC were shown between dogs and humans such as plasmacytoid and rhabdoid types.[16, 17] Further, canine-invasive BC replicates efficiently the local invasion and distant metastases (in frequencies and sites) of human muscle-invasive BC that are difficult in experimental models.[14] At the molecular level, similar basal and luminal subtypes were shown in humans[18, 19] and canine muscle-invasive BC[11, 20] Additionally, the tumor suppressor pathway (P53) is inactivated in most cases of humans[21] and canine muscle-invasive BC.[20, 22] EGFR overexpression occurs in 73%–75% of high-grade BC in humans[23] and dogs.[24] Cox-2 is also overexpressed in more than 80% of humans[25] and canine BC[14]. HER-2 (EGFR2/ERBB2/NEU) has been found to be significantly overexpressed in canine muscle-invasive BC,[26] as is the case in human muscle-invasive BC,[27, 28] when compared to non-neoplastic urothelium. The standard treatment options for invasive BC in humans and dogs are nearly similar, and both species showed similar responses to treatments.[13] Canine organoid models in this context can serve as a preclinical screening tool to identify the optimal drugs

DOI: 10.1201/9781003449942-61

or drug combinations for novel cancer drug candidates, which can then be tested in canine clinical trials before they are taken into human clinical trials.

54.2 Materials and Methods

54.2.1 Sample Collection and Preparation

The diagnosis of BC in dogs is made by a combination of signalment, history, diagnostic tests, and ultimately urine cytology or histology. A fresh tissue biopsy from the bladder mass or urine samples from BC-diseased dogs can be used to generate BC organoids.

Tissue biopsies can be obtained by surgery or cystoscopy, or in some cases using a urinary catheter. These techniques can confirm the state of BC as well as its subvariant, grade, invasiveness, and stage.[29] Tissue biopsies should be cleaned from necrotic and dead tissues as well as fats and blood clots and washed with a precooled PBS several times. Representative pieces (0.5–1 cm^3) from different tumor parts should be taken and kept in a cold preservation buffer or shipping medium (Table 54.1) and immediately transferred to a cold environment (4 °C) and then to the laboratory for processing.

Urine samples can be collected either by free catch or through the use of a catheter. A gentle scratch of the bladder mass with the catheter end can help collect additional tumor cells.[11, 30] Next, the urine sample is centrifuged at 300 g/5 min/4 °C to remove the urine supernatant. The cell pellet should then be suspended in a cold preservation buffer or shipping medium (Table 54.1)[31] and immediately transferred to a cold environment (4 °C) and then to the laboratory for processing.

54.2.2 Methods of Deriving UC Organoids from Urine-Derived Stem Cells

After receiving the sample, tubes are centrifuged at 300 g/5 min/4 °C to remove the supernatant and check the cell pellet conditions. If the pellet is contaminated with red blood cells, which is common in BC, 5 mL of red blood cell lysis buffer (Roche; 11814389001) can be pre-warmed at room temperature (RT) and mixed with the pellet on a nutator for 5 min at RT to remove RBCs. After this incubation, again centrifuge the tube at 300 g/5 min/4 °C and discard the supernatant. If the pellet is large, it may contain debris such as mucous and crystals. If this is the case, the pellet should be washed three times with cold PBS and centrifuged at 100 g/3 min/4 °C to remove as much debris as possible. After the final wash, an appropriate pellet size should be resuspended using an ECM such as Matrigel (Corning; 356255) (30 µl/well) on ice and seeded in 24-well plates (Corning; 3524). After solidifying the gel at 37 °C for ~30 min in the CO_2 incubator, the culture medium (Table 54.2) is added (500–750 µL/well) and changed three times weekly. Once organoids grow in size and density, BC organoids can be passaged at a ratio of 1:3–4 as previously described.[11, 32, 33]

This preservation buffer (Table 54.1) can preserve samples for up to five days in a cooled condition.

54.2.3 Methods of Preparing UC Organoids from BC Tissue-Derived Stem Cells

In a biosafety cabinet, small fragments from these parts are then dissected and transferred to a 6 cm culture dish. They are then washed three times with 5 mL of cold PBS. After washing, 2 mL of cold advanced Dulbecco's Modified Eagle's Medium (AdDMEM) (Gibco; 12634–010) is added to these fractions. The fragments are then mechanically cut into fine pieces as much as possible using sterile surgical blades or curved micro scissors. These fine fragments can either be plated in Matrigel, allowing for the cells to grow, or further enzymatically digested following the steps detailed later.

The dissected tissue parts are then transferred into a sterile 15 mL tube containing pre-warmed (37 °C) AdDMEM with a digestive enzyme such as Liberase TH (a mixture of collagenase type II and dispase II solution at 0.125 mg/mL of AdDMEM) and incubated in a shaking thermostatic water bath set at 37 °C for up to 45 min with pipetting every 15 min. This allows enzymatic digestion and the release of stem cell clusters from tissue fragments. The release of stem cells can be confirmed under a bright field microscope. After digestion, the tube is centrifuged at 600 g/5 min/4 °C, resuspended in 5 mL of PBS, and mixed before being centrifuged again. If the pellet contains many red blood cells, RBC lysis buffer can be used as mentioned earlier. Cell clusters are then dissociated by mixing the pellet with 3–5 mL of pre-warmed (37 °C) TrypLE™ Express (Gibco; 12604–021). The cell suspension is then passed through a 100 µm nylon cell strainer to a 50 mL tube containing 5 mL AdDMEM, which stops the digestion from Tryple Express. The strainer can be washed twice with 5 mL of PBS to collect more cellular fractions. The strained fraction is then centrifuged at 300 g/5 min/4 °C and the pellet

TABLE 54.1

Preservation Buffer for Shipping Samples

Name	Stock Concentration	Amount to Add	Final Concentration	Distributor	Reference Number
AdDMEM/F12 (1x)		500 ml		Gibco	12634–010
HEPES	1M	1 ml	2 mM	VWR Life Science	J848–500ML
GlutaMAX-1 (100x)		5 ml		Gibco	35050–061
Penicillin-streptomycin solution (100x)		5 ml		Gibco	15070063
Primocin	50 mg/ml	1 ml	100 µg/ml	InvivoGen	ant-pm-1

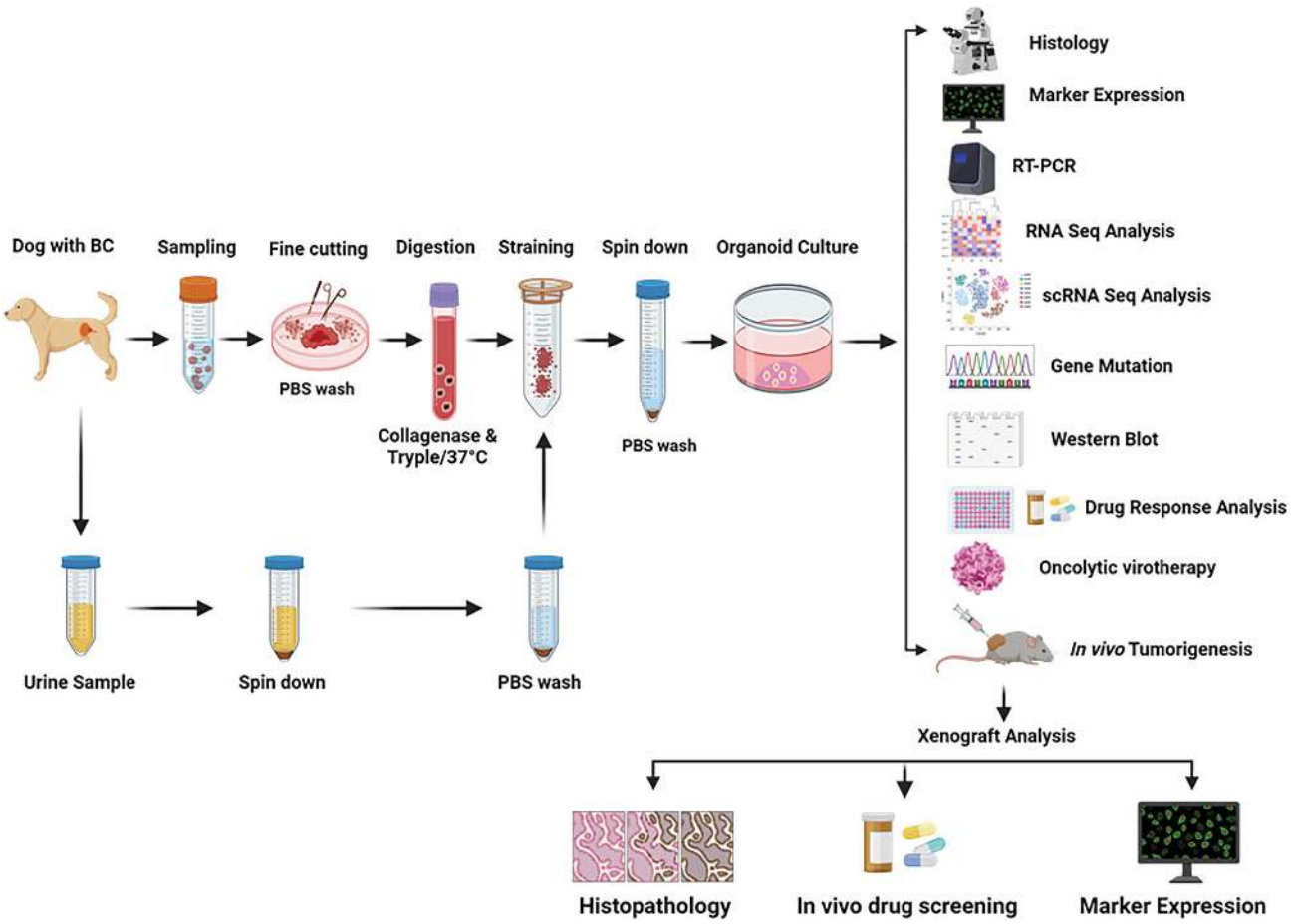


FIGURE 54.1 General workflow for generating canine bladder cancer organoids from tissue biopsies or urine. Potential analyses and applications of these patient-specific organoids are listed.

is washed twice with 5 mL of sterile PBS or AdDMEM. After the final wash, an appropriate pellet size is taken and mixed with an ECM such as Matrigel (30 µl/well) on ice, seeded in 24-well plates, and cultured as mentioned in the prior section. A general scheme for generating canine BC organoids from tissue biopsies or urine and their potential analyses and applications are shown in Figure 54.1.

54.2.4 Necessary Culture Medium Supplements and Growth Factors

When expanding canine BC organoids, the media composition may contain growth factors important for stem cell expansion, such as Wnt-3a, R-spondin, and Noggin. These growth factors can be obtained using 2D cell cultures (e.g., L-WRN cells) [34] and harvesting their conditioned media, which contains an approximate concentration of these secreted proteins. Alternatively, recombinant proteins can be purchased at a known weight (Table 54.2). Our laboratory utilizes the latter method to attain a more accurate concentration, along with a combination of antibiotics and additives that may help support stem cell growth for BC organoids as well as for other tissues.

Over time, we have modified and optimized our organoid media to include a variety of additional media components that support various organoid cultures including EGF, Nicotinamide, Gastrin, A-83–01, SB202190, Y27632, and CHIR99021. The specific pathways and functions of these growth factors have been previously described.[35] Further details and concentrations of the current media composition used to cultivate canine BC organoids are listed in Table 54.2.

54.2.5 Passaging of BC Organoids

After one to two weeks of culture, BC organoids can be passaged into new wells at a ratio of approximately 1:3–4 wells depending on the size and density of the organoids. To dissolve the ECM, such as Matrigel, 5 mmol/L EDTA/PBS is added to each well and the culture plate is placed on ice for 90 min, during which gel domes are detached using a sterile 1 mL tip after 30 min. Alternatively, 500 µL of Cell Recovery Solution (Corning; 354 270) can be added to each well, mixed by pipetting, and incubated on ice for 10–60 min to dissolve the Matrigel. The cell suspensions can then be collected into sterile 15 mL tubes and centrifuged at 100 g for 5 min at 4 °C. Cell pellets can be washed with PBS or AdDMEM before dissociation or left in Cell Recovery Solution. To dissociate the organoids, ~500 µL of TrypLE Express is added to the tubes at 37 °C for ~10 min depending on organoid plasticity. Vigorous

TABLE 54.2

Medium Components Necessary to Grow Canine Bladder Cancer Organoids

Component	Amount to Add	Solvent	Stock Conc.	Final Concentration (in 500 mL)	Distributor	Reference Number
Advanced DMEM/F12	500 mL	N/A	1X	N/A	Gibco	12634–010
FBS	40 mL	N/A	N/A	8.00%	Corning	35–010-CV
Glutamax	5 mL	N/A	200 mM	2 mM	Gibco	35050–061
HEPES	5 mL	N/A	1 M	10 mM	VWR Life Science	J848–500ML
Primocin	1 mL	N/A	50 mg/mL	100 µg/ml	InvivoGen	ant-pm-1
B27 supplement	10 mL	N/A	50X	1X	Gibco	17504–044
N2 supplement	5 mL	N/A	100X	1X	Gibco	17502–048
N-Acetyl-L-cysteine	1 mL	ddH_2O	500 mM	1 mM	Sigma	A9165–25G
Murine EGF	500 µL	0.1%BSA/PBS	50 µg/ml	50 ng/ml	PeproTech	315-09-1MG
Murine Noggin	500 µL	0.1%BSA/H_2O	100 µg/ml	100 ng/ml	PeproTech	250-38-250UG
Human R-Spondin-1	500 µL	0.1%BSA/PBS	500 µg/ml (500 µg/tube, add 1 mL solvent)	500 ng/ml	PeproTech	120-38-500UG
Murine Wnt-3a	1 mL (5 tubes)	0.1%BSA/PBS	50 µg/ml (10 µg/tube, add 200 µL solvent)	100 ng/ml	PeproTech	315-20-10UG
[Leu15]-Gastrin I human	50 µL	0.1%BSA/PBS	100 µM	10 nM	Sigma	G9145–.5MG
Nicotinamide	5 mL	ddH_2O	1 M	10 mM	Sigma	N0636–100G
A-83–01	50 µL	DMSO	5 mM	500 nM	Tocris	2939
SB202190	500 µL	DMSO	10 mM	10 µM	Biogems	1523072–25MG
Additional Supplements						
ROCK inhibitor (Y-27632)	1µL/1 mL	DMSO	10 mM	10 µM	Sigma	SCM075
Stemolecule CHIR99021 (GSK3β)	0.25 µL/1 mL	DMSO	10 mM	2.5 µM	Stemgent	04–0004-base

pipetting is performed to dissociate the organoids into single cells and 10% of FBS or 5 mL AdDMEM is added to the tube to neutralize trypsinization. Cell pellets are collected by centrifugation (100 g/5 min/4 °C) and mixed with new Matrigel on ice and then plated into a 24-well plate at 30 µL/well. The plate is then solidified in a CO_2 incubator at 37 °C for 30 min. Finally, 500 µL of warm culture medium (Table 54.2) is added to each well and changed three times weekly until the organoids become ready for passage (before organoids become too large or necrotic) or further analysis or harvesting.[11]

54.2.6 Cryopreservation and Reculture

To cryopreserve the organoids, remove the media and add 500 µL of Cell Recovery Solution (Corning; 354 270) to each well. Mix the organoids with a pipette, collect them, and transfer them to a 15 mL tube. Incubate on ice for 10 min to assist in the degradation of excess Matrigel before spinning at 100 g/5 min/4 °C. Discard the supernatant and resuspend the pellet in 6 mL of AdDMEM to remove the Cell Recovery Solution. Then, resuspend the pellet in 1 mL of cell banker solution, freezing medium,[(a)] or Cryostor CS10 (BioLife Solutions; 210 102) and immediately transfer it to a cryovial before placing it in a Mr. Frosty container (Nalgene; 5100–0001) filled with isopropanol and place in a 4 °C fridge for 10 min. After this short incubation, move the Mr. Frosty to −80 °C storage overnight with samples being transferred to vapor phase liquid nitrogen (−196 °C) indefinitely.

To subsequently thaw samples, recover cryovials from liquid nitrogen storage and immediately place them in a 37 °C water bath. After thawing the samples, transfer the cells in Cryostor to a tube containing 6 mL of AdDMEM and spin at 100 g/5 min/4 °C. Remove the supernatant, add Matrigel to the pellet, and plate 30 µL of sample in each well of 24 well plates. Transfer to a 37 °C and 5% CO_2 incubator for ~20–30 min, then add 500 µL of media to each well.

[(a)] A solution of 50% culture medium, 40% FBS, and 10% dimethyl sulfoxide. We found that the recovery of the organoids after thawing was superior when using Cryostor CS10.

54.3 Results

54.3.1 Characterization of Canine BC Organoids

To confirm the cell type of BC organoids, markers used to identify normal urothelium as well as markers used to identify urothelial carcinoma cells were utilized. Normal urothelium is composed of three distinct layers.[36] The basal or outermost layer is made up of two to three sublayers of small cuboid cells, while the intermediate or middle layers are made up of irregularly shaped cells, and the superficial layer that lines the apical side of the bladder epithelium is composed of umbrella cells. Each layer can be characterized using specific markers. For example, basal type BC organoids can be characterized using basal cell markers including cytokeratins (CK) 5, 6, and/or 14,

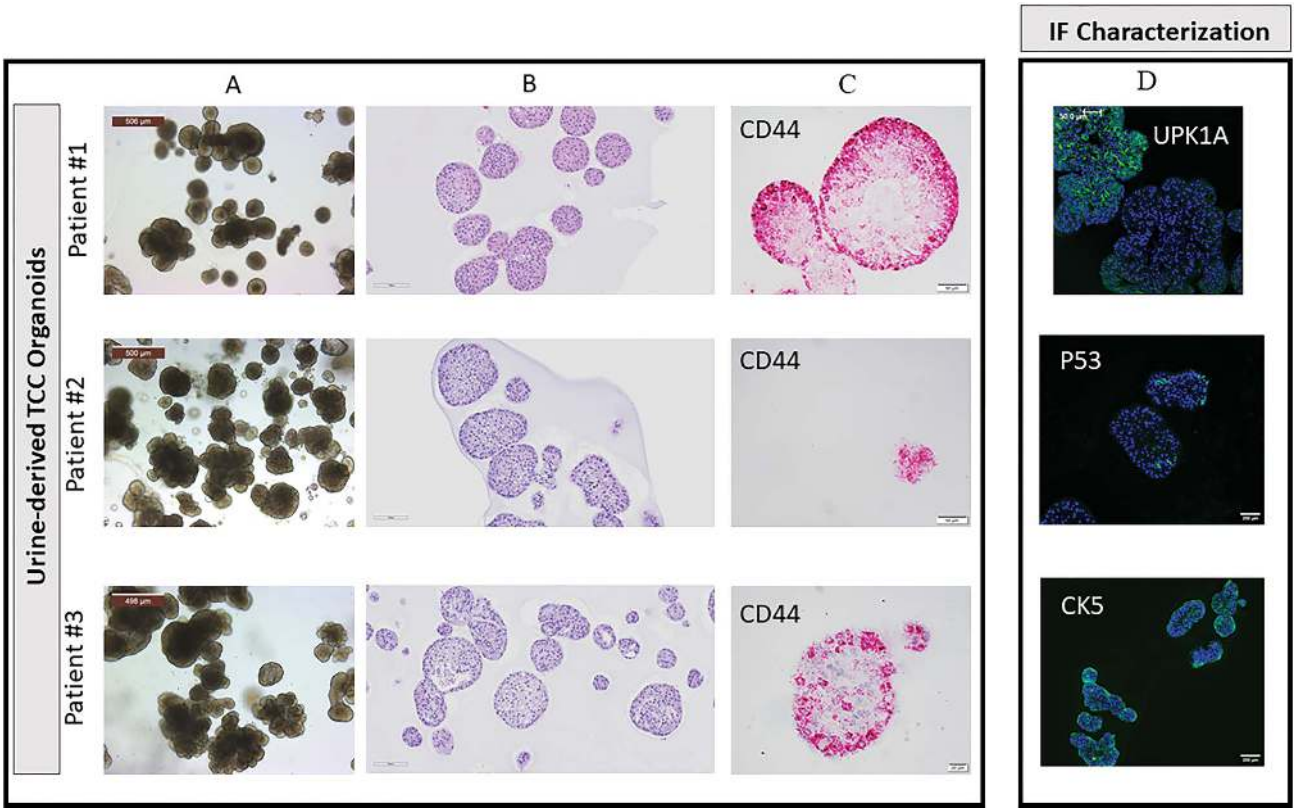


FIGURE 54.2 Morphological and histological characterization of three urine-derived canine bladder cancer organoid lines. (A) Bright field, (B) hematoxylin and eosin (H&E), and (C) RNA *in situ* hybridization (RNA-ish) corresponding to three urine-derived canine UC organoid lines. Additionally, representative (D) immunofluorescence (IF) taken from various organoid lines are shown (of note: P53 and CK5 did not utilize heat induced epitope retrieval [HIER] during processing). Organoid images were captured using a Leica Dmi1, a Stellaris confocal, or an Echo Revolution microscope. Image scalebars are in μm as indicated.

and/or P63 (a homolog of P53), and CD44.[11, 22, 37, 38] Luminal type BC organoids express markers of terminal urothelial differentiation, such as those seen in umbrella cells (Uroplakins Ia, Ib, II, and III, and KRT20).[11, 39] Additionally, loss or decrease of E-cadherin expression and an increase of vimentin and alpha smooth muscle actin expression in BC organoids can indicate the cells undergoing epithelial-mesenchymal transition (EMT), which can occur during tumor progression and metastasis.[40–42] Several methods can be used to characterize canine BC organoids including bright field imaging, histological staining using hematoxylin and eosin (H&E), RNA *in situ* hybridization (RNA-ish), and immunofluorescence (IF) (Figure 54.2).

54.3.2 Bright Field Microscopy

Bright field microscopy (BF) can be used to observe the organoid morphology, growth, and viability conditions during culture and to determine the morphological metrics of organoids such as diameter, length, and area.[43] BF microscopy also can be used to compare different culture conditions (such as healthy and disease models) and can help to carry out real-time and dynamic detection of live and dead cells.[44] Representative BF images of three canine BC organoids are shown in Figure 54.2A.

54.3.3 Histological Staining

To fix the organoids, the media is removed from each well, and 500 μL of formalin-acetic acid-alcohol is added as previously described.[45] Alternatively, tissue samples are fixed in 10% formalin, with approximately a 1:10 ratio of tissue to formalin. After 24 h at RT, samples are changed to 70% ethanol before processing and paraffin-embedding. Next, the samples are stained with H&E for histological assessment. After the H&E staining process, full slides are scanned on an appropriate microscope such as the Leica Aperio GT450 Scanner, and then analyzed using software, such as ImageScope (v12.4.3.5008). Representative H&E images of three canine BC organoids are shown (Figure 54.2B).

54.3.4 RNA *In Situ* Hybridization

RNA *in situ* hybridization relies on mRNA sequences and can be customizable to the gene of interest. When antibodies had low homology to canines, custom RNA-ish probes were used. RNA-ish is performed according to the manufacturer's protocol to characterize mRNA expression for CD44 (ACDBio; 838991) in the canine BC organoids using the RNAscope® 2.5 HD Detection Kit (ACDBio, Ref. 3223600). Once mounted and allowed to dry overnight, slides are imaged the next day

on an Olympus BX40 microscope and evaluated with a 60x objective. Representative ISH images of three canine BC organoids are shown (Figure 54.2C).

54.3.5 Immunofluorescence

For deparaffinization, microscopy slides should be submerged in xylene twice for 10 min. Then, they should be transferred to 100% ethanol twice for 1 min with regular agitation. Slides are then allowed to dry on tissue paper for 5 min. Next, the slides undergo heat induced epitope retrieval (HIER) using either of two buffers (citrate buffer—pH 6 or Tris/EDTA buffer—pH 9) heated to 75 °C in a HybEZ II Oven for 2 h. The slides were then removed from the oven for approximately 15 min with the lid off before being rinsed twice in PBS for 2 min each time before receiving a final 10 min wash in PBS. For permeabilization, slides should be placed in 0.25% Triton in PBS twice for 10 min each. After being rinsed in PBS three times, the slides are blocked in Casein in PBS for 1 h at RT. Then, the samples should have their primary antibody added and be incubated in a humidity tray overnight at 4 °C. Primary antibodies used included UPK1A (Novus Biologicals; NBP2–79733), P53 (Genetex; GTX102965) and CK5 (United States Biological; 222068). The following day, the slides should be rinsed in PBS, and the corresponding secondary antibody should be added (UPK1A [anti-mouse, Invitrogen; R37114], P53, and CK5 [anti-rabbit, Thermo; A32731]) at a ratio of 1:500–1000 in PBS for 1 h at RT. The slides should be rinsed again after this. Finally, the slides should be incubated with DAPI (Sigma; D9542–1MG) at a ratio of 1:500–1000 in PBS for 20 min and washed three times for 10 min each in PBS, and then switched to distilled water. For mounting, Fluoroshield (Sigma, Cat no: F6182–20ML) can be used and the slides should be allowed to dry overnight. The imaging can then be done on a Stellaris confocal microscope with scalebars added using the Leica LAS AF Lite (v. 2.6.0 build 7266) software. Representative IF images of three canine BC organoids are shown in Figure 54.2D.

54.4 Discussion

Canine organoid model has many applications in BC research. Canine BC organoids as a pivotal tool for preclinical drug screening represent a novel and promising approach, with the potential to streamline research and development of novel therapeutic modalities for BC in both human and canine patients. At the preclinical stage of drug research and development, canine organoids could help in identifying key therapeutic targets, which are important for designing new treatment options for BC using a relatively high-throughput system. Drug candidates evaluated in this model could later be tested *in vivo* for proof-of-concept studies in canine patients with BC, utilizing resources from the NIH Comparative Oncology Program. These studies could then demonstrate proof-of-concept efficacy and safety of drugs, enabling go/no go decisions before formal testing in human patients.

From a veterinary medicine perspective, organoids can accurately model the genetic heterogeneity and complexity inherent in BC, ensuring that therapeutic interventions are tailored to the specific genetic and molecular characteristics of individual tumors and enabling personalized medicine applications in canine patients.[11, 32, 33] Future development opportunities, including more complex versions of the canine model with the co-culture of immune cells, stromal cells, and endothelial cells incorporated into organoid cultures, will allow for emulation of the tumor microenvironment and will facilitate comparative studies of new modalities in BC immunotherapy.

Cancer-derived organoids have become a valuable tool in the field of oncology. They provide a more accurate representation of tumor biology and treatment response when compared to traditional two-dimensional cell cultures.[46] However, despite their benefits, there are still limitations and challenges when using organoids for these purposes. These include:

Tumor Heterogeneity and Subclonal Diversity. Tumors are composed of a diverse range of cell types, including cancer cells, stromal cells, immune cells, and endothelial cells. While cancer-derived organoids can capture some of this complexity, they may not fully replicate the tumor microenvironment and its various cell types, which are crucial for understanding tumor behavior and treatment response.[47, 48] Additionally, tumors often have multiple subclones with distinct genetic and phenotypic profiles.[49] Capturing this subclonal diversity can be challenging in organoids, which may limit their effectiveness in accurately modeling tumor evolution and resistance mechanisms.[47, 48]

Lack of Immune System Components. The tumor microenvironment plays a significant role in cancer progression and treatment response.[50] However, most organoid cultures do not include immune components, making it difficult to evaluate immunotherapies and understand immune evasion strategies and resistance to chemotherapy.[51]

Vascularization. Tumors have abnormal blood vessel networks that can affect drug delivery and hypoxia-induced pathways.[52] Organoids often lack a complete vascular network, which limits their ability to model the complex interactions between tumor cells and their surrounding vasculature.[53]

Drug Penetration and Metabolism. The dense extracellular matrix of organoids may hinder drug diffusion *in vitro*, potentially leading to an inaccurate assessment of drug efficacy and pharmacodynamics.[54] Furthermore, organoids may not fully replicate the metabolic activities of tumors, thereby affecting the metabolism of therapeutic agents and possibly skewing results regarding a drug's effectiveness or toxicity.[55]

54.5 Future Trends

To overcome these limitations, research is focused on:

Co-Culture Models. Research efforts are being made to integrate immune cells, stromal cells, and endothelial cells into organoid cultures in order to better mimic the complex tumor microenvironment.[56, 57] In particular, co-culturing organoids with other cell types or utilizing microfluidic devices can enhance

the modeling of tumor-stroma and tumor-immune interactions.[58]

Developing Vascularized Organoids. Advances in bioengineering aim to incorporate vascular networks within organoids in order to study drug delivery and hypoxia-related responses.[59]

Developing New Culture Technique. Application of microfluidic devices (Organ-On-A-Chip) capable of providing a gravity-driven flow that imitates a fluid stream in the bladder tumor and its perimetrical tissues can ease the oxygen supply and nutrient/waste exchange, resulting in a marked decrease of cell death, and support cell expansion as well as functional maturation, thereby recapitulating prominent features of BC in a much more precise and reproducible manner.[58]

Standardization of Protocols. Establishing standardized culture conditions and protocols can help improve the reproducibility and scalability of organoids derived from cancer.[60]

Despite these challenges, cancer-derived organoids represent a significant advancement in cancer research. They offer a more relevant model for studying tumor biology and testing existing therapeutics, as well as novel drug candidates (alone, or in combination with other therapeutic modalities, such as radiation therapy). As the field progresses, these models are expected to become even more sophisticated and valuable for translational research and personalized medicine.

Acknowledgments

The authors would like to thank Emily Rawlings and Maria Orbay-Cerrato, residents in Oncology Service, Small Animal Medicine and Surgery at Veterinary Teaching Hospital, University of Georgia, for their assistance in sample collection. We also thank Hannah Nicholson, Michael Catucci, and Bryan Melvin, graduate research assistants in Precision One Health Initiative, Department of Pathology, University of Georgia College of Veterinary Medicine, for their help culturing organoids and support throughout this chapter.

Abbreviations

Abbreviation	Details
AdDMEM	Advanced Dulbecco's Modified Eagle's Medium
ASC	Adult stem cells
BC	Bladder cancer
BF	Bright field
ECM	Extracellular matrix
H/E	Hematoxylin and eosin
IF	Immunofluorescence
PBS	Phosphate buffered saline
UC	Urothelial Carcinoma
RNA-ish	RNA *in situ* hybridization
RT	Room Temperature

REFERENCES

1. Artegiani, B. & Clevers, H. Use and application of 3D-organoid technology. *Hum Mol Genet* 27, R99–R107, doi:10.1093/hmg/ddy187 (2018).
2. Abugomaa, A. & Elbadawy, M. Patient-derived organoid analysis of drug resistance in precision medicine: Is there a value? *Expert Rev Precis Med Drug Dev* 5, 1–5, doi:10.1080/23808993.2020.1715794 (2020).
3. Gabriel, V. *et al*. Adult animal stem cell-derived organoids in biomedical research and the one health paradigm. *Int J Mol Sci* 25, 701 (2024).
4. Perlman, R. L. Mouse models of human disease: An evolutionary perspective. *Evol Med Public Health* 2016, 170–176, doi:10.1093/emph/eow014 (2016).
5. LeBlanc, A. K. *et al*. Perspectives from man's best friend: National academy of medicine's workshop on comparative oncology. *Sci Transl Med* 8, 324ps325, doi:10.1126/scitranslmed.aaf0746 (2016).
6. Coelho, L. P. *et al*. Similarity of the dog and human gut microbiomes in gene content and response to diet. *Microbiome* 6, 72, doi:10.1186/s40168-018-0450-3 (2018).
7. Schneider, B. *et al*. Model-based reverse translation between veterinary and human medicine: The one health initiative. *CPT Pharmacometrics Syst Pharmacol* 7, 65–68, doi:10.1002/psp4.12262 (2018).
8. Mochel, J. P. *et al*. Intestinal stem cells to advance drug development, precision, and regenerative medicine: A paradigm shift in translational research. *AAPS J* 20, 17, doi:10.1208/s12248-017-0178-1 (2017).
9. Chandra, L. *et al*. Derivation of adult canine intestinal organoids for translational research in gastroenterology. *BMC Biol* 17, 33, doi:10.1186/s12915-019-0652-6 (2019).
10. Allenspach, K., Zavros, Y., Elbadawy, M., Zdyrski, C. & Mochel, J. P. Leveraging the predictive power of 3D organoids in dogs to develop new treatments for man and man's best friend. *BMC Biol* 21, 297, doi:10.1186/s12915-023-01799-5 (2023).
11. Elbadawy, M. *et al*. Establishment of a novel experimental model for muscle-invasive bladder cancer using a dog bladder cancer organoid culture. *Cancer Sci* 110, 2806–2821, doi:10.1111/cas.14118 (2019).
12. de Brot, S. *et al*. The dog as an animal model for bladder and urethral urothelial carcinoma: Comparative epidemiology and histology. *Oncol Lett* 16, 1641–1649, doi:10.3892/ol.2018.8837 (2018).
13. Knapp, D. W. *et al*. Naturally-occurring invasive urothelial carcinoma in dogs, a unique model to drive advances in managing muscle invasive bladder cancer in humans. *Front Oncol* 9, 1493, doi:10.3389/fonc.2019.01493 (2019).
14. Knapp, D. W. *et al*. Urinary bladder cancer in dogs, a naturally occurring model for cancer biology and drug development. *ILAR J* 55, 100–118, doi:10.1093/ilar/ilu018 (2014).
15. Antoni, S. *et al*. Bladder cancer incidence and mortality: A global overview and recent trends. *Eur Urol* 71, 96–108, doi:10.1016/j.eururo.2016.06.010 (2017).
16. Lin, S. J. *et al*. Urothelial carcinomas of the urinary bladder with plasmacytoid or rhabdoid features and tendency of epithelial-mesenchymal transition in 3 dogs. *Vet Pathol* 55, 673–677, doi:10.1177/0300985818771151 (2018).
17. Humphrey, P. A., Moch, H., Cubilla, A. L., Ulbright, T. M. & Reuter, V. E. The 2016 WHO classification of tumours of

the urinary system and male genital organs-part B: Prostate and bladder tumours. *Eur Urol* 70, 106–119, doi:10.1016/j.eururo.2016.02.028 (2016).
18. Robertson, A. G. *et al*. Comprehensive molecular characterization of muscle-invasive bladder cancer. *Cell* 171, 540–556.e525, doi:10.1016/j.cell.2017.09.007 (2017).
19. Choi, W. *et al*. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. *Cancer Cell* 25, 152–165, doi:10.1016/j.ccr.2014.01.009 (2014).
20. Dhawan, D., Hahn, N. M., Ramos-Vara, J. A. & Knapp, D. W. Naturally-occurring canine invasive urothelial carcinoma harbors luminal and basal transcriptional subtypes found in human muscle invasive bladder cancer. *PLoS Genet* 14, e1007571, doi:10.1371/journal.pgen.1007571 (2018).
21. Watanabe, R., Tomita, Y., Nishiyama, T., Tanikawa, T. & Sato, S. Correlation of p53 protein expression in human urothelial transitional cell cancers with malignant potential and patient survival. *Int J Urol* 1, 43–48, doi:10.1111/j.1442-2042.1994.tb00007.x (1994).
22. Hanazono, K. *et al*. Immunohistochemical expression of p63, Ki67 and beta-catenin in canine transitional cell carcinoma and polypoid cystitis of the urinary bladder. *Vet Comp Oncol* 14, 263–269, doi:10.1111/vco.12095 (2016).
23. Chaux, A. *et al*. High epidermal growth factor receptor immunohistochemical expression in urothelial carcinoma of the bladder is not associated with EGFR mutations in exons 19 and 21: A study using formalin-fixed, paraffin-embedded archival tissues. *Hum Pathol* 43, 1590–1595, doi:10.1016/j.humpath.2011.11.016 (2012).
24. Dhawan, D. *et al*. Comparative gene expression analyses identify luminal and basal subtypes of canine invasive urothelial carcinoma that mimic patterns in human invasive bladder cancer. *PLoS One* 10, e0136688, doi:10.1371/journal.pone.0136688 (2015).
25. Bourn, J., Pandey, S., Uddin, J., Marnett, L. & Cekanova, M. Detection of tyrosine kinase inhibitors-induced COX-2 expression in bladder cancer by fluorocoxib A. *Oncotarget* 10, 5168–5180, doi:10.18632/oncotarget.27125 (2019).
26. Millanta, F. *et al*. Overexpression of HER-2 via immunohistochemistry in canine urinary bladder transitional cell carcinoma—a marker of malignancy and possible therapeutic target. *Vet Comp Oncol* 16, 297–300, doi:10.1111/vco.12345 (2018).
27. Eriksson, P., Sjodahl, G., Chebil, G., Liedberg, F. & Hoglund, M. HER2 and EGFR amplification and expression in urothelial carcinoma occurs in distinct biological and molecular contexts. *Oncotarget* 8, 48905–48914, doi:10.18632/oncotarget.16554 (2017).
28. Cimpean, A. M. *et al*. Critical overview of HER2 assessement in bladder cancer: What is missing for a better therapeutic approach? *Anticancer Res* 37, 4935–4942, doi:10.21873/anticanres.11903 (2017).
29. Brambilla, E. *et al*. Grading systems for canine urothelial carcinoma of the bladder: A comparative overview. *Animals (Basel)* 12, doi:10.3390/ani12111455 (2022).
30. Elbadawy, M. *et al*. Establishment of an experimental model of normal dog bladder organoid using a three-dimensional culture method. *Biomed Pharmacother* 151, 113105, doi:10.1016/j.biopha.2022.113105 (2022).
31. Sato, Y. *et al*. Establishment of an experimental model of canine malignant mesothelioma organoid culture using a three-dimensional culture method. *Biomed Pharmacother* 162, 114651, doi:10.1016/j.biopha.2023.114651 (2023).
32. Elbadawy, M. *et al*. Anti-tumor effect of trametinib in bladder cancer organoid and the underlying mechanism. *Cancer Biol Ther* 22, 357–371, doi:10.1080/15384047.2021.1919004 (2021).
33. Abugomaa, A. *et al*. Anti-cancer activity of Chaga mushroom (Inonotus obliquus) against dog bladder cancer organoids. *Front Pharmacol* 14, 1159516, doi:10.3389/fphar.2023.1159516 (2023).
34. Miyoshi, H. & Stappenbeck, T. S. In vitro expansion and genetic modification of gastrointestinal stem cells in spheroid culture. *Nat Protoc* 8, 2471–2482, doi:10.1038/nprot.2013.153 (2013).
35. Rauth, S., Karmakar, S., Batra, S. K. & Ponnusamy, M. P. Recent advances in organoid development and applications in disease modeling. *Biochim Biophys Acta Rev Cancer* 1875, 188527, doi:10.1016/j.bbcan.2021.188527 (2021).
36. Jafari, N. V. & Rohn, J. L. The urothelium: A multi-faceted barrier against a harsh environment. *Mucosal Immunol* 15, 1127–1142, doi:10.1038/s41385-022-00565-0 (2022).
37. Walz, S. *et al*. A protocol for organoids from the urine of bladder cancer patients. *Cells* 12, doi:10.3390/cells12172188 (2023).
38. Sahoo, D. K. *et al*. Abstract 3092: Urine-derived urinary carcinoma organoids: A novel tool for providing new insights into human and canine bladder cancer treatment. *Cancer Research* 82, 3092–3092, doi:10.1158/1538-7445.Am2022-3092 (2022).
39. Mullenders, J. *et al*. Mouse and human urothelial cancer organoids: A tool for bladder cancer research. *Proc Natl Acad Sci U S A* 116, 4567–4574, doi:10.1073/pnas.1803595116 (2019).
40. Tambunan, M. P., Saraswati, M., Umbas, R., Mochtar, C. A. & Hamid, A. R. A. H. E-cadherin expressions on bladder and its association with cancer progressivity: A retrospective cohort study. *Afr J Urol* 28, 14, doi:10.1186/s12301-022-00280-z (2022).
41. Monteiro-Reis, S. *et al*. Vimentin epigenetic deregulation in bladder cancer associates with acquisition of invasive and metastatic phenotype through epithelial-to-mesenchymal transition. *Int J Biol Sci* 19, 1–12, doi:10.7150/ijbs.77181 (2023).
42. Mezheyeuski, A. *et al*. Fibroblasts in urothelial bladder cancer define stroma phenotypes that are associated with clinical outcome. *Sci Rep* 10, 281, doi:10.1038/s41598-019-55013-0 (2020).
43. Borten, M. A., Bajikar, S. S., Sasaki, N., Clevers, H. & Janes, K. A. Automated brightfield morphometry of 3D organoid populations by OrganoSeg. *Sci Rep* 8, 5319, doi:10.1038/s41598-017-18815-8 (2018).
44. Fei, K., Zhang, J., Yuan, J. & Xiao, P. Present application and perspectives of organoid imaging technology. *Bioengineering (Basel)* 9, doi:10.3390/bioengineering9030121 (2022).
45. Gabriel, V. *et al*. Standardization and maintenance of 3D canine hepatic and intestinal organoid cultures for use in biomedical research. *J Vis Exp*, doi:10.3791/63515 (2022).

46. Kapalczynska, M. *et al*. 2D and 3D cell cultures—a comparison of different types of cancer cell cultures. *Arch Med Sci* 14, 910–919, doi:10.5114/aoms.2016.63743 (2018).
47. Fang, Z., Li, P., Du, F., Shang, L. & Li, L. The role of organoids in cancer research. *Exp Hematol Oncol* 12, 69, doi:10.1186/s40164-023-00433-y (2023).
48. Jiang, X. *et al*. Organoids: Opportunities and challenges of cancer therapy. *Front Cell Dev Biol* 11, 1232528, doi:10.3389/fcell.2023.1232528 (2023).
49. Jamal-Hanjani, M., Quezada, S. A., Larkin, J. & Swanton, C. Translational implications of tumor heterogeneity. *Clin Cancer Res* 21, 1258–1266, doi:10.1158/1078-0432.CCR-14-1429 (2015).
50. Wang, Q. *et al*. Role of tumor microenvironment in cancer progression and therapeutic strategy. *Cancer Med* 12, 11149–11165, doi:10.1002/cam4.5698 (2023).
51. Neal, J. T. *et al*. Organoid modeling of the tumor immune microenvironment. *Cell* 175, 1972–1988.e1916, doi:10.1016/j.cell.2018.11.021 (2018).
52. Dewhirst, M. W. & Secomb, T. W. Transport of drugs from blood vessels to tumour tissue. *Nat Rev Cancer* 17, 738–750, doi:10.1038/nrc.2017.93 (2017).
53. Zhao, X. *et al*. Review on the vascularization of organoids and organoids-on-a-chip. *Front Bioeng Biotechnol* 9, 637048, doi:10.3389/fbioe.2021.637048 (2021).
54. Karolak, A., Poonja, S. & Rejniak, K. A. Morphophenotypic classification of tumor organoids as an indicator of drug exposure and penetration potential. *PLoS Comput Biol* 15, e1007214, doi:10.1371/journal.pcbi.1007214 (2019).
55. Richiardone, E., Van den Bossche, V. & Corbet, C. Metabolic studies in organoids: Current applications, opportunities and challenges. *Organoids* 1, 85–105 (2022).
56. Luckett, K. A. & Ganesh, K. Engineering the immune microenvironment into organoid models. *Annu Rev Canc Biol* 7, 171–187, doi:10.1146/annurev-cancerbio-061421-040659 (2023).
57. Yuan, J., Li, X. & Yu, S. Cancer organoid co-culture model system: Novel approach to guide precision medicine. *Front Immunol* 13, 1061388, doi:10.3389/fimmu.2022.1061388 (2022).
58. Duzagac, F., Saorin, G., Memeo, L., Canzonieri, V. & Rizzolio, F. Microfluidic organoids-on-a-chip: Quantum leap in cancer research. *Cancers (Basel)* 13, doi:10.3390/cancers13040737 (2021).
59. Strobel, H. A., Moss, S. M. & Hoying, J. B. Vascularized tissue organoids. *Bioengineering (Basel)* 10, doi:10.3390/bioengineering10020124 (2023).
60. Zhao, Z. *et al*. Organoids. *Nat Rev Methods Primers* 2, 94, doi:10.1038/s43586-022-00174-y (2022).

55

Top-Down Proteomics of Cancer Cells by Advanced Capillary Electrophoresis–Mass Spectrometry

Tian Xu and Liangliang Sun

55.1 Introduction

Proteoforms derived from the same protein-coding gene can be heterogenous due to genetic variations/single nucleotide polymorphisms (SNPs), alternatively spliced ribonucleic acid (RNA) transcripts, post-translational modifications (PTMs), and truncations, leading to diverse functions and biological outcome [1, 2]. Investigating proteoform variations in the critical cancer process is crucial for interpreting specific cancer phenotypes and discovery of biomarkers. Conventionally, bottom-up proteomics (BUP) was adopted as a dominant proteomic strategy in cancer research [3–5]. However, it fails to capture the combinatorial PTMs or sequence variations in individual forms of proteins. Top-down proteomics (TDP) aims to characterize intact proteoforms, which provides a comprehensive picture of protein variations in cancer events. The Kelleher group previously compared TDP and BUP in studying patient-derived breast tumor xenografts [6]. Their work highlighted the significance of TDP in acquiring aberrations of cancers at the proteoform level. Another study from their group performed targeted immunoprecipitation (IP)-TDP for characterizing and quantifying KRAS4b proteoforms in colorectal cancer (CRC) cell lines and tumors. The study revealed crosstalk between genetically encoded mutations and proteoform PTMs [7]. Recently, they explored KRAS isoforms (KRAS4a and KRAS4b) using the same TDP workflow and disclosed distinct modifications between the two isoforms across different cells and tissues [8].

TDP greatly relies on mass spectrometry (MS)/tandem mass spectrometry (MS/MS) based approaches for proteoform mass measurement and sequence/PTM identification. For complex proteome samples, separation prior to MS or MS/MS is necessary for achieving good proteome coverage by reducing sample complexity. Various liquid chromatography, such as reversed-phase liquid chromatography (RPLC), size exclusion chromatography (SEC), and capillary electrophoresis (CE) methods, such as capillary zone electrophoresis (CZE), have been widely applied for protein separation [9–12]. Typically, CE has gained increasing applications in TDP in recent years because of high separation resolution and low loss of proteins in the capillary column compared to LC [13, 14]. Moreover, constructing multidimensional platforms by coupling different separation approaches can effectively boost proteoform identifications, directing TDP to deeper proteome coverage. For instance, our previous work coupling SEC and RPLC sample prefractionation with CZE-MS/MS attributed around 6000 proteoform IDs from *E. coli*, demonstrating a ten-fold ID number enhancement compared with one-dimensional CZE-MS/MS studies [11].

In this work, we performed a deep TDP study of non-metastatic (SW480) and metastatic (SW620) human CRC cell lines. We employed four different capillary zone electrophoresis–tandem MS approaches, performing 1D CZE-MS/MS, 2D SEC-CZE-MS/MS, 2D RPLC-CZE-MS/MS, and 3D SEC-RPLC-CZE-MS/MS analyses of the two cell lines for proteoform ID and label-free quantification (LFQ). Comparison of proteoform differences (PTMs, single amino acid variants, abundance) between the two cell lines allows us to discover the potential proteoform biomarkers involving critical cancer-related pathways and regulating cancer metastasis. The related work has previously been published [15].

55.2 Materials

55.2.1 Chemicals and Solutions

All the solutions and buffers are prepared fresh using MS-grade water or organic solvents.

1. Cell lysis buffer: 8 M urea, 50 mM Tris (pH 8.2), 1 mM β-glycerophosphate, 1 mM phenylmethylsulfonyl fluoride, 75 mM sodium chloride, 1 mM sodium fluoride, 1 mM sodium orthovanadate, 10 mM sodium pyrophosphate, and one complete protease inhibitor cocktail (EASYpacks, Roche).
2. Reduction reagent: 1 M dithiothreitol (DTT).
3. Alkylation reagent: 1 M iodoacetamide (IAA).
4. Sample buffer: ammonium bicarbonate, pH 8.
5. Background electrolyte (BGE): 5% acetic acid.
6. Sheath buffer: 0.2% formic acid (FA), 10% methanol.

55.2.2 Instrumentation

1. CESI 8000 Plus CE system (Beckman Coulter).
2. Q-Exactive HF mass spectrometer (Thermo Fisher Scientific).
3. Electrokinetically pumped sheath-flow CE-MS nanospray interface (CMP Scientific Corp.).

4.Agilent 1260 Infinity II HPLC system (Agilent).

DOI: 10.1201/9781003449942-62

55.3 Methods

55.3.1 Protein Extraction

1. Lysis buffer is added to the harvested SW480 and SW620 cells, respectively.
2. The cells are lysed for 1 minute using a homogenizer 150 (Fisher Scientific) and then sonicated on ice for 10 minutes with a Branson Sonifier 250 (VWR Scientific) at 15% amplitude.
3. The cell lysates are centrifuged at 15,00rpm for 10 minutes and the supernatants are collected.
4. Protein concentrations are measured using a bicinchoninic acid (BCA) protein assay kit (Thermo Scientific Pierce, Rockford, IL). The collected cell lysates are then stored at −80 °C until preparation for MS analysis.

55.3.2 Protein Denaturation, Reduction, and Alkylation

1. Denaturation: 500 μg of SW480 and SW620 proteins are denatured at 37 °C for 30 minutes.
2. Reduction: 1 μL of 1 M DTT is added to protein samples with incubation at 37 °C for 30 minutes.
3. Alkylation: 2.5 μL of 1M IAA is added to samples with incubation at room temperature for 20 min.
4. The reactions are quenched with 1 μL of 1 M DTT solution.

55.3.3 Sample Fractionation

55.3.3.1 Study Design

1. Experiment 1 (2D separation): 200 μg proteins (SW480 and SW 620), RPLC (13 fractions)-CZE-MS/MS (single run).
2. Experiment 2 (3D separation): 2 mg of proteins (SW480 and SW 620), SEC-RPLC (4×13 fractions)-CZE-MS/MS (triplicate runs).
3. Experiment 3 (2D separation): 420 μg of proteins (SW480 and SW 620), RPLC (6 fractions)-CZE-MS/MS (triplicate runs).
4. Experiment 4 (2D separation): 500 μg of proteins (SW480 and SW 620), SEC (6 fractions)-CZE-MS/MS (triplicate runs).
5. Experiment 5 (1D separation): 100 μg of proteins (SW480 and SW 620), CZE-MS/MS (triplicate runs).

55.3.3.2 RPLC Fractionation

1. Mobile phase A (MPA): 2% ACN, 0.1% FA in water; Mobile phase B (MPB): 80% ACN, 0.1% FA in water.
2. Gradient and flow rate: equilibration of the column (C4, 2.1 × 250 mm, Sepax Technologies) with 100% MPA for 10 minutes; running of MPB gradient from 0% to 80% over 90 minutes with a 0.25 mL/min flow rate.
3. Six fractions collection method for RPLC-CZE-MS/MS system (experiment 3): fractions are collected from 25 to 55 minutes (fraction 1), 50 to 70 minutes (5 minutes per fraction, fraction 2–5), and 70 to 95 minutes (fraction 6).
4. Thirteen fractions collection method for RPLC-CZE-MS/MS system (experiments 1 and 2) and SEC-RPLC-CZE-MS/MS system: fractions are collected from 15 to 22 minutes (fraction 1) and 22 to 70 minutes (4 minutes per fraction, fraction 2–13).

55.3.3.3 SEC Fractionation

1. Mobile phase: 0.05% trifluoroacetic acid (TFA) in water or 0.1% formic acid in water.
2. Flow rate: 0.35 mL/min of mobile phase.
3. Four fractions collection method for SEC-RPLC-CZE-MS/MS study (experiment 2): fractions are collected from 5 to 8 minutes (fraction 1) and 8 to 12.5 minutes (1.5 minutes per fraction, fraction 2–4).
4. Six fractions collection method for SEC-CZE-MS/MS study (experiment 4): fractions are collected from 5.6 to 8.6 minutes (fraction 1), 8.6 to 14.6 minutes with 1.5 minutes per fraction (fraction 2–5), and 14.6 to 19.0 minutes (fraction 6).

55.3.4 CZE-MS/MS Analysis

55.3.4.1 Preparation of Linear Polyacrylamide (LPA)-Coated Capillary for CZE

1. Capillary pretreatment:
   a. Flush the capillary with the following buffers: 1 M sodium hydroxide for 1 h, water for 30 min, 1 M hydrogen chloride for 30 min, water for 30 min, and methanol for 30 min.
   b. Flush the capillary with nitrogen gas for 10 min.
   c. Fill the capillary with 50% v/v 3-(trimethoxysilyl)propyl-methacrylate (in methanol) and seal both ends of the capillary with rubber.
   d. Incubate the capillary at room temperature for one to two days.
   e. Flush the capillary with methanol to remove the residual reagents from the capillary and dry the capillary with nitrogen gas overnight.
2. Capillary coating preparation:
   a. Prepare acrylamide solution by dissolving 40 mg of acrylamide in 1 mL of water.
   b. Prepare the initiator solution by dissolving 25 mg of APS in500 μL of water.
   c. Add 3 μL of initiator solution into 500 μL of the acrylamide solution.
   d. Degas the solution for 15 min with nitrogen gas to remove the oxygen from the solution.
   e. Introduce LPA coating solution into the capillary and seal both ends of the capillary with rubber.

f. Incubate the capillary in a 50 °C water bath for 50 min.
g. Flush the capillary with water to remove the residual reagents from the capillary and store the capillary at room temperature.
h. Etch one end of the capillary with hydrofluoric acid (HF) to reduce the outer diameter of the capillary to about 70–80 µm. (Caution: use appropriate safety procedures while handling hydrofluoric acid solutions.)

55.3.4.2 Setup of CE-MS System

1. Couple the CE to Q-Exactive HF mass spectrometer via an electrokinetically pumped sheath-flow CE-MS nanospray interface.
2. Install a glass spray emitter (orifice size: 20~30 µm) on the interface and the emitter with sheath buffer.
3. Adjust the spray voltage in the range of 2–2.3 kV to generate stable electrospray.
4. Install a 70 cm or 100 cm linear polyacrylamide (LPA)-coated fused silica capillary (50 µm i.d., 360 µm o.d.) for CZE separation. One end of the capillary is fixed in the CE cartridge, and the other end of the capillary is inserted in the CE-MS interface.
5. In experiments 1, 2, 4, and 5, the capillary (100 cm) is loaded with 500 nL of sample. In experiment 3, the capillary (70 cm) is filled with ~350 nL of sample.
6. After sample loading, the inlet of capillary is inserted into BGE and 30 kV voltage is applied to carry out separations.
7. MS settings: Data-dependent acquisition (DDA) mode is applied for collecting MS1 and MS2 DATA. The temperature of the ion transfer tube is320 °C, S-lens RF is 55.
   a. MS1 spectra are collected with: m/z range of 600–2000, mass resolution of 120,000 (at m/z 200), a microscan number of 3, AGC target value of 1E6, and maximum injection time of 100 ms.
   b. The top five most abundant precursor ions (charge state higher than 5, or charge state unassigned and intensity threshold 2E4) in the MS1 spectra are isolated with a window of 4 m/z and fragmented via HCD with NCE of 20%. Other settings for MS2 spectra are resolution of 120,000 (at m/z 200), a microscan number of 3, AGC target value of 1E5, and maximum injection time of 200 ms. The dynamic exclusion is set to a duration of 30 s and the isotopic peaks are excluded.

55.3.5 Data Analysis

All the RAW files are converted into mzML files with msconvert and further processed on TopPIC Suite (version 1.4.0) pipeline integrated with TopFD, TopPIC, and TopDiff tools for proteoform identification and quantification.

55.3.5.1 Proteoform Identification

1. Spectral deconvolution: All the mzML files are processed with TopFD (version 1.4.0) to convert precursor and fragment isotope clusters into neutral monoisotopic masses and find proteoform features. The resulting mass spectra with monoisotopic neutral masses are stored in msalign files and the proteoform feature information are stored in text files.
2. Database search: (a) msalign files are searched on TopPIC (version 1.4.0) using the target-decoy approach. (b) The UniProt human proteome database (UP000005640) is used for database search. (c) Other searching settings: cysteine carbamidomethylation as a fixed modification; maximum number of unexpected modifications of 1; precursor and fragment mass error tolerances of 15 ppm; maximum mass shift of unknown modifications of 500 Da. The identified proteoforms are combined and filtered with 5% spectrum-level FDR and 1% proteoform-level FDR.

55.3.5.2 Proteoform Quantification

1. TopDiff (version 1.4.0) is applied to match the proteoform IDs across the SW and SW620 replicates and compare the feature intensities.
2. The quantitative results are analyzed using Perseus software. (a) The intensities of each proteoform in triplicate CZE-MS/MS runs of SW480 and SW620 are normalized to the intensity of the corresponding proteoform from the first run of SW480, converting proteoform intensity to proteoform ratio. (b) Proteoform ratios of each run are divided by the corresponding median to make sure the ratios are centered at 1. (c) All the data is further performed log2 transformation and the significantly differentially expressed proteoforms are determined by performing *t*-test analysis (FDR threshold: 0.05, S0: 1).

55.4 Results

55.4.1 Proteoform Identifications from CRC cells Using Four CZE-MS/MS-Based Strategies

We applied TDP workflow (Figure 55.1A) to study a pair of isogenic human non-metastatic and metastatic CRC cell lines (SW480 and SW620) collected over 400 MS raw files using four CZE-MS/MS-based strategies (Figure 55.1B). We identified 23,622 proteoforms of 2332 proteins from the SW480 and SW620 cell lines with a 1% proteoform-level FDR. The number of proteoform IDs from the CRC cells is about five- to eight-fold higher than that reported in previous TDP studies of human cancer cells [16–19]. (23,622 vs. 3000–5000 proteoforms).Identified from SW480 and SW620 cell lines were 17,316 and 14,504 proteoforms (on average 15,910 proteoforms), respectively.

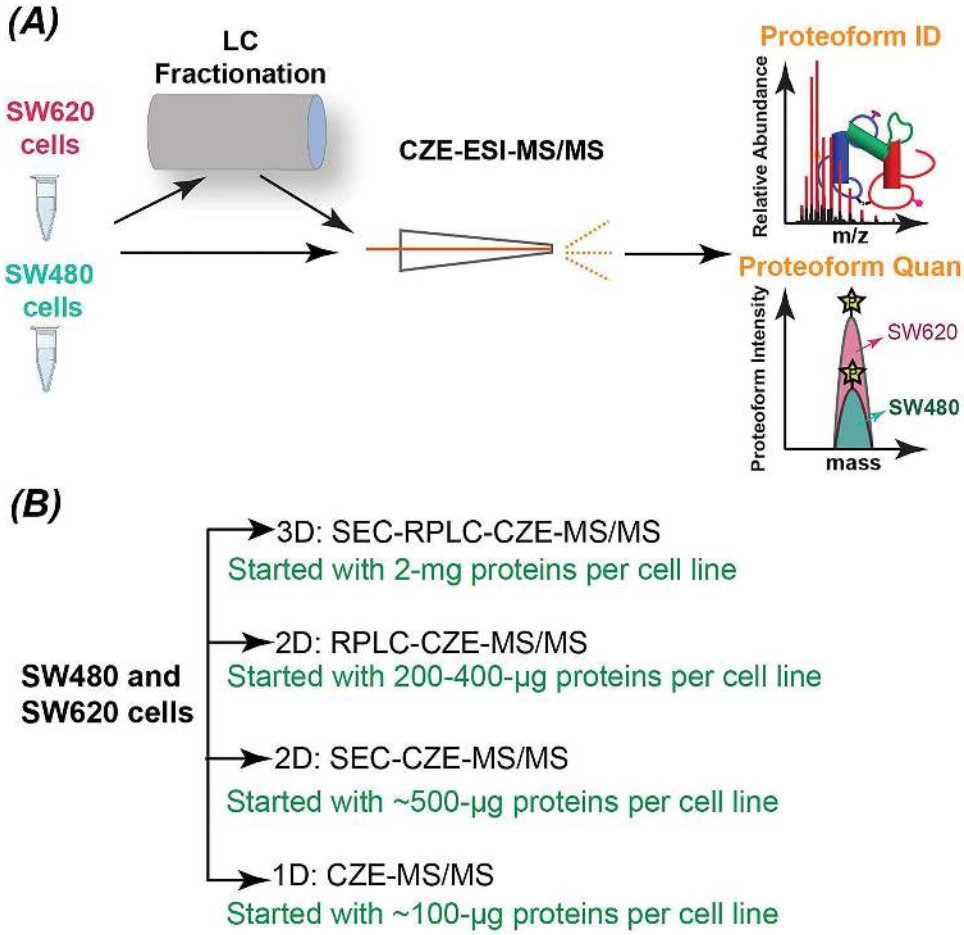


FIGURE 55.1 Schematic of the study design. (A) Workflow of the TDP of metastatic (SW620) and non-metastatic (SW480) CRC cells. (B) Four CZE-MS/MS-based strategies with the amounts of protein starting materials. (The figure was adapted from reference [15] with permission.)

We compared the four CZE-MS/MS strategies on the number and efficiency of proteoform IDs from the SW480 cells. SEC-RPLC-CZE-MS/MS outperformed SEC-CZE-MS/MS, RPLC-CZE-MS/MS, and CZE-MS/MS in terms of the number of proteoform IDs (~6000 vs. 200~5000) due to better LC fractionation (2D LC vs. 1D or no LC) and many more CZE-MS/MS runs (52 vs. 6 and 13). In terms of the proteoform identification efficiency (the number of proteoform IDs per CZE-MS/MS run), the SEC-CZE-MS/MS (6 LC fractions) produced nearly 700 proteoform IDs per run, which is nearly six-fold and four-fold higher than those from SEC-RPLC-CZE-MS/MS and CZE-MS/MS, respectively. Therefore, multidimensional separation is beneficial for large-scale TDP analysis of human cell lysates by effectively reducing the sample complexity. In particular, SEC-CZE-MS/MS and RPLC-CZE-MS/MS are powerful techniques for deep TDP of human cell lysates with high throughput.

Moreover, we compared the proteoforms and proteins identified from the SW480 and SW620 cells by SEC-CZE-MS/MS. About 60%–70% of proteoforms identified in one technical replicate of SW480 or SW620 cells were also identified in another replicate of the same cell line, indicating reasonable reproducibility of proteoform ID using SEC-CZE-MS/MS. Of proteoforms identified in one replicate of SW480 cells, 40%–50% were identified in one replicate of SW620 cells. (The proteoform overlaps between the two cell lines are statistically significantly lower than that within each cell line (44±4% vs. 67±4%, $p < 10^{-14}$, two-tailed student's *t*-test.) The data demonstrate that the pair of isogenic human non-metastatic (SW480) and metastatic (SW620) CRC cell lines have significantly different proteoform profiles. The two cell lines are also significantly different at the protein level. The difference in protein overlaps between the two cell lines and within each cell line is statistically significant (69±8% vs. 83±3%, $p < 10^{-6}$, two-tailed student's *t*-test).

55.4.2 Discovery of Proteoforms in Well-Known CRC-Related Pathways

QIAGEN Ingenuity Pathway Analysis (IPA) analysis of the proteoform-related genes identified in this work disclosed several significantly enriched and well-known CRC-related pathways (Figure 55.2A), such as WNT/β-catenin signaling (p-value: 10^{-3}), PI3K/AKT Signaling (p-value: 10^{-4}), mTOR

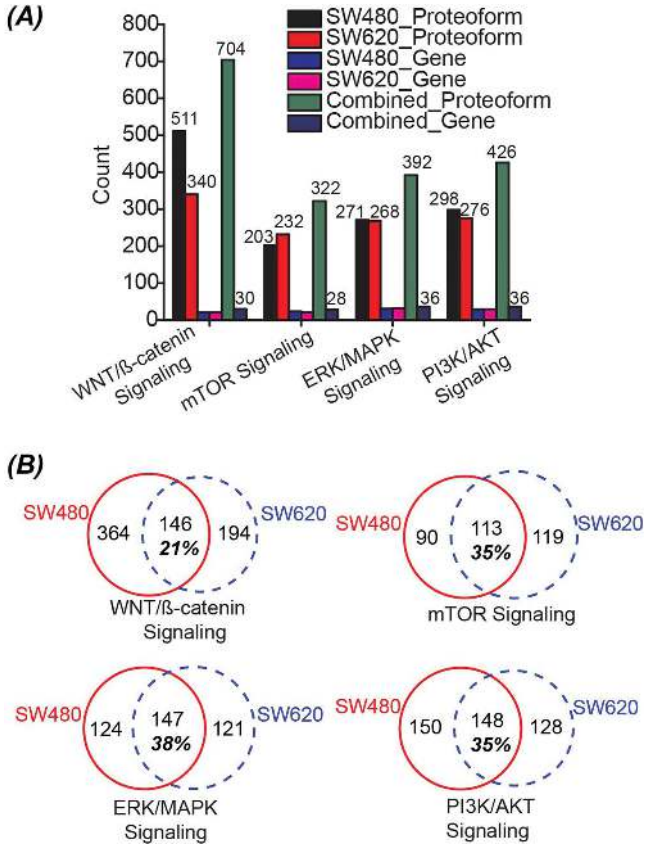


FIGURE 55.2 Summary of proteoform-related genes involved in well-known CRC-related pathways. (A) The number of proteoforms and genes in four CRC-related pathways identified from SW480 and SW620 cells. (B) Overlaps of identified and pathway-related proteoforms between SW480 and SW620 cells. (The figure was adapted from reference [15] with permission.)

Signaling (p-value: 10^{-14}), and ERK/MAPK Signaling pathways (p-value: 10^{-4}). Those pathways participate in CRC progression by regulating cell proliferation, apoptosis, survival, and so on. Hundreds of proteoforms from dozens of genes can be assigned to each pathway (Figure 55.2A). We found a much higher number of proteoforms of SW480 cells compared to SW620 cells for the WNT/β-catenin signaling pathway (511 vs. 340). In total, the results from SW480 and SW620 cells produced about 40% more proteoforms related to the four CRC pathways compared to one cell line alone, indicating the potential differences in proteoform profiles for the well-known CRC-related pathways between the non-metastatic and metastatic CRC cell lines. Figure 55.2B shows the shared proteoforms between SW480 and SW620 cells for each pathway is only about 21%–38% of the total proteoforms identified from the two cell lines. The data suggests that proteoforms in those pathways could potentially play important roles in driving CRC progression and metastasis.

Here, we highlighted some proteoforms of important genes (MARK2, SOX9, EIF4B, and EIF4EBP1) related to the WNT/β-catenin signaling, mTOR signaling, and PI3K/AKT signaling pathways. MARK2 plays a vital role in modulating directional cancer cell migration, which is crucial for cancer metastasis [20]. SOX9 is a high mobility group (HMG) box transcription factor and plays essential roles in regulating CRC progression [21]. Expression of SOX9 is closely associated with the five-year overall survival rate of CRC patients [21]. EIF4B regulates cancer cell proliferation and has been reported as a potential target for developing anti-cancer therapies [22]. Phosphorylation of EIF4EBP1 has been reported as an important regulator of cancer progression [23].

55.4.3 Proteoforms with PTMs and Single Amino Acid Variants (SAAVs)

Our large-scale TDP study of SW480 and SW620 identified 4872 proteoforms with N-terminal acetylation (+42 Da mass shift), 319 proteoforms with phosphorylation [+80 Da (single phosphorylation) or +160 Da (double phosphorylation) mass shift], 321 proteoforms with methylation (+14 Da mass shift), and 241 proteoforms with oxidation (+16 Da mass shift). TDP is powerful for the characterization of combinations of various PTMs on proteoforms. We identified 54 proteoforms with two phosphorylation sites and 90 proteoforms with both acetylation and phosphorylation PTMs. For example, we found 28 kDa heat- and acid-stable phosphoprotein (PDAP1) and calmodulin-1 (CALM1) proteoforms carry either two phosphorylation sites or the combination of N-terminal acetylation and one lysine trimethylation. Those PTMs of the two proteins agree with the literature data [24, 25]. The two proteoforms were identified with high confidence and were well characterized in terms of PTMs. PDAP1 and CALM1 are both prognostic markers of cancer according to the Human Protein Atlas (https://www.proteinatlas.org/). However, the potential roles of those specific proteoforms of PDAP1 and CALM1 in cancer still need to be investigated.

In addition, we applied TopPG, which has a human protein database incorporated with transcriptomic variants information based on the RNA-seq data for searching the SAAVs in SW480 and SW620. We identified 111 proteoforms containing SAAVs of 82 genes from both cell lines with a proteogenomic approach with a 5% proteoform-level FDR. Only 20% of the 111 proteoforms were identified from both cell lines, indicating potentially different SAAV profiles between the two cell lines. For example, TP53 is an important tumor suppressor closely related to CRC development, and it is an essential member of WNT/β-catenin signaling and PI3K/AKT signaling pathways. We identified one TP53 proteoform containing an AAV at position 72 (P→R) due to the codon 72 polymorphism (Figure 55.3A). Studies have shown the functional differences between the P72 and R72 proteoforms of TP53 [26, 27]. The R72 proteoform does a markedly better job of inducing apoptosis compared to the P72 proteoform [26]. Another study indicated that the expression of P72 proteoform increased CRC metastasis, and that the R72 proteoform does not exist in the non-metastatic CRC cell line (SW480) based on the nucleic-acid data [27]. Interestingly, we only identified the R72 proteoform of TP53 in the SW620 cell line, not in the SW480 cell line, from the top-down MS data. MSH6 is one of the DNA mismatch repair genes and its mutations play a crucial role in Lynch syndrome, which is an inherited form of CRC. We identified one MSH6 proteoform containing a SAAV due to polymorphism at position 39 (G→E) (Figure 55.3B).

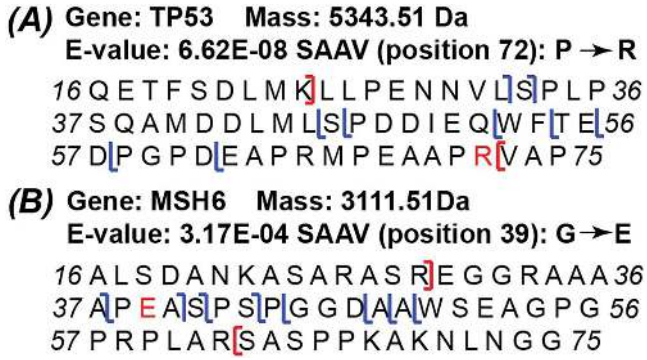


FIGURE 55.3 Sequences and fragmentation patterns of TP53 (A) and MSH6 (B) proteoform containing SAAVs. (The figure was adapted from reference [15] with permission.)

The G39E SAAV has been associated with an increased risk of CRC according to the nucleic-acid data [28]. We identified the G39 proteoform of MSH6 in both SW480 and SW620 cells, but identified the E39 proteoform only in the SW480 cells, not in the SW620 cells.

55.4.4 Quantitative TDP of Metastatic and Non-Metastatic Human CRC Cell Lines

SEC fractionated samples were mainly used for quantitative TDP study of SW480 and SW620. The fractions were analyzed by CZE-MS/MS in technical triplicate. We identified roughly 4000 proteoforms per replicate per cell line with a 1% proteoform-level FDR. About 1500 proteoforms were quantified in all six samples (three replicates per cell line and two cell lines) by TopDiff. The SEC-CZE-MS/MS system shows reasonably good reproducibility regarding the intensities of shared proteoforms, as evidenced by the strong linear correlations of proteoform intensities between technical replicates of SW480 or SW620 cells (Pearson correlation coefficients: 0.86–0.93). The Pearson correlation coefficients of proteoform intensity between SW480 and SW620 cells are statistically significantly lower than that between technical replicates of one cell line (0.71±0.01 vs. 0.90±0.03, $p < 10^{-10}$, two-tailed student's *t*-test), indicating significant differences between the two cell lines in terms of proteoform intensity. We used the Perseus software for further data analysis [29]. The two cell lines can be easily distinguished using the proteoform quantification profiles, Figure 55.4A. Two clusters of differentially expressed proteoforms across the six samples were revealed.

The volcano plot in Figure 55.4B showed 460 proteoforms of 248 proteins with statistically significant differences in abundance between the two cell lines (FDR < 0.05). Specifically, 244 proteoforms of 152 proteins had higher abundance in the SW480 cell line and 216 proteoforms of 132 proteins had higher expression in the SW620 cell line. Figure 55.4B shows that one HMGN1 proteoform and one RBM8A proteoform have the most significant abundance changes between SW480 and SW620 cells. HMGN1 regulates gene expression and PTMs of core histones, affecting DNA repair and tumor progression [30]. It has been reported that RBM8A promotes tumor cell migration and invasion in the most common type of primary liver cancer [31].

Comparing the overexpressed and underexpressed proteoforms in the two cell lines revealed that 36 genes (e.g., DAP, CALM1, HDGF, JPT1, and NPM1) have both overexpressed and underexpressed proteoforms in one cell line, suggesting that different proteoforms of the same gene had completely different expression patterns in the two cell lines. Figure 55.4C shows two differentially expressed proteoforms of death-associated protein 1 (DAP), one of those 36 genes. It has been reported that DAP modulates cell death and correlates with the clinical outcome of CRC patients [32]. Interestingly, we revealed that one phosphorylated proteoform of DAP (~7607 Da, phosphorylation site S51 or T56) had a higher abundance in SW480 cells and another phosphorylated proteoform (~4605 Da, phosphorylation site S51) showed higher expression in SW620 cells. Both the S51 and T56 are known to be phosphorylated according to PhosphoSitePlus, with S51 being the most common phosphorylation site of DAP. We noted that the differentially expressed proteoforms in this study include phosphorylated proteoforms of several important genes related to CRC, such as RALY [33], NPM1 [34], DAP [32], and HDGF [35]. The functions of phosphorylated forms of those four proteins in modulating CRC development are still unclear. However, the differential expressions of those phosphorylated proteoforms in the metastatic and non-metastatic CRC cells suggest their potential roles in regulating CRC metastasis.

We highlight several differentially expressed proteoforms of CALM1, JPT1 (HN1), and EPCAM. CALM-dependent systems play important roles in cancer metastasis [36]. JPT1 (HN1) promotes cancer metastasis via activating the NF-ƙB signaling pathway [37]. EPCAM is a human cell surface glycoprotein and plays crucial roles in tumor biology, especially CRC [38]. EPCAM has been recognized as an important therapeutic target for cancer. We discovered two CALM1 proteoforms having significantly higher abundance in SW620 cells compared to SW480 cells; one of them contains K116 trimethylation. We revealed one CALM1 proteoform showing higher abundance in SW480 cells and the proteoform carries N-terminal acetylation and a 58-Da mass shift between amino acid residues 73 and 89. The 58-Da mass shift can be explained as a trimethylation/acetylation plus oxidation. Three of JPT1 proteoforms have higher abundance in SW480 cells, and one of them contains a 167-Da mass shift between the amino acid residues 66 and 89, where seven serine residues can be phosphorylated according to the PhosphoSitePlus database (https://www.phosphosite.org/). The 167-Da mass shift most likely represents a combination of phosphorylation and other PTMs. Interestingly, one JPT1 proteoform shows higher abundance in SW620 cells. We also observed two EPCAM proteoforms having higher abundance in SW480 cells.

We then performed IPA analyses of the genes of those differentially expressed proteoforms between SW480 and SW620 cells. Those genes are heavily involved in cancer-related diseases, for example, tumorigenesis of tissue and metastasis, Figure 55.4D. Five of those proteins (EIF4E, EPCAM, FKBP1A, GAA, and HSP90AB1) are drug targets. IPA network analyses revealed that 26 proteins (highlighted in purple) whose proteoforms showed higher abundance in SW480 compared

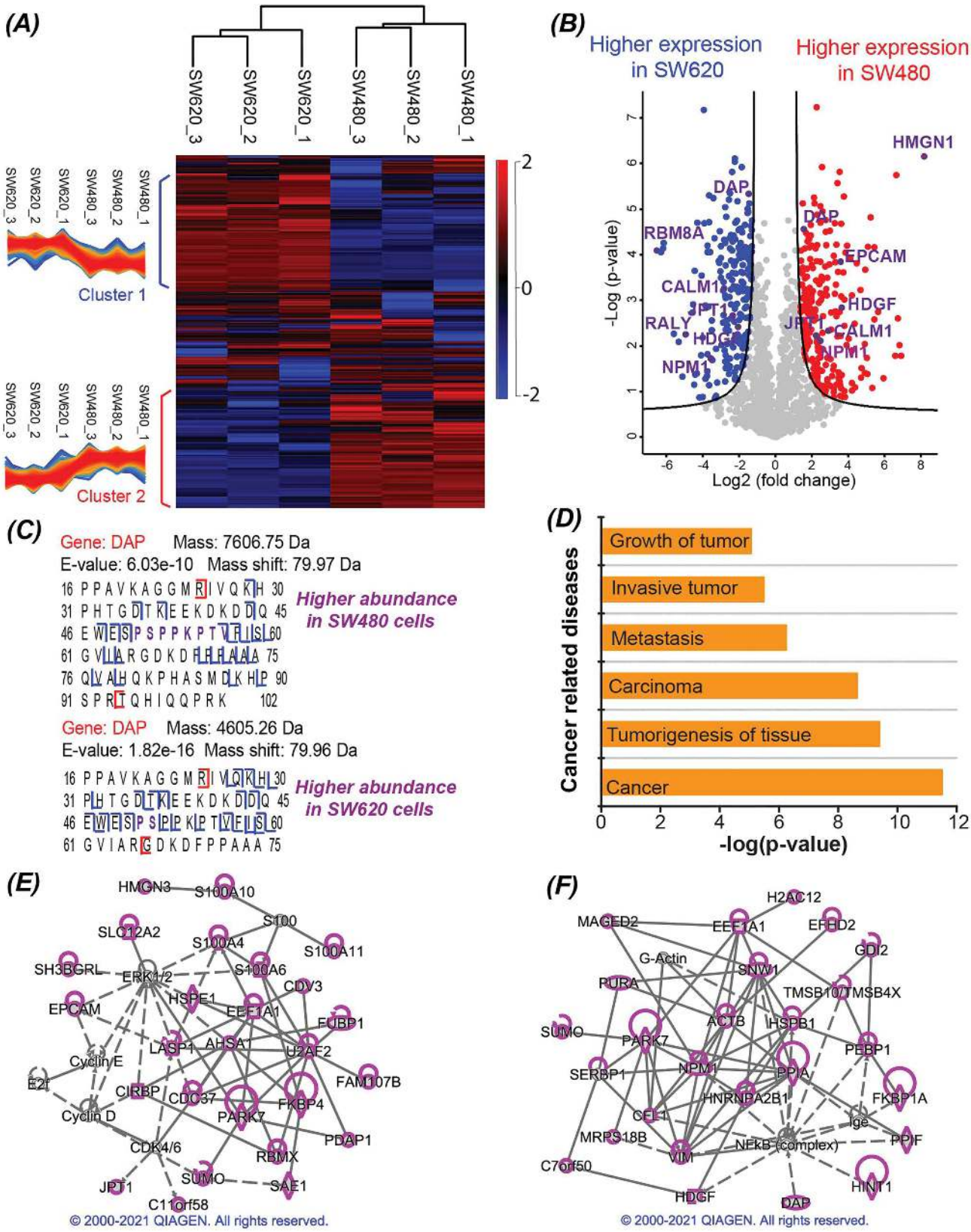


FIGURE 55.4 Summary of the LFQ data of SW480 and SW620 cells. (A) Heatmap and cluster analysis of the quantified proteoforms regarding LFQ intensities. (B) Volcano plot showing differentially expressed proteoforms between the two cell lines (S0 = 1 and FDR = 0.05). Red dots and blue dots represent proteoforms having statistically significantly higher abundance in SW480 and in SW620, respectively. (C) Sequences and fragmentation patterns of two phosphorylated proteoforms of the gene DAP. (D) An IPA analysis reported some cancer-related diseases that are related to the differentially expressed genes in the two cell lines. Proteoforms with higher abundance in SW480 cells (E) or higher abundance in SW620 cells (F) correspond to genes that are involved in cancer-related networks with high scores. (The figure was adapted from reference [15] with permission.)

to SW620 were involved in a cancer-related network (score 51), Figure 55.4E. Those proteins belong to several families, including enzyme (diamond shape, e.g., PARK7 and FKBP4), transcription regulator (oval shape, e.g., FUBP1), translation regulator (hexagon shape, e.g., CIRBP and EEF1A1), transporter (trapezium shape, e.g., SLC12A2 and LASP1), and others (circle shape, e.g., EPCAM and JPT1). Most of those proteins have direct (solid line) and indirect (dotted line) interactions with one another. We also carried out network analysis for the proteins whose proteoforms had higher expression in SW620 cells, and observed high-scores for cancer-related networks. Figure 55.4F shows one cancer-related network (score 54), and 26 of those proteins are involved in the network (highlighted in purple). Those proteins include several CRC-related important proteins, NPM1 (oval shape, transcription regulator, located in nucleus), DAP (transcription regulator, located in cytoplasm), and HDGF (square shape, growth factor, located in extracellular space). NPM1 is a crucial protein in the network and many of the highlighted proteins have direct interactions (solid line) with NPM1, for example, PARK7, VIM, and PPIA. NPM1 also has indirect interaction (dotted line) with the NFkB complex, which plays crucial roles in modulating DNA transcription and cell survival. Human NPM1 boosts the activation of NFkB according to Ingenuity relationships from the IPA analysis. Besides NPM1, several other highlighted proteins (e.g., HDGF and DAP) also have indirect interactions with the NFkB complex. For example, NFkB regulates the transcription of HDGF, and DAP deactivates the NFkB according to the IPA network analysis results. The IPA analysis also revealed that 13 proteoforms of three genes (EIF4B, EIF4E, EIF4EBP1) in the mTOR signaling pathway had statistically significant differences in abundance between the SW480 and SW620 cells.

55.5 Discussion

TDP of metastatic and non-metastatic cells is crucial for discovering new proteoform biomarkers and providing a more accurate understanding of molecular mechanisms of cancer metastasis. In this study, we advanced TDP of human cancer cells via coupling LC fractionations to CZE-MS/MS and drastically boosted the number of proteoform IDs per human cell line compared to previous LC-MS/MS-based studies (~16,000 vs. ~3000).

Our qualitative and quantitative TDP of SW480 and SW620 cells come to several conclusions about CRC metastasis. First, CRC cells have a significant transformation in proteoforms and SAAVs after metastasis, evidenced by differences of proteoform and SAAV profiles between SW480 and SW620 cells. Second, different proteoforms from the same cancer-related gene (e.g., DAP, CALM1, HDGF, JPT1, RALY, and NPM1) may have potentially varied biological functions in modulating CRC metastasis, because they show opposite expression profiles between the SW480 and SW620 cells. Some proteoforms of those genes have higher abundance in SW480 cells; some of their proteoforms show higher expression in SW620 cells. Third, PTMs (i.e., phosphorylation) of important cancer-related genes (i.e., DAP, HDGF, JPT1, RALY, NPM1, MARK2, SOX9, EIF4B, and EIF4EBP1) could play important roles in regulating CRC metastasis, evidenced by the significant abundance differences of phosphorylated proteoforms from those genes between the SW480 and SW620 cells. The differentially expressed proteoforms, especially those with PTMs, of important cancer-related genes could be novel proteoform biomarkers of CRC metastasis. Fourth, proteoforms of genes in well-known CRC-related pathways (WNT/β-catenin signaling, PI3K/AKT signaling, mTOR signaling, and ERK/MAPK signaling) are different between SW480 and SW620 cells, and those proteoforms could play vital roles in modulating CRC metastasis.

55.6 Future Trends

Although TDP has gained great attention in understanding heterogeneity and mechanism of cancer metastasis at the proteoform level, it still confronts technical challenges in identifying large and low abundant proteoforms. In this study, we mainly focused on proteoforms smaller than 30kDa, considering the bottleneck of liquid phase separation tools and MS instrumentations (e.g., limitation on resolving power, sensitivity, and fragmentation of large proteoforms). Recently, we have coupled field asymmetric waveform ion mobility spectrometry (FAIMS) to CZE to enable efficient gas phase prefractionation of proteoforms on their sizes [39]. We achieved 20-fold boosts in sensitivity and six-fold enhancement in large proteoform IDs for a yeast model. The platform is highly transferable to cancer studies to delve into the roles of large proteoforms in cancer progression. Additionally, CZE was limited in sample-loading capacity compared to RPLC (nL vs. μL), which greatly restricts the sensitivity of detecting large and low abundant proteoforms. Improvement in online sample stacking/pre-concentration methods for CZE, such as incorporating solid phase microextraction (SPME) into CZE for in-capillary pre-concentration, is a potential exploring area to advance the sample-loading capacity and sensitivity. Furthermore, on the MS aspect, leveraging individual ion mass spectrometry (I2MS) for TDP is the future trend for cancer applications. I2MS has shown extraordinary sensitivity in measuring single ions by charge detection and is adaptable to proteins up to MDa [40, 41]. The recent studies have demonstrated the power of I2MS to study highly complex samples and extremely diluted immunoprecipitated (IP) protein targets [40–43]. The I2MS has also been applied with various fragmentation techniques, such as HCD and ETD, to deepen the proteoform sequence coverage and accurate PTMs localization of large proteoforms [40, 42]. The innovation of I2MS technology directs cancer biomarkers discovery to large functional proteoforms and potentially provides better interpretations of key players in cancer events.

Acknowledgments

The work was funded by the National Cancer Institute (NCI) through the grant R01CA247863. We also thank the support from the National Institute of General Medical Sciences

(NIGMS) through grants R01GM125991 and R01GM118470 and the support from the National Science Foundation (CAREER Award, grant DBI1846913). We thank MSU AgBioResearch and Michigan State University for providing access to the QIAGEN IPA platform.

REFERENCES

1. Smith, L. M. and Kelleher, N. L., 2013. Proteoform: A single term describing protein complexity. *Nature Methods*, 10, pp. 186–187.
2. Smith, L.M. and Kelleher, N.L., 2018. Proteoforms as the next proteomics currency. *Science*, 359, pp. 1106–1107.
3. Tyanova, S., Albrechtsen, R., Kronqvist, P., Cox, J., Mann, M. and Geiger, T., 2016. Proteomic maps of breast cancer subtypes. *Nature Communications*, 7(1), p. 10259.
4. Besson, D., Pavageau, A.H., Valo, I., Bourreau, A., Bélanger, A., Eymerit-Morin, C., Moulière, A., Chassevent, A., Boisdron-Celle, M., Morel, A. and Solassol, J., 2011. A quantitative proteomic approach of the different stages of colorectal cancer establishes OLFM4 as a new nonmetastatic tumor marker. *Molecular & Cellular Proteomics*, 10(12).
5. Choudhary, C. and Mann, M., 2010. Decoding signalling networks by mass spectrometry-based proteomics. *Nature Reviews Molecular Cell Biology*, 11(6), pp. 427–439.
6. Ntai, I., LeDuc, R.D., Fellers, R.T., Erdmann-Gilmore, P., Davies, S.R., Rumsey, J., Early, B.P., Thomas, P.M., Li, S., Compton, P.D. and Ellis, M.J., 2016. Integrated bottom-up and top-down proteomics of patient-derived breast tumor xenografts. *Molecular & Cellular Proteomics*, 15(1), pp. 45–56.
7. Ntai, I., Fornelli, L., DeHart, C.J., et al., 2018. Precise characterization of KRAS4b proteoforms in human colorectal cells and tumors reveals mutation/modification cross-talk. *Proceedings of the National Academy of Sciences of the United States of America*, 115(16), pp. 4140–4145.
8. Adams, L.M., DeHart, C.J., Drown, B.S., et al., 2023. Mapping the KRAS proteoform landscape in colorectal cancer identifies truncated KRAS4B that decreases MAPK signaling. *Journal of Biological Chemistry*, 299(1), p. 102768.
9. Mohr, J., Swart, R., Samonig, M., Böhm, G. and Huber, C.G., 2010. High-efficiency nano-and micro-HPLC–high-resolution Orbitrap-MS platform for top-down proteomics. *Proteomics*, 10, pp. 3598–3609.
10. Tucholski, T., Knott, S.J., Chen, B., Pistono, P., Lin, Z. and Ge, Y., 2019. A top-down proteomics platform coupling serial size exclusion chromatography and Fourier transform ion cyclotron resonance mass spectrometry. *Analytical Chemistry*, 91, pp. 3835–3844.
11. McCool, E.N., Lubeckyj, R.A., Shen, X., Chen, D., Kou, Q., Liu, X. and Sun, L., 2018. Deep top-down proteomics using capillary zone electrophoresis-tandem mass spectrometry: Identification of 5700 proteoforms from the Escherichia coli proteome. *Analytical Chemistry*, 90(9), pp. 5529–5533.
12. Lubeckyj, R.A., Basharat, A.R., Shen, X., Liu, X. and Sun, L., 2019. Large-scale qualitative and quantitative top-down proteomics using capillary zone electrophoresis-electrospray ionization-tandem mass spectrometry with nanograms of proteome samples. *Journal of the American Society for Mass Spectrometry*, 30(8), pp. 1435–1445.
13. Chen, D., McCool, E.N., Yang, Z., Shen, X., Lubeckyj, R.A., Xu, T., Wang, Q. and Sun, L., 2023. Recent advances (2019–2021) of capillary electrophoresis-mass spectrometry for multilevel proteomics. *Mass Spectrometry Reviews*, 42(2), pp. 617–642.
14. Shen, X., Yang, Z., McCool, E.N., Lubeckyj, R.A., Chen, D. and Sun, L., 2019. Capillary zone electrophoresis-mass spectrometry for top-down proteomics. *TrAC Trends in Analytical Chemistry*, 120, p. 115644.
15. McCool, E.N., Xu, T., Chen, W., Beller, N.C., Nolan, S.M., Hummon, A.B., Liu, X. and Sun, L., 2022. Deep top-down proteomics revealed significant proteoform-level differences between metastatic and nonmetastatic colorectal cancer cells. *Science Advances*, 8(51), p. eabq6348.
16. Tran, J.C., Zamdborg, L., Ahlf, D.R., Lee, J.E., Catherman, A.D., Durbin, K.R., Tipton, J.D., Vellaichamy, A., Kellie, J.F., Li, M. and Wu, C., 2011. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*, 480(7376), pp. 254–258.
17. Catherman, A.D., Durbin, K.R., Ahlf, D.R., Early, B.P., Fellers, R.T., Tran, J.C., Thomas, P.M. and Kelleher, N.L., 2013. Large-scale top-down proteomics of the human proteome: Membrane proteins, mitochondria, and senescence. *Molecular & Cellular Proteomics*, 12(12), pp. 3465–3473.
18. Anderson, L.C., DeHart, C.J., Kaiser, N.K., Fellers, R.T., Smith, D.F., Greer, J.B., LeDuc, R.D., Blakney, G.T., Thomas, P.M., Kelleher, N.L. and Hendrickson, C.L., 2017. Identification and characterization of human proteoforms by top-down LC-21 tesla FT-ICR mass spectrometry. *Journal of Proteome Research*, 16(2), pp. 1087–1096.
19. Tran, J.C., Zamdborg, L., Ahlf, D.R., Lee, J.E., Catherman, A.D., Durbin, K.R., Tipton, J.D., Vellaichamy, A., Kellie, J.F., Li, M. and Wu, C., 2011. Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature*, 480(7376), pp. 254–258.
20. Pasapera, A.M., Heissler, S.M., Eto, M., Nishimura, Y., Fischer, R.S., Thiam, H.R. and Waterman, C.M., 2022. MARK2 regulates directed cell migration through modulation of myosin II contractility and focal adhesion organization. *Current Biology*, 32(12), pp. 2704–2718.
21. Lü, B., Fang, Y., Xu, J., Wang, L., Xu, F., Xu, E., Huang, Q. and Lai, M., 2008. Analysis of SOX9 expression in colorectal cancer. *American Journal of Clinical Pathology*, 130(6), pp. 897–904.
22. Shahbazian, D., Parsyan, A., Petroulakis, E., Hershey, J.W. and Sonenberg, N., 2010. eIF4B controls survival and proliferation and is regulated by proto-oncogenic signaling pathways. *Cell Cycle*, 9(20), pp. 4106–4109.
23. Chen, Y., Wang, J., Fan, H., Xie, J., Xu, L. and Zhou, B., 2017. Phosphorylated 4E-BP1 is associated with tumor progression and adverse prognosis in colorectal cancer. *Neoplasma*, 64(5), pp. 787–794.
24. Zhou, H., Di Palma, S., Preisinger, C., Peng, M., Polat, A.N., Heck, A.J. and Mohammed, S., 2013. Toward a comprehensive characterization of a human cancer cell phosphoproteome. *Journal of Proteome Research*, 12(1), pp. 260–271.

25. Sasagawa, T., Ericsson, L.H., Walsh, K.A., Schreiber, W.E., Fischer, E.H. and Titani, K., 1982. Complete amino acid sequence of human brain calmodulin. *Biochemistry*, 21(10), pp. 2565–2569.
26. Jeong, B.S., Hu, W., Belyi, V., Rabadan, R. and Levine, A.J., 2010. Differential levels of transcription of p53-regulated genes by the arginine/proline polymorphism: p53 with arginine at codon 72 favors apoptosis. *The FASEB Journal*, 24(5), pp. 1347–1353.
27. Katkoori, V.R., Manne, U., Chaturvedi, L.S., Basson, M.D., Haan, P., Coffey, D. and Bumpers, H.L., 2017. Functional consequence of the p53 codon 72 polymorphism in colorectal cancer. *Oncotarget*, 8(44), p. 76574.
28. Zelga, P., Przybyłowska-Sygut, K., Zelga, M., Dziki, A. and Majsterek, I., 2017. Polymorphism of Gly39Glu (c. 116G> A) hMSH6 is associated with sporadic colorectal cancer development in the Polish population: Preliminary results. *Advances in Clinical and Experimental Medicine*, 26(9).
29. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and Cox, J., 2016. The Perseus computational platform for comprehensive analysis of (prote) omics data. *Nature Methods*, 13(9), pp. 731–740.
30. Postnikov, Y. and Bustin, M., 2010. Regulation of chromatin structure and function by HMGN proteins. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1799(1–2), pp. 62–68.
31. Liang, R., Lin, Y., Ye, J.Z., Yan, X.X., Liu, Z.H., Li, Y.Q., Luo, X.L. and Ye, H.H., 2017. High expression of RBM8A predicts poor patient prognosis and promotes tumor progression in hepatocellular carcinoma. *Oncology Reports*, 37(4), pp. 2167–2176.
32. Jia, Y., Ye, L., Ji, K., Toms, A.M., Davies, M.L., Ruge, F., Ji, J., Hargest, R. and Jiang, W.G., 2014. Death associated protein 1 is correlated with the clinical outcome of patients with colorectal cancer and has a role in the regulation of cell death. *Oncology Reports*, 31(1), pp. 175–182.
33. Sun, L., Wan, A., Zhou, Z., Chen, D., Liang, H., Liu, C., Yan, S., Niu, Y., Lin, Z., Zhan, S. and Wang, S., 2021. RNA-binding protein RALY reprogrammes mitochondrial metabolism via mediating miRNA processing in colorectal cancer. *Gut*, 70(9), pp. 1698–1617.
34. Grisendi, S., Mecucci, C., Falini, B. and Pandolfi, P.P., 2006. Nucleophosmin and cancer. *Nature Reviews Cancer*, 6(7), pp. 493–505.
35. Sun, B., Gu, X., Chen, Z. and Xiang, J., 2015. MiR-610 inhibits cell proliferation and invasion in colorectal cancer by repressing hepatoma-derived growth factor. *American Journal of Cancer Research*, 5(12), p. 3635.
36. Villalobo, A. and Berchtold, M.W., 2020. The role of calmodulin in tumor cell migration, invasiveness, and metastasis. *International Journal of Molecular Sciences*, 21(3), p. 765.
37. Chen, J., Qiu, J., Li, F., Jiang, X., Sun, X., Zheng, L., Zhang, W., Li, H., Wu, H., Ouyang, Y. and Chen, X., 2020. HN1 promotes tumor associated lymphangiogenesis and lymph node metastasis via NF-κB signaling activation in cervical carcinoma. *Biochemical and Biophysical Research Communications*, 530(1), pp. 87–94.
38. Armstrong, A. and Eck, S.L., 2003. EpCAM: A new therapeutic target for an old cancer antigen. *Cancer Biology & Therapy*, 2(4), pp. 320–325.
39. Xu, T., Wang, Q., Wang, Q. and Sun, L., 2023. Coupling high-field asymmetric waveform ion mobility spectrometry with capillary zone electrophoresis-tandem mass spectrometry for top-down proteomics. *Analytical Chemistry*, 95(25), pp. 9497–9504.
40. Kafader, J.O., Durbin, K.R., Melani, R.D., Des Soye, B.J., Schachner, L.F., Senko, M.W., Compton, P.D. and Kelleher, N.L., 2020. Individual ion mass spectrometry enhances the sensitivity and sequence coverage of top-down mass spectrometry. *Journal of Proteome Research*, 19(3), pp. 1346–1350.
41. Kafader, J.O., Melani, R.D., Durbin, K.R., Ikwuagwu, B., Early, B.P., Fellers, R.T., Beu, S.C., Zabrouskov, V., Makarov, A.A., Maze, J.T. and Shinholt, D.L., 2020. Multiplexed mass spectrometry of individual ions improves measurement of proteoforms and their complexes. *Nature Methods*, 17(4), pp. 391–394.
42. Drown, B.S., Gupta, R., McGee, J.P., Hollas, M.A., Hergenrother, P.J., Kafader, J.O. and Kelleher, N.L., 2024. Precise readout of MEK1 proteoforms upon MAPK pathway modulation by individual ion mass spectrometry. *Analytical Chemistry*, 96(11), pp. 4455–4462.
43. Des Soye, B.J., McGee, J.P., Hollas, M.A., Forte, E., Fellers, R.T., Melani, R.D., Wilkins, J.T., Compton, P.D., Kafader, J.O. and Kelleher, N.L., 2024. Automated immunoprecipitation, sample preparation, and individual ion mass spectrometry platform for proteoforms. *Analytical Chemistry*, 96(34), pp. 13879–13887.

56

X-ray Diffraction Imaging for Cancer Detection and Diagnosis

Joel A. Greenberg and Anuj J. Kapadia

56.1 Introduction

X-ray imaging was the first medical imaging modality and has been in continuous use for over 125 years as its performance, architecture, and impacts continue to grow and evolve [1]. Transmission-based X-ray imaging methods, including radiography and computed tomography (CT), are ubiquitously used for diagnostic cancer detection in a variety of organs, treatment planning, image-guided interventions, and monitoring treatment response and recurrence. Despite the success and widespread use of transmission X-ray imaging, the approach is fundamentally limited to providing image contrast based on the density and the effective atomic number of the tissue being imaged. While this is sufficient to identify the presence of calcifications or concretions, determine bone density changes, or detect anomalies in the breast, lung, or brain, the fact that nearly all soft tissue has similar densities and chemical compositions renders conventional, transmission-based X-ray imaging inadequate to confirm or differentiate between benign and malignant tissue [2].

In this chapter, we discuss an alternative approach to X-ray imaging that incorporates molecular sensitivity into the X-ray imaging paradigm to provide critical additional information for cancer detection and biospecimen analysis. Our approach, which we refer to as X-ray diffraction imaging (XRDI), combines the groundbreaking X-ray diffraction (XRD) work of Von Laue and father-son Bragg team in the early 1900s with modern computational imaging methods. Figure 56.1a shows a schematic of the generalized XRDI paradigm that captures both transmission and XRD measurements. In transmission imaging, an X-ray source illuminates a sample, and a detector placed behind the sample records the attenuation (i.e., the reduction in signal strength) of the X-rays passing directly through the sample. This results in the grayscale transmission image, where higher or lower attenuation results in darker or brighter regions of the image. In XRDI, however, the X-rays interact with the sample and scatter (i.e., undergo deflection) that is dependent on the molecular structure of the sample (i.e., the scatterer). These scattered X-rays are recorded and, after a computational post-processing step, generate a spatially resolved image with an XRD form factor, or "signature," at each location [3]. The detailed shape of this XRD form factor can provide information about the molecular structure of the sample at lengths scales ranging from 0.1 nm to more than 100 nm and, in this way, can capture important information about the composition of the tissue and disease state throughout the sample.

While a complete understanding of the relationship between the XRD tissue form factor and the underlying cancer biology is still an area of active research, it is well established that the XRD form factor is directly related to the biomechanical properties and structure of the tissue. These properties are largely influenced by the structure and organization of the extracellular matrix (ECM), which functions as a cellular scaffold and supports cell communication, migration, adhesion, and signaling [4]. The development of metastases has been shown to be dependent on the interaction between cancer cells and the ECM, resulting in a reorganization of the ECM [5, 6]. For example, previous XRD studies have demonstrated that collagen assemblies in the ECM undergo changes during early invasion. Relatedly, it is currently thought that modifications to the alignment of collagen fibers facilitates tumor cell migration and invasion during tumor progression. These and related changes in the extracellular tumor microenvironment have been shown to modify the composition of concretions and calcifications [7, 8], which can provide additional information about the process and progression of tumor development. As a result of these structural tissue changes at the molecular and supramolecular levels, the XRD scatter signal can vary in significant and meaningful ways with cancer progression, and in turn can provide novel insights for cancer biology, biomarker studies, and even indicators of dynamic changes in tumor microenvironment following treatment for real-time patient monitoring [9].

Numerous studies have explored the application of XRD in cancer detection. While the majority of the literature targets breast cancer [10, 11], other studies have presented XRD explorations on liver/colorectal cancer [12], brain/hemorrhage [13], and microcalcifications in breast [14] and prostate. Most of these studies have focused on XRD-based form factor analyses, and none have performed high-resolution XRD-based imaging of tissue samples. As will be discussed in this chapter, we have extended this body of literature to include full XRD images of tissue specimens in both 2D and 3D, enabling the differentiation of ductal carcinoma *in situ* (DCIS) from invasive ductal carcinoma (IDC), identification of white matter, gray matter, and glioblastoma (GBM) in the brain, and evaluation of prostate soft tissue. An important strength of XRDI resides in its ability to provide spatially resolved analysis of tissue across a broad range of cancers.

56.2 Materials

56.2.1 XRDI Scanning System

As shown in Figure 56.1, XRDI systems involve an X-ray source for generating a primary beam of X-rays to illuminate

DOI: 10.1201/9781003449942-63

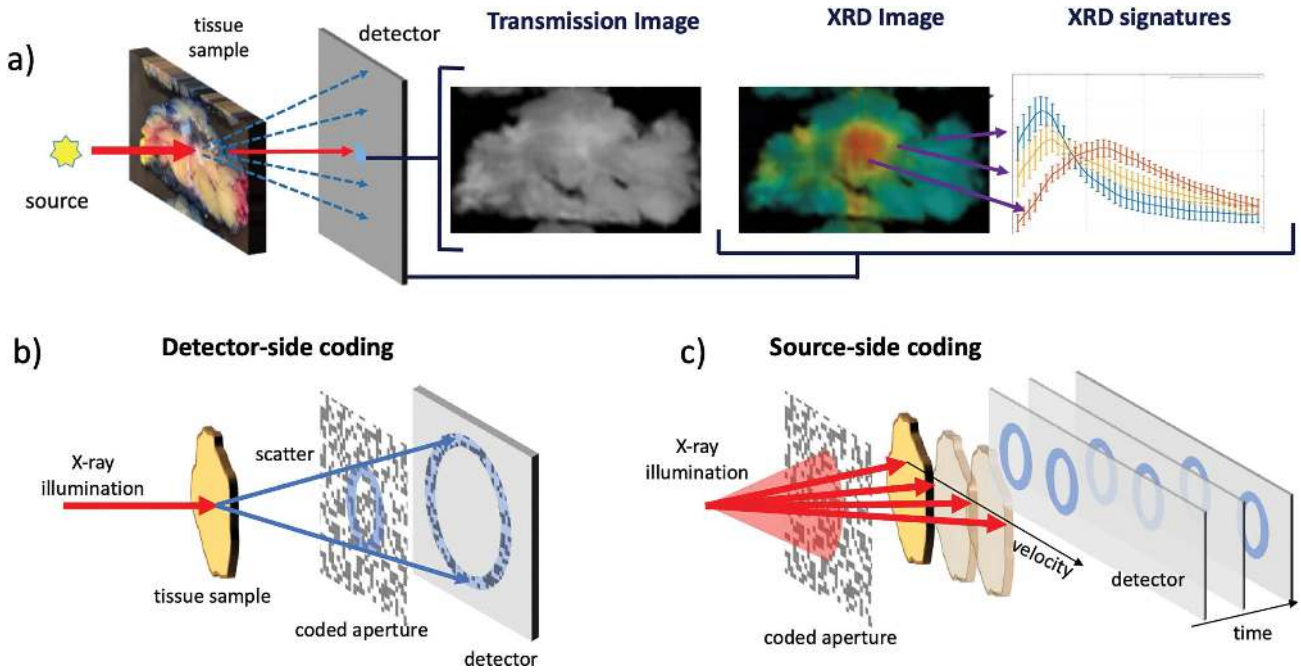


FIGURE 56.1 (a) Schematic of a multi-modality X-ray imaging system, in which an X-ray source illuminates a tissue sample and measurement of the direct-beam attenuation results in a transmission image and measurement of the scattered X-rays results in an X-ray diffraction image. (b) and (c) show specific XRDI implementations using static, detector-side and dynamic, source-side coding configurations, respectively.

the sample as well as a detector to record the scattered X-rays. The tissue is placed between the source and detector, and the system can be configured for either *ex vivo* analysis (e.g., surgically resected tissue) or *in vivo* use (e.g., XRD-enhanced mammography) for diagnostic imaging. The challenge with X-ray scatter imaging is that each location, or voxel, of the object that is illuminated by X-rays produces X-ray scatter that spreads out in nearly all directions. While the XRD signal of interest tends to be localized to small deflection angles relative to the incident primary beam's direction, the spatial extent of the scattered XRD signal means that X-rays from many different voxels can overlap on the detector.

Multiple architectures and methods exist for disentangling the scatter signal from multiple voxels, known as demultiplexing, which generally involve the use of a combination of physical elements and measurement sequences [3]. For example, physical elements, such as collimators and coded apertures, can be used in the system to limit the region of the sample that is illuminated by X-rays and/or to limit the region of the sample whose scatter is visible to the detector. Figure 56.1b shows a schematic of an architecture in which a source-side collimator first creates a pencil beam so that only a line of voxels in the tissue generates a scatter signal [15, 16]. A thin, patterned opaque modulating element, known as a coded aperture, is placed between the sample and detector to encode the location of the scatterer into the measured XRD signal. This method allows for imaging of the entire illuminated volume in a single, snapshot measurement. Alternatively, one can employ motion of the sample relative to the source and detector and a sequence of measurements over time to perform XRDI [17]. Figure 56.1c shows an architecture in which a coded aperture placed between the source and sample creates a spatially structured illumination pattern through which the sample is moved, and a multitude of scatter images are recorded at the detector. The choice of source- and detector-side structuring and sample trajectory should be optimized to the particular measurement constraints of the imaging task.

Regardless of the measurement architecture selected, a computational post-processing step is then used to generate XRD images from the complex, multiplexed scatter measurement data. In many approaches, this postprocessing step involves a reconstruction algorithm that seeks to use knowledge of the system geometry, component responses, scatter physics, and noise sources to provide a best estimate of the object's scattering properties. Specifically, the dependence of the scatter signal as a function of X-ray deflection angle and/or energy is directly related to the XRD form factor, and simulation tools such as QsimRT [18–20], MCGPU [18–20], and GEANT4 [21] can be useful in creating relevant system models used in image reconstruction [22, 23]. The final result of this measurement and estimation process is the XRD image, which is best described as a hyperspectral datacube consisting of a spatially resolved estimate of the XRD form factor at each voxel in the sample. This large, rich datacube can then be used by either a human or a computer to analyze the sample. For example, we have developed tools for visualizing the images by mapping the full XRD form factor to a color for interpretation and evaluation by a human (see Figure 56.2 for an example) [24]. Additionally, we have developed computer-based detection and classification algorithms for accomplishing specific tasks, such as distinguishing between benign and malignant tissue, recognizing tissue subtypes, and identifying anomalous XRD form factors to help isolate biomarkers (see Figure 56.3 for an example) [25–27]. While simple, deterministic models can be

useful for creating interpretable classifiers when the form factor variability is small and a complete library of form factors can be created, machine learning- (ML) and artificial intelligence- (AI) based approaches (e.g., support vector machine (SVM), random forest (RF), neural networks (NN), etc.) coupled with datasets of sufficient size, breadth, and labeling have demonstrated the best results for analyzing the XRDI datacube.

56.2.2 Samples

In addition to the XRDI scanner and associated computational tools required to make the XRD images, sample selection and preparation are critical to XRDI-based cancer studies. For *ex vivo* studies, we have previously shown that both fresh and formalin-fixed tissue provide self-consistent XRD form factors [28] (i.e., formalin fixation does not affect the XRD form factor of a tissue). While paraffin-embedded samples, which are readily available and have a very long shelf life, are an attractive alternative, additional research is necessary to determine whether these samples can be used without deparaffinization. Similarly, additional research should be conducted in the future regarding the impact of flash freezing on XRD form factors of tissue. Regardless of the processing method used, tissue samples for XRDI can be in the form of biopsies, slices, or whole resections, depending on the details of the XRDI system employed. However, because it is often difficult to know the ground truth for a biological sample, it is advantageous to also employ known phantoms for system calibration, quality assurance, and algorithm development [25, 29–31]. Anthropomorphic phantoms, in particular, represent an excellent way of bridging the gap between *ex vivo* and *in vivo* studies, where extrapolations to diagnostic practice and radiological applications can be performed safely and without risk to patients before clinical trials involving humans.

56.3 Methods

Ex vivo analysis of tissue is most relevant to basic researchers, surgeons, and pathologists. In this case, the excised tissue samples are first photographed using an optical camera, which records information about the layout of the samples as well as visible features and any indications based on inking to preserve the spatial orientation within the patient. The samples are then placed in the XRDI scanner and transmission and XRD images are recorded. Depending on the size of the sample and intended image resolution, a complete scan can take anywhere from a few minutes to a few hours. Once the scan is finished, the tissue samples then undergo standard histological preparation, which includes either flash freezing or formalin fixation and paraffin embedding. When the tissue is in a more solid form, it is microtomed, stained, and thin (e.g., 10–50 micron) microscope slides are created.

Pathologists can then image the slides using either conventional microscopy or whole-slide imaging techniques and provide annotations based on cellular and subcellular structures. Examples of annotation include identifying regions of DCIS or IDC in breast, assigning Gleason scores and grading in prostate, and glioblastoma staging in the brain tissue. It is worth noting that the rise of artificial intelligence (AI) histopathology annotation models may enable a future path to more efficient and robust annotations with minimal reliance on pathologists [32]. The annotated tissue maps can then be co-registered with the X-ray transmission and XRD images and serve as a pathology-based classification (or "ground truth") map for the tissue samples. Using this pathology-based classification in combination with the transmission and XRD images, one may exploit data analytics techniques to relate the X-ray-derived features to diagnostic and prognostic information. Examples of this analysis include building ML and AI models for performing cancer detection and grading as well as mining the data for XRD-based biomarker discovery.

While we have focused initially on *ex vivo* tissue analysis as a method to refine the XRD imaging technology and provide a safe and controlled understanding of the relationship between XRD form factors and the underlying disease state, we have also explored the potential for *in vivo* diagnostic XRD imaging. Our initial studies were performed *in silico* using numerical tools to simulate different XRD imaging architectures [21, 33]. In this approach, virtual X-ray sources, detectors, and tissue samples were constructed, and a particular measurement architecture was modeled. As an example, Lakshmanan and colleagues used GEANT4 to model a coherent scatter computed tomography (CSCT) architecture utilizing tissue cross sections from the National Institute of Standards and Technology (NIST) XCom database, tissue chemical compositions available from literature, and in-house measurements of the XRD form factor for various tissues [21]. In addition to GEANT4, we have also developed models with MCGPU [34], which is a GPU-accelerated Monte Carlo simulation tool based on Penelope physics that provides considerably faster calculations compared to GEANT4. Using this approach, we have studied coded aperture-based XRDI systems to understand the detectability of tumors of different sizes surrounded by different volumes and densities of other tissues [35, 36]. When possible, we have complemented these simulation-based studies with empirical measurements using known, controllable phantoms to validate the results and further inform and refine the simulations [37].

56.4 Results

Figure 56.2 shows representative results obtained for a variety of different cancerous tissues using the materials and methods described in Sections 56.3 and 56.4. From top to bottom, we show breast tissue (DCIS and IDC), brain tissue (glioblastoma, GBM), and prostate tissue (adenocarcinoma of the prostate) imagery, respectively. From left to right we show an optical image (or photograph), an H&E stained microscopy image with pathologist-drawn annotations, an X-ray transmission image, and an XRD image visualized with RGB colorization, respectively. Before considering these images in detail, it is worth observing that each imaging modality provides distinct information. For example, the optical and microscopy images provide information from only a very thin layer of the sample, whereas the X-ray images provide an integrated representation

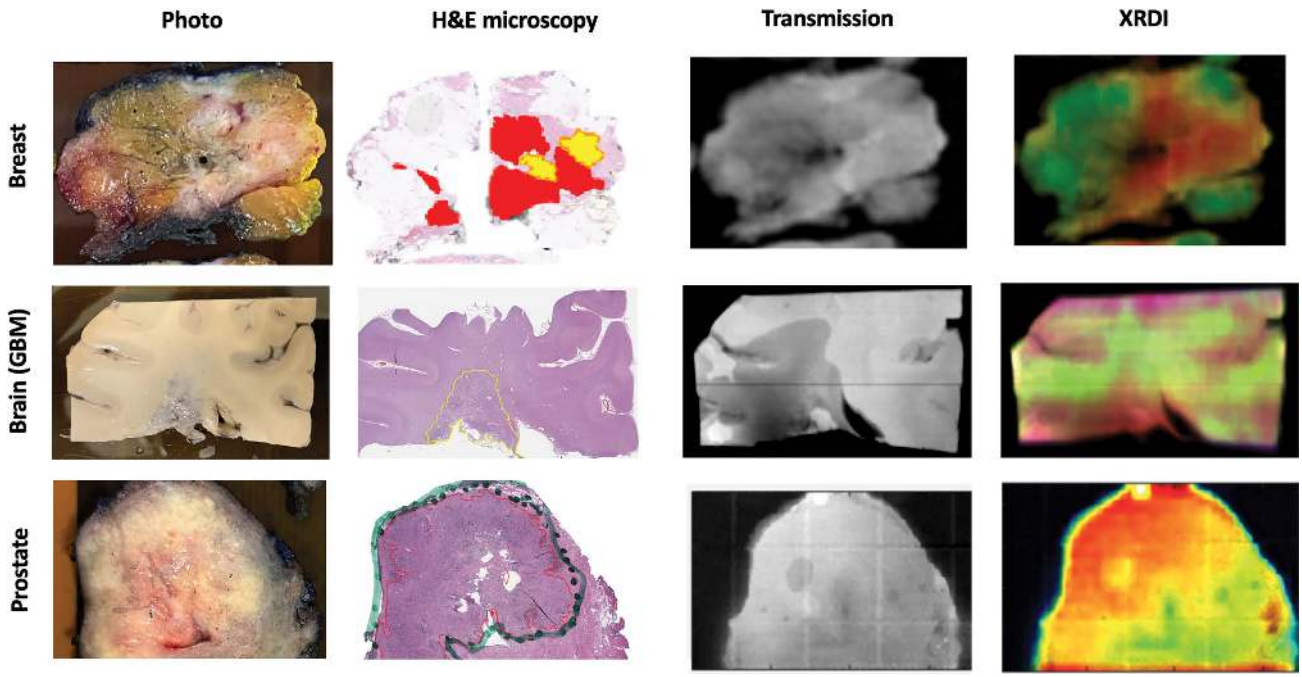


FIGURE 56.2 Images of different cancers via different modalities. (left to right) Optical image (photograph), H&E stained microscopy image (with pathologist-drawn annotations indicating the presence of cancer within the outlined region), X-ray transmission image, and XRD image, respectively. (top to bottom) Images correspond to breast (DCIS and IDC), brain (glioblastoma), and prostate (adenocarcinoma) tissue, respectively.

of the tissue composition throughout the entire thickness of the sample. Also, the transmission and XRD images provide very different contrast from one another, but certain characteristics of the H&E stained microscopy images are maintained in the XRD images, as discussed later in more detail.

Looking first at the breast tissue data, the annotated microscopy images are divided up into three categories: healthy tissue (encompassing adipose, fibroglandular, and mixtures thereof), DCIS (yellow), and invasive carcinoma (red) [31]. While this annotation is a significant oversimplification because it does not consider ductal hyperplasia, does not differentiate between lobular and ductal cancer, and is agnostic of the grade of DCIS, it provides a first and relevant test regarding the efficacy of the various imaging modalities. By first considering the optical image, one can identify the inking as well as regions of discoloration and fat that may indicate a region of interest but is only qualitatively related to the regions of DCIS and IDC observed in the microscopy image. Similarly, the X-ray transmission image shows some degree of contrast and structure but does not show readily observable features that correspond to the DCIS or invasive carcinoma regions, as density alone is not able to capture tissue malignancy or composition. The XRD image, however, shows different regions that generally correspond to the cancerous regions identified by pathology. Here we have chosen a color scheme in which green corresponds to healthy tissue, red corresponds to carcinoma, and a continuum of colors in between correspond to the relative likelihood of each. While this RGB visualization of the XRD image does not represent the full information present in the XRD tissue form factors, its sensitivity to the subtle, molecular differences between the different tissues is evident.

Considering next the brain tissue, one sees both similarities and differences with the observations around the breast tissue. An important difference is that, unlike in breast, the optical image allows one to differentiate easily between regions of white matter (smooth light regions), gray matter (smooth gray regions) and glioblastoma (darker, mottled regions that include necrosis). The XRD image color scheme is chosen to best identify each of these tissues, and one can see that the white matter (in green), gray matter (in purple), and cancer (dark red) can be readily distinguished from one another and agree with the optical image. The annotation on the microscopy image confirms that the region in the bottom middle of the brain sample is GBM and agrees with the findings of the XRDI system [31]. In contrast, the transmission image shows that the density of white and gray matter is nearly identical and therefore is unable to provide meaningful differentiation between these two tissue types. In addition, while a large higher-density (presented as darker) region can be observed in the image, it does not correspond to the GBM and, as such, further highlights the ability for XRDI to differentiate between the tissue based on their unique molecular structures rather than on chemical or density features.

Finally, the bottom row in Figure 56.2 shows a set of images of prostate tissue. Prostate gland analysis is a unique problem in pathology as it is often challenging to identify the locations of prostate tumors at prosection and perform quantitative analysis based on established biomarkers at or around point of care. This difficulty can be observed based on the fact that neither the optical nor X-ray transmission images show much contrast indicative of a differentiation between benign and malignant tissue, let alone the ability to provide a meaningful

cancer grading. Nevertheless, changes in the molecular structure of the tissue can be observed in the XRD image, where the red color indicates a stronger presence of prostate cancer while the green color indicates a stronger presents of normal, healthy prostate tissue. The findings agree generally with the annotated microscopy image, indicating the potential utility of XRDI in providing additional understanding of the molecular changes accompanying prostate cancer progression.

56.5 Discussion

The colorized XRD images are useful for providing humans an interpretable, analog method for viewing the XRD datacube; however, quantitative use of XRDI for tissue analysis requires use of the full XRD form factor. Figure 56.3a shows the average form factors of healthy, DCIS, and IDC (cancerous) breast tissue. A change can be seen in the peak location and amplitude of the XRD form factor, with a progression of peak location from left to right for healthy, DCIS, and IDC tissue, respectively. We differentiate between the case of XRD form factor for DCIS adjacent to IDC and distant from IDC. This helps account for some of the subjectivity in pathologist annotations as well as furthers the understanding that progression from healthy to malignant breast tissue occurs along a continuum and that the ECM modifications to which XRD is sensitive may not be directly correlated with the intracellular changes presented in the H&E images. By training a neural network classifier to distinguish between IDC and non-IDC tissue across tens of thousands of independent XRD form factors spanning dozens of patients, we can see in Figure 56.3b that regions of cancer can be automatically identified and that the results are in good agreement with the annotations provided by pathologists (referred to here as an "H&E classification" of the tissue). While the results vary based on the specific classifier model and training used, we typically find an area under the curve (AUC) of over 0.95.

We have applied this same methodology to the case of glioblastoma in the brain. Figure 56.3c shows the average XRD form factors for white matter, gray matter, and glioblastoma. While the differences in the XRD form factors are slight, they are statistically relevant and consistent across the eight brain samples that we have evaluated to date. By training a SVM classifier to distinguish between GBM and non-GBM tissue, we find good agreement between the XRDI classification result and the "ground truth" based on a visual inspection of the optical image (see Figure 56.3d).

In the *in vivo*, diagnostic domain, the results of our simulation-based studies show that XRDI may be a useful complement to standard MRI and X-ray transmission-based imaging. With respect to breast imaging, tumors on the order of a centimeter in extent can be evaluated within a reasonable

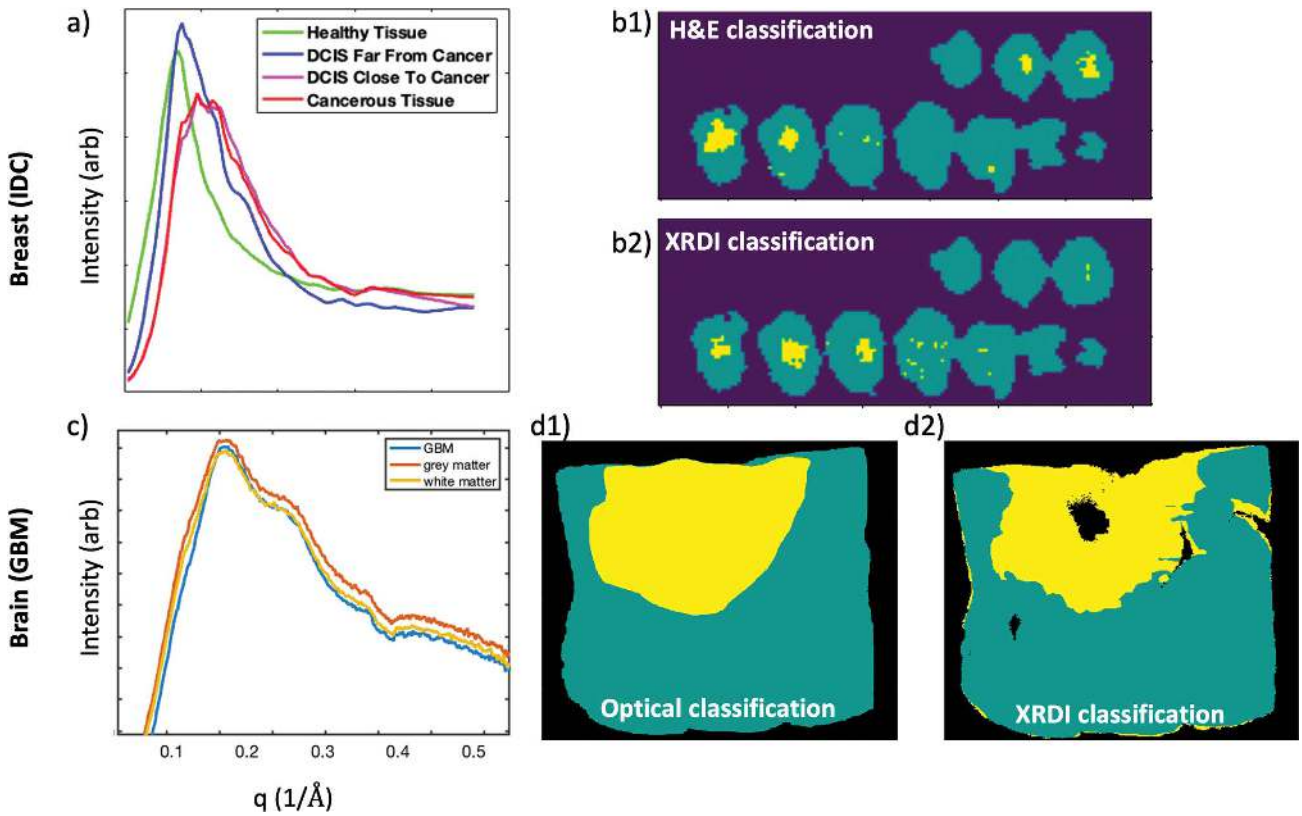


FIGURE 56.3 Quantitative analysis of XRD form factors. (a) and (c) show XRD form factors for breast tissue (healthy, DCIS, and IDC) and brain tissue (white matter, gray matter, and GBM), respectively. (b) and (d) show a comparison between ground truth classification maps (obtained by H&E or optical imagery and annotation) and XRDI-based classification maps, respectively. Here yellow indicates a cancerous region and teal indicates a non-cancerous region.

dose using conventional X-ray equipment and AI-based data analytics models [36, 38]. This could potentially enable XRDI to become a secondary screening tool in detecting and diagnosing breast cancer. Additional research is needed to understand its applicability to the detection and analysis of cancers in other organs, especially in the case of non-superficial imaging of the brain or prostate.

56.6 Future Trends

XRDI represents an exciting new medical imaging modality because of its ability to penetrate through centimeters of tissue, provide sum-millimeter scale spatial resolution, and offer novel molecular sensitivity at the nanometer scale. While the tissue form factors are not sensitive to single cells or intracellular structure, they do provide access to extracellular molecular and structural changes that accompany tumor initiation and progression and aid in studying the microtumor environment. We have shown that XRDI is relevant to the evaluation of a range of different cancers found in a variety of organs, including analysis of soft tissue in the breast, brain, and prostate as well as hard tissue in the bones and microcalcifications. The technology is relevant to basic research and biospecimen analysis in 2D and 3D [30, 39, 40], clinical *ex vivo* applications in pathology and surgery for performing tissue triaging and margin assessment, and, potentially, for *in vivo* diagnostic imaging.

Despite the rich information contained in the XRD form factor, XRDI currently is more complex, slower, and lower resolution than X-ray transmission imaging. We therefore view both X-ray transmission and diffraction imaging as being valuable for tissue analysis in clinical and research settings, and we have developed multi-modality systems that include both technologies. Figure 56.4a–d shows the development of XRDI systems from an initial testbed on an optical table through multiple generations of maturation culminating in Quadridox's XR4D scanner (for 3D biospecimen analysis) and BioSpecx scanner (for clinical assessment of tissue). A future implementation of the technology is shown schematically in Figure 56.4e, where XRDI could be combined with diagnostic radiology imaging (e.g., mammography or breast CT) to provide pinpointed secondary evaluation of suspicious regions. Other applications of this technology are in the analysis of concretions, such as kidney and bladder stones [41], and as a complement to dual energy X-ray imaging for bone densitometry and mineral analysis [42] in veterinary and human medicine. We have demonstrated XRDI utility and shown that the technology is mature and readily available to researchers and clinicians; as such, deeper and more fundamental studies of tissue structure and disease can be

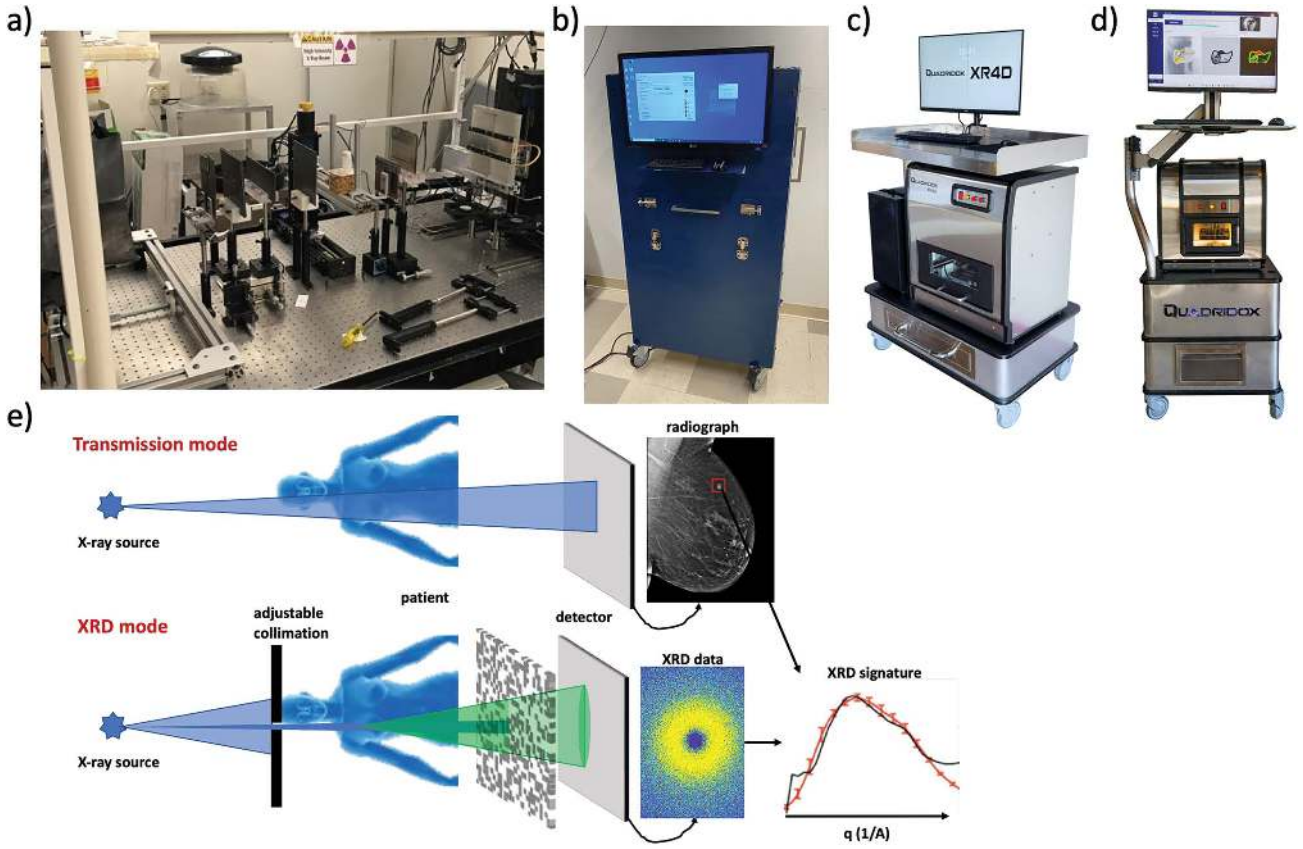


FIGURE 56.4 Development of XRDI technology. (a) An optical testbed system initially developed for exploring detector-side coded aperture-based XRDI; (b) the XR3D system developed at Duke University; (c) the XR4D system developed at Quadridox to perform 3D XRDI biospecimen analysis; and (d) the BioSpecx system developed at Quadridox to perform clinical XRDI tissue analysis. (e) A potential future implementation of *in vivo* diagnostic imaging using detector-side coded apertures.

now conducted to potentially aid in the detection, diagnosis, and treatment monitoring of cancer.

REFERENCES

1. M. Berger, Q. Yang, and A. Maier, "X-ray imaging," in *Medical Imaging Systems: An Introductory Guide*, A. Maier, S. Steidl, V. Christlein, and J. Hornegger, Eds., Cham: Springer International Publishing, 2018, pp. 119–145. doi: 10.1007/978-3-319-96520-8_7.
2. A. Pulumati, A. Pulumati, B. S. Dwarakanath, A. Verma, and R. V. L. Papineni, "Technological advancements in cancer diagnostics: Improvements and limitations," *Cancer Rep.*, vol. 6, no. 2, p. e1764, Feb. 2023, doi: 10.1002/cnr2.1764.
3. J. Greenberg, "X-ray diffraction imaging: Technology and applications," in *Devices, Circuits, and Systems*, CRC Press, 2018 [Online]. Available: https://books.google.com/books?id=dz58DwAAQBAJ.
4. J. Wu, S. Sheng, X. Liang, and Y. Tang, "The role of tumor microenvironment in collective tumor cell invasion," *Future Oncol.*, vol. 13, no. 11, pp. 991–1002, May 2017, doi: 10.2217/fon-2016-0501.
5. M. Khoonkari *et al.*, "Physics of brain cancer: Multiscale alterations of glioblastoma cells under extracellular matrix stiffening," *Pharmaceutics*, vol. 14, 2022.
6. S. Kaushik, M. W. Pickup, and V. M. Weaver, "From transformation to metastasis: Deconstructing the extracellular matrix in breast cancer," *Cancer Metastasis Rev.*, vol. 35, no. 4, pp. 655–667, Dec. 2016, doi: 10.1007/s10555-016-9650-0.
7. S. Gosling *et al.*, "Calcification microstructure reflects breast tissue microenvironment," *J. Mammary Gland Biol. Neoplasia*, vol. 24, no. 4, pp. 333–342, Dec. 2019, doi: 10.1007/s10911-019-09441-3.
8. J. Nallala *et al.*, "Abstract 5561: Molecular imaging of calcification and soft tissue defines prognosis of DCIS to invasive breast cancer," *Cancer Res.*, vol. 83, no. 7_Supplement, pp. 5561–5561, Apr. 2023, doi: 10.1158/1538-7445.AM2023-5561.
9. Alex J. Walsh, Rebecca S. Cook, Jae H. Lee, Carlos L. Arteaga, and Melissa C. Skala, "Collagen density and alignment in responsive and resistant trastuzumab-treated breast cancer xenografts," *J. Biomed. Opt.*, vol. 20, no. 2, p. 026004, Feb. 2015, doi: 10.1117/1.JBO.20.2.026004.
10. G. Kidane, R. D. Speller, G. J. Royle, and A. M. Hanby, "X-ray scatter signatures for normal and neoplastic breast tissues," *Phys. Med. Biol.*, vol. 44, no. 7, p. 1791, Jul. 1999, doi: 10.1088/0031-9155/44/7/316.
11. P. C. Johns, and M. P. Wismayer, "Measurement of coherent x-ray scatter form factors for amorphous materials using diffractometers," *Phys. Med. Biol.*, vol. 49, no. 23, p. 5233, Nov. 2004, doi: 10.1088/0031-9155/49/23/003.
12. C. Theodorakou, and M. J. Farquharson, "The classification of secondary colorectal liver cancer in human biopsy samples using angular dispersive x-ray diffraction and multivariate analysis," *Phys. Med. Biol.*, vol. 54, no. 16, p. 4945, Jul. 2009, doi: 10.1088/0031-9155/54/16/007.
13. P. Lazarev, M. Paukshto, N. Pelc, and A. Sakharova, "Human tissue X-ray diffraction: Breast, brain, and prostate," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Cat. No.00CH37143)*, vol. 4, pp. 3230–3233, Jul. 2000, doi: 10.1109/IEMBS.2000.901591.
14. R. Scott, N. Stone, C. Kendall, K. Geraki, and K. Rogers, "Relationships between pathology and crystal structure in breast calcifications: An in situ X-ray diffraction study in histological sections," *Npj Breast Cancer*, vol. 2, no. 1, p. 16029, Sep. 2016, doi: 10.1038/npjbcancer.2016.29.
15. K. MacCabe, K. Krishnamurthy, A. Chawla, D. Marks, E. Samei, and D. Brady, "Pencil beam coded aperture x-ray scatter imaging," *Opt. Express*, vol. 20, no. 15, pp. 16310–16320, Jul. 2012, doi: 10.1364/OE.20.016310.
16. J. A. Greenberg, K. Krishnamurthy, and D. Brady, "Snapshot molecular imaging using coded energy-sensitive detection," *Opt. Express*, vol. 21, no. 21, pp. 25480–25491, Oct. 2013, doi: 10.1364/OE.21.025480.
17. J. A. Greenberg, M. Hassan, K. Krishnamurthy, and D. Brady, "Structured illumination for tomographic X-ray diffraction imaging," *Analyst*, vol. 139, no. 4, pp. 709–713, 2014, doi: 10.1039/C3AN01641B.
18. David Coccarelli, Michael E. Gehm, and Joel A. Greenberg, "Flexible rapid x-ray simulation for large-scale synthetic data generation across a variety of x-ray modalities," in *Presented at the Proc. SPIE*, p. PC1210403, Jun. 2022, doi: 10.1117/12.2618945.
19. Q. Gong, R.-I. Stoian, D. S. Coccarelli, J. A. Greenberg, E. Vera, and M. E. Gehm, "Rapid simulation of X-ray transmission imaging for baggage inspection via GPU-based ray-tracing," *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.*, vol. 415, pp. 100–109, Jan. 2018, doi: 10.1016/j.nimb.2017.09.035.
20. Q. Gong, J. A. Greenberg, R.-I. Stoian, D. Coccarelli, E. Vera, and M. E. Gehm, "Rapid simulation of X-ray scatter measurements for threat detection via GPU-based ray-tracing," *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.*, vol. 449, pp. 86–93, Jun. 2019, doi: 10.1016/j.nimb.2019.03.006.
21. M. N. Lakshmanan, B. P. Harrawood, E. Samei, and A. J. Kapadia, "Volumetric x-ray coherent scatter imaging of cancer in resected breast tissue: A Monte Carlo study using virtual anthropomorphic phantoms," *Phys. Med. Biol.*, vol. 60, no. 16, p. 6355, Aug. 2015, doi: 10.1088/0031-9155/60/16/6355.
22. Amit Ashok, Shu Yang, and Joel Greenberg, "Experimental validation of spatio-spectral regularized x-ray coherent diffraction reconstruction," in *Presented at the Proc. SPIE*, p. 1253107, Jun. 2023, doi: 10.1117/12.2664384.
23. Turner Richmond *et al.*, "Improved reconstruction and classification performance in a hybrid CT+XRD explosives detection system," in *Presented at the Proc. SPIE*, p. 1304306, Jun. 2024, doi: 10.1117/12.3013531.
24. S. Stryker, J. A. Greenberg, S. J. McCall, and A. J. Kapadia, "X-ray fan beam coded aperture transmission and diffraction imaging for fast material analysis," *Sci. Rep.*, vol. 11, no. 1, p. 10585, May 2021, doi: 10.1038/s41598-021-90163-0.
25. S. Stryker, A. J. Kapadia, and J. A. Greenberg, "Application of machine learning classifiers to X-ray diffraction imaging with medically relevant phantoms," *Med. Phys.*, vol. 49, no. 1, pp. 532–546, Jan. 2022, doi: 10.1002/mp.15366.
26. Bi Zhao, Scott Wolter, and Joel A. Greenberg, "Application of machine learning to x-ray diffraction-based classification," in *Presented at the Proc. SPIE*, p. 1063205, May 2018, doi: 10.1117/12.2304683.

27. R. Thamvichai, Dan Pike, J. Greenberg, and A. Ashok, "Classifiers for material discrimination with dual modality x-ray transmission and coherent x-ray diffraction system," in *Presented at the Proc. SPIE*, p. 1253104, Jun. 2023, doi: 10.1117/12.2664196.
28. J. Spencer, "X-ray diffraction spectral imaging for breast cancer assessment," Master's thesis, Duke University, 2017.
29. B. Ghammraoui, A. Zidan, A. Alayoubi, A. Zidan, and S. J. Glick, "Fabrication of microcalcifications for insertion into phantoms used to evaluate x-ray breast imaging systems," *Biomed. Phys. Eng. Express*, vol. 7, no. 5, p. 055021, Aug. 2021, doi: 10.1088/2057-1976/ac1c64.
30. Zachary Gude, Anuj J. Kapadia, and Joel A. Greenberg, "3D x-ray diffraction imaging of optically thin samples," in *Presented at the Proc. SPIE*, p. 1210405, Jun. 2022, doi: 10.1117/12.2618938.
31. Joel A. Greenberg *et al.*, "X-ray diffraction imaging of glioblastoma and ductal carcinoma in-situ," in *Presented at the Proc. SPIE*, p. 129250I, Apr. 2024, doi: 10.1117/12.3006600.
32. M. Y. Lu *et al.*, "A multimodal generative AI copilot for human pathology," *Nature*, vol. 634, no. 8033, pp. 466–473, Oct. 2024, doi: 10.1038/s41586-024-07618-3.
33. S. Stryker, A. J. Kapadia, and J. A. Greenberg, "Simulation based evaluation of a fan beam coded aperture x-ray diffraction imaging system for biospecimen analysis," *Phys. Med. Biol.*, vol. 66, no. 6, p. 065022, Mar. 2021, doi: 10.1088/1361-6560/abe779.
34. A. Badal, and A. Badano, "Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit," *Med. Phys.*, vol. 36, no. 11, pp. 4878–4880, Nov. 2009, doi: 10.1118/1.3231824.
35. M. Japzon, "Optimization of X-ray diffraction imaging of medical specimens by Monte Carlo," Master's thesis, Duke University, 2019.
36. K. Ferguson, "Analysis of X-ray diffraction imaging for thick tissue imaging using a GPU-accelerated Monte Carlo code," Master's thesis, Duke University, 2023.
37. C. Miller, "Quantifying the sensitivity in X-ray diffraction measurements in thick tissue samples," Master's thesis, Duke University, 2023.
38. O. Fasina, "Implementation and validation of a GPU-based X-ray diffraction Monte Carlo simulator for in-vivo breast imaging applications," Master's thesis, Duke University, 2021.
39. Z. Gude, A. J. Kapadia, and J. A. Greenberg, "Resolution analysis of a volumetric coded aperture X-ray diffraction imaging system," *J. X-Ray Sci. Technol.*, vol. 32, no. 3, pp. 809–822, 2024, doi: 10.3233/XST-230244.
40. Manu N. Lakshmanan, Joel A. Greenberg, Ehsan Samei, and Anuj J. Kapadia, "Design and implementation of coded aperture coherent scatter spectral imaging of cancerous and healthy breast tissue samples," *J. Med. Imaging*, vol. 3, no. 1, p. 013505, Feb. 2016, doi: 10.1117/1.JMI.3.1.013505.
41. S. (Porikli) Durdağı, A. H. H. Al- Jalawee, P. Yalçin, A. S. Bozkurt, and S. Salcan, "Morphological characterization and phase determination of kidney stones using X-ray diffractometer and scanning electron microscopy," *Chin. J. Phys.*, vol. 83, pp. 379–388, Jun. 2023, doi: 10.1016/j.cjph.2022.08.025.
42. L. C. Bonar, A. H. Roufosse, W. K. Sabine, M. D. Grynpas, and M. J. Glimcher, "X-ray diffraction studies of the crystallinity of bone mineral in newly synthesized and density fractionated bone," *Calcif. Tissue Int.*, vol. 35, no. 1, pp. 202–209, Dec. 1983, doi: 10.1007/BF02405032.

57

Restriction Spectrum Imaging (RSI) Models for the Detection of Female Cancers

Aida Haddad, Summer J. Batasin, Sabrina Kohanzad, Emily Barnard, Ana E. Rodríguez-Soto, Stéphane Loubrie, and Rebecca Rakow-Penner

57.1 Introduction

Diffusion weighted imaging (DWI) is a magnetic resonance (MR) imaging technique which makes use of water's diffusion properties to elucidate information about tissue cellularity [1–3]. Diffusion phenomena are typically described in three modes: free, hindered, and restricted [4–7].

Free diffusion refers to the unobstructed movement of water molecules *in vivo* in free fluid and demonstrates Gaussian diffusion (Figure 57.1) [4]. Hindered diffusion of water molecules in tissue describes extracellular water where distance traveled is limited by obstacles in the extracellular compartment, also demonstrating Gaussian diffusion. On the other hand, restricted diffusion is non-Gaussian and is best described as the movement of water within a cell (the intracellular compartment). Of note, all modes of water molecule diffusion are dependent on cell membrane permeability and diffusion time. While rates of free and hindered diffusion depend on the space between cells in the extracellular compartment, restricted diffusion is affected by the size of the intracellular compartment [4, 8].

Diffusion-weighted MRI can be quantitatively represented with the apparent diffusion coefficient (ADC), which serves as a diffusion imaging biomarker. ADC is the average area occupied by a water molecule per unit time (mm^2/s) and is calculated using signal measurements from diffusion weighted imaging acquisitions. Measuring the diffusion-weighted signal intensity of restricted and hindered modes of diffusion plays an integral role in assessing the presence of hypercellular tissue. This relationship is described in the following equation:

$$S_D = S_0 e^{-b \cdot ADC} \tag{57.1}$$

where S_D is the diffusion weighted signal intensity, S_0 is the signal intensity without diffusion weighting, and b is the diffusion sensitization factor in s/mm^2 [2, 3]. This equation provides a measurement of ADC of the individual tissues in the imaging volume, or voxels. Of note, ADC and Equation 57.1 both account for free, hindered, and restricted diffusion. Though ADC is quantitative, the ADC threshold values for benign versus malignant lesions often overlap.

Physiologically, free diffusion is observed in vesicular organs such as the gallbladder and bladder, or in the peritoneum where free fluid is plentiful. Healthy solid organs exhibit hindered diffusion, which makes sense when we consider the intracellular compartment predominance of the liver versus the fluid contents of a full bladder. Pathologically, hindered and restricted modes of diffusion are most salient in hypercellular tissues. For example, the increased cellularity of smaller malignant cells impacts the size of the intracellular and extracellular compartments, resulting in increased obstacles for water molecule diffusion in both compartments. However, separating true intracellular restricted diffusion from hindered diffusion is challenging with standard diffusion imaging and requires advanced modeling.

Restriction spectrum imaging (RSI) has been shown to separately assess free, hindered, and restricted diffusion weighted signal intensities in heterogeneous tissues, historically within the brain and prostate [9, 10]. Separate maps, derived from independent compartments with fixed ADCs, are produced using an input of multiple b-values collected in multiple diffusion directions. The model is tissue-specific. For example, there are different models (and fixed ADCs for different tissue-specific compartments) for breast, cervical, and ovarian cancer imaging. The maps produced represent restricted diffusion (C1), restricted/hindered diffusion (C2), and free/flow diffusion (C3) (Figure 57.1) [4]. These separate maps allow the reader to not only visually observe diffusion characteristics of a given tissue, but also offer definitive quantitative measures for free, hindered, and restricted diffusion with increased conspicuity. Doing so allows for the distinction between restricted and hindered diffusion, surrogates for tissue cellularity, and free diffusion, or a surrogate for acellular edema or necrosis [11].

The generic RSI equation is as follows:

$$S_{diff}(b,N) = \sum_{i=1}^{N} C_{i,N} e^{-bD_{i,N}} = S_0 \sum_{i=1}^{N} e^{-bD_{i,N}} \tag{57.2}$$

where N is the total number of exponential decays (here 2 and 3), $C_{i,N}$ are the signal contributions of each exponential component, b are the b-values, and $D_{i,N}$ are the ADCs of each exponential component, and $D_{1,N} < D_{2,N} < D_{3,N}$ [12].

RSI-MRI is being developed as a useful tool in breast and gynecologic cancer screening, staging, and surveillance. Our work has shown that RSI provides better visual tumor conspicuity, as well as offers a more accurate and distinguishable imaging biomarker when compared to standard clinical DWI- or dynamic contrast-enhanced (DCE)-MRI.

DOI: 10.1201/9781003449942-64

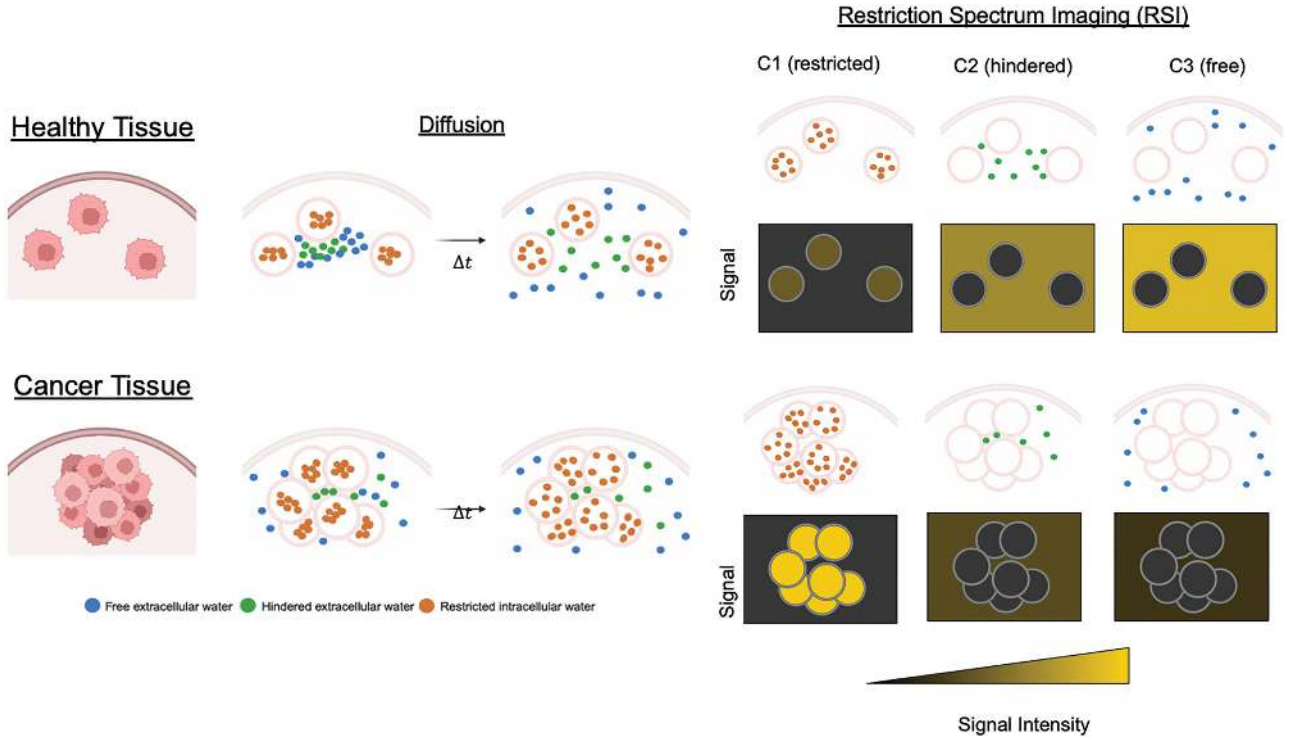


FIGURE 57.1 *Principles of diffusion and restriction spectrum imaging.* Healthy and cancerous tissue display different spatial arrangements, impacting their diffusion properties and signal representations in restriction spectrum imaging (RSI). RSI stratifies diffusion into different compartments: restricted (C1), restricted/hindered (C2), hindered/free (C3), and free/flow (C4). Note, only C1–3 are visualized here. The signal in each compartment corresponds to the distribution of each diffusion type in a tissue. Healthy tissue, with its organized cell structure, has more extracellular water (blue and green) compared to intracellular water (orange). This corresponds to higher signals in C2 and C3. Conversely, cancerous tissue has a higher nuclear-to-cytoplasm ratio, resulting in higher C1 signal. Additionally, it is more densely packed and has less extracellular space, leading to moderate hindered water (green) signal and reduced C3 signal. (Created with BioRender.com.)

57.1.1 Breast Cancer

Breast cancer is the leading cause of cancer in women in the US [13], aside from skin cancer. The American College of Radiology recommends that women at high risk of developing breast cancer should undergo annual dynamic contrast-enhanced MRI (DCE-MRI) exams starting at ages 25–30; MRI is also used to stage women with a new diagnosis of breast cancer or evaluate response to neoadjuvant chemotherapy [14, 15]. Diffusion-weighted imaging has been shown to evaluate response to treatment earlier than DCE-MRI and is being evaluated for potential use as a non-contrast screening tool [16–18]. As a result of these factors, a breast-specific RSI model was created to provide a complementary imaging technique for use in high-risk screening, diagnosis, and follow-up.

Rodríguez-Soto and colleagues developed an RSI model for breast, and the application of this model by Andreassen and colleagues has described the different modes of diffusion of the three primary tissue constituents within breast tissue: fat, fibroglandular tissue, and neoplasm [12, 19–21]. C1 represents restricted diffusion, largely in cancer cells, and C2 demonstrates restricted/hindered diffusion in cancer and normal cells. Water molecules in fibroglandular tissue demonstrated increased restricted/hindered diffusion weighted signal (C2), implying a more extensive extracellular compartment in this healthy tissue when compared to healthy fatty tissue. While cancer demonstrates both restricted and hindered diffusion, the product of C1 and C2 results in a clinically significant differentiation of healthy breast tissue from breast cancer [19].

57.1.2 Cervical Cancer

Cervical cancer is the fifth most common cause of cancer death in women in the US [13]. Radiation therapy, a mainstay in cervical cancer treatment, causes significant edema of the treated area and delays surveillance of treatment efficacy using PET/CT by three to six months. While the pathophysiology of radiotherapy edema is not well understood, the principles of RSI still take advantage of its effects. Knowing a patient's response to treatment earlier allows the clinician and patient to adjust the treatment plan earlier rather than delaying a decision by three to six months. RSI offers an alternative to the clinical challenge of delayed treatment response monitoring compared to PET/CT [22]. Furthermore, PET/MRI shows promise in future staging utility, given its ability to more sensitively assess primary tumor burden than its CT counterpart [23]. RSI models for cervical cancer were first developed in 2018, and the tetra-exponential RSI model for cervical cancer was developed in 2022 [22, 24, 25]. Preliminary studies, addressed in detail later in this chapter, show promising data of not only the model differentiating cancerous from healthy cervical tissue,

but also its role in distinguishing cancer from edema better than PET/CT.

57.1.3 Ovarian Cancer

Ovarian cancer is the sixth leading cause of cancer-related deaths among women in the United States [13]. Those with high-risk BRCA 1 and 2 mutation carriers aged 35–40 are recommended to undergo prophylactic salpingo-oophorectomy as the standard care, despite its association with premature menopause and the associated consequences [26–29]. Current screening methods, including ultrasound and CA-125 levels, have shown no mortality benefit for average-risk women and remain limited for high-risk patients, emphasizing the critical need for advanced imaging techniques to improve early detection [26, 27, 29, 30]. RSI-MRI shows a promising prospective application for ovarian cancer screening, particularly in high-risk patients with BRCA 1 and 2 mutations, potentially delaying or eliminating the need for prophylactic salpingo-oophorectomies.

57.2 Materials

The parameters in RSI Equation 57.2 are organ-specific and must be tuned to each organ of interest due to variations in tissue configurations that affect the estimated ADCs per model compartment. These values are determined empirically using data acquired from a multi-shell DWI acquisition, where multiple non-zero b-values (corresponding to the number of compartments intended for the model) are collected at multiple diffusion directions. The parameters of the MRI acquisition, which includes the choice of b-values, number of diffusion directions, field-of-view size, and so on, are specific to the organ of interest [8]. Acquired data is then corrected for noise, eddy currents, and distortions [31–33]. For the applications discussed later, all studies were performed at 3T MRI (GE Healthcare, Waukesha WI). The technique itself can be performed on any vendor's instrument, as the technique is a post-processing technique assuming the data was collected appropriately.

57.3 Methods

For cancer detection using RSI, all diffusion directions are averaged at each b-value. This results in a pre-processed diffusion volume, which can be used for RSI model determination.

RSI modeling is multi-exponential, and bi-, tri-, and tetra-exponential models are commonly explored with each indication. This is mathematically limited by the number of non-zero b-values available in the multi-shell DWI dataset used in the RSI framework. The RSI models, along with their interpreted corresponding diffusion modes, are outlined in Table 57.1 [24].

Compartmental ADCs for the organ-specific RSI model are determined empirically utilizing regions-of-interest outlined on the diffusion images, and signal contributions for each compartment are calculated in each voxel. Individual compartment maps that separate signal into their respective diffusion modes are the final outputs of the RSI model [12]. These maps, and their extracted values, have been useful in improving tumor conspicuity across female cancers such as in breast, cervical, and ovarian cancers compared to standard metrics.

57.4 Results

57.4.1 Breast-Specific RSI Model

$$S_{diff}(b) = C_{1,3} + C_{2,3}e^{-b\cdot 1.5\times 10^{-3}} + C_{3,3}e^{-b\cdot 10.8\times 10^{-3}} \quad (57.3)$$

As described earlier, the breast-specific RSI model was made with the following three RSI diffusion components: C1, which represents restricted/hyper-restricted diffusion

TABLE 57.1

Restriction Spectrum Imaging Model Types and Interpreted Water Pools

RSI model	Compartment	Interpreted Water Pool
Bi-exponential	C_1	Restricted-Hindered
	C_2	Free-Flow
Tri-exponential	C_1	Restricted
	C_2	Hindered
	C_3	Free-Flow
Tetra-exponential	C_1	Restricted
	C_2	Hindered
	C_3	Free
	C_4	Flow

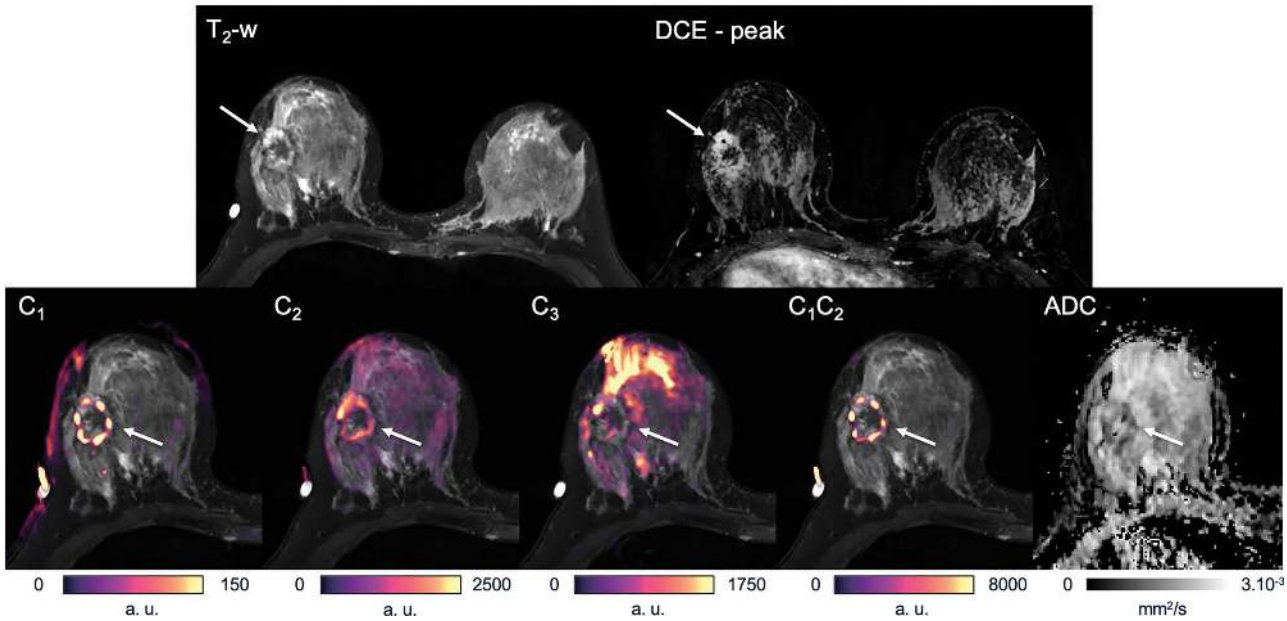


FIGURE 57.2 *Example of RSI outputs from a breast-specific tri-exponential model demonstrating similar tumor conspicuity as DCE in a patient with an invasive ductal carcinoma*. The lesion, indicated by the white arrow, is seen on the peak enhancement image from the DCE. The cancer has intense signal on C1, moderate signal on C2, and almost no signal on C3. The cancer is well delineated on C1C2 and is also seen on the ADC map.

in cancer and healthy fatty tissue; C2, which represents restricted/hindered diffusion in cancer and healthy fibroglandular tissue; and C3, which represents fast diffusion and vasculature [12]. The C1C2 parameter of the three-component model discriminates cancer from healthy breast tissue voxels. Voxels low on C1 or C2 corresponded to healthy breast tissue and increased voxels in C1 and C2 corresponded to increased cancer probability [19]. The C1C2 parameter demonstrates comparable diagnostic accuracy to DCE-MRI for cancer detection in lesions greater than 2 cm (Figure 57.2) [19]. The high diagnostic accuracy of this combined C1C2 component is because of its combination of signals from the two slowest components, allowing for suppression of most of the fatty and fibroglandular tissue while still maintaining signal contributions from T2. The result is improved specificity in differentiating cancer from healthy breast tissue while maintaining comparable sensitivity [19].

Furthermore, because the RSI model aims to characterize the tumor microenvironment based on tissue-specific diffusion patterns, RSI-MRI is less affected by medication-induced alterations in edema or pregnancy-induced alterations in lactation, serving great clinical utility in ability to differentiate active breast malignancy from edema and lactation [11, 34].

Additionally, RSI-MRI can differentiate between benign and malignant lesions, as well as differentiating between average-risk and high-risk benign lesions. We found that the C1 signal was higher in cancer than both benign lesions and fibroglandular tissue; the C3 signal was higher in average-risk benign lesions versus high-risk benign lesions; and the C2 signal was higher in both cancers and benign lesions when compared to fibroglandular tissue [20].

The RSI model was also used to evaluate response to neoadjuvant therapy in breast cancer patients. Traditionally, tumor size in DCE-MRI and mean ADC values from DWI are used to determine treatment response; however, the RSI classifier with elastic imaging registration can predict response to treatment after only three weeks and showed performance comparable to DCE [35, 36]. Earlier detection of treatment response via RSI is important because it allows for tailoring of treatment regimens, and more specifically, identifies those who may be poor responders to treatment and need escalated treatment.

57.4.2 Cervix-Specific RSI Model

Promising preliminary results demonstrate that the restricted diffusion compartment (C1) may identify cancerous cervical tissue from neighboring healthy tissue (Figure 57.3). DWI images of healthy volunteers (n = 11) were matched with patients with stages IA2 and IIIC1 cervical cancer (n = 15). Regions of interest (ROIs) were delineated by body radiologists, and bi-, tri-, and tetra-exponential restriction spectrum imaging compartments (ROIs) were then applied to these ROIs. The tetra-exponential, cervix-specific RSI model with compartmental ADCs is as follows:

$$S_{diff}(b) = C_{1,4} + C_{2,4}e^{-b\cdot 1.2\times 10^{-3}} + C_{3,4}e^{-b\cdot 3.5\times 10^{-3}} + C_{4,4}e^{-b\cdot 548.5\times 10^{-3}} \quad (57.4)$$

RSI differentiates restricted (C1) and restricted/hindered compartments (C2), while also resulting in a more visually distinct difference between neoplastic and healthy tissue when compared to standard diffusion/ADC [24, 25].

Building on this work, the cervical cancer-specific RSI model was applied to monitoring chemoradiotherapy treatment response using PET/MRI in a pilot cohort. Images were obtained pre- (n = 7) and four weeks, three months, and six

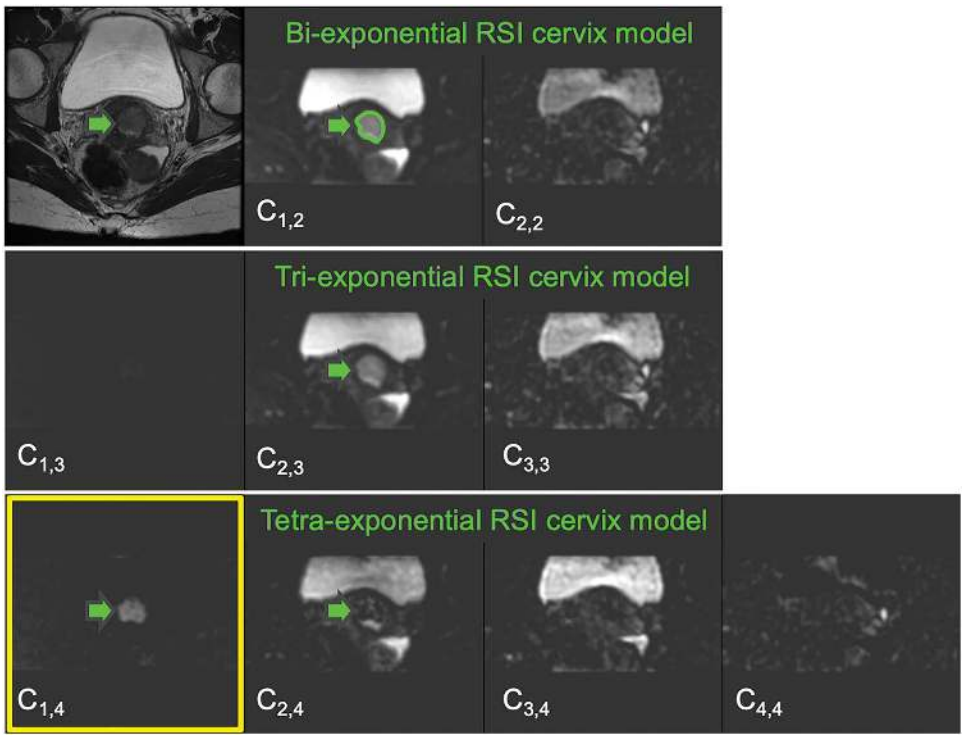


FIGURE 57.3 *Example of RSI outputs from cervix-specific bi-, tri-, and tetra-exponential models in a cancer patient.* The tumor, indicated by the green arrows and contour, is seen on C1 in the bi-exponential model, C2 in the tri-exponential model, and in both C1 and C2 in the tetra-exponential model. The tumor is most conspicuous in the C1 compartment of the tetra-exponential model (yellow box).

months post-therapy (n = 3). Two healthy volunteer counterparts underwent single pelvic MRI. Pre-therapy patient imaging demonstrated increased conspicuity of tumor versus healthy tissue in the restricted and restricted/hindered sequences. For one post-therapy patient, C1 and C1C2 RSI maps distinguished tumor burden from edematous post-treatment changes when compared to PET/CT. Another post-therapy patient exhibited complete morphologic response to treatment using C1 and C1C2 RSI maps, which was concordant with PET/CT findings of complete metabolic response. RSI enhancement was concordant with conventional PET/MRI, pre- and post-therapy [23].

The cervical cancer-specific RSI model is repeatable in the setting of patient repositioning, as demonstrated by a test-retest paradigm showing similar model outputs in the hindered and free compartments of healthy tissue (n = 10) [37].

57.4.3 Ovarian-Specific RSI model

We evaluated the diffusion properties of normal ovaries in premenopausal women across the menstrual cycle using a bi-exponential diffusion model, to explore advanced MRI model's use for detecting and characterizing ovarian cancer. Our preliminary data, which include that from six healthy women, show no statistically significant difference in the estimated ADC values across four weekly scans. We anticipated natural variation in the diffusion properties of the ovaries throughout the menstrual cycle; however, the discrepancies may be attributed to the imaging parameters and the model used [38].

To further develop an advanced MRI screening protocol for high-risk patients, we optimized a multi-shell diffusion-weighted pulse sequence with multi-slice excitation to capture high resolution in various patients including those with high-risk BRCA 1 and 2, and patients with possible ovarian cancer [39]. For MRI screening utility, the protocol has high in-plane and through-plane resolution with a total scan time within 45 minutes. Preliminary results demonstrate high specificity for discriminating between normal ovarian tissue and malignant lesions. Of the 14 participants involved in the study, three were high-risk BRCA 1 and 2 who were scheduled for prophylactic oophorectomy and the other 11 had suspicious ovarian lesions based on previous imaging. With this patient cohort, a preliminary ovarian cancer RSI model was developed, allowing separation of signal in pools of diffusion types (Figure 57.4). Our preliminary ovarian cancer RSI model and high-resolution imaging with high b-value DWI characterized diffusion properties of normal ovaries from ovarian lesions. Notably, the signal-to-noise ratio was consistent between benign and borderline lesions while up to three times greater in the pathology-confirmed malignant lesions.

57.5 Discussion

57.5.1 Breast

Overall, RSI is a non-contrast MRI technique that can be used for screening, detection, and monitoring of treatment response in women's cancer imaging. RSI can be time-effective for both

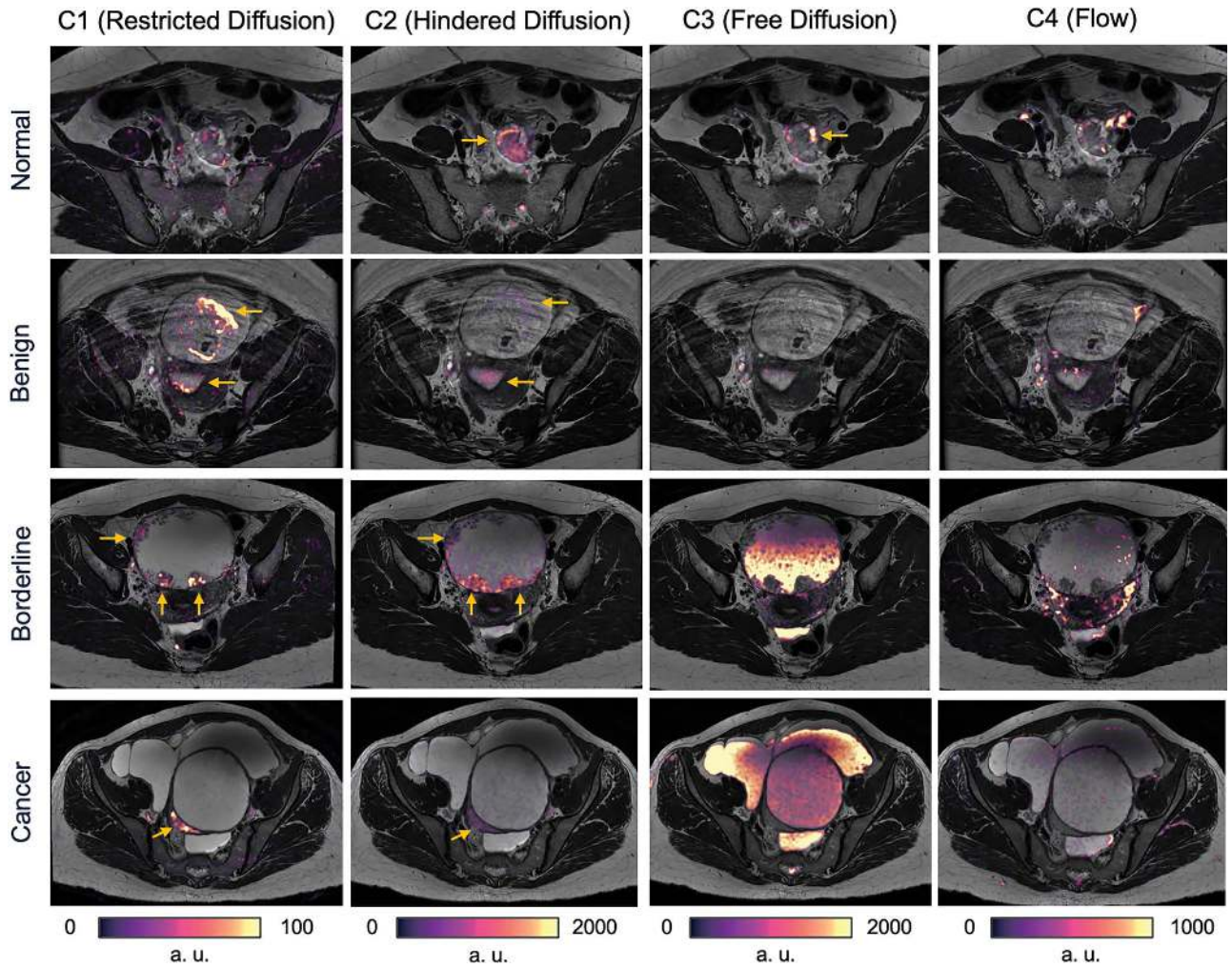


FIGURE 57.4 *Examples of RSI outputs from a tetra-exponential model in normal ovaries and benign, borderline, and cancer ovarian patients.* Normal ovaries (O-RADS 1, row 1) show minimal C1, C3, and C4 signal with moderate C2 signal. The benign mature cystic teratoma (O-RADS 2, row 2) displays high C1 in fat-containing cells, moderate C2 around the normal endometrium, and minimal C3 and C4 throughout the pelvis. The serous borderline lesion (O-RADS 4, row 3) displays high C1 and C2 in the solid components of the lesion and high C3 in the cystic component and free fluid, with scattered C4 signal. The cancer lesion (row 4), an endometrioid carcinoma of the ovary, shows high C1 and minimal but present C2 in solid tumor components, high C3 in the cystic component, and minimal C4 signal.

patients and the healthcare system. In addition, while RSI is comparable to the standard of care in detecting malignancies, it offers numerous advantages over the current standard of care, including less susceptibility to background edematous and lactation changes in the breast, as well as earlier detection of treatment response to neoadjuvant chemotherapy.

57.5.2 Cervical

The cervical cancer-specific RSI model shows promise in distinguishing malignant tissue from healthy tissue, even in the setting of post-chemoradiotherapy edematous changes. Our pilot work shows significant difference in tumor conspicuity between RSI and global ADC, as well as concordance between RSI and conventional PET/MRI for post-chemoradiotherapy; however, it is limited by the patient sample size and minimal signal of the restricted diffusion compartment in healthy patients.

57.5.3 Ovarian

While ovarian cancer detection and characterization techniques demonstrate promising potential, we aim to apply our preliminary ovarian RSI model on a larger cohort to optimize protocols, validate findings, and comprehensively assess clinical utilities. However, our current work emphasizes the potential of emerging RSI-MRI techniques to address critical challenges in ovarian cancer imaging and management.

57.6 Future Trends

57.6.1 Breast

Looking towards the future, another useful advantage of RSI is being developed for breast lesion classification that regroups lesions into the following two groups: “no biopsy required,” consisting of the control groups and average-risk benign

lesions, and "biopsy required," consisting of high-risk benign lesions and cancer. Our lab found that RSI has a statistically significant accuracy of 85% in predicting these populations [40]. In essence, RSI can be used to minimize patients' unnecessary biopsies with the hope that RSI can essentially become a "virtual biopsy" for benign and malignant breast disease.

57.6.2 Cervical

Future work with this emerging technology would include applying the model to assess response to chemoradiation treatment in a pre- and early post-therapy population when compared to conventional PET/CT, with the hope that the assessments happen sooner post-therapy and without additional radiation. Additional research directions may involve comparing PET-MRI/RSI output versus conventional PET/CT for pre-therapy staging, as well as investigating PET-MRI/RSI as a viable replacement for PET/CT.

57.6.3 Ovarian

Future directions for ovarian cancer detection and characterization require further development of the ovarian cancer RSI model, which could provide a more comprehensive characterization of the inherent diffusion heterogeneity within ovarian lesions, potentially conferring improved diagnostic accuracy. Prospective applications include improving screening accuracy, enabling earlier detection and characterization of ovarian cancers, differentiating borderline from invasive tumors, monitoring treatment response, and characterizing various ovarian pathologies.

REFERENCES

1. Le Bihan, D., Molecular diffusion, tissue microdynamics and microstructure. *NMR in Biomedicine*, 1995. 8(7–8): p. 375–386.
2. Le Bihan, D., et al., MR imaging of intravoxel incoherent motion: Application to diffusion and perfusion in neurologic disorders. *Radiology*, 1986. 161(2): p. 401–407.
3. Stejskal, E.O., and J.E. Tanner, Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 1965. 42(1): p. 288–292.
4. White, N.S., et al., Diffusion-weighted imaging in cancer: Physical foundations and applications of restriction spectrum imaging. *Cancer Research*, 2014. 74(17): p. 4638–4652.
5. Camps-Herrero, J., Diffusion-weighted imaging of the breast: Current status as an imaging biomarker and future role. *BJR|Open*, 2019. 1(1): p. 20180049.
6. Basser, P.J., Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 1995. 8(7–8): p. 333–344.
7. Basser, P.J., and C. Pierpaoli, Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance. Series B*, 1996. 111(3): p. 209–219.
8. White, N.S., et al., Probing tissue microstructure with restriction spectrum imaging: Histological and theoretical validation. *Human Brain Mapping*, 2013. 34(2): p. 327–346.
9. Rakow-Penner, R.A., et al., Novel technique for characterizing prostate cancer utilizing MRI restriction spectrum imaging: Proof of principle and initial clinical experience with extraprostatic extension. *Prostate Cancer and Prostatic Diseases*, 2015. 18(1): p. 81–85.
10. White, N.S., et al., Improved conspicuity and delineation of high-grade primary and metastatic brain tumors using "restriction spectrum imaging": Quantitative comparison with high B-value DWI and ADC. *American Journal of Neuroradiology*, 2013. 34(5): p. 958–964, S1.
11. Mendez, A.M., et al., Diffusion breast MRI: Current standard and emerging techniques. *Frontiers in Oncology*, 2022. 12: p. 844790.
12. Rodríguez-Soto, A.E., et al., Characterization of the diffusion signal of breast tissues using multi-exponential models. *Magnetic Resonance in Medicine*, 2022. 87(4): p. 1938–1951.
13. Siegel, R.L., A.N. Giaquinto, and A. Jemal, Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 2024. 74(1).
14. Lee, C.S., D.L. Monticciolo, and L. Moy, Screening guidelines update for average-risk and high-risk women. *American Journal of Roentgenology*, 2020. 214(2): p. 316–323.
15. Monticciolo, D.L., et al., Breast cancer screening for women at higher-than-average risk: Updated recommendations from the ACR. *Journal of the American College of Radiology*, 2023. 20(9): p. 902–914.
16. Chu, W., et al., Diffusion-weighted imaging in identifying breast cancer pathological response to neoadjuvant chemotherapy: A meta-analysis. *Oncotarget*, 2018. 9(6): p. 7088–7100.
17. Gao, W., N. Guo, and T. Dong, Diffusion-weighted imaging in monitoring the pathological response to neoadjuvant chemotherapy in patients with breast cancer: A meta-analysis. *World Journal of Surgical Oncology*, 2018. 16(1): p. 145.
18. Gu, Y.L., et al., Role of magnetic resonance imaging in detection of pathologic complete remission in breast cancer patients treated with neoadjuvant chemotherapy: A meta-analysis. *Clinical Breast Cancer*, 2017. 17(4): p. 245–255.
19. Andreassen, M.M.S., et al., Discrimination of breast cancer from healthy breast tissue Using a three-component diffusion-weighted MRI model. *Clinical Cancer Research*, 2021. 27(4): p. 1094–1104.
20. Loubrie, S., et al., Discrimination between Benign and Malignant lesions with restriction spectrum imaging MRI in an enriched breast cancer screening cohort. *Journal of Magnetic Resonance Imaging*, 2025. 61(4): p. 1876–1887. doi: 10.1002/jmri.29599. Epub 2024 Sep 18. PMID: 39291552; PMCID: PMC11896923.
21. Besser, A.H., et al., Tri-compartmental restriction spectrum imaging breast model distinguishes malignant lesions from Benign Lesions and healthy tissue on diffusion-weighted imaging. *Cancers (Basel)*, 2022. 14(13).
22. Yamin, G., et al., Cervical cancer staging and surveillance using restriction spectrum imaging (RSI)-MRI in comparison to PET-CT: Pilot clinical application. In *2018 Proceedings of the International Society for Magnetic Resonance in Medicine (ISMRM)*. 26(2018). 1177.
23. Lundström, E., et al. Evaluation of a cervical cancer-specific restriction spectrum imaging model on an independent pilot cohort using PET/MRI: Pre- and post-therapy. In *2024 ISMRM & ISMRT Annual Meeting & Exhibition*. 2024.

24. Rodríguez-Soto, A.E., et al. Multi-exponential model of diffusion signal with fixed ADCs in healthy and cancerous cervix tissues. In *Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2023.
25. Rodríguez-Soto, A.E., et al. Restriction spectrum imaging (RSI) models for cervical cancer. In *Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2024.
26. Finch, A.P.M., et al., Impact of oophorectomy on cancer incidence and mortality in women with a BRCA1 or BRCA2 mutation. *Journal of Clinical Oncology*, 2014. 32(15): p. 1547–1553.
27. Metcalfe, K., et al., Effect of oophorectomy on survival after breast cancer in BRCA1 and BRCA2 mutation carriers. *JAMA Oncology*, 2015. 1(3): p. 306–313.
28. Rebbeck, T.R., N.D. Kauff, and S.M. Domchek, Meta-analysis of risk reduction estimates associated with risk-reducing salpingo-oophorectomy in BRCA1 or BRCA2 mutation carriers. *Journal of the National Cancer Institute*, 2009. 101(2): p. 80–87.
29. Marchetti, C., et al., Risk-reducing salpingo-oophorectomy: A meta-analysis on impact on ovarian cancer risk and all cause mortality in BRCA 1 and BRCA 2 mutation carriers. *BMC Women's Health*, 2014. 14: p. 150.
30. Menon, U., et al., Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: Results of the prevalence screen of the UK collaborative trial of ovarian cancer screening (UKCTOCS). *The Lancet Oncology*, 2009. 10(4): p. 327–340.
31. Conlin, C.C., et al., Improved correction of B_0 inhomogene ity-induced distortions in diffusion-weighted images of the prostate. *medRxiv*, 2024: p. 2024. https://doi.org/10.1101/2024.03.26.24304935.
32. Gudbjartsson, H., and S. Patz, The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 1995. 34(6): p. 910–914.
33. Irfanoglu, M.O., et al., Evaluating corrections for Eddy-currents and other EPI distortions in diffusion MRI: Methodology and a dataset for benchmarking. *Magnetic Resonance in Medicine*, 2019. 81(4): p. 2774–2787.
34. Nissan, N., et al., Breast MRI during lactation: Effects on tumor conspicuity using dynamic contrast-enhanced (DCE) in comparison with diffusion tensor imaging (DTI) parametric maps. *European Radiology*, 2020. 30(2): p. 767–777.
35. Andreassen, M.M.S., et al., Restriction spectrum imaging with elastic image registration for automated evaluation of response to neoadjuvant therapy in breast cancer. *Frontiers in Oncology*, 2023. 13: p. 1237720.
36. Tong, M.W., et al., Longitudinal registration of T(1)-weighted breast MRI: A registration algorithm (FLIRE) and clinical application. *Magnetic Resonance Imaging*, 2024. 113: p. 110222.
37. Sawhney, J., et al. Repeatability of cervix restriction spectrum imaging outputs in healthy cervix. In *Proceedings of the Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2024.
38. Rodríguez-Soto, A.E., et al. Diffusion properties of ovaries in pre-menopausal women using a bi-exponential model. In *Proceedings of the Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2021.
39. Batasin, S.J., et al. High-resolution MRI protocol for ovarian cancer screening. In *Proceedings of the Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2024.
40. Loubrie, S., et al. Predictive model for breast lesion assessment using restriction spectrum imaging. In *Proceedings of the Annual Meeting for The International Society for Magnetic Resonance in Medicine (ISMRM)*. 2023.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com

Index

Note: Page numbers in *italics* indicate a figure and page numbers in **bold** indicate a table on the corresponding page.

3D cancer models (canine bladder cancer organoids), 508–509
  bright field microscopy, 509
  cryopreservation, 508
  culture medium supplements and growth factors, 507
  future trends, 510–511
  histological staining, 509, *509*
  immunofluorescence, 510
  passaging of, 507–508
  preservation buffer, 506, **506**
  RNA *in situ* hybridization, 509–510
  sample collection and preparation, 506
  from tissue-derived stem cells, 506–507
  from urine-derived stem cells, 506
  workflow for, 507, *507*
3D mammography, *see* digital breast tomosynthesis (DBT)

A

accuracy, *see* diagnostic visual search (DVS)
actionable alterations, 280
  in post-CAR T therapy in multiple myeloma, 290
  for targeted therapies in multiple myeloma, 285–287
acute lymphoblastic leukemia, 2
  B-cell, 297, 298, 300–301
  T-cell, 297
acute myeloid leukemia (AML), 300–301
  cancer cell differentiation, 223
  genomic subtype classification, 298
  leukemia lineage classification, 297
Afatinib-TRITC (Afa-TRITC), 67, 69, 71, *71*, 74
aging
  cancer risk factors, 2, 23, 27
  C. elegans, 184
  for CRC screening, 442
  at death, 232
  hematopoiesis, 228
  for LDCT screening, 143
  somatic mutations, 1–2
Alexa Fluor (AF) dyes, 10
alternative lengthening of telomeres (ALT), 239
  biomarkers for detection, 239–240
  biomarkers of telomere maintenance mechanisms (TMM), 239–240
  exploration of ctDNA, 244
  method for C-circle assay, *see* C-circle assay
  real-time PCR preparation, 242–243
  TMM, biomarkers of, 239
  UTF assay, 241, 243, 244
American Type Culture Collection (ATCC), 12, 44, 214, 249, 357
amplification
  cDNA, 224–225, 251
  conditions for ddPCR assays, **208**
  genomic library, **274**
  micro-channel plate (MCP), 13
  partitioned PCR droplet assays, 206, 207
animals
  in colonic epithelium, 198–199
  in near-infrared (NIR), 12–14
  in subharmonic imaging, 414, 415, 418, 421
antibodies
  anti-tumor, 351
  circulating tumor cells (CTCs), 451
  for enrichment of HCC EVs, 469
  goat, 87
  missense mutant p53 (mtp53), 262
  monoclonal, *see* monoclonal antibodies (mAbs)
  against oncogenic receptors, 11
  present in TBs, 111
  target engagement, *11*, 11–12
  TCO conjugates, 250, *471*, 473, 475
  TEX proteins, 125
  tumor-immune microenvironment (TIME), 191
AQ-FLEX CLE miniprobe, *28*, *29*
artificial intelligence (AI), 33, 428, *see also* deep learning (DL), for colorectal polyps detection; large language model (LLM), for prostate cancer
  complex imaging data interpretation, 54
  in confocal laser endomicroscopy, 31–32
  in digital pathology, 110, 114
  enhancement and conversion of FIBI images, 65
  for FLI, 17
  histopathology annotation models, 526
  in MRI and radiomics, 161
  and organoids, 500
  in personalized treatment planning, 3, 54, 101
  for risk-stratification of IPMNs, 32–33
  SAM-Adapter, 115
  for TB segmentation, 112–113, **114**, 115
  using HM-MRI, 160–161
artificial intelligence diffraction (AIDA), 388
aryl hydrocarbon nuclear translocator (ARNT), 139, *140*
aryl hydrocarbon receptor (AHR), 139–140, *140*
aryl hydrocarbon receptor repressor (AHRR) gene, 139–140, *140*
asymmetric least squares (ALS) algorithm, 184
avalanche photodetector (APD), 22, *22*, 126, *126*

B

basal cell carcinoma (BCC), 20, 23, *24*, 36, 40–41
B-cell acute lymphoblastic leukemia (B-ALL), 297
  pediatric cancers, 298, 300–301
beam splitter, 13, 20, *22*, 37
benign tumor, 4, 276, 420
  mixed, 5
binding affinity, 67, 69, 71, *72*, 74
biochip, Exo-PROS biosensor, 125, *126*, 128–129, *131*
biomarkers, 3, 358–359, 362
  for ALT activation, 239–240
  in cell-based tissue microarray and tissue section, 407–408, 410, *410*
  clinical trials, 411
  in exfoliative buccal cell samples, 407, 408, 410, *410*
  future trends, 411–412
  DNA abundance as, 358–359, 362
  DNA methylation, *see* DNA methylation biomarkers
  exosomal, 125
  immunohistochemistry, 408
  NGS and real-time PCR, 409
  multiplexing based detection, 327–328
  prognostic markers, 410–411
  protein β-sheet richness, 359–361, 362
  proteolytic, 344
  quantitative multiplex immunofluorescence, 408–409
  of Ra-223 response in mCRPC, *454*
  samples for, 3
  for screening, 407
  for telomerase detection, 239–240
  TEX, 125–127
  for thyroid nodules, 270–277
  TMM, 239–240
  total RNA from formalin-fixed paraffin-embedded tissues, 408
  tumor-derived nanoparticles, 321
biosamples, 407–412
  buccal cells, 407
  cell-based tissue microarray, 407–408, 410
  exfoliative buccal cell collection, 408
  future trends, 411–412
  immunohistochemistry, 408
  materials, 407–408
  next-generation sequencing, 409, 410
  overview, 407
  quantitative multiplex immunofluorescence, 408–409
  real-time PCR, 409, 410
  research methods, 408–409
  research results, 411

screening markers in exfoliative buccal cell samples, 410
tissue section, 407–408, 410
total RNA from formalin-fixed paraffin-embedded tissues, 408
biosensor, *see* Exo-PROS biosensor
bladder cancer, *see* 3D cancer models (canine bladder cancer organoids)
blood-based tests for CRC screening, 442–443
analysis, 443
for BC biology and treatment, 336, *337*
biological reagents, 493
data collection and recruitment, 443
FIT-DNA testing and FIT-RNA testing, 446–447
future trends, 447
as living biomarker for BC patients' response, 336, *338*
materials, 443
patient and provider perceptions, 443, 444–446
research results, 444–446, **445**–**446**
study, 443
blood tumor mutational burden (bTMB), 3
BODIPY analogs, 10
brain metastases (BM) using quantitative CEST-MRF, 163
acquisition schedule optimization, *165*, 165–166
advantages of, 167, *168*, *169*
clinical scenarios, 163, *164*
complex data analysis, 164
future research, 168
limitations, 167–168
long image acquisition times, 163–164
with MRF, 164
MRI studies, 164
neural network definition and training, 164, 166
pseudo-progression, 163
pulse sequence, 164–165, *165*
qualitative CEST contrast, 163
in vivo human subject studies, 166–167, *167*
brain tumors, challenges of drug delivery to, 213
branch duct (BD) IPMNs, 30
breast cancer
diagnosis, 131, *132*
diagnosis, Exo-PROS biosensor, *132*
biological materials, 389
chemicals materials, 389
drug delivery, *11*, 12
future trends, 394
genomic-based diagnosis, 3
image processing pipeline, 392–393
LFIA, *see* lateral flow immunoassay (LFIA)
in low- and middle-income countries, 387
metastatic, *164*
molecular subtyping of, 387–389, *388*, 394
point-of-care technologies, 387–388, 389
PROVIDS reader system, *391*, 391–392, *392*
RSI models, 532–538
subharmonic-aided pressure estimation (SHAPE), 414, 415–416, *416*, 418–421
subtyping, 387–389, *388*
TMEM doorway activity in murine, 86–87
triple negative, 264–267
tumor models, *11*
breast circulating tumor cells (bCTCs), 353
breast narrow-beam CT (NBCT), *174*, *175*, 177–178
breast CT imaging, 171–173
future trends, 178–179
image acquisition methodology, 173
radiation dose, 175–176, *176*
scatter buildup, 175, *175*
scientific premise, 171–173
spatial resolution, 176–177, *177*
use of "gantry-free" system, 178–179
broadband coherent anti-Stokes Raman scattering (BCARS) microscopy, 181–182
in biological samples, 184–186, *185*
bone metastasis of prostatic adenocarcinoma, 186, *186*
C. elegans, 182, 184, 186
cell lines, 182
corresponding spectra, 184, *185*
future trends, 187
materials, 182–183
measurements, 183
nonresonant background, 187
research methods, 183–184
research results, 184–187
Raman spectral processing and extraction, *183*, 183–184
SPADE algorithm, 183, *183*, 184

C

C. elegans, 182, 184, 186
cancer
biomarkers, *see* biomarkers
defined, 1, 4
detection and diagnosis, 2–3
development factors, 2
economic impact, 4
epigenetic changes, 2
genome-wide perspective of, 2
in low HDI countries, 1
monitoring technologies, 2–3
rates in developed countries, 1
cancer metabolism detection
future development, 81
hyperpolarization, 77
hyperpolarized NMR experiment with cells, 78–79
nanocoil, 76, *77*
nanocoil circuit, 76, 78
nanocoil fabrication, 77, *78*
quality of nanocoil, 79
quantifying metabolic flux in cells, 79, *80*
CancerSEEK, 3
canine bladder cancer organoids, *see* 3D cancer models (canine bladder cancer organoids)
carbohydrate microarrays
anti-tumor antibodies, 351
carbohydrate antigens, 351
ELISA, 352, *353*
future trends, 353–354
identification of blood group precursor SSEA-0 tumor biomarker, 352, *352*, *354*
methods, 351–352
target validation using conventional immunoassays, 352–353
carcinomas
basal cell, *see* basal cell carcinoma (BCC)
breast, 87
defined, 4
head and neck squamous cell, 67, 110
hepatocellular, *see* hepatocellular carcinoma (HCC)
invasive ductal, 524, 527
non-small cell lung, 67
oral, 327
ovarian, recurrent, 43
renal cell, 287
squamous cell, *see* squamous cell carcinoma (SCC)
urothelial, 28, 338, 508
C-circle assay
equipment for, 241
interpretation of, 243
isothermal PCR preparation, 242, **242**
method for, 241–242
reagents for, 241
supplies for, 241
telomeric DNA, 241
use of, 243–244
CellDive platform, 196
cell phenotyping, 192, *193*, 195
cell-free RNA (cfRNA)
future trends, 316
identification of biomarkers, 313–315, *314*, *315*
isolation and quantification, 310–311, 312, *313*
qPCR, identification of housekeeping and gene signals, 312, *314*
sample collection, 310
sequencing and transcriptomic analysis, 311–313
study design, 310
target region selection and qRT-PCR, 311
workflow for, 310, *311*, 315
Centers for Medicare and Medicaid Services, 37
central nervous system tumors (CNSTs), 213
cervical cancer
HPV assay for, 481–483, 486
in LMICs, 480, 481
prevalence of, 483
risk for women living with HIV, 480
RSI models, 533–538
screening, 480, *481*
Cg05575921 methylation
in Copenhagen City Heart Study, 140–141
in National Lung Cancer Screening Trial, **141**, 141–142
smoking intensity, 140
charged coupled device (CCD) cameras, 45
chemical exchange saturation transfer MRF (CEST-MRF), 163, *164*
acquisition schedule optimization, 165–166
advantages, 167
box and whiskers plots for, *168*
clinical translation, 163–164

complex data analysis, 164
derived water relaxation, *169*
effective in cancer imaging, 163
future trends, 168
healthy subjects, 166
limitations, 167–168
long image acquisition times, 163–164
materials, 164
MRI studies, 164
neural network training, 164, 166
overview, 163
patients with brain metastases, 166–167, *167*
pulse sequence, 164–165, *165*
qualitative imaging, 163, 164
research methods, 164–165
in vivo human subject studies, 166, *166–167*
chemical exchange saturation transfer MRI (CEST MRI)
cell culture, 214
challenges of delivering drugs to brain tumors, 213
chemicals and reagents, 214
for image-guided drug delivery and response assessment, 216, *216*
imaging DNA alkylating agents with, 213–214, *214*
patient stratification and decision-making, 218, *218*
pH-dependent signal of melphalan, 215, 216, *216–217*
chemical imaging (CI) techniques, 459
colon tissue segmentation, 461, *463*
deepfake model application on colon tissue, 464, *464*
future trends, 465–466
materials, 460
method, 460–461
results, 461–464
virtual staining, 461, *463*
chemophototherapy (CPT), 43, 44
circulating tumor cells (CTCs)
breast, 353
defined, 4
interactions, 4
liquid biopsy-based tests, 247, 310, 338, 497
circulating tumor cells (CTC), prostate cancer
antibodies, 451
buffer, 451
cancer cell lines and human blood samples, 451
chip design and microfabrication, 452
CTC-iChip, 449, *450*, 451, 452, *453*
DNA methylation analyses, 452, 453–456
droplet digital PCR of bulk CTC mixture, 452
future trends, 456
immunofluorescence imaging, 452, 453
isolation yield, 452–453
leukopak processing, 451, 452
^{LP}CTC-iChip, 451, 452
magnetic beads, 451
RNA-based digital CTC signatures, 453, *454*
single-cell RNA-sequencing, 452, 453, *455*
circulating tumor DNA (ctDNA)
automated DNA extraction, 207–208
based cancer diagnosis, 3
blood sample collection, 207
control materials, 207
ddPCR, *see* digital droplet PCR (ddPCR)
DNA quantification and QC, 208
droplet, 208
exploration, 244
genomic characterization of cancer, 210
integrity, 208–209
laboratory, 207
next-generation sequencing, 210
PCR, 208
plasma isolation, 207
in positive mutant calls, *211*
pre-extraction preparation, 207
research methods, 207–208
research results, 208–210
Tapestation analysis, 208, *209*
Clinical Interpretation of Variants in Cancer (CIViC), 280
co-culture models, 510–511
coherent anti-Stokes Raman scattering (CARS), 181, *see also* broadband coherent anti-Stokes Raman scattering (BCARS) microscopy
coherent Raman imaging (CRI) techniques, 181
colonoscopy, 67
adenoma detection by, 198, 199
for CRC screening, 442–447, 491, 497, 499
OCT data, 202
optical, 433
screening, 198, 203
virtual, 149–151
colon segmentation, *see* deep learning (DL), for colorectal polyps detection
colorectal cancer (CRC)
aging, 2
blood-based tests for, *see* blood-based tests for CRC screening
cell lines and tumors, 514
CLE application in, 28
CTC for polyp detection, *see* CT colonography (CTC), for colorectal polyp detection
CZE-MS/MS-based strategies, 516–517, *517*
deep learning for polyp detection, *see* deep learning (DL), for colorectal polyps detection
discovery of proteoforms, 517–518, *518*
FGS, 103
interval, 491
KRAS4b proteoforms, 514
metastatic and non-metastatic human, 519–521, *520*
mortality rates, 110, 198, 491
OCT measurements from mouse models, 200, 203
optical colonoscopy, 433
organoid models for, *see* organoid models for CRC early detection
risk of, 198
samples from, 312, 313, 315
TB from H&E images in, *see* tumor budding (TB), from H&E images in CRC
tumor EVs in, 247
colorectal polyps, *see* deep learning (DL), for colorectal polyps detection
computed tomography (CT) scan, 9, 27, 43, 524, 526, *see also* breast narrow-beam CT (NBCT); CT colonography (CTC), for colorectal polyp detection
for lung cancer, 92, 125
x-ray, 171
conditional reprogramming (CR) technology, for urological cancer
for BC patients' response, 336, *338*
challenges in clinical setting, 340
clinical background, 331
co-culture and separation of feeders, 335
comparison of technology and approaches, 332, **333**
complete F medium, 334
conditionally reprogrammed cells (CRCs), 331–332
cryopreservation medium, 334
cryopreservative tissue, 334
differential trypsinization, 335
equipment, 334, **334**
establishment of first patient-derived NEPC cell line, 336, *339*
feeder cells, 335–336
feeder layer tissue culture and treatment, 335
fresh tissue handling, transportation, and shipping, 334
as functional platform for BC biology and treatment, 336, *337*
future trends, 340
isolate cells from urine samples, 336
isolation of primary cells, 334–335
liquid biopsies, 337–338
prostate patient-derived cells, 340
reagents, 332–334, **333**
tissue biopsies, 336–337
confocal microscopy *see* confocal laser endomicroscopy (CLE), pancreatic cystic lesions; EUS-guided confocal laser endomicroscopy (EUS-CLE); portable confocal microscopy, (PCM) skin cancer; reflectance confocal microscopy (RCM), skin cancer
confocal laser endomicroscopy (CLE), pancreatic cystic lesions, 27–30
and artificial intelligence integration, 31
for differentiation of pancreatic cysts, 31
endoscopic ultrasound with, *31*
functional artificial intelligence algorithm for, 32–33
management of pancreatic cysts, 29–30
for risk stratification of IPMNs, 30–32
continuous-wave (CW)
acquisition, 12
data, 12, 14, 16, 46
convolutional neural networks (CNN), *see also* artificial intelligence (AI)
colon segmentation, 436–439
eye tracking, 93
polyp detection in CT, 147
tumor budding detection, 110, 147
Copenhagen City Heart Study, 138

CosmoDA, 3
CosMx platform, 196
CTC-iChip
  LPCTC-iChip, 451
  in prostate cancer, 449, *450*, 451
  whole blood processing with, 452
CT colonography (CTC), for colorectal polyp detection
  automatic polyp detection in, 147, 148, **149**
  for Classifier Stage, 152–153
  for CT-based detector stage, 151–152, **152**
  dataset, 149
  early diagnosis, 147
  hardware, 149
  location and, 147
  methods, 149
  multi-view fusion network, 151, *151*
  polyp detection approach, 149, *150*
  shape-based algorithms, 147, *148*
  software and packages, 149
  for VC-based detector stage, 152, **152**
  VC-based polyp detection approach, 150–151
Cultrex BME, 13
Cy5p53Tet in mice, *see also* mutant p53-based diagnostics and therapeutics
  cellular uptake, live cell imaging, 262
  IVIS spectrum imaging of, 263
  penetration into TNBC cells, *264*, 264–265
  penetration into TNBC PDTOs, 265, *265*
  stapled peptides SPCy5p53Tet, 267
  in TNBC mtp53 xenograft tumors, 265–267, *266*
cystic neuroendocrine neoplasm (cystic-NET), 27
CZE-MS/MS studies
  analysis, 515–516
  chemicals and solutions, 514
  in CRC-related pathways, 517–518
  data analysis, 516
  future trends, 521
  instrumentation, 514
  LPA-coated capillary, 515–516
  proteins, 515
  proteoform identifications from CRC cells, 516–517, *518*
  proteoform quantification, 516
  with PTMs and SAAVs, 518–519
  quantitative TDP of metastatic and non-metastatic human CRC cell lines, 519, *520*, 521
  RPLC fractionation, 515
  SEC fractionation, 515
  setup of CE-MS System, 516
  study design, 515, *517*
  TP53 proteoform, 518–519, *519*

D

DAPHNI platform, 280, *281*
decision errors, 93, 94
deep learning (DL), for colorectal polyps detection
  CNN-based segmentation approach, 436–439
  colon segmentation, 440, **440**
  colon segmentation, classical approach, 436, *437*
  for colorectal polyps detection, 433–434
  CTC platform, 433, *434*
  dataset, 436
  encoder and decoder, 436, *438*, 439–440
  future trends, 440
  hardware, 436
  for image analysis, 54
  MRF-based segmentation approach, 436
  segmentation challenges, 433, *435*
  software and packages, 436
  volume segmentations for nnUNet, Base Model, 440, **440**
deep reconstruction network (DRONE), 164
dendritic cells (DCs), 4
dermoscopy, 23, 24, 25, 36, 40
diagnostic visual search (DVS), 92–93
  eye-tracking data (ETD), 93–94, 96, 99, 521
  machine learning to detect abnormalities, 93–94
  participants, 94
  in radiology, 92
  stimuli and apparatus, 94
dielectrophoresis (DEP), 321–322
  cancer-derived nanoparticles, *322*, 326–327
  electrode array design, 322–323
  of EVs, 326–327
  fluorescence immunostaining, 324, *325*
  force, 322–323
  future trends, 329
  materials, 323–324
  methods, 324–326
  multiplexing biomarker detection, 327–328
  nanoparticle-based diagnostics, 328–329, *329*
  process of conducting cyclic immunofluorescence, 327, *328*
  synthetic nanoparticles collection, *324*, 326
  theory and applications, 321–322
  tumor-derived nanoparticles as cancer biomarkers, 321
  using reporter systems, 327
differential abundance (DA) test, 3
digital breast tomosynthesis (DBT), 171
digital chemical pathology, 459–461
digital droplet PCR (ddPCR), *see also* circulating tumor DNA (ctDNA)
  amplification conditions for, 208, **208**
  ctDNA extraction, 208, 210
  droplet clusters, 209–210, *210*
  prostate CTCs, 453, *454*, 456
  quality control, 209
  reagents, 206–207
  reverse-transcription, 469, 470, *471*, 472, 475
  using Bio-Rad QX200 platform, 206, 452
DNA
  alkylating agents, 213–214, *214*, 216
  automated extraction, 207–208
  extraction, 206–208, 358, 362
  quantification and QC, 208
DNA methylation, 3
  CTC in prostate cancer, 452, 453–456
  LDCT screening, 139, 142–143
  of lung cancer, 142–143
  smoking intensity assessment, 139
  "turning off" genes, 2
DNA methylation biomarkers, thyroid cancer diagnostics, 272
  in benign thyroid nodules, 274–275, *275*
  bisulfite conversion, 273
  data analysis, 274
  DDMS diagnostics, 275–277
  end repair and A-tailing, 273
  gel purification, 274
  genomic library amplification, 274
  for indeterminate thyroid FNA, **271**, 271–272
  library preparation and NGS sequencing, 272
  linker ligation, 273
  molecular diagnostic tests, 276
  MspI digestion, 272–273
  NGS sequencing, 274
  preoperative diagnosis, 270, 276
  samples, 272
doxorubicin (Dox)
  encapsulated in PoP-liposomes, 44
  excitation and emission spectra, 48
  fluorescence, 44
  liposomal, 43
  release of, 45, *52*, 52–53
drug binding, 9, 16, 67, 69, 72
  affinity, 71, *72*, 74
drug delivery, 9, 15
  to brain tumors, 213
  in breast cancer tumor models, *11*, 12
  of DNA alkylating agent, 216, *216*
  imaging-guided, 43, *44*, 53–54, 216, 218
drug penetration, 9, 16, 264, 265, 510
dynamic nuclear polarization (DNP), 76

E

elliptical dichroism (ED)
  spectroscopy for detecting β-sheet richness, *360*, *361*, 361–362
  stereochemical analysis, 358
endoscopic spatial frequency domain imaging (eSFDI)
  dual-channel strategy, 50, *50*
  for imaging and drug release setup, 50–51
  instrumentation, 50
  for ovarian cancer, 53–54
  to quantify Dox concentration, 53
endoscopic ultrasound (EUS)
  with fine-needle aspiration (EUS-FNA), 28, *28*, 31
  guided confocal laser endomicroscopy (EUS-CLE), 28, *28*, 29
  needle-based CLE (EUS-nCLE), *29*, 29–33, *30*, *32*
  for visualizing pancreatobiliary system, 27–28
enzyme-linked immunosorbent assay (ELISA), 125, 132–133
  carbohydrate-specific, 352, *353*
  inhibition assays, 353, *353*
  mass spectrometry version of, 373
epidermal growth factor receptor (EGFR)
  cetuximab (CTM), 12
  Erlotinib, 67
  HER-2, 11, 505
  phosphoactivated, 408, 410
  phospho-EGFR, 70

protein binding affinity and cytotoxicity, 71
protein with fluorophore-TKI conjugates, 71–72, *72*, *73*
receptor quantification, 69
TEX, 126, 129–131, *131*
tyrosine kinase inhibitors (TKIs), 67, 68–71, *70*, 74
epigenetics, 1–2, 276, 305
assessments of cg05575921, 143
changes, 2, 491
negative modifications, 2
proteomics and, 344
of smoking, 139
epithelial-mesenchymal transition-positive (EMT-positive), 4
Erlotinib-tetramethylrhodamine isothiocyanate (Erl-TRITC), 67–69, *70–73*, 71, 74
EUS-guided confocal laser endomicroscopy (EUS-CLE)
differentiating IPMNs, *32*
equipment, *29*, 29–30, *30*
needle-based CLE (EUS-nCLE), *29*, 29–33, *30*, *32*
risk stratification, 32
EV Click Beads
click chemistry mediated enrichment, 249–250, 252, *253*, 256, 257
$DSPE\text{-}PEG_{1000}$-TCO preparation of, 250
Ewing Sarcoma EVs or PDAC EVs, enrichment of, 250
preparation of, 248, 249–250
EV Click Chips
blood sample processing, 252
click chemistry-mediated enrichment, 250, 256, 257
control cohorts, 251–252
enrichment of HCC EVs or PCa EVs, 250
Ewing sarcoma patients, 251
extracting RNA from, 250–251
fabrication of, 250
HCC and liver cirrhosis patients, 251
PCa patients, 251
PDAC patients, 251
PDMS Chaotic Mixers[48], fabrication of, 250
RT-digital PCR, 251
RT-droplet digital PCR, 251
statistical analysis, 252
TCO-antibody conjugates, preparation of, 250
Tz-grafted SiNWS, fabrication of, 250
EV isolation through immunoprecipitation (EVIP)
qPCR-based assay of EV DNA (EVIPqPCR), 358, 359, *360*, 363
ThT-based assay of EV proteins (EVIPThT), 358, *360*, 361–363
excisional biopsy, 36, 388
exclusive feature bundling (EFB), 95
Exo-PROS biosensor, 125–127
biomolecules, 125
detection of TEX protein-microRNA pairs, 129
exosomal biomarkers, 125
fluorescence signals, 126
instrumentation, 127
isolation from cell culture media, 128
isolation from human serum samples, 128–129
materials and supplies, 127–128
optical components of, 127
overview of, 125–126, *126*
preparation of, 128
SPR and SPEFS, 125
surface modification of, 128
TEX biomarkers, 125–127
using cell-derived exosomes, 129–130, *130*
using human serum samples, 130–131, *131*, *132*
extracellular matrix (ECM) interactions, 3, 4
extracellular vesicles (EVs)
cell lines, 357
chemicals, 357
DNA abundance as cancer biomarker, 358–359, *359*, 362
ED spectroscopy for detecting β-sheet richness, *360*, *361*, 361–362
equipment, 357
future trends, 362–363
materials, 357
mitochondrial DNA abundance, 359, 362
nucleic-acid-to-protein-ratio (NPr), 358
protein/DNA/RNA extraction, 358
protein β-sheet richness as cancer biomarker, 359–361, 362
statistics, 358
ThT-based assay of EVIPThT, 358
eye-tracking data (ETD), 93, 94, 96, 99, 521, *see also* diagnostic visual search (DVS)

F

Family Health History Tool (FHHT)
clinic-based and individual-initiated models, 235, *235*
community oncology practices, 232–233
completion rate of, *236*
dashboard and patient view, 233, *233–234*
deployment of, 232–235
development and technical implementation of, 231–232
dissemination and implementation of, 236–238
germline genetic testing, 237, *237*
MiGHT study, 232
fecal immunochemical testing (FIT)
for CRC screening, 198, 442–444, **445–446**
DNA testing, 442, 446–447, 497
RNA testing, 446
fibro-adenoma, 4
fine-needle aspiration (FNA)
biopsy, 344–345
endoscopic ultrasound (EUS-FNA) with, 28, *28*, 31
HER2 extraction, 388
mock protocol, 345, 348, *348*
of thyroid nodules, 270, 271–272, 276
flow cytometry, 81, 125, *298*, 326
of bone marrow tumor microenvironment (TME), 280
cell-surface staining, 353
of MCF7 and MDA-MB-468, *264*, 265
spectral, 290, 292
fluorescence imaging
macroscopic, 12
NIR, 10, 265
quantitative, *see* quantitative fluorescence imaging (qFI)
fluorescence imitating brightfield imaging (FIBI)
advantages of, 60
automated image processing and analysis, 65
challenges, 63
deployment to low-resource countries, 63
functional assembly, *59*, 60
future iterations of, 64
hardware, 64
histology, 61, *61–62*
limitation of, 106–107
optical design and components, 60
rapid biopsy evaluation, 63
research, 63
software and scanning, 60
specimen preparation, 60
surgical guidance and intraoperative margin detection, 62–63
tissue preparation, 64–65
tradeoffs of, *64*
validation study, 61
versatility and accessibility, 61–62
fluorescence lifetime imaging (FLI/FLIM)
autofluorescence molecular imaging, 106
near-infrared (NIR), 9–10
fluorescence molecular tomography (FMT), 10, 12
fluorescence reflectance imaging, 10
fluorescent-guided surgery (FGS), 103, 106
fluorescently labeled EGFR TKIs
binding affinity characterization, 69
cell culture, 67–68
chemicals, 67
design and synthesis of, *70*, 70–71
EGFR protein, live-cell imaging of, 71–72, *73*
live-cell imaging studies, 69–70
MTS assay inhibition, 69
phospho-EGFR imaging studies, 70
photophysical properties of, 71, *71*
photophysical property characterization, 69
protein binding affinity and cytotoxicity, 71, *72*
receptor quantification, 69
structure confirmation and characterization, 68
synthesis of, 68–69
Förster resonance energy transfer (FRET) imaging
measurements, 10
MFLI, 12
NIR FLI, *see* NIR FLI FRET imaging
FoundationOne®LiquidCDx, 3
full field digital mammography (FFDM), 171

G

GALAD score and HCC EVs
antibodies for enrichment, 469
collection of patient plasma samples, 472
digital scoring assay, *see* HCC EV digital scoring assay

early detection in US Populations, 473, *474*, 475
measurement of, 469
methods, 472
phase 2 biomarker studies, 472
preparation of artificial plasma samples, 471–472
statistical analysis, 473
surface protein assay, *see* HCC EV surface protein assay
galvanomirror (GM), 13
gene expression analysis, 276, 280, 290
generative adversarial network (GAN) framework, 460–461, 464
genetic mutation, 2, 308
genetic susceptibility, 231–232
community oncology practices, 232–233
digital tools, 231
FHHT, *see* Family Health History Tool (FHHT)
MiGHT study, 232
genome-wide perspective of cancer, 2
genomic-based cancer diagnosis, 3
genomics, 3, 298, 387
glycopeptides preparation
application, 377–378
Byologic-quantification of, 375
Byonic-assignment of, 374–375
data analysis, 377
double enzymatic digestion, 376
filter processes, 376
HILIC enrichment, 376
of Hp N-glycopeptide markers with AFP, 380
LC-Stepped HCD-DDA-MS/MS, 377
LC-Stepped-HCD-PRM-MS/MS, 377, 378
removal of serum high-abundance proteins, 375–376
site-specific glycopeptides, 372–373, 378–380
GPU-based fast Monte Carlo model, 53–54
gradient-based one-side sampling (GOSS), 95
GSK 2849330, anti-^{89}Zr-labeled HER-3 monoclonal antibody, 9

H

HCC EV digital scoring assay, *471*
EV Click Chip, 470, 472, 475
quantification of 10 HCC-specific mRNA markers, 472, *476*
RT-ddPCR, 470, 475
HCC EV surface protein assay, *471*
Click Bead, 470, 475–477
click chemistry-mediated capture of subpopulations, 473
phase 2 biomarker studies for clinical validation of, 473
real-time immuno-PCR, 470, 473, 475–477, *476*, *478*
sequentially labeling each subpopulation in plasma, 472
head and neck squamous cell carcinoma (HNSCC), 67, 110
hematoxylin and eosin (H&E), 58
organoid models for, 493
slides, 110
staining, 14
TB from H&E images, 113–115
hepatocellular carcinoma (HCC)
and cirrhosis, 379, *379*, **380**, 382
early detection, 372, 380, 469
EV digital scoring assay, *see* HCC EV digital scoring assay
EVs and GALAD score, *see* GALAD score and HCC EVs
EV surface protein assay, *see* HCC EV surface protein assay
MASLD, 373–374
NASH, 374, **374**, 378
high-grade dysplasia/invasive adenocarcinoma (HGD-Ca), 30
Human Development Index (HDI) countries, 1
human epidermal growth factor receptor 2 (HER2) receptor, *see also* breast cancer
anti-HER2 monoclonal antibody, 10
fine-needle aspiration (FNA), 388
HER2/Neu signaling, 82
lateral flow immunoassay (LFIA), 391, *393*, 393–394
macroscopic fluorescence lifetime (MFLI), *15*
NIR FLI FRET imaging, *11*
drug delivery in breast cancer tumor models, *11*, 12
human papillomavirus (HPV) assay
clinical study in Cape Town, South Africa, 483
materials, 482–483
operational issues, 482
results, 485–486
screen-and-treat, 480–481
statement of problem, 480
statistical methods, 483–485
Xpert HPV, 481–482
hybrid multi-dimensional MRI (HM-MRI), of prostate cancer
analysis, 156
biopsy equipment, 156
data, 155
feasibility study, 157, *158*
future trends, 160–161
histology, 155–157
limitations, 160
MRI, 155–156
observer studies, 157, 159
patient recruitment, 156
prospective clinical trial, 157, 159, *159*
prostate tissue composition, 159–160
sampling of, 155, *156*, 160
software, 156
validation against quantitative histology, 157, *158*
hydrophilic interaction liquid chromatography (HILIC), 376–377
hyperpolarized micromagnetic resonance spectroscopy (HMRS) platform, 76, *77*, *80*

I

image foundational model, 110–112
AI model for tumor buds (TBs) segmentation, 112–113, **114**
implementation, 113
mask refinement, 113
materials, 111–112
overview, 110–111
research methods, 112–113
research results, 113–114, **114**
segmentation anything model (SAM), 112, *112*
task-specific prompt, 113
image-guided drug delivery, CEST MRI, 216–217
imaging technology, advancements in, 29, 53
immunofluorescence (IF), 65, 72, 493
cyclic, 327, *328*
imaging of CTCs, 452, 453, 456
multiplexed, *see* multiplexed immunofluorescence (MxIF) spatial imaging
organoid, 493, 494, 510
immunohistochemical (IHC) approach, 11, 407, 408, 410, 464
immunohistochemistry (IHC) validation, 12–15, 388, 408
immunosuppression, in TCCN, 193–194, *194*
indocyanine green (ICG), 10, 103–106
infrared (IR) imaging, 459, 461, 464–466
instrument response function (IRF), 13, 14
intensified charged coupled device (ICCD), 13, 17
interobserver agreement (IOA), 29–32
interobserver reliability (IOR), 29–30
interstitial fluid pressure (IFP)
estimation in rats with breast cancer, 415–416, 418–420
estimation in swine with melanoma, 415, 418, **418**, *418*
pilot study, 419
prospective clinical study, 420
study results, 420, **420**
subharmonic amplitudes, 419, *419*
in vitro exploration, 417–418
interval colorectal cancer (iCRC), 491, 499
intraductal oncocytic papillary neoplasms (IOPN), 32
intraductal papillary mucinous neoplasms (IPMNs), 27, 29–33
intravital mesoscopic fluorescence molecular tomography (IMFMT), 12
invasive biochemical approach, 11
invasive ductal carcinomas, 524, 527, *535*
inverse spectroscopic OCT (ISOCT), 200
IRDye 800CW, 10, 11

L

large language model (LLM), for prostate cancer
clinical guideline compliance, 401
clinical information marking, 398–399
clinical retrieval performance, 402, **402**
cohort characteristics for, 398, **398**
compliance detection, 403–404, **404**
domain-specific, *399*, 399–400
domain vocabulary coverage, 401
future trends, 405–406
general purpose language understanding, 399–400
in-domain clinical applications, 400–401
masked clinical information prediction, 400, 401–402

materials, 398
question-answering, 400, **402**, 402–403, **403**, *405*
text pre-processing, 398
tokenizers, 399, 401, *402*
treatment prediction cohort, 400–401, **401**, 403, **403**
laser ablation, 25
lateral flow immunoassay (LFIA)
assembly and preparation, 389–390
for breast cancer subtyping, 388–389
fluorometric multiplex design, 389, *390*
singleplex design, 391, *393*, 393–394
light emitting diode (LED), 37, 40, 61, 63
array microscope, 119
near-infrared, 38–39
wavelengths of, 48, 50, 58
light gradient boosting machine classifier (LightGBM), 95–96, 99
linear polyacrylamide (LPA)-coated capillary for CZE, 515–516
liquid biopsy
advancement, 326
biomarkers in, 125
cell-based, 456
cell-free RNA (cfRNA) from, *see* cell-free RNA (cfRNA)
circulating tumor cells (CTCs), 4, 247, 310, 338, 497
clinical diagnostics, 328, 329
CR technology, for urological cancer, 337–338
CtDNA, 3, 247
sampling and realtime monitoring, 310
liquid chromatography-mass spectrometry (LC-MS), 125, 306, 345, 377, 521
liquid tumor, defined, 4
long-circulating Dox in PoP liposomes (LC-Dox-PoP), 44
low coherence interferometry (LCI), 20, *21*
low dose computerized tomography (LDCT) screening
biology of AHR pathway, 139–140
biomarker guided smoking cessation, 143
Cg05575921 methylation, 140–142
conflict of interest, 144
costs of, 138
drawbacks of, 137–138
estimating smoking intensity, 139
exposure to radiation, 137
integration with genetic information, 143
lung cancer screening strategies, 137
medical complications from biopsy, 138
psychological toll of, 138
selection of nodules for biopsy, 143
using DNA methylation, 139
low-resource countries (LRCs), 63
lung cancer
aging, 2
CR technology, 332
CT screening, 92, 125
detection of, 100–101
diagnosed, 92
DNA methylation of, 142–143
Exo-PROS assay, 129–131, 133
LDCT screening, 137
missed lung nodules, 96, *97*, *98*, 99, **99**
mortality rates, 92, 100
non-small cell, 126, 259
$PLCOM_{2012}$ lung cancer prediction model, 138, 141
risk for, 140–143
suicide rate of, 138

M

machine learning (ML)
advent of, 459
BCARS signal processing, 187
data pre-processing, 96
of DVS, 93–94, 100–101
ensemble, 94–95
features of model, 95–96
innovative applications of, 428
LightGBM, 95, 96
optimized hyperparameters, 96
macrophages, 4, 84, 85, 190, 192–195
macroscopic fluorescence lifetime (MFLI)
FRET imaging, *see* MFLI FRET imaging
of HER2 tumor xenografts, 15, *15*
imager setup, 13
NIR, 10
and TD-MFMT measurements, 16, 17
magnetic resonance fingerprinting (MRF), 163, *164*, 165–167, *167*, *169*
magnetic resonance imaging (MRI), 9, 27, 43, 79, 83
AI integration with, 161
breast, 171
CEST, *see* chemical exchange saturation transfer MRI (CEST MRI)
diffusion-weighted, 532
dynamic contrast-enhanced, 533, 535
gadolinium-based, 89
hybrid multi-dimensional, *see* hybrid multi-dimensional MRI (HM-MRI), of prostate cancer
RSI, 535, 537
TMEM doorway activity, 89
mammography, 67, 93
for breast cancer, 125
contrast-enhanced, 179
digital breast tomosynthesis (DBT), 171
full field digital, 171
screening, 171, 177
XRD-enhanced, 525
Markov random field (MRF) models, 433, 436, 439, 440, **440**
melanoma, *see also* skin cancer
defined, 2
IFP estimation in swine with, 415, 418, **418**, *418*
quantitative phase imaging (QPI), 120
TIME of, *see* tumor-immune microenvironment (TIME), of melanoma
melphalan therapy, *214*, 214–218, *216*, *217*
mesoscopic fluorescence molecular tomography (MFMT), 11
of AU565 tumor, 15, *16*
intravital, 12
time domain, 12–17, *16*
metastasis
BCA, 351
CRC, 518–519, 521
defined, 4
in distant organs, 82, 87, 110, 196
lymph node, 110
middle frontal gyrus, 167, *169*
of prostatic adenocarcinoma, 186
skeletal, 405
tumor microenvironment of, *see* tumor microenvironment of metastasis (TMEM) doorways
metastasis-initiating cells (MICs), 4
MFLI FRET imaging, 12
and data analysis, 14
drug-target engagement in human breast cancer, *11*, 11–12
of EGFR-positive tumors in live mice, 12
intracellular delivery efficacy of EGFR CTM, 12
measurements, 12
NIR, 10, 12
micro-channel plate (MCP), 13
micro-electro-mechanical systems (MEMS), 21–23, *199*
microsatellite instability-high (MSI-H), 3
Mie scattering, 45
minimally invasive biopsy (needle aspirations), 9
mitochondrial alteration enrichment from single-cell transcriptomes to establish relatedness (MAESTER), 223–228, *226*, *227*
mitochondrial DNA (mtDNA) barcoding
analysis, 225
cDNA amplification and preparation, 224–225
MAESTER, 223–228, *227*
PCR product quantification and quality assessment, 223
PCR1 reaction and purification, 225, **225**
PCR2 reaction and purification, 225
PCR reactions, 223
primer mix volumes for PCR1, 224, **224**
sequencing and analysis, 223, 225
single-cell sequencing tools, 223
starting cDNA and PCR primers, 223
variant acquisition and propagation, 223, *224*
mixed benign tumor (pleomorphic adenoma), 5
mixed cancers (mixed malignant tumors), 5
mixed germ cell tumor, 5
mixed tumor, 5
modulation transfer function (MTF), 176–177, *177*
molecular imaging, 3, 106, 213, 268
molecular markers, 3, 28, 411
monoclonal antibodies (mAbs), 33, 351
anti-CD68 mouse, 83, 84
anti-melphalan mouse, 215
phospho-EGFR, 70
trastuzumab (TZM), 10, 13, 15
mortality-to-incidence ratio (MIR), 1
mouse model
of breast cancer, 84, 87
of colon polyps, 200
genetically engineered, 497
of ovarian cancer, 44–45
$UVRAG^{FS}$, 491, 499
xenograft, 104, 261
mucinous cystic neoplasms (MCNs), 27, 29–31
mucinous PCLs, 27, 31

multi-cancer early detection (MCED) Galleri, 3
multi-DNA stool testing, 198
multiple instance learning (MIL), 111
multiple myeloma (MM)
  CAR T cell and bispecific antibody therapies, *289*
  clonality modeling in, *291*
  defined, 280
  newly diagnosed, 290
  PSN classification, 285
  RADAR for, *see* RADAR for multiple myeloma (MM)
  risk stratification in, 285
multiplexed immunofluorescence (MxIF) spatial imaging
  complexity of, 190–191
  data acquisition and analysis workflow for, 190, *191*
  metastatic melanoma, 191
  for resolving patterns of paracrine interactions, 195
multi-view fusion network (MVFN), 149, *151*, 151–153
murine breast cancer
  image analysis, 87
  TMEM doorway scoring, 86–87
  tumor and sample preparation, 86–87
mutant p53-based diagnostics and therapeutics
  antibodies, 262
  cell culture, 262
  Cy5p53Tet (Patent Pending), 262, 264–267
  ethics, 264
  IVIS spectrum imaging of Cy5p53Tet in mice, 263
  organoids, 263
  peptides, 262–263
  protein extraction, 263

N

nanocoil, 76, *77*, *see also* cancer metabolism detection; hyperpolarized micromagnetic resonance spectroscopy (HMRS) platform
  circuit, 76, 78
  fabrication, 77, 78
  quality of, 79
NanoDrop Lite spectrophotometer, 12–13, 14
National Lung Screening Trial (NLST), 138, 141–142
near-infrared (NIR)
  FLI, 9–12, 15, 17
  FLI/FLIM, 9–10
  FLI FRET imaging, *see* NIR FLI FRET imaging
  fluorescent probes in clinical studies, 104, **105**, *106*, 107
  light, 43, 45
  MFLI, 10
needle-based CLE (nCLE), 29, *29*, 29–33, *30*, *32*
neuroblastomas, defined, 4
neutral density (ND) filter, 13
Next-Generation Cellvizio®, 29
next-generation sequencing (NGS), 125, 210
  cyst fluid, 28
  marker studies, 409–411
  for preoperative thyroid cancer diagnostics, 272
NIR FLI, 9–12, 15, 17
NIR FLI FRET imaging
  AF 700 and AF 750, 10, 11
  antibody-based targeted therapy, 11
  drug-target engagement, 10
  HER2 targeted drug delivery, *11*
  indocyanine green (ICG), 10
  IRDye QC-1 (QC-1), 11
  measurements, 10
  occurrence, 10
nonlinear least square (NLSF) fitting algorithm, 14
  Levenberg-Marquardt, 14
non-malignant cells, 4, 193, 239
non-melanoma skin cancers (NMSCs), 20, 21, 36
non-small cell lung carcinoma (NSCLC), 67, 126, *126*, 129–130, *131*
nPLEX sensor, 125
nuclear Overhauser enhancement (NOE) effects, 164

O

Olympus BX40 microscope, 15
optical coherence tomography (OCT)
  colonoscopy, 202
  inverse spectroscopic (ISOCT), 200
  measurements from mouse models, 200, 203
  and RCM, 21, *22*, **25**
  spectroscopic, *see* spectroscopic optical coherence tomography (SOCT)
optical colonoscopy, 433
optical density (OD) filter, 49
optical imaging technology
  cellular entry of p28 *in vitro*, 104, *104*
  confocal microscopy, 103
  ICG conjugated p28 (ICG-p28), 104, *105*
  limitation of fluorescence imaging, 106–107
  mice and cell culture, 103
  NIR fluorescent probes in clinical studies, 104, **105**, *106*, 107
  in vivo animal imaging, Odyssey CLx, 104
optical interferometer, 20
oral carcinomas (OCs), 327
organoid models for CRC early detection
  detection and diagnosis in, 491, *492*
  functional characterization of murine colon organoids, 494
  future trends, 500
  H&E and IF, 493
  image processing and analyses, 493–494
  murine colon organoid culture, establishment of, 494, *495*, *496*
  organoid culture, 493
  organoid passaging, 493
  primer names and sequences for qPCR, 494, **494**
  RT-qPCR analysis, 494
  sample preparation for organoid culture, 493
  UVRAG^{FS} murine colon organoids, 494–497
ovarian cancer
  advanced, 43, 53
  drug delivery in, 53
  folate-targeting for, 53
  mortality rates, 534
  mouse model of, 44–45
  RSI models, 534, 536–538

P

pancreatic cancer
  artificial intelligence for, 33
  EVIPThT assay, 362–363
  EVs from, 247
  KRAS mutations in, 256, 327
  MIRs for, 1
  mortality rates, 27
  by TDAM, 89
pancreatic cyst, 27–28, *28*
pancreatic cystic lesions (PCLs), *see also* confocal laser endomicroscopy (CLE), pancreatic cystic lesions,
  differentiation, 27
  EUS-nCLE, 29
  imaging modalities, 27–28
  IOA and IOR in, 29–30
  mucinous and non-mucinous, 31
  non-mucinous, 27
  risk stratification of, 33
pancreatic ductal adenocarcinoma (PDAC), 67, *360*, 361
  EV Click Beads, 250
  EV Click Chips, 251
  EVs, enrichment of, 250, 357–359
  oncogenic gene alterations in, 252–254, *253*
  patients, 251
  RNA-based digital scoring, 249–250
parallel reaction monitoring (PRM) quantitative analysis
  current methods in early detection, 372
  glycopeptides preparation, 375–378
  HCC and cirrhosis, 379, *379*, **380**
  hepatocellular carcinoma, 372
  LC-Stepped HCD-DDA-MS/MS, 374
  LC-Stepped-HCD-PRM-MS/MS, 374
  in mass spectrometry, 373, *373*
  NASH, 374, **374**
  reagents, 373
  receiver operating characteristic (ROC) curves, 381, *381*
  serum samples, 373–374
  statistical analysis and software, 375
  technology-parallel reaction monitoring, 373
  using 3-marker panels, 380–381, **381**
patient-derived tumor organoids (PDTOs), 261, 263, 265, *265*, 267
PD168393-TRITC (PD-TRITC), 67
pediatric cancers
  accuracy and prediction probability, 300–301
  advantages and limitations, 301
  B-ALL and AML genomic subtype classification, 298
  computational analysis and classification, 297
  evaluation and validation, 297
  future trends, 301–302
  informatics and classification, 300
  Leukemia Lineage Classification, 297–298

library preparation and sequencing, 297
lymphoma classification, 298
nanopore technology, 295
novel diagnostic approaches, 295
RNA extraction, 296, 297
sequencing and computational hardware, 296
solid tumor classification, 298, 300
specimen collection and storage, 296–297
specimens, 295–296, **296**
personalized AI, 101
personalized medicine
advancements in, 53, 291
for breast cancer treatment, 394
in canine organoid model, 491, 510–511
phosphate buffered saline (PBS)
buccal cell collection, 407–408, 451
canine organoid model, 510
cell culture, 335–336
Exo-PROS biosensor, 125, 127–131, *132*
melphalan therapy, 215, *216*
peptide synthesis, 262–263
wash, 14, 191, 249, 323, 473
photodynamic therapy (PDT), 43, 51, 53
polymerase chain reaction (PCR)
ALT, 242–243
C-circle assay, 242, **242**
clonal analysis of cancer, 223, 225
ctDNA, 208
digital, 206
digital droplet, *see* digital droplet PCR (ddPCR)
real-time, 242–243, 409–411
RT-digital, 251
testing, 12
porphyrin-based compounds (photosensitizers), 43
porphyrin conjugate (PoP), 43–45, 48
porphyrin-phospholipid doped (PoPD) liposomes, 43
portable confocal microscopy (PCM), skin cancer
advantages, 40
alignment process, 38
benefits, 39
cost and size of, 41
device components, 37
enhanced version of, 41
expensive optoelectrical components, 37
images of skin lesions, *40*
imaging performance, 38
imaging procedure, 39, *39*
interpretation of, 41
limitations, 41
low-cost implementation, 40
morphologic features, 41
performance of, 40
schematic of, 38, *38*
skin lesions with, 39, *40*
training, 41
positron emission tomography (PET), 9, 143, 268, 421
probe-based and needlebased confocal laser endomicroscopy (pCLE/nCLE) system, 29
prostate cancer (PCa)
aging, 2
androgen deprivation therapy, 331
biomarker detection, 327
CR technology, 332, 340
CTC in, *see* circulating tumor cells (CTC), prostate cancer
HM-MRI of, *see* hybrid multi-dimensional MRI (HM-MRI), of prostate cancer
increased fatty acid synthesis in, 186
LLM for, *see* large language model (LLM), for prostate cancer
mortality rates, 92, 155
patient-derived cells, 340
PSA, *see* prostate-specific antigen (PSA)
TMEM doorway score, 89
tumor margins during surgery, 103
XRDI, 528
Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, 138
prostate-specific antigen (PSA) test, *see also* prostate cancer (PCa)
assay process, 425
biosensor setup, 425
future trends, 428
in personalized QR code, 424–425
smartphone App development, 425
proteomics, 3, 196, 344, 348, 387
bottom-up, 514
top-down, 514
punch biopsy, 36

Q

qPCR-based assay of EV DNA (EVIPqPCR), 358
quantitative fluorescence imaging (qFI)
for assessing fluorescence concentration, 48–50
corrected fluorescence, 49
correction approach, *49*, 49–50
Dox fluorophore, 48–49
fluorescence generation and detection, 48, 49
liposomal formulation of "nanoballoon," 48
raw fluorescence, 50
in vitro detection and delineation of tumors, 51
in vivo detection and delineation of tumors, 51–52
quantitative phase imaging (QPI)
clinical translation, 122–123
device materials and instrumentation, 119
future trends in, 123
kinetic assays, 120
kinetic phenotyping, 122
during melanoma adaptation, 120
multiparametric analysis of cell growth dynamics, 120, 121, *122*
rate of phenotypic state transition, 120
reagents, 119
spheroid outgrowth, 120–121, *121*
tumor cell adaptability, 118, *119*
quantitative reverse transcription polymerase chain reaction (qRTPCR), 125, 133

R

rabbit monoclonal anti-TZM, 13, 15
RADAR for multiple myeloma (MM)
in advancing precision medicine, 291–292
CAR T cell and bispecific antibody therapies, *289*
case study on clonality modeling, *291*
clinical actionability annotation, 283
copy number alterations, 282
future trends, 292
gene expression, 283–284
gene fusions, 283
generation of personalized patient reports, 284
genomic metrics, 282, 285, *286*, 287
germline variant discovery, 281–282
hallmark aberrations, 285, *286*, 287
for immune-based interventions, 288, 290
incorporating germline screening, 290
informing therapy selection with RNA expression signatures, 287–288
materials, 281
merging DNA and RNA primary results, 283
modeling of intra-tumor heterogeneity, 283
multiomics patient stratification, 284
pathways activity, calculation of, 284
in post-CAR T settings, 290
precision oncology, 280
for precision prognostication, 287
pre-processing of DNA data, 281
processing of RNA data, 282
quality control (QC), 281
risk stratification in, 285, *286*, 287
somatic mutation calling, 282
to subtype-specific therapy recommendations, 290–291
for targeted therapies, 285
workflow, 284–285
radiotherapy (RT), 12, 25, 161, 166, 213, 511, 533
Raman spectroscopy, 181, 187
rapid lifetime determination (RLD), 9
rapid on-site evaluation (ROSE), 63
recognition errors, 93, 94, 100
reflectance confocal microscopy (RCM), skin cancer, 20, *21*, 24, 36
benefits of, 37
images of skin lesions, *40*
and OCT imaging, **25**
and OCT instrumentation, 21, *22*
principle, 20, *21*
schematic of, *37*
reflectance geometry (2D imaging), 13
reflectance imaging, 10, 45
regenerative AI, 33
renal cell carcinomas, 287
restriction spectrum imaging (RSI) models
breast cancer, 533–538
cervical cancer, 533–538
materials, 534
methods, 534
ovarian cancer, 534
retinoblastomas, 4
reverse-transcription digital droplet PCR (RT-ddPCR), 469, 470, *471*, 472, 475

S

sarcomas, 4
scanning lens (SL), 13

sCMOS camera, 13
screening colonoscopy, 198, 203
search errors, 93
SEER data, 2
segmentation anything model (SAM), 112, *112*
sequencing data, 225, 295, 409
  cfRNA, 312
  genomic-based cancer diagnosis, 3
  nanopore, 297
  RADAR's workflow, 281
  RNA, 243
serous cystadenoma (SCA), 27
SFDI reflectance imaging
  demodulation and detection of, 46, *47*
  model (diffuse light propagation), 45–46
  projection, 46, *47*
  quantification of optical parameters, *47*, 47–48, *48*
  tissue-mimicking phantom, 47, *47*
shave biopsy, 36
signaling
  HER2/Neu, 82
  and tumor cell interaction, 4
single amino acid variants (SAAVs), 518–519
single-cell RNA sequencing (scRNA-seq) technologies, 223
skin biopsy, 36
  excisional biopsy, 36
  punch biopsy, 36
  shave biopsy, 36
skin cancer, 36
  dermoscopy, 23, *24*
  materials for testing and imaging, 37–38
  portable confocal microscopy, *38*, 38–39
  RCM-OCT setup, 21–23, *22*, **25**
  skin imaging procedure, 39–40
slide-free microscopy, 58
smoking, 2, 137, *see also* lung cancer
  cessation, 142, 143
  cg05575921 methylation, 139, 140
  to improve prediction, 139
  quitting, 142, 143
  using DNA methylation, 139
sodium valproate, use of, 2
solid pseudopapillary neoplasm (SPN), 27
solid-state (SS) nanopore assay, 305
  biologicals, 304–305
  chemicals, 305
  device fabrication, 305
  device materials and instrumentation, 304
  future trends, 308
  of 5-hydroxymethylcytosine (5hmC), 305–306
  measurements, 305
  nucleic acid motif extraction, 305
  of other DNA base modifications, 306, *306*
  silicon nitride window, *304*
  of single base modifications and lesions in DNA, 305–306
  of target microRNA sequences, 306–307, *307*
solid tumor
  advanced, 103, 104
  biopsies, 296
  classification, 298
  CNSTs, 213
  defined, 4
  EVs from, 247–248, 254, 257
  pediatric, *300*, 301
somatic mutations, 1–2, 282, 285, 314–315, 345
SPAD detector, 13
spanning-tree progression analysis of density-normalized events (SPADE) algorithm, 183–184, *185*, 186, 460
spatial frequency domain imaging (SFDI) projects, 43–45, 47–51
spectrally encoded imaging, 38
spectrophotometer DU 640, 12, 14
spectroscopic optical coherence tomography (SOCT)
  animal studies, 198–199
  block diagram of, 199, *199*
  of human colon polyps, 201–203, *202*
  human tissue study, 199
  of mouse colon polyps, 200–201, *201*
  signal analysis scheme, 200, *200*
squamous cell carcinoma (SCC), 20, 28, 36
standardization of protocols, 511
stimulated Raman scattering (SRS), 181
subclonal diversity, 287, 510
subharmonic-aided pressure estimation (SHAPE)
  animal specimens, 414
  clinical subjects, 414
  first in-human pilot study, *416*, 416–417
  future trends, 421–422
  IFP, *see* interstitial fluid pressure (IFP)
  prospective clinical study, 417
  ultrasound contrast agent (UCA), 413–414, *414*
  in vitro setup and data analysis, 414–415
sun exposure, 2
surface plasmon enhanced fluorescence spectroscopy (SPEFS), 125
surface plasmon resonance (SPR), 125
surgical resection, 27, 31, 43, 372, 413
SwissSPAD2 detectors, 17

T

T-cell acute lymphoblastic leukemia (T-ALL), 297
TD-MFMT FRET measurements, 16
telomere maintenance mechanisms (TMM), 239–240, 243–244
temporal point spread function (TPSF), 14
ThT-based assay of EV proteins (EVIPThT), 358
thyroid cancer
  benign and tumor samples, 348, *348*
  degradome proteins, 346, *346*
  diagnosis, 344–345
  donor tissue samples, 345, **345**
  fluorescence assays, 346
  future trends, 348–349
  materials, 345
  NAT and tumor samples, 346, *347*
  protease specificity profiling, 346
  proteolytic biomarkers, 344
  proteolytic extract preparation, 345
  proteomic analysis, 345–346
  ultrasound, 270, 276
time-correlated single photon counting (TCSPC), 9, 12–14
time-domain (TD) acquisition, 12
time domain MFMT (TD-MFMT), *16*
  and data analysis, 14
  Imager, 13–14
  protocol, *13*
time-resolved (TD), 9, 10, 12–16
Ti:Sapphire pulse laser, 13
tissue sectioning, 20
tobacco control program implementers in schools, India, 365
  carousel-style display, 367, *367*
  characteristics of study participants, 365
  digital storytelling, 365
  elements of training app, 365
  funding, 370
  future directions, 369
  gamification, 365
  home page features, map of MP, 367, *367*
  login screen with idyllic school image, *366*
  potential applications, 369–370
  TFT-TFS, *see* TFT-TFS, *see* Tobacco-Free Teachers, Tobacco-Free Society (TFT-TFS) program
  cultural relevance, 366
  dashboard, 368, *368*
  formative research, 365
  identifying factors affecting program implementation, 369
  kite theme, 365–366
  mixing qualitative and quantitative data, 369
  principal's journey, 366–368
  teachers' journey, 368
  technology, 366
  testing in a comparative effectiveness trial, 368–369
  for training app development, 365–369
  user roles, 366
Tobacco-Free Teachers, Tobacco-Free Society (TFT-TFS) program, 364–370
  applications, 369–370
  app user roles, 366
  characteristics of study participants, 365
  cultural relevance, 366
  dashboard, 368, *368*
  described, 366–368
  digital storytelling, 365
  elements of, 365
  formative research, 365
  future directions, 369
  future trends, 369–370
  gamification, 365
  identifying factors affecting program implementation, 369
  kite theme, 365–366
  materials, 365
  overview, 364
  principal's role, 366–368
  qualitative data, 369
  quantitative data, 369
  research methods for app development, 365–369
  research results for app development, 369
  teachers' role, 368
  technology, 366
  testing, 368–369
trastuzumab (TZM)
  acceptor, 12
  anti-HER2 monoclonal antibody, 10
  donor, 12

labeling, 14
resistance mechanisms, 11–12
tumor budding (TB), from H&E images in CRC
future developments, 115
implementation details, 113
mask refinement, 113
segmentation, 113–114, **114**
segmentation anything model, 112, *112*
task-specific prompt, 113, *113*
TB segmentation, 113–114, **114**
visualizations, 114, *114*
tumor cell adaptability, 118
tumor cell interaction, 3
with CTCs, 4
with ECM interactions, 3
with immune system, 4
with macrophages and DCs, 4
with non-malignant cells, 4
and signaling, 4
with TIMaE, 3
with TME, 3
tumor-centric cellular neighborhood (TCCN) analysis, 191
tumor, defined, 4
tumor-derived EVs for RNA-based digital scoring
cell line culture, 249
characterization of cell line-derived EVs, 249
chemicals and reagents, 248
DLS, 249
EV-based mRNA digital scoring assays, 256, **256**
by EV Click Chips, 250–252
of Ewing sarcoma or PDAC EVs, 249–250
future plan and limitations, 257
harvesting cell line-derived EVs, 249
HCC EV digital scoring assay, *254*, 254–255
immunogold staining, 249
PCa EV digital scoring assay, *255*, 255–256
preparation of synthetic EV samples, 249
SEM, 249
TEM, 249
tumor-derived exosomes (TEX), 125
tumor heterogeneity, 247, 256, 510
tumor immune macroenvironment (TIMaE), 3, 4
tumor-immune microenvironment (TIME), of melanoma, 2, 36, *see also* skin cancer
antibody panel design, 191
cell neighborhood analysis, 192
cell segmentation and quantification, 192
cell subtyping, 192
cyclic MxIF enables cell detection, 191, 192, *193*
DAPI-referenced QC, 192
evaluation of spatial tumor-centric community, 193
image acquisition and preprocessing, 191
image quality checking, 192
immunosuppression in context of TCCN, 193–194
macro-architecture assessment of, 193
MxIF spatial imaging, 190–191, *191*
myeloid cells impact functionality of, 194–195, *195*
phenotypic and spatial paracrine interactions, 192, 193
slide preparation, 191
spatial evaluation of, 194
spatial image analysis of MxIF data, 192
study cohort, 191
TCCN analysis, 191
tumor microenvironment (TME), 3, 16, 17
tumor microenvironment of metastasis (TMEM) doorways
automated method, 85
chemicals, 83–84
clinicopathological factors, 82
as companion diagnostic, 87
digital pathology infrastructure supporting automation of, 85–86
doorway score, 82–83, *83*
equipment, 84
future efforts, 89–90
HER2/Neu signaling, 82
human breast cancer protocol, 84
manual method, 85
methodology for, 84–86
mouse mammary cancer protocol, 84–85
in murine breast cancer, 86–87
quantification, 85
in racial disparities of metastatic potential, 87, *88*, 89
scanning, 85
selection of tissue for, 84
use and activity in preclinical research, 87
tumor xenografts
models, 11
in vivo MFLI of, 15
tyrosine kinase inhibitors (TKIs), 67
EGFR, 67, 68–71, *70*, 74
with EGFR mutations, 67
protein with fluorophore conjugates, 71–72, *72*, *73*

U

ultrabright telomere foci (UTF) assay, 241, 243, 244
ultrasonography, 9, 344
ultrasound, 20, 23, 25, 43
assisted drug delivery, 213
breast, 171
endoscopic, *see* endoscopic ultrasound (EUS)
HCC early detection, 372
liver, 473, *474*, 477
with serum AFP, 469, 479
thyroid cancer early detection, 270, 276
transcutaneous, 28
ultrasound contrast agent (UCA), 413–414, *414*
urological cancer
clinical background, 331
CR technology for, *see* conditional reprogramming (CR) technology, for urological cancer
urothelial carcinomas, 28, 338, 508
US Centers for Disease Control and Prevention (CDC), 1
US Food and Drug Administration (FDA), 171, 267
US National Cancer Institute (NCI), 1, 3
UVRAG^{FS} murine colon organoids
age-related hyperplastic changes and spontaneous cancers, 491
early-stage tumor cell phenotype *in vivo*, 494–497
forecast disease progression, 497
patient-derived colon organoids (PDCOs), 497, *498*

V

vascularization, 510
vascularized organoids, 511
virtual colonoscopy (VC), 149–151
Vivascope 1500 and 3000, 37

W

Warburg effect, 76, 186
western blot, 79, *264*, 265, *266*, 267
whole-genome sequencing (WGS), 223, 281, *281*, 301
women
adjuvant chemotherapy, 82
breast cancer *see* breast cancer
living with HIV, risk for cervical cancer, 480, 481, 485–487
ovarian cancer in, 534, 536
thyroid cancer in, 270

X

Xpert HPV
advantages, 481–482
in fluorescent channels, 481
operational issues, 482
rationale for Ct cut-offs, 482
rationale for type restriction, 482
in screen-and-treat, 481
X-ray diffraction imaging (XRDI)
development of, 529, *529*
of different cancers, 527, *527*
future trends, 529–530
methods, 526
quantitative analysis of, 528, *528*
samples, 526
scanning system, 524–526, *525*

Z

^{89}Zr-lumretuzumab, 9