



sensors

Special Issue Reprint

Biomedical Electronics and Wearable Systems

Edited by
José Machado Da Silva

mdpi.com/journal/sensors



Biomedical Electronics and Wearable Systems

Biomedical Electronics and Wearable Systems

Guest Editor

José Machado Da Silva



Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editor

José Machado Da Silva
INESC TEC
Faculdade de Engenharia
Universidade do Porto
Porto
Portugal

Editorial Office

MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Sensors* (ISSN 1424-8220), freely accessible at: https://www.mdpi.com/journal/sensors/special_issues/AL8061X6W5.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. <i>Journal Name</i> Year , Volume Number, Page Range.
--

ISBN 978-3-7258-3537-9 (Hbk)

ISBN 978-3-7258-3538-6 (PDF)

<https://doi.org/10.3390/books978-3-7258-3538-6>

© 2025 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

About the Editor vii

Daniel Carvalho, Ana Margarida Rodrigues, João Santos, Dulce Geraldo, Armando Ferreira, Marcio Assolin Correa, et al.
Evaluation of Performance and Longevity of Ti-Cu Dry Electrodes: Degradation Analysis Using Anodic Stripping Voltammetry
Reprinted from: *Sensors* **2024**, 24, 7477, <https://doi.org/10.3390/s24237477> 1

Guilherme Correia, Michael J. Crosse and Alejandro Lopez Valdes
Brain Wearables: Validation Toolkit for Ear-Level EEG Sensors
Reprinted from: *Sensors* **2024**, 24, 1226, <https://doi.org/10.3390/s24041226> 18

Mesut-Ömür Özden, Giuseppe Barbieri and Martina Gerken
A Combined Magnetolectric Sensor Array and MRI-Based Human Head Model for Biomagnetic FEM Simulation and Sensor Crosstalk Analysis
Reprinted from: *Sensors* **2024**, 24, 1186, <https://doi.org/10.3390/s24041186> 54

Sergio J. Ibáñez, Carlos D. Gómez-Carmona, Pablo López-Sierra and Sebastián Feu
Intensity Thresholds for External Workload Demands in Basketball: Is Individualization Based on Playing Positions Necessary?
Reprinted from: *Sensors* **2024**, 24, 1146, <https://doi.org/10.3390/s24041146> 74

Fernando Villalba-Meneses, Cesar Guevara, Paolo A.Velásquez-López, Isaac Arias-Serrano, Stephanie A. Guerrero-Ligña, Camila M.Valencia-Cevallos, et al.
BackMov: Individualized Motion Capture-Based Test to Assess Low Back Pain Mobility Recovery after Treatment
Reprinted from: *Sensors* **2024**, 24, 913, <https://doi.org/10.3390/s24030913> 90

Zhuoxiu Jin, Alice E. Thackray, James A. King, Kevin Deighton, Melanie Davies and David J. Stensel
Analytical Performance of the Factory-Calibrated Flash Glucose Monitoring System FreeStyle Libre2™ in Healthy Women
Reprinted from: *Sensors* **2023**, 23, 7417, <https://doi.org/10.3390/s23177417> 112

Sarah De Guzman, Andrew Lowe, Cylie Williams, Anubha Kalra and Gautam Anand
Comprehensive Understanding of Foot Development in Children Using Capacitive Textile Sensors
Reprinted from: *Sensors* **2022**, 22, 9499, <https://doi.org/10.3390/s22239499> 126

Sarah De Guzman, Andrew Lowe, Cylie Williams, Anubha Kalra and Gautam Anand
The Development of a Built-In Shoe Plantar Pressure Measurement System for Children
Reprinted from: *Sensors* **2022**, 22, 8327, <https://doi.org/10.3390/s22218327> 146

Olivier Chételat, Michaël Rapin, Benjamin Bonnal, André Fivaz, Josias Wacker and Benjamin Sporrer
Remotely Powered Two-Wire Cooperative Sensors for Biopotential Imaging Wearables
Reprinted from: *Sensors* **2022**, 22, 8219, <https://doi.org/10.3390/s22218219> 168

Marta Freitas, Francisco Pinho, Liliana Pinho, Sandra Silva, Vânia Figueira, João Paulo Vilas-Boas and Augusta Silva
Biomechanical Assessment Methods Used in Chronic Stroke: A Scoping Review of Non-Linear Approaches
Reprinted from: *Sensors* **2024**, 24, 2338, <https://doi.org/10.3390/s24072338> 186

Aron Syversen, Alexios Dosis, David Jayne and Zhiqiang Zhang
Wearable Sensors as a Preoperative Assessment Tool: A Review
Reprinted from: *Sensors* **2024**, *24*, 482, <https://doi.org/10.3390/s24020482> **213**

About the Editor

José Machado Da Silva

José Machado da Silva received his licentiate and Ph.D. degree in Electrical and Computer Engineering from the Faculty of Engineering, University of Porto (FEUP), Portugal, in 1984 and 1998, respectively. He is currently an Associate Professor at the Department of Electrical and Computer Engineering, FEUP, and Research Manager and Project Leader at INESC TEC, with teaching and research responsibilities in analog and mixed-signal electronics, VLSI design and testability, signal processing, and biomedical electronics and instrumentation, in which he has been a principal investigator or group leader in fifteen national and international projects. He has supervised or is supervising eight Ph.D. students, and more than forty M.Sc. theses. He is a co-editor of one book, and co-author of six book chapters, has more than sixty papers published in international and national journals and conferences, and holds three patents (two pending).

Article

Evaluation of Performance and Longevity of Ti-Cu Dry Electrodes: Degradation Analysis Using Anodic Stripping Voltammetry

Daniel Carvalho ¹, Ana Margarida Rodrigues ², João Santos ², Dulce Geraldo ^{2,*}, Armando Ferreira ^{1,3}, Marcio Assolin Correa ^{1,4,5}, Eduardo Alves ⁶, Nuno Pessoa Barradas ^{6,7}, Claudia Lopes ^{1,3,*,†} and Filipe Vaz ^{1,3,†}

¹ Physics Centre of Minho and Porto Universities (CF-UM-UP), University of Minho, 4710-057 Braga, Portugal; dani.fdcarvalho26@gmail.com (D.C.); armando.f@fisica.uminho.pt (A.F.); marciocorrea@fisica.ufrn.br (M.A.C.); fvaz@fisica.uminho.pt (F.V.)

² Chemistry Centre, University of Minho, 4710-057 Braga, Portugal; anamaggy1989@gmail.com (A.M.R.); jpedro.fonsecasantos@gmail.com (J.S.)

³ LaPMET—Laboratory of Physics for Materials and Emergent Technologies, University of Minho, 4710-057 Braga, Portugal

⁴ Graduate Program in Materials Science and Engineering, Federal University of Rio Grande do Norte (UFRN), Natal 59078-970, RN, Brazil

⁵ Department of Physics, Federal University of Rio Grande do Norte (UFRN), Natal 59075-000, RN, Brazil

⁶ Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal; ealves@ctn.tecnico.ulisboa.pt (E.A.); nunoni@ctn.tecnico.ulisboa.pt (N.P.B.)

⁷ Centro de Ciências e Tecnologias Nucleares, Instituto Superior Técnico, Universidade de Lisboa, Estrada Nacional 10 Bobadela LRS, 2695-066 Lisboa, Portugal

* Correspondence: gdulce@quimica.uminho.pt (D.G.); claudialopes@fisica.uminho.pt (C.L.)

† These authors contributed equally to this work.

Abstract: This study aimed to investigate the degradation of dry biopotential electrodes using the anodic stripping voltammetry (ASV) technique. The electrodes were based on Ti-Cu thin films deposited on different polymeric substrates (polyurethane, polylactic acid, and cellulose) by Direct Current (DC) magnetron sputtering. TiCu_{0.34} thin films (chemical composition of 25.4 at.% Cu and 74.6 at.% Ti) were prepared by sputtering a composite Ti target. For comparison purposes, a Cu-pure thin film was prepared under the same conditions and used as a reference. Both films exhibited dense microstructures with differences in surface topography and crystalline structure. The degradation process involved immersing TiCu_{0.34} and Cu-pure thin films in artificial sweat (prepared following the ISO standard 3160-2) for different durations (1 h, 4 h, 24 h, 168 h, and 240 h). ASV was the technique selected to quantify the amount of Cu(II) released by the electrodes immersed in the sweat solution. The optimal analysis conditions were set for 120 s and −1.0 V for time deposition and potential deposition, respectively, with a quantification limit of 0.050 ppm and a detection limit of 0.016 ppm. The results showed that TiCu_{0.34} electrodes on polyurethane substrates were significantly more reliable over time compared to Cu-pure electrodes. After 240 h of immersion, the TiCu_{0.34} electrodes released a maximum of 0.06 ppm Cu, while Cu-pure electrodes released 16 ppm. The results showed the significant impact of the substrate on the electrode's longevity, with cellulose bases performing poorly. TiCu_{0.34} thin films on cellulose released 1.15 µg/cm² of copper after 240 h, compared to 1.12 mg/cm² from Cu-pure films deposited on the same substrate. Optical microscopy revealed that electrodes based on polylactic acid substrates were more prone to corrosion over time, whereas TiCu thin-film metallic glass-like structures on PU substrates showed extended lifespan. This study underscored the importance of assessing the degradation of dry biopotential electrodes for e-health applications, contributing to developing more durable and reliable sensing devices. While the study simulated real-world conditions using artificial sweat, it did not involve in vivo measurements.

Keywords: Ti-Cu dry electrodes; ASV; electrochemical performance

Citation: Carvalho, D.; Rodrigues, A.M.; Santos, J.; Geraldo, D.; Ferreira, A.; Correa, M.A.; Alves, E.; Barradas, N.P.; Lopes, C.; Vaz, F. Evaluation of Performance and Longevity of Ti-Cu Dry Electrodes: Degradation Analysis Using Anodic Stripping Voltammetry. *Sensors* **2024**, *24*, 7477. <https://doi.org/10.3390/s24237477>

Academic Editor: José Machado Da Silva

Received: 13 October 2024

Revised: 14 November 2024

Accepted: 21 November 2024

Published: 23 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The demand for long-term care increases with the ageing population. Consequently, developing e-health physiological monitoring solutions becomes crucial to providing person-centred healthcare solutions and ensuring healthy lives and well-being for all ages. This necessity drives the development of durable, comfortable, and cost-effective physiological sensors, particularly in wearable technologies involving remote monitoring [1–6].

This study investigates the integrity of Ti-Cu dry biopotential electrodes intended for use in wearables for remote monitoring applications [1,7]. The electrodes were designed to be placed over the skin for long-term monitoring periods, which inevitably implies their contact with sweat, a very corrosive environment. Therefore, the impact of sweat on the electrode's degradation and reliability over time was examined, drawing insights from earlier research [8]. The biopotential electrodes were prepared through magnetron sputtering by depositing nanometric Ti-Cu thin films onto different polymeric substrates: polylactic acid (PLA), polyurethane (PU), and cellulose. Biodegradable PLA was selected for its biocompatibility and biodegradability, making it ideal for skin-contact applications and minimising environmental impact. Cellulose was chosen for similar reasons, with its similarity to lyocell fabric enhancing the electrode's integration into wearable systems. PU was selected for its flexibility, allowing easy conformity to the skin, and for its durability in harsh, high-sweat conditions. This selection of polymers with distinct characteristics and flexibility enables a comprehensive exploration of the electrode's performance in wearable applications. The specific properties of each substrate lead to a complex relationship between the microstructure, morphology, and topography developed by the film [9], which are all critical characteristics of the electrode's performance under real-world conditions. The degradation inflicted on the Ti-Cu biopotential dry electrodes can be easily assessed by electrochemical techniques. These techniques are distinguished by their ability to measure and control electrical quantities, offering a multifaceted characterisation of analytes, surpassing mere quantification. The versatility of electroanalytical techniques, particularly in providing selective, accurate, and highly sensitive results, positions them as indispensable tools across diverse domains, including environmental monitoring, industrial control, and biomedical analysis [10,11]. The essence of electrochemical techniques lies in measuring and controlling electrical quantities—current, potential, or charge—using an electrochemical cell typically consisting of three electrodes. Compared to other spectroscopic techniques, such as Atomic Absorption Spectroscopy (AAS), Inductively Coupled Plasma Mass Spectrometry (ICP-MS), and Inductively Coupled Plasma-Optical Emission Spectrometry (ICP-OES), electroanalytical techniques offer notable advantages in terms of selectivity and sensitivity, especially for the analysis of metal ions where pre-concentration steps can be performed (e.g., Cu, Sn, Zn, Cd, Au, or Cr) [10,12]. Furthermore, voltammetry techniques are both economical and suitable for integration into portable devices. This allows for real-time and continuous monitoring without the need for elaborate pretreatment of samples, making them an efficient choice for field applications and on-site analysis [13–17].

Voltammetric techniques measure current as a function of the potential applied to the working electrode, with cyclic voltammetry and pulse voltammetry techniques being the most widely used methods. Cyclic voltammetry involves dynamically altering the potential of the working electrode between two fixed values. This cyclic alteration allows for a comprehensive analysis of electrochemical systems, facilitating the study of redox compound behaviour and enabling the determination of reaction rate constants through detailed current–potential relationship data. Pulse voltammetry techniques, including Differential Pulse Voltammetry (DPV) and Square Wave Voltammetry (SWV), employ incremental potential steps, tailored to the specific potential profile of the working electrode. These techniques offer enhanced sensitivity and superior resolution compared to cyclic voltammetry, making them especially advantageous for analysing complex samples with low analyte concentrations [18].

Stripping voltammetry is distinguished for its ability to reach lower quantification limits. This is accomplished through a pre-concentration phase, where the analyte is

first deposited on the electrode, followed by its dissolution. This process amplifies the measured current due to the increased analyte concentration at the electrode surface. This characteristic makes stripping voltammetry exceptionally suitable for detecting trace levels of metals in environmental and biomedical samples. Anodic stripping voltammetry (ASV), a subset of this technique, is known for its high sensitivity and minimal susceptibility to sample interference, yielding precise and reliable results even in complex matrices. It is particularly useful for determining heavy metal ions in biological solutions at trace levels [19–22]. For even lower detection limits, the stripping phase in ASV can be effectively combined with pulse techniques like DPV or SWV, harnessing their sensitivity to enhance the overall analytical performance [22–24].

Building upon a previous study that evaluated the lifespan of TiAg dry electrodes [8], with a particular focus on the role of ASV in understanding degradation dynamics, this research extends those findings to the Ti-Cu electrodes designed for biopotential monitoring. ASV is highlighted due to its unique ability to elucidate the complex degradation processes of Ti-Cu electrodes. Through systematic evaluation of Cu(II) released into sweat solutions over various periods, valuable insights into the durability and structural integrity of the dry electrodes are provided. Such insights are crucial, given copper's susceptibility to corrosion in aggressive environments such as human sweat, which could potentially compromise the electrodes' long-term functionality.

2. Materials and Methods

2.1. Preparation of the Ti-Cu Electrodes

The polymeric substrates for the electrodes were produced via 3D fused deposition modelling (FDM), using a ZMorph 3D Printer (Model VX printer Wrocław, Wrocław, Poland). Three different filaments with 1.75 mm diameter were selected: (i) biodegradable polylactic acid (PLA) (Eastman Amphora™ AM3300, Kingsport, TN, USA); (ii) flexible thermoplastic polyurethane (PU) (SMARTFILL, FLEX Lot Code 155264002953, Alcalá la Real, Jaén, Spain), and (iii) sustainable cellulose nanofibrils (FILAMENTIVE, Batch 840620, New Lane, Bradford, UK). The disc-shaped substrates (radius of 7.5 mm and height of 1 mm) were designed with a flat surface to be in contact with the skin and a snap button on top for the electrode's mechanical and electrical connections into wearables. Further details can be found in [1]. The effective coated area of each electrode was determined as 433.10 mm².

Immediately before the depositions, the polymeric substrates were cleaned with ethanol and activated through plasma treatments [7,25]. It is widely recognised that polymers exhibit low surface energy with reduced wettability, leading to challenging adhesion problems with metallic thin films [26]. However, plasma treatments in reactive atmospheres offer the possibility to significantly change the chemical and physical characteristics of the polymer's surface, thereby promoting stronger adhesion at the polymer/thin film interface. During the plasma process, the polymeric substrate is activated by different mechanisms: (i) cleaning; (ii) etching; and/or (iii) activation promoted by the creation of new reactive sites. Plasma effectively removes organic contaminants and impurities from the surface (cleaning), modifies the surface texture/roughness through bombardment with energetic ions and reactive species (etching), and promotes the scission of the polymeric chains and the introduction of functional groups at reactive sites (activation), OH, -OOH, or dangling oxygen bonds for oxygen reactive atmospheres, as well as -NH₃ or (R₂C=NR) groups for nitrogen-rich atmospheres [1]. The optimal experimental parameters for plasma treatment, to activate PLA, PU, and cellulose substrates, were investigated in a previous work [1] and employed in the current study as outlined in Table 1.

The plasma treatments were conducted in a Diener plasma cleaner system (Plasma System Zepto, Diener electronic GmbH & Company KG, Ebhausen, Germany) equipped with a 13.56 MHz generator connected to a rotary pump operating at a low base pressure of 20 Pa.

Table 1. Experimental parameters for plasma treatment used on the activation of the polymeric substrates. (bold is: Polymeric Substrate).

	Atmosphere	Exposure Time (min)	Power (W)	Working Pressure (Pa)
PU	Ar	5	50	100
PLA	Ar + N ₂	15		
Cellulose	O ₂	5		

After activation, the substrates were deposited by Direct Current (DC) magnetron sputtering. The samples were positioned on a three-dimensional grounded sample holder at the centre of the vacuum chamber. The depositions were carried out using a metallic target with 99.99 at.% of purity and dimensions of 200 × 100 × 6 mm³. Two different targets were used: pure Cu for the thin films used as a reference in this study, and a Ti target modified with 25 cylindrical pellets of Cu (area of 16 mm² and thickness of 0.5 mm) glued onto its erosion zone for the preparation of the Ti-Cu thin films. The depositions were conducted by applying a current density of 75 A/m² to the target, at room temperature and in rotation mode (5 rpm), to guarantee the integrity of the polymer’s characteristics and the homogeneity of the as-deposited electrodes. The deposition time was carefully adjusted to achieve a consistent thickness of approximately 200 nm. The base pressure of the chamber was kept below 7.0 × 10^{−4} Pa, while the working pressure was set to 3.0 × 10^{−1} Pa, using an Ar flow rate of 25 sccm.

2.2. Characterisation of the Ti-Cu Electrodes

The as-deposited Cu and Ti-Cu thin films were characterised in terms of their chemical composition, crystallinity, and morphology. To perform the (micro)structural analysis, silicon substrates (100 P-type, (100) orientation) were used.

The chemical composition of both Cu and Ti-Cu thin films was determined by Rutherford Backscattering Spectrometry (RBS). The measurements were conducted using a 2 MeV beam of ions ⁴He⁺ at normal incidence, with three detectors for the backscattered ions: one silicon surface barrier detector located at a 140° scattering angle and two-pin diode detectors symmetrically positioned relative to each other at 165°. The recorded data were analysed using the NDF software 10.1 [27,28].

The surface morphology and the growth microstructure of the films were examined through scanning electron microscopy (SEM) using a high-resolution microscope (NanoSEM-FEI Nova 200: FEI Company, part of Thermo Fisher Scientific, Hillsboro, OR, USA) equipped with a field emission gun. The cross-section micrograph analysis was used to determine the films’ thickness.

The assessment of the crystallinity and phase distribution of Ti-Cu thin films was carried out through X-ray diffraction (XRD), using a Bruker D8 diffractometer (Bruker Corp., Billerica, MA, USA). The analysis was conducted in grazing incidence geometry with Cu-Kα radiation (λ = 1.5406 Å), maintaining the step size of 0.04° within the 2θ range from 25° to 90°.

Additionally, the optical surface characteristics of the as-deposited and corroded electrodes were observed in detail using a Dino-Lite digital microscope (AM7013MZT4: AnMo Electronics Corporation, New Taipei City, Taiwan) with 5-megapixel resolution and 450× magnification. For each sample, four different photographs were taken (immediately before and after immersion in artificial sweat for different periods). Each captured image was analysed by MATLAB (version R2018a, The MathWorks, Inc., Natick, MA, USA), producing binary images through a thresholding approach to evaluate the optical defects promoted on the surface of the electrode by corrosion processes [29].

2.3. Degradation of the Ti-Cu Electrodes

Considering their intended real-world application involving prolonged contact with sweaty skin, Ti-Cu dry electrodes were subjected to immersion in artificial sweat to assess

their degradation over time. Sweat, known for its corrosive nature, has the potential to corrode the metallic thin films, thereby impacting the electrodes' performance [30]. Each type of electrode, consisting of the same thin film and polymer, was immersed into 15 mL of a physiological sweat solution prepared according to ISO standard 3160-2 [31].

All electrodes were placed in an incubator (LabCompanion IST-4075R) regulated at body temperature (37.0 ± 0.1 °C), with a rotation speed set at 30 rpm, for durations of 1 h, 4 h, 24 h, 168 h, and 240 h. The impact of sweat degradation on the electrodes was evaluated by measuring the amount of Cu(II) released into the sweat solution using ASV. Two electrodes from each condition were analysed for comparative purposes.

2.4. ASV Analysis

Anodic stripping voltammetry (ASV) was performed using an Autolab potentiostat (PGSTAT30, Eco Chemie B.V., Utrecht, The Netherlands) controlled by GPES 4.9 software. The experiments were performed in a three-electrode cell, comprising (i) an Ag/AgCl reference electrode (1 M KCl; CH Instruments Inc.: IJ Cambria Scientific Ltd., Llanelli, UK); (ii) a secondary platinum wire electrode (CHI115 CH Instruments: IJ Cambria Scientific Ltd., UK); and (iii) a glassy carbon working electrode (GCE; 3 mm diameter; BAS M-2012: West Lafayette, IN 47906, USA), as illustrated in Figure 1. Before the voltammetric experiments, the GCE was polished with alumina powder (0.05 µm; Bulher: Lake Bluff, IL, USA) on a polishing cloth. All electrodes were carefully rinsed with ultrapure water and dried with absorbent paper before each voltammogram recording. Immediately after the experiments, the reference electrode was washed and immersed in a 1 M KCl solution.

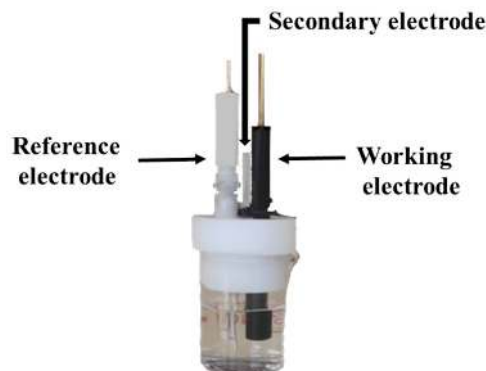


Figure 1. Electrochemical cell used for the ASV experiments.

Voltammograms were acquired using the Square Wave Voltammetry (SWV) technique, following established parameters from the literature [32]: frequency of 80 Hz, pulse amplitude of 100 mV, and potential step of 2 mV. All experiments were conducted at room temperature (25 ± 2 °C) with stirring set at 300 rpm for the deposition step and an equilibrium time of 5 s.

3. Results

3.1. Thin Film Characterisation

The use of Ti-Cu thin films as biopotential electrodes has been investigated in previous work [1]. The chemical composition revealed that the thin films prepared with a Ti-composed target, containing 25 pellets glued on the erosion zone, resulted in a Cu/Ti chemical proportion of 1:3. The films were composed of 25.6 at.% Cu and 74.4 at.% Ti, yielding a Cu/Ti atomic ratio of 0.34 ($\text{TiCu}_{0.34}$) in contrast to the pure-Cu films used as reference and prepared under the same conditions. The RBS results were also fundamental for determining the films' densities, which were found to be approximately similar,

despite the Cu film having a slightly higher density (9.27×10^{22} at/cm³), compared to 8.02×10^{22} at/cm³ for the TiCu_{0.34} thin films.

The microstructure of the TiCu_{0.34} thin films was analysed with the Cu-pure thin film as a reference. Figure 2 shows the topography and cross-section micrographs of both films, as well as their crystalline structures.

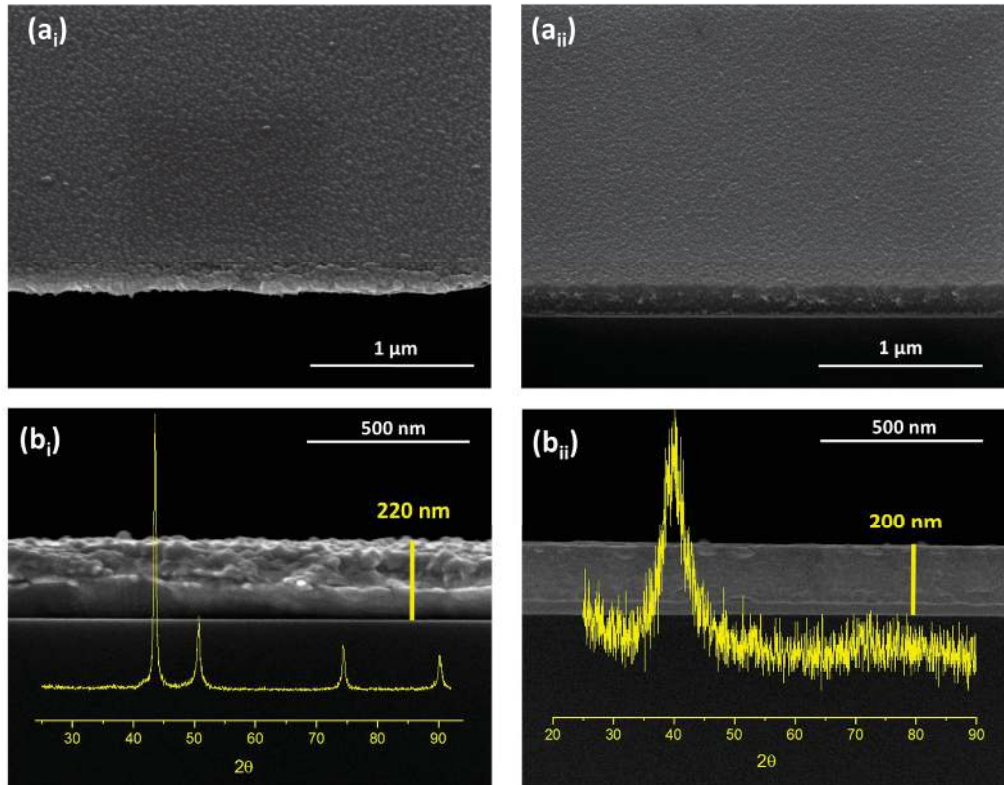


Figure 2. SEM images representative of (a) the surface morphology and (b) cross-section view of the film's growth with the respective X-ray diffractograms for (i) Cu-pure film and (ii) TiCu_{0.34} thin film.

Compared to the Cu-pure thin film (Figure 2a_i), the TiCu_{0.34} thin films (Figure 2a_{ii}) exhibited a smoother surface with a fine-grained microstructure, which is a common feature observed for Ti-based binary systems [1,7,8]. Regarding the microstructure, there is no evidence of columnar growth; instead, both films prepared with similar thicknesses (≈ 200 nm) revealed very dense and featureless microstructures, consistent with the high atomic density determined by RBS.

The Cu-pure reference film displays the typical microstructure of ductile materials, characterised by a high-density compact growth and a well-defined atomic face-centred cubic (fcc) structural arrangement. As shown in Figure 2b_i, the XRD pattern of the Cu-pure film evidenced a well-defined fcc polycrystalline structure, indicating the absence of significant voids or defects in the films' structure [33,34]. Four distinct Bragg peaks are noticeable at the positions 43.56° , 50.72° , 74.36° , and 90.24° , corresponding to the (111), (002), (022), and (113) orientations, respectively, according to the ICSD card #426938. Moreover, the cross-sectional image of the as-deposited Cu thin film brings some evidence of poor adhesion to the substrate. Several studies have reported the significant influence of residual stresses on the extent of plastic deformation during debonding processes of ductile thin-film structures on brittle substrates [35–37]. This phenomenon was particularly

studied in Cu thin films used in electronic devices. The sputtering discharge parameters, including target potential, substrate temperature, and the use of substrate biasing, primarily influence residual stresses. The energy of the impacting particles is also crucial for creating implantation sites on the substrate's surface, thereby improving interfacial adhesion. However, in this work, due to the specific characteristics of the substrates, the depositions were carried out at room temperature without BIAS, reducing the ballistic energy of the sputtered Cu atoms impacting the substrate and compromising adhesion. Additionally, the high deposition rate of the reference Cu-pure film (0.9 nm/s), which was 3.6 times higher than the deposition rate of the TiCu_{0.34} thin film, could affect the surface diffusion of the Cu ad-atoms, contributing to reduced adhesion [38,39].

In contrast, the TiCu_{0.34} thin film, despite its dense microstructures (Figure 2b_{ii}), revealed visible shear striations and vein patterns, typical features of thin-film metallic glasses (TFMGs). This evidence was confirmed by the amorphous diffraction pattern exhibited by the TiCu_{0.34} thin film, with a broad diffraction hump at 2θ ranging from 35° to 47°, indicating the film's amorphous structure. This suggests the presence of one or more metastable Ti-Cu intermetallic phases [1,7,40], precipitating in minor traces, below the XRD limit detection. Unlike the crystalline Cu-pure film of reference, the disordered arrangement shown by the TiCu_{0.34} thin film and the absence of crystalline grains led to the smoother surface morphology observed in Figure 2a_{ii}, which is also a typical feature of TFMGs [1,7,41].

In fact, the atomic density determined for the TiCu thin films by RBS measurements shows a high atomic packing density of the film, increasing from 5.62×10^{22} at/cm³ (theoretical calculation for a pure-Ti hexagonal close-packed structure) [42,43] to 8.02×10^{22} at/cm³, close to that of the Cu thin film. High atomic packing density is also a characteristic of amorphous metallic structures forming glassy structures [44–46]. As the Cu dopant element (atomic radii 128 pm) is smaller than the Ti matrix element (atomic radii 146 pm), it can easily occupy interstitial sites or substitute matrix atoms in the hcp Ti structure. According to the model developed by O.N Senkov and D.B. Miracle for bulk amorphous alloys [42], interstitial and substitutional atoms attract each other, producing short-range order atomic clusters. These clusters stabilise the amorphous state and create glassy structures, which seems to be evident in the behaviour of the TiCu_{0.34} thin films.

3.2. Voltammetry Analysis

3.2.1. ASV Optimisation

The experimental parameters for the anodic stripping voltammetry coupled with square wave modulation (ASV-SWV) were refined to accurately quantify trace amounts of Cu(II) in artificial sweat following electrodes' immersion, aiming to develop a more sensitive method. Consequently, the optimal parameters of deposition time (t_{dep}) and deposition potential (E_{dep}) were investigated using a 5.00 ppm solution of Cu(II) prepared by diluting a Cu(II) stock solution (1000 ppm; Merck Company: Merck Certipur—Darmstadt, Germany) in artificial sweat. For the t_{dep} , a range of 0 to 210 s with 30 s intervals was tested, while maintaining E_{dep} at -1.0 V [22,47]. Regarding the E_{dep} , potentials ranging from -1.20 V to -0.70 V with increasing increments of 0.10 V were evaluated, with t_{dep} fixed at 120 s. Each test was conducted in triplicate to ensure precision and reliability. The obtained results are depicted in Figure 3.

As expected, the peak current (I_p) increased with the t_{dep} , as shown in Figure 3a. Longer deposition times allow for a greater deposition of Cu(II) ions on the working electrode, consequently yielding higher peak currents during the subsequent stripping step. Despite the higher I_p values associated with longer deposition durations, the t_{dep} of 120 s was selected to determine the release of Cu(II) ions by the Ti-Cu electrodes. This choice was made considering that it enables fast analyses, and longer times do not properly benefit the I_p/t_{dep} ratio.

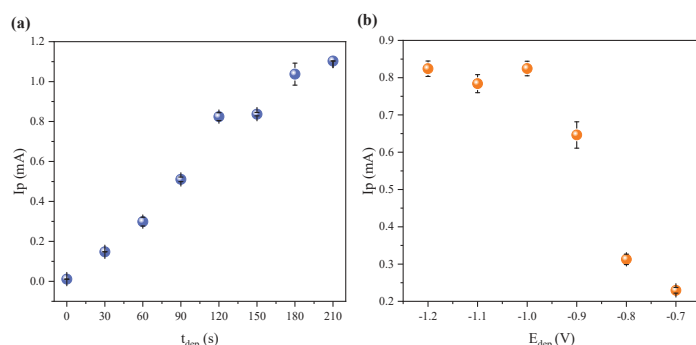


Figure 3. Variation in the peak current (I_p) of a 5.00 ppm solution of Cu(II) in artificial sweat as a function of (a) deposition time (t_{dep}) at a deposition potential of -1.0 V and (b) deposition potential (E_{dep}) with a deposition time of 120 s. Data were obtained using the ASV-SWV technique with the established parameters.

Conversely, the extension of the electrochemical reduction decreases as the deposition potential E_{dep} becomes less negative (Figure 3b). More negative E_{dep} values promote higher rates of Cu(II) ion reduction, resulting in greater deposition on the electrode surface and consequently higher peak currents during the stripping step. The I_p value reached its maximum at a potential of -1.0 V, which was selected as the optimal deposition potential for all subsequent tests performed on the electrodes.

3.2.2. Determination of Cu(II) into Artificial Sweat

After setting the fundamental parameters for the optimisation of ASV-SWV technique, a calibration curve was obtained by varying the concentration of Cu(II) from 0.050 to 1.5 ppm. Standard solutions were prepared through the Cu(II) stock solution referred to in Section 3.2.1. Figure 4 shows the voltammograms obtained using ASV-SWV for the standard solutions, with the inset featuring the respective calibration curve, delineating the relationship between copper concentration and peak current response.

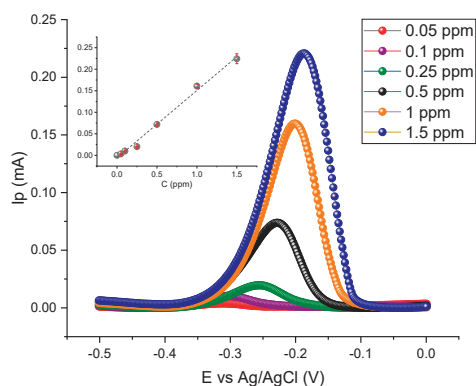


Figure 4. Anodic stripping voltammograms of copper in artificial sweat obtained at the GCE electrode for the standard Cu(II) solutions with different concentrations. To ascertain the limit of quantification (LOQ) and the limit of detection (LOD) of the method, 6 replicate analyses of the standard Cu(II) solution with the lowest concentration (0.05 ppm) were performed. The relative standard deviation was less than 10%, satisfying the acceptance criterion, thus establishing the LOQ at 0.05 ppm. The LOD was estimated to be one-third of the LOQ, yielding a value of 0.016 ppm. Although this LOD is relatively low, it is higher than the values reported in the literature (0.0002 ppm) [22]. This suggests that while the current method is effective, there is still room for improvement.

Figure 4 evidences a linear relationship between the stripping peak current and Cu(II) concentration in the range of 0.05 to 1.50 ppm. The calibration curve is described by Equation (1), with uncertainties calculated using Student's t-factor for a 95% confidence level with four degrees of freedom.

$$I_p(\text{mA}) = (0.16 \pm 0.02)C(\text{ppm}) - (0.01 \pm 0.01) \quad (1)$$

The correlation coefficient of 0.996 confirms the linearity of the method. Thus, the calibration curve described by Equation (1) was used to quantify the Cu(II) released by the Ti-Cu electrodes immersed in artificial sweat solutions during distinct periods.

3.3. Analytical Assessment of Ti-Cu Electrodes in Artificial Sweat

After establishing the calibration curve under optimised conditions, the copper concentration was determined in 15 mL of the artificial sweat samples where the Ti-Cu electrodes underwent degradation. As this study aims to assess the impact of sweat on the degradation of the TiCu_{0.34} electrodes prepared with different polymeric substrates (PLA, PU, and cellulose), two replicas of each type of electrode were used. For comparison, electrodes based on Cu-pure films deposited on the same substrates and submitted to the same degradation conditions were also tested. The results are depicted in Figure 5. Despite the same trend, in some cases, it is possible to note some discrepancies between the replicas. Since the polymeric bases were 3D-printed, the possibility of significant structural variations in the substrates must be considered. These variations can affect the deposition of the thin film and, consequently, the final performance of the electrode.

The concentration of Cu(II) released by the TiCu_{0.34} electrodes into artificial sweat is approximately three times lower than that released from Cu-pure electrodes (reference), across all the substrate types, under the same degradation conditions. The superior performance of the TiCu_{0.34} electrodes is closely related to the thin films' characteristics discussed in Section 3.1. In fact, the corrosion behaviour is strongly influenced by the morphology of the thin film [48].

The dense and featureless microstructure of the TiCu_{0.34} thin film, typical of TFMGs, results in smoother surfaces free from defects, maintaining the electrode's integrity over extended periods in corrosive environments. Additionally, Ti is well known for its impressive corrosion resistance due to the formation of a stable passivation layer of titanium dioxide (TiO₂) on its surface. Therefore, both the microstructure and the TiO₂ passivation layer in the film containing almost 75 at.% of Ti could play a crucial role in resisting degradation over time [44,49]. In contrast, the poor performance of the Cu-pure electrodes, used as a reference, indicates their susceptibility to dissolution. The Cu thin film, when exposed to the saline and corrosive environment of sweat, readily oxidises, leading to the formation of the cupric ion Cu(II). As a consequence, the rate of the copper released increases, which is a prime factor for its dissolution [49–51]. Additionally, the active–passive corrosion behaviour, along with the ductile microstructures discussed in Section 3.1, further weakens the poor adhesion of the film to the substrate, facilitating corrosion.

The release of Cu(II) from the TiCu_{0.34}/PU electrodes (Figure 5a_{ii}) remains very low and relatively constant, regardless of the degradation time. For degradation periods of less than 24 h, the concentration of Cu(II) released remains under the LOQ, and after 240 h immersed in artificial sweat, it does not exceed 0.06 ppm. The result suggests that no significant degradation occurs for these electrodes after the first hour of contact with artificial sweat. In turn, for the reference Cu-pure electrodes, the concentration of Cu(II) in the artificial sweat solution significantly increases from approximately 2 to 8 ppm in the first 24 h to reach the maximum value of 15.4 ppm after 240 h. The results leave no doubts about the severity of the corrosion processes in the electrodes of pure Cu.

Except for the TiCu_{0.34}/PU electrodes, the integrity of all other electrodes prepared in this work highly depends on the degradation time. For TiCu_{0.34} electrodes based on PLA (Figure 5b_{ii}), the amount of Cu(II) released nearly doubles (from 0.07 to 0.16 ppm) after 24 h, always exceeding the LOQ. There is a clear increase in the amount of metal released into

the sweat solution over time, showing a tendency to stabilise from 168 h onwards. In turn, the reference electrodes using PU substrates (Figure 5b_i) exhibit a similar behaviour to that reported for TiCu_{0.34}/PU electrodes. After 24 h, the Cu(II) concentration in the solution exceeds 8 ppm, exhibiting a constant Cu(II) dissolution for longer immersion times.

For TiCu_{0.34} electrodes based on cellulose (Figure 5c_{ii}), the Cu(II) concentration released into the solution doubles (from 0.11 to 0.21 ppm) after the first 4 h of immersion in artificial sweat solution. After the first hour of immersion, the Cu(II) released by these electrodes was already over the LOQ. This behaviour significantly worsens for Cu-pure electrodes (Figure 5c_i) where the results seem to indicate that for longer periods (>168 h), the entire thin film was dissolved and released to the sweat solution, due to the high concentrations detected (≈ 75 ppm).

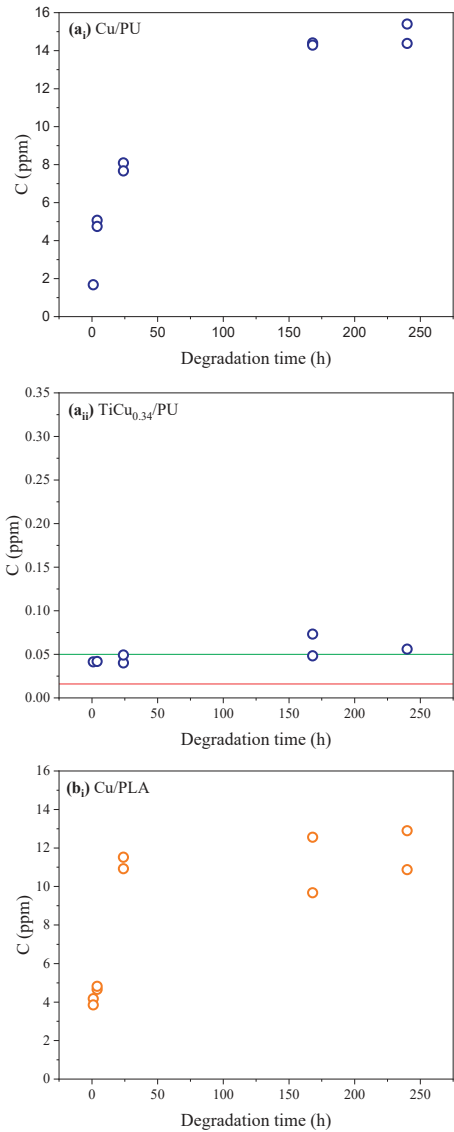


Figure 5. Cont.

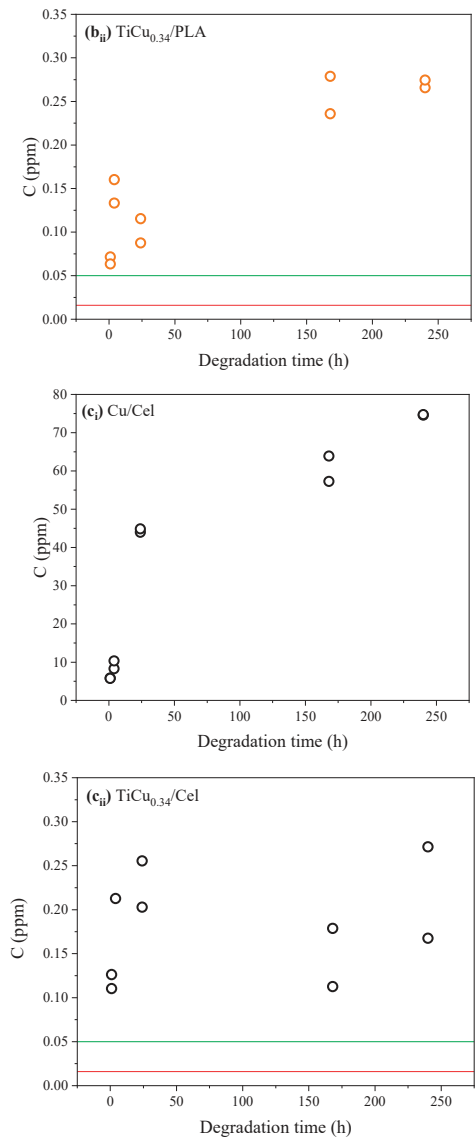


Figure 5. Copper concentration determined in artificial sweat released from the electrodes of (i) Cu-pure and (ii) TiCu_{0.34}, immersed in artificial sweat using different substrates: (a) PU, (b) PLA, and (c) cellulose. The red line represents the LOD (limit of detection), and the green line represents the LOQ (limit of quantification).

The results clearly demonstrate the significant influence of the substrate on the electrode's degradation. To facilitate a comprehensive analysis, the concentration of Cu(II) in the sweat solution was converted to mass released per unit area ($\mu\text{g}/\text{cm}^2$), considering the active area of the electrodes. Figure 6 enables a comparison of the substrate's impact on the degradation behaviour of Cu-pure and TiCu_{0.34} electrodes over time.

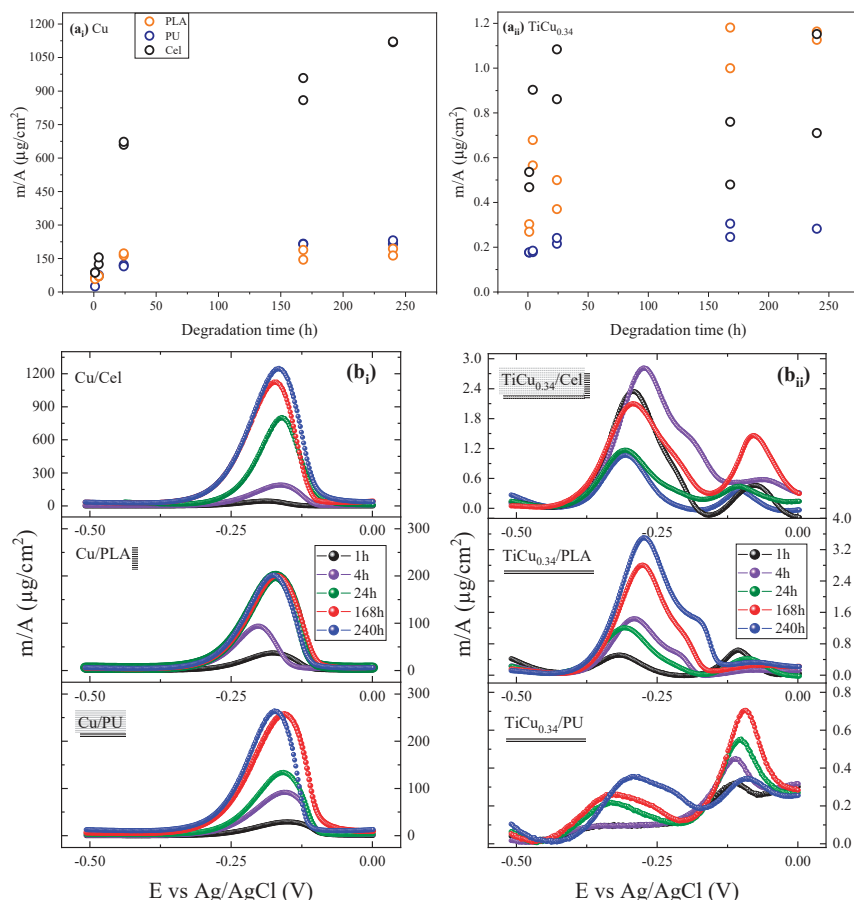


Figure 6. Mass of copper released per unit area in the artificial sweat solution from thin films of (a_i) Cu-pure and (a_{ii}) TiCu_{0.34} deposited on the different substrates, along with the experimental voltammograms to determine the copper release for (b_i) Cu-pure electrodes and (b_{ii}) TiCu_{0.34} pure electrodes.

Cellulose substrates have the greatest negative impact on electrode performance. This effect is particularly evident for Cu-pure electrodes but is also consistent for TiCu_{0.34} electrodes. Except during the 168 h immersion period, when TiCu_{0.34}/PLA electrodes show the poorest performance, the high porosity of cellulose significantly affects the characteristics and properties of the deposited films, reducing the electrode's lifetime. Also, for the Cu-pure electrodes on cellulose, the copper mass determined in sweat reaches milligrams per cm² levels of 1.12 mg/cm² after 240 h of immersion.

The results also show that, even if the amount of Cu(II) in the TiCu_{0.34} electrodes were four times greater (100 at.% of Cu as in the Cu-pure film), the release would not exceed 5.00 $\mu\text{g}/\text{cm}^2$ (after 240 h of immersion). This value is still lower than the smallest amount released by the Cu-pure electrodes just 1 h after degradation (25 $\mu\text{g}/\text{cm}^2$), underscoring the quality of the TFMG-like microstructure developed in the TiCu_{0.34} film, which hinders sweat penetration and preserves the electrode integrity for extended periods.

Considering all the results, TiCu_{0.34} electrodes prepared with PU substrates emerge as the preferable option for real applications. The amount of copper released to sweat was the lowest, under or close to the LOQ, with no significant variation over time. After 240 h, only 0.28 $\mu\text{g}/\text{cm}^2$ of copper was released into the sweat solution, indicating the high

resistance of these electrodes to sweat-induced degradation and an improved lifespan. In comparison to Ti-Ag dry electrodes deposited on PTFE substrates, analysed by the authors in a previous study [8], the $\text{TiCu}_{0.34}$ electrodes demonstrate greater stability. For similar chemical compositions ($\text{TiAg}_{0.23}$ electrodes), a higher release rate of Ag was detected, suggesting that $\text{TiCu}_{0.34}$ electrodes may be more suitable for prolonged applications.

However, understanding the poor performance of $\text{TiCu}_{0.34}$ /PLA electrodes over extended periods is crucial. Therefore, the electrodes were subjected to optical characterisation immediately before and after immersion. Due to the specific experimental parameters and mechanisms involved in the sputtering process, there is a deep connection between substrate characteristics and film growth, a factor that strongly conditions the final properties of the electrode and amplifies the substrate defects. Since the polymeric substrates were prepared using 3D FDM, they inherently exhibit significant surface defects, including heterogeneities in surface topography, voids, porosity, and roughness. These defects are replicated and amplified during film deposition, increasing the susceptibility to degradation.

3.4. Optical Characterisation

Optical characterisation helps understand the influence of defects on the degradation process and how the electrode/thin film surface evolves after various periods of immersion in physiological sweat. Surface defects on the electrode provide open paths for the electrolyte's entry through the film, increasing corrosion kinetics. For the optical analysis, four images of each electrode's surface were taken before and after sweat immersion. Figure 7 shows the surface of three randomly selected $\text{TiCu}_{0.34}$ electrodes, each deposited on a different substrate. Images were captured with an optical microscope (Figure 7a), along with corresponding images created by a MATLAB algorithm, which quantifies the area of imperfections/heterogeneities marked by the dark field (Figure 7b).

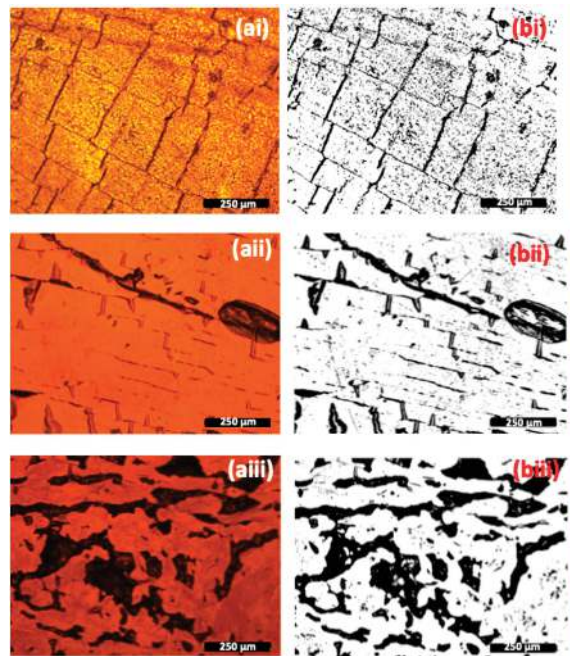


Figure 7. Optical characterisation of a (i) TiCu /PU, (ii) TiCu /PLA, (iii) TiCu /cellulose electrode's surface before degradation: (a) are the images acquired by the Dino-Lite digital microscope and (b) the respective binary images, processed by MATLAB to calculate the optical defects (black areas in the image).

The optical pictures provide a measure of the defects' progression on the film surface during degradation, emphasising the influence of the substrate on the film's topography and on the overall behaviour of the electrode. The percentage of the damaged area is calculated by dividing the damaged area (black regions) by the total area of the image (Equation (2)). Since four images of each electrode's surface were taken, the number of defects was given by the average number obtained for all images. The results, presented in Figure 8, show the defect's progression as a function of the degradation time, where

$$\Delta \text{Defects (\%)} = \text{Defects (\%)}(\text{after degradation}) - \text{Defects (\%)}(\text{before degradation}) \quad (2)$$

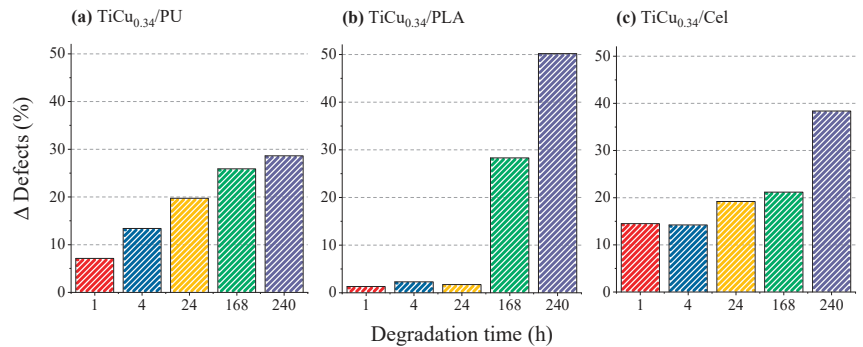


Figure 8. Variation in the number of defects on the surface of the TiCu electrodes prepared by the functionalisation of different polymeric bases (a) PU, (b) PLA, and (c) cellulose for 1 h, 4 h, 24 h, 168 h, and 240 h of degradation.

The homogeneities in the surface topography (Figure 8) of the electrode correlate closely with the concentration of Cu(II) released into the artificial sweat solution during different immersion periods (Figure 6). For the TiCu_{0.34}/PU electrodes, a slow and gradual increase in heterogeneities is observed over extended immersion periods, indicative of a light degradation process (Figure 8a). This behaviour is similarly noted in the TiCu_{0.34} electrodes based on cellulose substrates during shorter degradation periods (Figure 8b). However, after 240 h, the percentage of defects significantly increases, reflecting the higher Cu(II) amount released into the sweat solution, as shown in Figure 6b. The most notable deviation is observed with the TiCu_{0.34}/PLA electrodes. Here, the percentage of defects rises sharply during prolonged immersion periods (168 h and 240 h), highlighting the impact of the electrolyte on the electrode surface. As a consequence, the degradation of the film is accelerated leading to a higher release of Cu(II) content into the sweat solution.

The results for Cu electrodes are not presented, since after 1 h of immersion in physiological sweat, almost all the thin films had been released from the polymer into the solution, regardless of the type, due to dissolution phenomena promoted by corrosion processes. This evidence extensively confirms the poor resistance of the Cu-pure reference electrodes to degradation processes, shortening their service life. The active dissolution of Cu into sweat solutions has already been reported in several studies [49–51]. The anodic reaction process is accelerated by the sodium chloride (NaCl) in the sweat solution, resulting in the formation of reactive and soluble dichlorocuprate(I) (CuCl_2^-) ions, as well as cuprous oxide (Cu_2O) and copper chloride hydroxide ($\text{Cu}_2(\text{OH})_3\text{Cl}$), through which the Cu ions diffuse [50].

4. Conclusions

This study emphasises the critical importance of assessing the lifespan of dry biopotential electrodes, which have significantly contributed to the expansion of e-health applications in recent years. Based on the strong performance of Ti-Cu electrodes, this work evaluated their degradation and corrosion resistance using anodic stripping voltamme-

try. The study compared Ti-Cu thin film-based electrodes deposited on three polymeric substrates (PU, PLA, and cellulose) with Cu-pure electrodes, all prepared under identical conditions. The primary focus was on quantifying the degradation of thin-film electrodes by monitoring copper release into an artificial sweat solution over various immersion times.

The ASV analysis showed that TiCu_{0.34} electrodes released significantly less copper compared to Cu-pure electrodes. The type of substrate polymer had a substantial impact on the overall performance of the electrodes, with TiCu_{0.34}/PU electrodes exhibiting exceptional performance. Notably, after 1 h of immersion, Cu-pure electrodes released approximately 2 ppm of Cu(II), which was 25 times higher than the 0.06 ppm released by TiCu_{0.34} electrodes, even after 240 h. The influence of the substrate on the electrode's performance was also evidenced by the TiCu_{0.34} electrodes prepared with PLA and cellulose substrates, showing contrasting behaviours for different immersion times.

The findings indicate that the degradation of the electrodes is closely linked to the microstructural features of the thin films prepared by sputtering. The TFMG-like characteristics of the TiCu_{0.34} thin films, along with the excellent corrosion resistance of titanium, contribute to the outstanding performance of these electrodes. Therefore, TiCu_{0.34} electrodes are promising candidates for long-term health monitoring applications, offering enhanced durability and reliability even in aggressive environments.

Author Contributions: Conceptualization, D.G., C.L. and F.V.; Methodology, D.G. and E.A.; Software, N.P.B.; Validation, D.G., E.A., N.P.B., C.L. and F.V.; Formal analysis, A.F., M.A.C., E.A. and C.L.; Investigation, D.C., A.M.R., J.S., E.A. and N.P.B.; Data curation, D.C., J.S., D.G. and N.P.B.; Writing—original draft, D.C. and A.M.R.; Writing—review & editing, D.G., A.F., M.A.C., C.L. and F.V.; Visualization, D.G.; Supervision, D.G., C.L. and F.V.; Project administration, C.L. and F.V.; Funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Portuguese Foundation for Science and Technology (FCT) in the framework of the Strategic Funding UIDB/04650/2020 and UIDP/00686/2020; the Bilateral Projects: 2023.10186.CBM—Re-Muscles and 2023.09567.CBM—FLEX-HEALTH; and the project 2022.15686.UTA—NanoPower (DOI: 10.54499/2022.15686). Cláudia Lopes acknowledges IAPMEI for her researcher contract from the project Agenda Drivolution—Transição para a Fábrica do Futuro (C644913740-00000022). MAC acknowledges the support of the INCT of Spintronics and Advanced Magnetic Nanostructures (INCT-SpinNanoMag), CNPq 406836/2022-1. A. Ferreira also thanks the FCT founding CT0077 (DOI: <https://doi.org/10.54499/DL57/2016/CP1377/CT0077>).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lopes, C.; Veloso, H.; Hayes, M.; Cullinan, M.A.; Vaz, F. Nanostructured (Ti,Cu)N dry electrodes for advanced control of the neuromuscular activity. *IEEE Sens. J.* **2022**, *23*, 3629–3639. [CrossRef]
2. Poon, C.C.; Liu, Q.; Gao, H.; Lin, W.-H.; Zhang, Y.-T. Wearable Intelligent Systems for E-Health. *J. Comput. Sci. Eng.* **2011**, *5*, 246–256. [CrossRef]
3. Niu, X.; Gao, X.; Liu, Y.; Liu, H. Surface bioelectric dry Electrodes: A review. *Measurement* **2021**, *183*, 109774. [CrossRef]
4. Kalasin, S.; Surareunghai, W. Challenges of Emerging Wearable Sensors for Remote Monitoring toward Telemedicine Healthcare. *Anal. Chem.* **2023**, *95*, 1773–1784. [CrossRef] [PubMed]
5. Majumder, S.; Mondal, T.; Deen, M.J. Wearable Sensors for Remote Health Monitoring. *Sensors* **2017**, *17*, 130. [CrossRef]
6. Al-Ayyad, M.; Abu Owida, H.; De Fazio, R.; Al-Naami, B.; Visconti, P. Electromyography Monitoring Systems in Rehabilitation: A Review of Clinical Applications, Wearable Devices and Signal Acquisition Methodologies. *Electronics* **2023**, *12*, 1520. [CrossRef]
7. Lopes, C.; Fiedler, P.; Rodrigues, M.S.; Borges, J.; Bertollo, M.; Alves, E.; Barradas, N.P.; Comani, S.; Haueisen, J.; Vaz, F. Med-Doped Ti-Me Intermetallic Thin Films Used for Dry Biopotential Electrodes: A Comparative Case Study. *Sensors* **2021**, *21*, 8143. [CrossRef]

8. Carvalho, D.; Marques, S.; Siqueira, G.; Ferreira, A.; Santos, J.; Geraldo, D.; Castro, C.R.; Machado, A.V.; Vaz, F.; Lopes, C. Enhancing the Longevity and Functionality of Ti-Ag Dry Electrodes for Remote Biomedical Applications: A Comprehensive Study. *Sensors* **2023**, *23*, 8321. [CrossRef]
9. Lopes, C.; Rodrigues, M.S.; Ferreira, A.; Macedo, F.; Borsan, I.; Gabor, C.; Pop, M.-A.; Alves, E.; Barradas, N.P.; Munteanu, D.; et al. The influence of the nanostructure design on the optical, electrical and thermal properties of TiN_x thin films prepared by reactive magnetron sputtering. *Mater. Chem. Phys.* **2023**, *306*, 127981. [CrossRef]
10. Bersier, P.M.; Howell, J.; Bruntlett, C. Tutorial review. Advanced electroanalytical techniques versus atomic absorption spectrometry, inductively coupled plasma atomic emission spectrometry and inductively coupled plasma mass spectrometry in environmental analysis. *Analyst* **1994**, *119*, 219–232. [CrossRef]
11. Harvey, D. *Modern Analytical Chemistry*, 1st ed.; Bensink, J.L., Ed.; McGraw-Hill: Boston, MA, USA, 2000.
12. Meucci, V.; Battaglia, F.; Marchetti, V.; Gori, E.; Intorre, L. Rapid and simultaneous electrochemical method to measure copper and lead in canine liver biopsy. *MethodsX* **2020**, *7*, 101154. [CrossRef] [PubMed]
13. Wang, J. *Analytical Electrochemistry*, 2nd ed.; Wiley-VCH Verlag GmbH: New York, NY, USA, 2000.
14. Borrill, A.J.; Reily, N.E.; Macpherson, J.V. Addressing the practicalities of anodic stripping voltammetry for heavy metal detection: A tutorial review. *Analyst* **2019**, *144*, 6834–6849. [CrossRef]
15. Wygant, B.R.; Lambert, T.N. Thin Film Electrodes for Anodic Stripping Voltammetry: A Mini-Review. *Front. Chem.* **2022**, *9*, 809535. [CrossRef]
16. Pungjunun, K.; Chaiyo, S.; Jantrahong, I.; Nantaphol, S.; Siangproh, W.; Chailapakul, O. Anodic stripping voltammetric determination of total arsenic using a gold nanoparticle-modified boron-doped diamond electrode on a paper-based device. *Microchim. Acta* **2018**, *185*, 324. [CrossRef]
17. Planková, A.; Jampilek, J.; Švorc, L.; Hanko, M.; Mikuš, P. Anodic stripping voltammetry: Affordable and reliable alternative to inductively coupled plasma-based analytical methods. *Monatshefte Für Chem.–Chem. Mon.* **2018**, *149*, 913–920. [CrossRef]
18. Southampton Electrochemistry Group. *Instrumental Methods in Electrochemistry*, Repr. in Ellis Horwood Series in Physical Chemistry; Horwood: Chichester, UK, 2011.
19. Brainina, K.Z.; Malakhova, N.A.; Stojko, N.Y. Stripping voltammetry in environmental and food analysis. *Fresenius J. Anal. Chem.* **2000**, *368*, 307–325. [CrossRef]
20. Bakirhan, N.K.; Uslu, B.; Ozkan, S.A. Sensitive and Selective Assay of Antimicrobials on Nanostructured Materials by Electrochemical Techniques. In *Nanostructures for Antimicrobial Therapy*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 55–83. [CrossRef]
21. Wang, T.; Yue, W. Carbon Nanotubes Heavy Metal Detection with Stripping Voltammetry: A Review Paper. *Electroanalysis* **2017**, *29*, 2178–2189. [CrossRef]
22. Bobrowski, A.; Krlicka, A. Anodic Stripping Voltammetric Determination of Copper Traces in Carbonate Minerals and Fly Ash Extracts Using a Screen-Printed Electrode Modified In Situ with Antimony Film. *Insights Anal. Electrochem.* **2015**, *1*, 1–6. [CrossRef]
23. Stankovic, D.; Roglic, G.; Andjelkovic, I.; Skrivanj, S.; Mutic, J.; Manojlovic, D. SW Anodic Stripping Voltammetry Determination of Copper with New Tape of Modified GC Electrode. *Anal. Bioanal. Chem.* **2011**, *3*, 556–564.
24. Lieu, N.T.; Trung, P.Q.; Trang, L.T.T.; Giang, L.T. Simultaneous effect of pH, deposition time, deposition potential, and step potential on the stripping peak current of copper on platinum nanoflowers modified glassy carbon electrode (PtNFs/GCE) using response surface methodology. *Vietnam. J. Chem.* **2020**, *58*, 302–308. [CrossRef]
25. Rodrigues, M.S.; Fiedler, P.; Küchler, N.; Domingues, R.P.; Lopes, C.; Borges, J.; Haueisen, J.; Vaz, F. Dry Electrodes for Surface Electromyography Based on Architected Titanium Thin Films. *Materials* **2020**, *13*, 2135. [CrossRef] [PubMed]
26. Vesel, A.; Mozetic, M. New developments in surface functionalization of polymers using controlled plasma treatments. *J. Phys. D Appl. Phys.* **2017**, *50*, 293001. [CrossRef]
27. Barradas, N.; Alves, E.; Jaynes, C.; Tosaki, M. Accurate simulation of backscattering spectra in the presence of sharp resonances. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **2006**, *247*, 381–389. [CrossRef]
28. Barradas, N.; Jaynes, C. Advanced physics and algorithms in the IBA DataFurnace. *Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At.* **2008**, *266*, 1875–1879. [CrossRef]
29. Proença, M.; Borges, J.; Rodrigues, M.S.; Meira, D.I.; Sampaio, P.; Dias, J.P.; Pedrosa, P.; Martin, N.; Bundaleski, N.; Teodoro, O.M.; et al. Nanocomposite thin films based on Au-Ag nanoparticles embedded in a CuO matrix for localized surface plasmon resonance sensing. *Appl. Surf. Sci.* **2019**, *484*, 152–168. [CrossRef]
30. Chen, M.; Wang, X.; Hou, X.; Qiu, J.; Zhang, E.; Hu, J. Enhanced mechanical, bio-corrosion, and antibacterial properties of Ti-Cu alloy by forming a gradient nanostructured surface layer. *Surf. Coat. Technol.* **2023**, *465*, 129609. [CrossRef]
31. International Organization for Standardization (ISO). *Watch-Cases and Accessories—Gold Alloy Coverings—Part 2: Determination of Fineness, Thickness, Corrosion Resistance, and Adhesion*; International Organization for Standardization (ISO): Geneva, Switzerland, 2015.
32. Ma, D.; Jing, P.; Gong, Y.; Wu, B.; Deng, Q.; Li, Y.; Chen, C.; Leng, Y.; Huang, N. Structure and stress of Cu films prepared by high power pulsed magnetron sputtering. *Vacuum* **2019**, *160*, 226–232. [CrossRef]
33. Le, M.-T.; Sohn, Y.-U.; Lim, J.-W.; Choi, G.-S. Effect of Sputtering Power on the Nucleation and Growth of Cu Films Deposited by Magnetron Sputtering. *Mater. Trans.* **2010**, *51*, 116–120. [CrossRef]

34. Zhao, D.; Letz, S.; Jank, M.; März, M. Adhesion strength of ductile thin film determined by cross-sectional nanoindentation. *Int. J. Mech. Sci.* **2024**, *270*, 109103. [CrossRef]
35. Lassnig, A.; Putz, B.; Hirn, S.; Többs, D.; Mitterer, C.; Cordill, M. Adhesion evaluation of thin films to dielectrics in multilayer stacks: A comparison of four-point bending and stressed overlayer technique. *Mater. Des.* **2021**, *200*, 109451. [CrossRef]
36. Chang, S.-Y.; Chang, T.-K.; Lee, Y.-S. Nanoindenting Mechanical Responses and Interfacial Adhesion Strength of Electrochemically Deposited Copper Film. *J. Electrochem. Soc.* **2005**, *152*, C657–C663. [CrossRef]
37. Lim, J.D.; Lee, P.M.; Chen, Z. Understanding the bonding mechanisms of directly sputtered copper thin film on an alumina substrate. *Thin Solid Films* **2017**, *634*, 6–14. [CrossRef]
38. Volinsky, A.A.; Tymiak, N.I.; Kriese, M.D.; Gerberich, W.W.; Hutchinson, J.W. Quantitative Modeling and Measurement of Copper Thin Film Adhesion. *MRS Proc.* **1998**, *539*, 277. [CrossRef]
39. Lopes, C.; Gabor, C.; Cristea, D.; Costa, R.; Domingues, R.; Rodrigues, M.; Borges, J.; Alves, E.; Barradas, N.; Munteanu, D.; et al. Evolution of the mechanical properties of Ti-based intermetallic thin films doped with different metals to be used as biomedical devices. *Appl. Surf. Sci.* **2019**, *505*, 144617. [CrossRef]
40. Tang, J.-F.; Huang, P.-Y.; Lin, J.-H.; Liu, T.-W.; Yang, F.-C.; Chang, C.-L. Microstructure and Antimicrobial Properties of Zr-Cu-Ti Thin-Film Metallic Glass Deposited Using High-Power Impulse Magnetron Sputtering. *Materials* **2022**, *15*, 2461. [CrossRef]
41. Senkov, O.; Miracle, D. Effect of the atomic size distribution on glass forming ability of amorphous metallic alloys. *Mater. Res. Bull.* **2001**, *36*, 2183–2198. [CrossRef]
42. Aqua-Calc. n.d. Available online: <https://www.aqua-calc.com/page/density-table/substance/titanium> (accessed on 26 July 2024).
43. Rajan, S.T.; Arockiarajan, A. Thin film metallic glasses for bioimplants and surgical tools: A review. *J. Alloys Compd.* **2021**, *876*, 159939. [CrossRef]
44. Apreutesei, M.; Steyer, P.; Billard, A.; Joly-Pottuz, L.; Esnouf, C. Zr–Cu thin film metallic glasses: An assessment of the thermal stability and phases' transformation mechanisms. *J. Alloys Compd.* **2015**, *619*, 284–292. [CrossRef]
45. Jain, M.; Sharma, A.; Pajor, K.; Wiczerzak, K.; della Ventura, N.M.; Maeder, X.; Kruzic, J.J.; Gludovatz, B.; Michler, J. Mechanical properties and thermal stability of thin film metallic glass compared to bulk metallic glass from ambient to elevated temperatures. *J. Alloys Compd.* **2023**, *960*, 170728. [CrossRef]
46. Agra-Gutiérrez, C.; Hardcastle, J.L.; Ball, J.C.; Compton, R.G. Anodic stripping voltammetry of copper at insonated glassy carbon-based electrodes: Application to the determination of copper in beer. *Analyst* **1999**, *124*, 1053–1057. [CrossRef]
47. Abdeen, D.H.; El Hachach, M.; Koc, M.; Atieh, M.A. A Review on the Corrosion Behaviour of Nanocoatings on Metallic Substrates. *Materials* **2019**, *12*, 210. [CrossRef] [PubMed]
48. Fovet, Y.; Gal, J.-Y.; Toumelin-Chemla, F. Influence of pH and fluoride concentration on titanium passivating layer: Stability of titanium dioxide. *Talanta* **2001**, *53*, 1053–1063. [CrossRef] [PubMed]
49. Porcayo-Calderon, J.; Rodríguez-Díaz, R.A.; Porcayo-Palafox, E.; Martínez-Gómez, L. Corrosion Performance of Cu-Based Coins in Artificial Sweat. *J. Chem.* **2016**, *2016*, 9542942. [CrossRef]
50. Zhou, P.; Ogle, K. The Corrosion of Copper and Copper Alloys. In *Encyclopedia of Interfacial Chemistry*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 478–489. [CrossRef]
51. Hostýnek, J.J. Corrosion Chemistry of Copper: Formation of Potentially Skin-Diffusible Compounds. *Exog. Dermatol.* **2004**, *3*, 263–269. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Brain Wearables: Validation Toolkit for Ear-Level EEG Sensors

Guilherme Correia ¹, Michael J. Crosse ^{2,3} and Alejandro Lopez Valdes ^{3,4,5,6,*}¹ Department of Physics, NOVA School of Science and Technology, 2829-516 Caparica, Portugal; lourenog@tcd.ie² Segotia, H91 HE9E Galway, Ireland; mick@segotia.xyz³ Trinity Centre for Biomedical Engineering, Trinity College Dublin, D02 R590 Dublin, Ireland⁴ Global Brain Health Institute, Trinity College Dublin, D02 X9W9 Dublin, Ireland⁵ Trinity College Institute of Neuroscience, Trinity College Dublin, D02 X9W9 Dublin, Ireland⁶ Department of Electronic and Electrical Engineering, Trinity College Dublin, D02 Dublin, Ireland

* Correspondence: alejandro.lopez@tcd.ie

Abstract: EEG-enabled earbuds represent a promising frontier in brain activity monitoring beyond traditional laboratory testing. Their discrete form factor and proximity to the brain make them the ideal candidate for the first generation of discrete non-invasive brain–computer interfaces (BCIs). However, this new technology will require comprehensive characterization before we see widespread consumer and health-related usage. To address this need, we developed a validation toolkit that aims to facilitate and expand the assessment of ear-EEG devices. The first component of this toolkit is a desktop application (“EaR-P Lab”) that controls several EEG validation paradigms. This application uses the Lab Streaming Layer (LSL) protocol, making it compatible with most current EEG systems. The second element of the toolkit introduces an adaptation of the phantom evaluation concept to the domain of ear-EEGs. Specifically, it utilizes 3D scans of the test subjects’ ears to simulate typical EEG activity around and inside the ear, allowing for controlled assessment of different ear-EEG form factors and sensor configurations. Each of the EEG paradigms were validated using wet-electrode ear-EEG recordings and benchmarked against scalp-EEG measurements. The ear-EEG phantom was successful in acquiring performance metrics for hardware characterization, revealing differences in performance based on electrode location. This information was leveraged to optimize the electrode reference configuration, resulting in increased auditory steady-state response (ASSR) power. Through this work, an ear-EEG evaluation toolkit is made available with the intention to facilitate the systematic assessment of novel ear-EEG devices from hardware to neural signal acquisition.

Citation: Correia, G.; Crosse, M.J.; Lopez Valdes, A. Brain Wearables: Validation Toolkit for Ear-Level EEG Sensors. *Sensors* **2024**, *24*, 1226. <https://doi.org/10.3390/s24041226>

Academic Editor: José Machado Da Silva

Received: 18 December 2023

Revised: 9 February 2024

Accepted: 12 February 2024

Published: 15 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: in-ear EEG; hearables; brain–computer interfaces; non-invasive; electroencephalography; EEG phantom

1. Introduction

Conventional electroencephalography (EEG) is an invaluable tool for assessing neurological disorders, with measurements performed in controlled clinical environments, utilizing full cap systems over the scalp with wet electrodes that provide low impedance for quality measurements with high temporal resolution. However, researchers and clinicians have been interested in measuring EEG outside the laboratory for better assessing neurological disorders [1].

The development of technology has led to the evolution of ambulatory EEG, resulting in research on smaller EEG measurement options. This has given rise to a new category of wearable EEG devices that are wireless, have aesthetic designs limited to only the area around the head, and use dry electrodes. The convenience of these devices enables individuals to incorporate EEG measurement into their everyday lives, making it easier to assess certain medical conditions such as epilepsy and sleep disorders and allowing for brain–computer interface (BCI) applications outside of a laboratory environment. While

these types of devices are commercially available and their capabilities acceptable for some current applications, the overall design of these wearables still constitutes a barrier to daily usage, being uncomfortable over long periods of time and making it obvious when a person is utilizing such an EEG device [2,3].

Quoting Looney et al., the next generation of wearable devices must be “discreet, unobtrusive, robust, user friendly and feasible”. An in-ear EEG approach checks all these boxes, trading the wide coverage across the scalp with an inconspicuous EEG recording ability based on the ear [4]. The ability to acquire reliable brain recordings from inside the human ear is critical to accelerating the development of next-generation EEG-enabled earbuds. While ear-EEG is one of the best candidates for consumer BCI, being referred to as “beyond wearable” [2], real-world brain recordings are affected by multiple environmental factors that are not present in the laboratory. A systematic, well-established characterization of in-ear EEG hardware and signal quality is crucial to understand the limitations and applications of newly developed devices.

When characterizing a novel EEG system, it is important to be able to evaluate all components of the system. This includes the biological neural signals, derived from validated EEG paradigms like the alpha block or the various event-related potentials (ERPs) and the hardware and signal chain components (electrical and mechanical), which require testing in a controlled setting through a test-bench or phantom model [5].

In many areas of research and instrumentation, phantoms are utilized for testing, validating, and calibrating acquisition systems, namely in medical imaging like positron emission tomography (PET) and magnetic resonance imaging (MRI) scans where standardized phantoms are available [6]. In EEG, a standardized phantom does not exist, although many prototypes from different materials and designs have been proposed.

A common approach used by research groups to test novel EEG systems or new types of electrodes consists of using salt-doped ballistic gelatin (BG) head models onto which scalp-EEG sensors can be mounted for testing [7,8]. Nonetheless, phantoms can also be made from different organic (e.g., agar) or synthetic (e.g., carbon-doped thermoplastics) materials [9,10]. The basic principle for an EEG phantom prototype is to obtain a conductive physical model with the shape of the human head. This is usually achieved through image processing of anatomical scans (i.e., MRI or CT) and 3D printing of the resulting phantom’s cast for molding [11].

An EEG phantom allows for the playback of a previously recorded known EEG signal. This recognizable signal is referred to as the “ground truth”. Then, measurements can be made on the phantom for comparison with this known signal, which constitutes a way of characterizing the acquisition system without the need to account for the inherent variability and lack of repeatability of neural signals present when recording from human subjects, as well as identifying external sources of noise [12,13]. For a known previously recorded signal to be played out of the phantom, electrode antennas are driven inside it. These can be simple screws, exposed wire tips, or a coaxial cable to create dipoles [10,14,15]. The driven antennas can either be attached to the interior side of the scalp layer in hollow-shaped phantoms or be put in place during a phase of assembly where the constitution of the filling material allows for this procedure [10,16]. Additionally, EEG phantoms allow for a controlled analysis of electrode contact impedance measurements and signal noise floor characteristics [16,17].

Despite the existence and use of EEG phantoms for evaluating scalp-based EEG systems, there is a lack of literature for ear-EEG validation via appropriate phantom models. To the best of our knowledge, appropriate EEG phantoms for this purpose do not exist and all previous EEG phantoms neglect the structures of the outer ear and ear canal systematically.

The feasibility of measuring brain signals through ear-EEG devices has been validated in recent publications [4,18,19]. For an in-depth review of the field and overview of the technological state-of-the-art around ear-EEG we please refer to [20,21]. A commonality across the ear-EEG literature is that the validation paradigms utilized tend to be known

ERP paradigms, namely, the alpha modulation (or alpha blocking) paradigm, the auditory steady-state response (ASSR), the steady-state visual evoked response (SSVEP), auditory evoked potentials (AEPs), visual evoked potentials (VEPs), and oddBall-type paradigms to elicit responses linked to higher processes of the brain, like the P300 and mismatch negativity (MMN) responses. Similarly, the interest of utilizing electro-oculography (EOG) measured by ear-EEG as a possible input for BCI applications has also been explored [19].

Given the predominance of this select group of paradigms utilized to derive ear-EEG responses and the lack of a dedicated phantom model suitable for ear-based systems, here we propose a complete ear-EEG validation toolkit to contribute to and expand the characterization of this technology through:

- 1. A software framework (“EaR-P Lab”) that allows the user to readily make a validation test battery for the characterization of ear-EEG devices at the neural signal acquisition level.
- 2. The design and prototyping of an ear-EEG-suitable physical phantom for systematic characterization of in-ear sensors, allowing controlled comparison of fit form factors for ear-EEG acquisition.

This contribution to the field aims to allow for a more reliable assessment of out-of-the-box ear-EEG devices by providing proper benchmarking tools for comparing systems.

2. Materials and Methods: Ear-EEG Toolkit Design and Validation

2.1. EaR-P Lab—Design and Validation

A custom application, programmed with nine commonly utilized EEG recording paradigms, was created for the assessment of ear-EEG devices at the neural signal level. This application was written in the Python programming language and consists of a collection of scripts (requiring a local Python installation), pre-synthesized stimuli, and other auxiliary files and folders wrapped in an usable graphical user interface (GUI), based on the Tkinter Python package. The stimuli delivery and precise experiment timings are handled via PsychoPy, a Python-based open source package for experiment control [22].

We dubbed this software script “EaR-P Lab” from the amalgamation of the words “ear” and “ERP”, as per the motivation of streamlining the measuring of ERPs (as well as other paradigms) from the ear. The organizational structure and main functionalities of EaR-P Lab are depicted in Figure 1.

EaR-P Lab - validation stimuli and markers

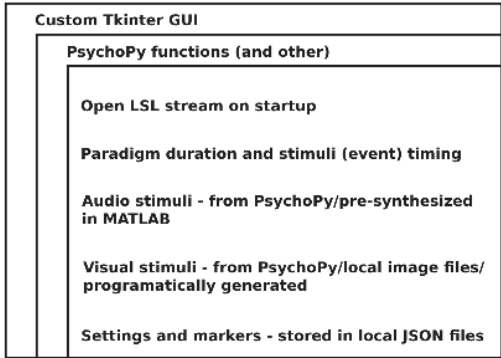


Figure 1. EaR-P Lab—structure and main attributes.

When initialized, EaR-P Lab provides a user-friendly GUI to present the necessary stimuli and insert the relevant markers that delimit different paradigms or indicate the start of a given stimulus in the validation test battery. The marker’s timestamps are synchronized with the EEG data through the Lab Streaming Layer (LSL) protocol that handles the synchronization between the EEG data and maker streams through a local

network (an EEG amplifier compatible with LSL must be used) [23,24]. The marker and EEG data streams can then be recorded with another software (i.e., LabRecorder) into a single .xdf file on the hard disk of the computer for offline analysis, as illustrated in Figure 2.

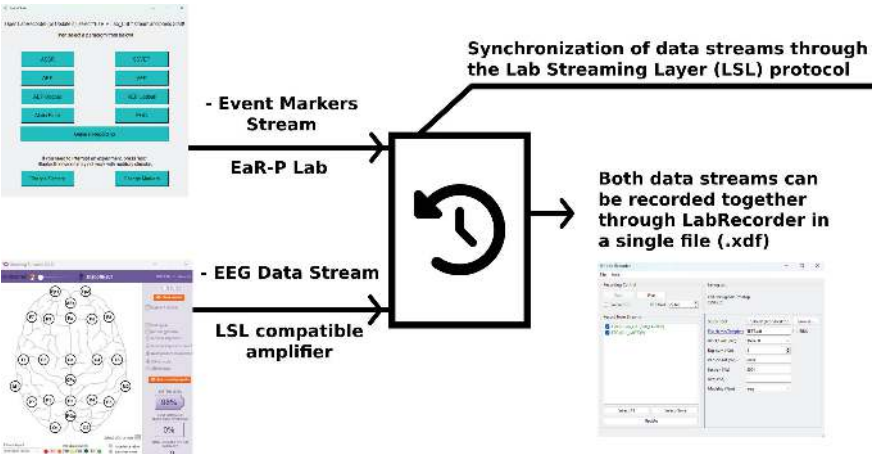


Figure 2. EaR-P Lab—schematic of functional framework.

2.1.1. GUI—Main Menu

When EaR-P Lab is launched, the main menu in Figure 3 appears. This menu gives the user access to the nine pre-loaded EEG acquisition paradigms, with the script automatically opening the LSL marker stream using the *pylsl StreamOutlet* function at startup. Eleven buttons are presented: nine that direct to the different paradigms and two (Change Settings and Change Markers) that allow users to modify specific experiment parameters. Some disclaimer information on how EaR-P Lab operates is also shown in this window. Refer to Appendix A for a full overview of the features and functions of each button and the settings menus.

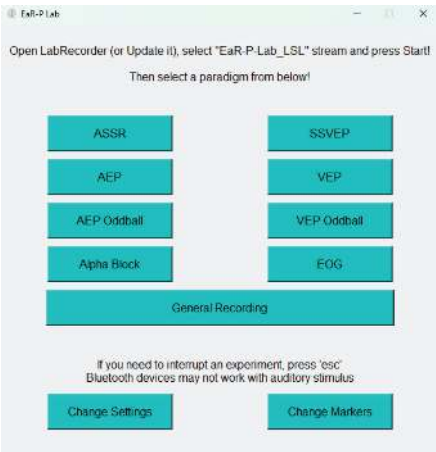


Figure 3. EaR-P Lab—main menu.

2.1.2. Stimuli and Trigger Latency

To accurately evaluate event-related potentials (ERPs), good synchronization between data and markers sent at stimulus onset (triggering) is crucial. The PsychoPy library allows sounds to be pre-scheduled with the high precision “PTB” settings and the *callOnFlip* method to start visual stimuli promptly, both used in EaR-P Lab [25]. However, unavoidable

latency and jitter issues related to screen refresh rates and monitor syncing may still arise [26].

To measure the timing characteristics of EaR-P Lab, we used an amplifier-specific accessory (i.e., mBrainTrain Delay/Jitter (DJ) box). This accessory includes an audio input, output, and photo-diode sensor that allowed us to test the marker synchronization with the audio and visual stimuli. The device was placed at the latest stage possible in the audio delivery setup, with the photo-diode centered and 30 cm away pointing to the center of the screen.

During the early testing phase, we observed an increasing cascading effect on the latency delay when recording different blocks of transient responses within the same *.xdf* file, illustrated in Figure 4.

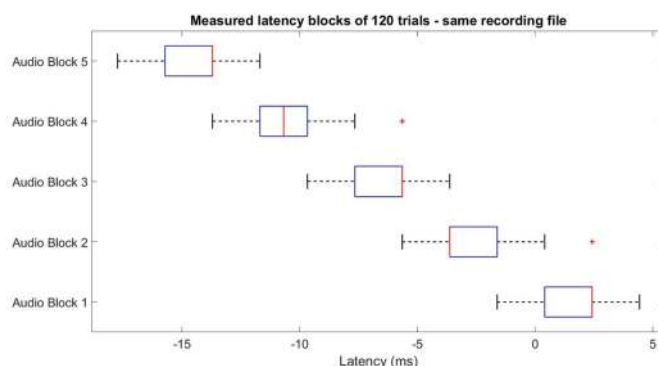


Figure 4. Latency variation when recording multiple event-related potential (ERP) blocks on the same file, exemplified for auditory stimuli—a similar effect happens for visual stimuli.

Our troubleshooting showed that to correct this effect, the data streaming from the amplifier software (mBrainTrain Streamer) had to be restarted after each EEG recording block. Hence, each block of a given paradigm must be recorded in a separate *.xdf* file to ensure reliable trigger latencies, as shown in Figure 5.

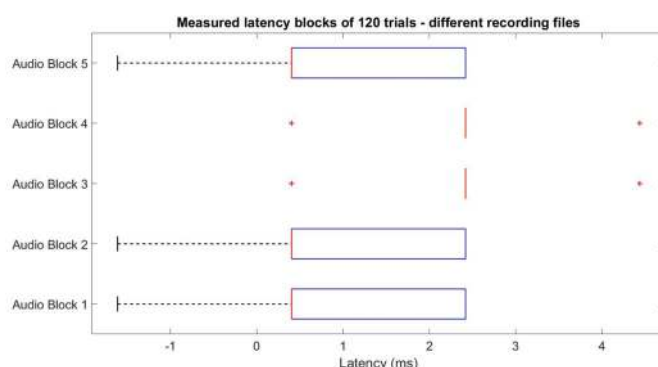


Figure 5. Nullified cascading effect is when recording multiple ERP blocks in different files after restarting data streaming, exemplified for auditory stimuli—a similar effect happens for visual stimuli.

However, to the best of our knowledge, this may only apply to the mBrainTrain Smarting Mobi recording platform. We recommend testing the system's latency before recording and having separate *.xdf* files for the AEP, VEP, EOG, and oddBall-type blocks to avoid latency synchronization issues.

Once our latency issues were resolved, the latency delay was measured for 1000 trials of both auditory and visual events, showing a mean delay of 0 ms (2.2 ms std./jitter) and 21 ms (2.8 ms std./jitter), respectively. The processing of the EEG paradigms accounted for these measured latencies. These values are specific to the laptop used to present the stimuli and are likely to vary between systems.

2.1.3. Data Acquisition and Test Battery

In order to validate EaR-P Lab as a tool for assessing ear-EEG (and EEG in general), standard scalp-EEG data were acquired for each of the EEG paradigms. EEG data were collected from five test subjects with normal or corrected-to-normal vision and normal hearing (age avg. 27.0 ± 4.3 std.). A custom EEG cap with 12 channels was used for the benchmark condition. The cap uses conventional Ag/AgCl electrodes at the F3, Fz, F4, T7, C3, Cz, C4, T8, P3, Pz, P4, and Oz positions. The reference electrode (CMS) and ground (DRL/GND) were located at FCz and AFz, respectively. Conductive gel (Signagel, Parker) was applied to each scalp electrode, typically achieving impedances of <5 k Ω , measured by the mBrainTrain Streamer. EEG data were recorded using the mBrainTrain Smarting Mobi amplifier—capable of high-quality wireless recordings—at a sampling rate of 500 Hz with wireless data transmission, providing 24-bit resolution in the analog-to-digital converter (ADC).

A laptop (ASUS VivoBook 15, Windows 11 OS, Intel Core i7-10750H Processor, NVIDIA GeForce GTX 1650 Ti Graphics Card, 12 GB RAM) ran the essential recording software (mBrainTrain Streamer and LabRecorder) and stimulus presentation software (EaR-P Lab). Data were transmitted via Bluetooth from the amplifier secured at the back of the subject's head to a USB dongle connected to the laptop. Auditory stimuli were delivered using research-grade ER-2 earphones fitted with foam ear tips (Etymotic Research, INC., Elk Grove Village, IL, USA), shielded with copper tape to reduce interference. To ensure that the stimuli were audible without being distorted, an external sound card (TASCAM US-100) and digital-to-analog converter (DAC, FiiO Alpen2) were used (see Figure 6). The laptop and DAC were set to the maximum volume, with adjustments made to the sound card only. Subjects were instructed to adjust the volume to a comfortable, audible level. The laptop screen brightness setting was set to the maximum for visual stimuli.

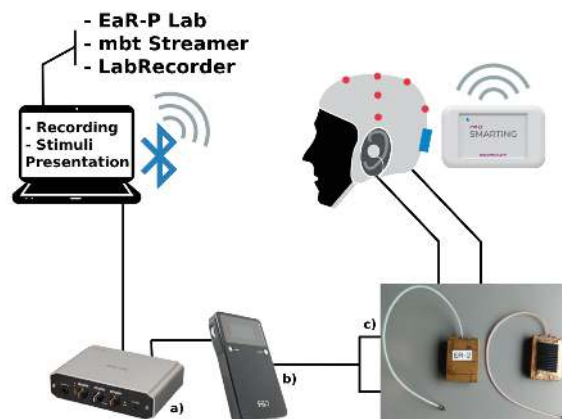


Figure 6. EEG acquisition setup schematic and equipment: (a) USB audio interface TASCAM US-100; (b) digital-analog converter (DAC) amplifier FiiO Alpen 2; (c) ER2 etymotic tubal-insert research-grade earphones.

EEG recordings were conducted in a dedicated sound-attenuated room with only the participant and the technician present. Participants sat in a comfortable chair and were instructed to relax and minimize motor and eye movements during recordings. Participants

were given a brief explanation of the recording process at the start of the session, and detailed instructions were provided at the beginning of each task.

The Ear-P Lab EEG paradigms employed were as follows:

1. **Resting State:** 4 min of resting-state EEG recording (Figure A1)
2. **Alpha Block:** 4 min of eyes open/closed, 1 min per block (Figure A3)
3. **ASSR:** 4 min of continuous auditory stimulation, 1 kHz carrier signal with 40 Hz amplitude modulation (Figure A1)
4. **SSVEP:** 4 min of continuous visual stimulation, 10 Hz flickering radial checkerboard, subjects seated at a distance of 60 cm from the center of the screen with room lights turned off (Figure A2)
5. **AEP:** 200 trials of discrete auditory events, 1 kHz pure tone of 200 ms duration with 10 ms rise/fall time, interstimulus interval (ISI) between 1200 and 1800 ms, total duration of 7–8 min (Figure A4)
6. **VEP:** 200 trials of discrete visual events, pattern-reversal radial checkerboard of 500 ms with 500 ms ISI, total duration of 5 min (Figure A5)
7. **AEP OddBall:** 200 trials of discrete standard/deviant auditory events (standard: 440 Hz pure tone; deviant: 880 Hz pure tone), 100 ms duration with 10 ms rise/fall time and 1200–1800 ms ISI, total duration of about 15 min
8. **VEP OddBall:** 200 trials of discrete standard/target visual events (standard: blue square; target: red circle), 500 ms duration with 600–700 ms ISI, total duration of about 18 min, subjects instructed to respond to target with button press (Figure A6)
9. **EOG:** 80 trials of discrete visual events, 500 ms duration dot movements with 1000–1600 ms ISI, total duration of about 10 min, subjects seated at a distance of 30 cm from the center of the screen (i.e., visual angle of 16.2°); subject's head was stabilized using an adjustable chin rest and the monitor was centered with the subject's eyes (Figure A7).

Each acquisition session took about 2 h to complete from start to finish, including EEG cap setup. Subjects were entitled to short breaks (2–5 min) at the end of each paradigm but were also allowed longer breaks when requested. For some subjects, the protocol's order was not strictly followed.

2.1.4. Data Processing and Statistical Analysis

All data preprocessing and analysis was performed via custom scripts in MATLAB (MATLAB (R2020a), The MathWorks Inc., Natick, MA, USA). EEG data were first highpass-filtered (Butterworth IIR), notch-filtered (50 Hz second-order IIR), and then lowpass-filtered (Butterworth IIR). Frequency-based paradigms (ASSR, SSVEP, Alpha Block) were bandpass-filtered between 1 and 100 Hz, and transient responses (AEP, VEP, MMN, P300, EOG) were bandpass-filtered between 1 and 20 Hz. To isolate the EOG blinks in the respective three-second window, these data were bandpass-filtered between 0.2 and 3 Hz. All EEG data were re-referenced to Cz.

ASSR and SSVEP responses were extracted by computing the power spectral density via Welch's method using an 8 s window with 50% overlap, and the alpha block spectrograms were obtained by computing the short-time Fourier transform (STFT) using a 2 s window with 50% overlap.

All transient responses were baseline-corrected based on the 100 ms pre-stimulus interval and then averaged across trials. To obtain the MMN and P300 responses, the average standard cue waveform was subtracted from the average target cue waveform for each subject, with the standard events immediately before the oddBalls being considered [27].

To measure the signal-to-noise ration (SNR) of the ASSR responses, the power at 40 Hz was compared with the surrounding frequency bins, according to Equation (1) [18]:

$$40 \text{ Hz SNR} = \frac{P(40 \text{ Hz})}{P_{\text{average}}(35 - 45 \text{ Hz})^*}, \text{ *excluding 40 Hz} \quad (1)$$

The same formula was adapted for the 10 Hz SSVEP responses as per Equation (2):

$$10 \text{ Hz SNR} = \frac{P(10 \text{ Hz})}{P_{\text{average}}(5-15 \text{ Hz})^*}, \text{ *excluding } 10 \text{ Hz} \quad (2)$$

An *F*-test was also conducted on these ratios using a one-way analysis of variance (ANOVA).

Equation (3) was used to calculate alpha power by quantifying the power ratio of the eyes-closed and eyes-open conditions between 8 and 12 Hz.

$$R_{AM} = \frac{P_{\text{average}}(\text{Alpha Band}_{\text{Eyes Closed}})}{P_{\text{average}}(\text{Alpha Band}_{\text{Eyes Open}})}, \text{ Alpha Band } (8-12 \text{ Hz}) \quad (3)$$

Significant modulations in the average ERP amplitude were detected using one-sample (two-tailed) *t*-tests and are indicated in each of the plots by the green highlighted segments ($p < 0.05$, not corrected for multiple comparisons) [28].

EOG profiles were quantified by the amplitude and polarity of each movement type. For the EOG blinks, an amplitude factor between the averaged peak-to-peak amplitude of the intentional blinks over the regular blinks was established [18].

2.1.5. EaR-P Lab Validation

The following section describes the scalp-EEG data used to validate EaR-P Lab's EEG paradigms. Each validation analysis was conducted at both the typical scalp site for a given test (e.g., electrode Oz for a visual stimulus), as well as a location close to the ear (electrode T8) for ear-EEG validation.

Alpha Block

As seen in Figure 7, a distinct effect on alpha wave blocking (with suppression of the 8–12 Hz band occurring during eyes open) was achieved with the highest modulation at Oz (6.1 dB), dropping to 3.7 dB near the ear at T8.

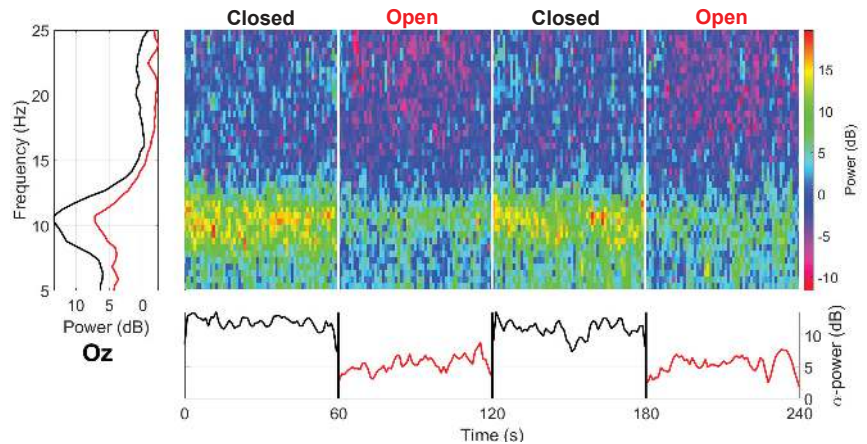


Figure 7. Grand average spectrogram for the alpha block paradigm at Oz (Cz referenced). The bottom horizontal plot shows the mean alpha power (8 Hz) as a function of eye state, while the left vertical plot shows the frequency response for the two conditions.

ASSR

For the ASSR (Figure 8), the SNR at the 40 Hz modulation frequency was measured, with the highest mean SNR at P4 (9.9 dB), dropping to 4.7 dB at T8.

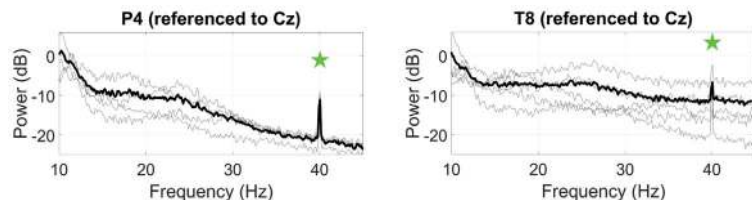


Figure 8. Grand average ASSR responses (black line) to a 40 Hz AM auditory stimulus at P4 (left) and T8 (right). Statistically significant peaks are highlighted by the green star token, based on an *F*-test ($p < 0.05$), grey lines represent individual responses.

SSVEP

The SSVEP at the 10 Hz modulation frequency was most pronounced at Oz with a mean SNR of 11.0 dB (Figure 9), as well as harmonics of the 10 Hz fundamental frequency up to the eighth harmonic present. The SSVEP was weaker at T8, with a mean SNR of 7.5 dB, with harmonic responses only distinguishable up to 40 Hz.

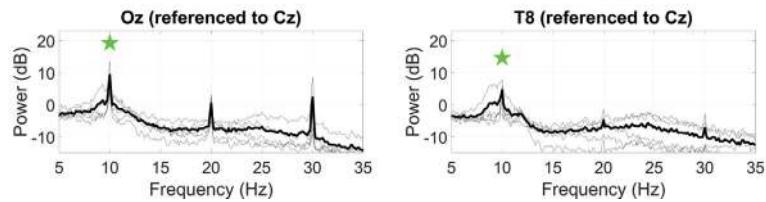


Figure 9. Grand average SSVEP responses (black line) to a 10 Hz visual stimuli at Oz (left) and T8 (right). Statistically significant peaks are highlighted by the green star token, based on an *F*-test ($p < 0.05$), grey lines represent individual responses. Only the first harmonic was statistically evaluated.

AEP

The AEP response to the 1 kHz tone is shown in Figure 10, comprising three significant deflections corresponding to the P1, N1, and P2 AEP components. The largest deflection was around 100 ms with an amplitude of 5 μ V, corresponding to the N1 component (positive due to re-referencing), and a peak-to-peak amplitude of 8.9 μ V between the N1 and P2 components.

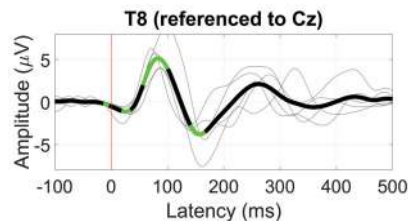


Figure 10. Grand average AEP waveform (black line) at T8. Statistically significant segments are highlighted in green, based on *t*-tests ($p < 0.05$, not corrected for multiple comparisons), grey lines represent individual responses.

VEP

The VEP response to the pattern-reversal radial checkerboard is shown in Figure 11. At Oz, a significant component is present between 150 ms and 300 ms with a peak amplitude of 15 μ V at 200 ms, corresponding to the P2 VEP component. At T8 (near the ear), the P2 component was not significantly different from zero, instead showing two smaller positive deflections of 2 and 3 μ V at around 90 and 300 ms, respectively.

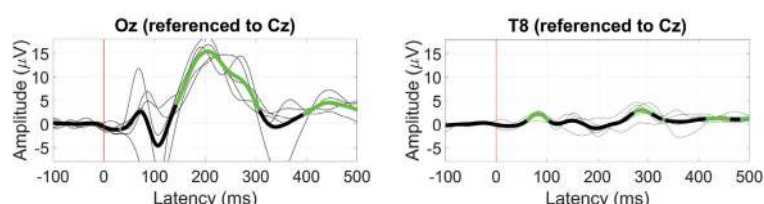


Figure 11. Grand average VEP waveform (black line) at Oz (**left**) and T8 (**right**). Statistically significant segments are highlighted in green, based on t -tests ($p < 0.05$, not corrected for multiple comparisons), grey lines represent individual responses.

AEP OddBall (MMN)

The MMN response was the strongest at electrode Pz, shown in Figure 12. At 200 ms, the expected negative differential between the standard and deviant ERPs was observed with an amplitude of 2 μ V and was also present at electrode T8 with a similar amplitude.

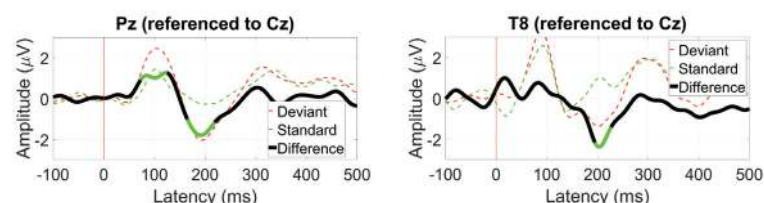


Figure 12. Grand average mismatch negativity (MMN) waveform at (**left**) Pz and (**right**) T8. Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons).

VEP OddBall (P300)

The P300 response to a target visual stimulus was better observed at P4 at around 280 ms with an amplitude of 6 μ V (Figure 13). In contrast, the P300 response at T8 only peaked at 4 μ V with that amplitude achieved at around 250 ms from the onset.

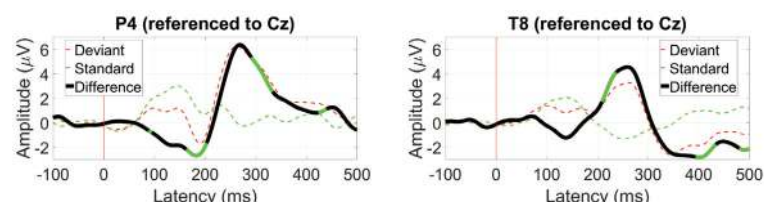


Figure 13. Grand average P300 waveform at (**left**) P4 and (**right**) T8. Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons).

EOG (Blinks and Saccades)

While not cortical in origin, EOG responses to blinking and saccade movements were also assessed as possible inputs for BCI applications. EOG responses are naturally more prominent at frontal scalp electrodes such as Fp1 and Fp2, as they are near the eye muscles responsible for these movements. In our scalp-EEG validation analysis, we considered F3 and F4 as approximate locations for Fp1 and Fp2.

Figure 14 shows the peak-to-peak EOG amplitude at electrode F3 for hard blinking (8 mV) and soft blinking (1.5 mV) for an example subject, with a lower absolute peak-to-peak EOG amplitude at T8 (but a similar soft/hard ratio).

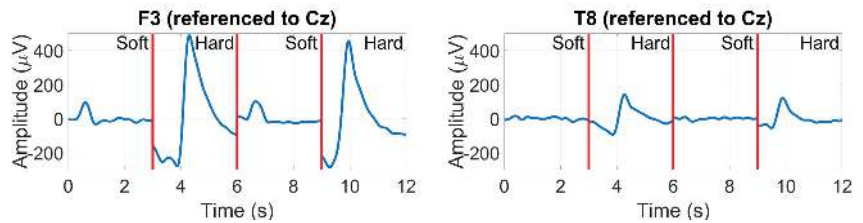


Figure 14. EOG amplitudes for soft and hard blinks in an example subject recorded at F3 (left) and T8 (right).

Figure 15 shows the saccade profiles in different directions. The saccades were assessed at around 200 ms after stimulus onset, where the separation between directions is greatest. Between T7 and T8, it is possible to see a 10 μV difference between the right (blue trace) and left (green trace) saccades, with their polarity inverting, depending on the side of the sensing electrode. The vertically oriented saccades are less differentiated, with only around 2 μV separating their traces.

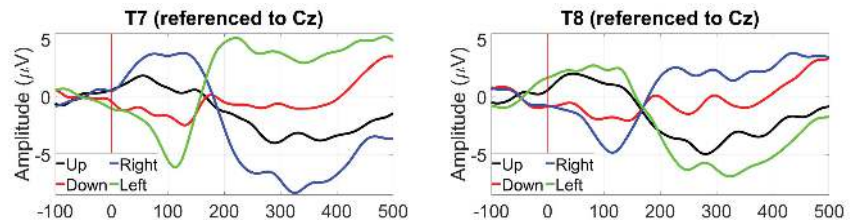


Figure 15. Grand average saccade profiles in the four cardinal directions, at T7 (left) and T8 (right).

2.2. Ear-EEG Phantom—Design and Validation

A physical test bench to evaluate ear-EEG devices was developed. The basis for the prototype was similar to the most common type of EEG phantoms: a mold that creates the proper shape, is fit for use with an ear-EEG device, and can be filled with a conductive mixture.

In contrast to most common EEG phantoms, this work focused on developing a compact design solely focused on the ear anatomy instead of a complete head-shaped phantom. We also set out to create a modular ear section for the mold, allowing for testing different ear form factors without having to rebuild an entirely new mold, as well as having a way to secure the signal antennas in place during the setting process of the filling material.

The phantom was designed using the Fusion 360 computer-aided design (CAD) software and then exported as an *.stl* file for 3D printing. The base design is an 18 cm modular cylinder split into two parts along its length with open ends. The bottom half has two longitudinal railings where the top fits, leaving two round holes centered in the middle for inserting antennas. The cylinder sides can then be closed with the lids. The top has two square openings for pouring of the conductive mixtures. A 0.1 mm gap was used for all fitting parts to account for the 3D printing resolution. The rendered prototype is shown in Figure 16, and the phantom prototype's full dimensions are presented in Appendix B.

The inside of these lids can then accommodate the outer ear canal and concha shape to create an imprint on the material inside, where an ear-EEG device can be placed for testing. The current methodology used ear canal scans by a professional audiologist, digitally scanned on-site, and delivered as an *.stl* file (as shown in Figure 17), which can be imported as a mesh to Fusion 360 and easily fused with the lid mesh by using the *Mesh Menu - Modify - Combine - Join Operation*.

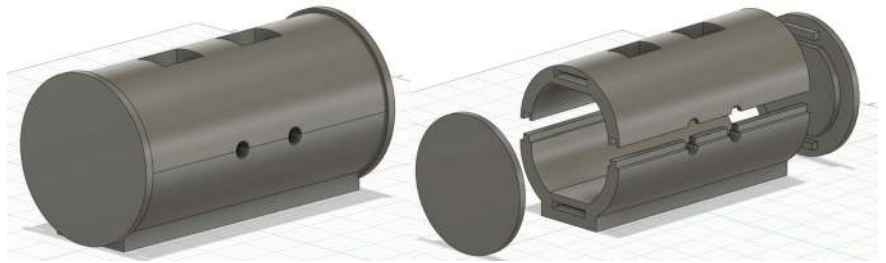


Figure 16. CAD drawings of ear-EEG phantom mold casing. Closed render of the mold (left). Exploded render of the mold (right).

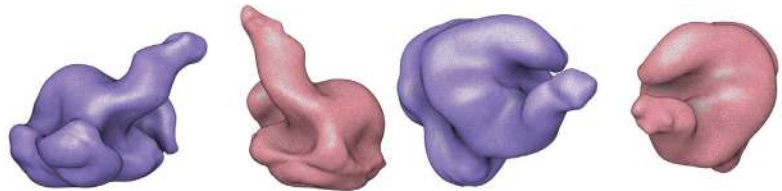


Figure 17. Outer ear scans (blue: left ear, red: right ear) from an example subject shown in elevation (left) and plan (right) view. The scans were obtained by an expert audiologist and digitized as .stl files.

Ear canal scans were centered on the lid's mesh with the concha cymba outline parallel to the mesh's top/bottom. Meshes intersected to an adequate depth as long as the ear canal and concha structures were visible and protruding, similar to Figure 18.

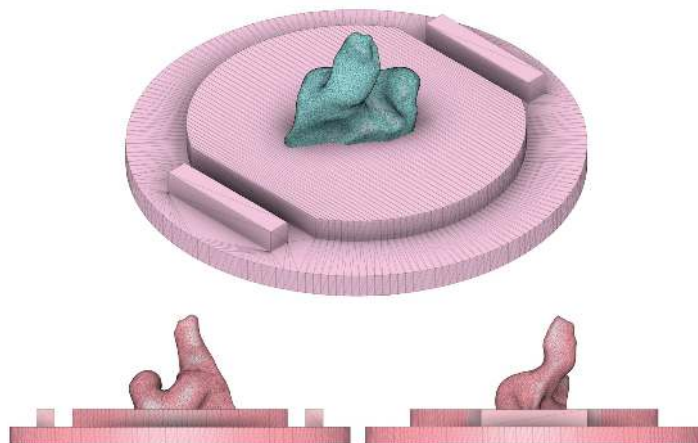


Figure 18. Example of a left ear scan being centered and oriented with the phantom's lid mesh. Different views of the alignment and depth of the ear mesh and the lid mesh into a single rendered object are shown below.

The ear-EEG phantom prototype was 3D-printed on a Prusa i3 MK3 3D printer using polylactic acid (PLA). PLA was chosen due to its ease of use, fast printing times, affordability, non-warping nature, and lack of post-processing requirements. The printing settings on the Prusa software were set to a 0.15 mm *QUALITY* printing resolution, using *Prusament PLA* filament with a default structural infill of 15%. The software automatically added necessary support structures where required. The combined printing time of the phantom body and

lids was approximately 33 h. To minimize difficult-to-remove supports, the cylinder halves were printed vertically. If the phantom's form-fitting factor needs to be changed, printing just a set of lids with the ear impression takes only 6 h. Figure 19 shows the different modular parts after printing and support removal.



Figure 19. Disassembled ear-EEG phantom: bottom half (yellow), top half (white), and two lids with a left and right ear imprint from one of the test subjects.

2.2.1. Phantom Assembly and Bulk Materials

To assemble the phantom, the antennas (pair of 3.5 mm AUX cables) were adjusted to fit the diameter of the cylinder holes by inserting one side of each cable inside PVC tubing and sealing it with duct tape. The cables were placed on the bottom half, then joined by the top and the lids with ear imprints on the side, holding the prototype together. Three strips of plumbing tape were used on the railing fittings, cables, and inner circumference of the phantom lids to prevent leaks, as in Figure 20. The phantom could then be filled with a conductive substance.



Figure 20. Ear-EEG phantom assembly—antennas and railing fittings were sealed with tape.

Here, three materials were tested with our design: agar doped with salt, BG doped with salt, and silicone doped with carbon fiber (CF), as a non-perishable option.

Agar

Agar is a gelatinous vegan substance derived from certain types of seaweed, making a stable gel structure when cooled down after being boiled. An agar phantom can be created

by mixing agar, water, and salt to produce the recommended weight percentages of 4% and 0.5% for agar and salt, respectively, as per Equation (4) [9]:

$$\begin{cases} \frac{A}{A+w+s} = 4.0\% \\ \frac{s}{A+w+s} = 0.5\% \end{cases} \Rightarrow \begin{cases} A = \frac{8}{191}w \\ s = \frac{1}{191}w \end{cases} \quad (4)$$

where A is the weight of agar, s the weight of salt, and w the weight of water.

The phantom was estimated to have a capacity of 700 mL (equivalent to 700 g of water). By substituting w for that value in the above equations, we obtain the necessary quantities of 30 g of agar (*Hoosier Hill Farm*, €10.20 for 115 g from Amazon) and 4 g of salt, rounded up to the nearest unit. The mixture was prepared as follows:

1. Boil 700 mL of regular tap water (or deionized water)
2. Add 30 g of agar slowly while stirring the mixture
3. Add 4 g of table salt while stirring until no granules are present (keep mixing while letting it cool down at room temperature for 10 min)
4. Pour the mix into the assembled phantom through the top vents until the liquid reaches half the vent's height and let it sit in a refrigerator until it fully solidifies (minimum 2 h, preferably overnight)

The resulting agar-based phantom is shown in Figure 21, left.

Ballistic Gelatin (BG)

BG is a common material used for making EEG phantoms, typically consisting of cattle/pork gelatin. Various weight percentage combinations of both gelatin and NaCl are shown to be an appropriate conductive medium [16]. Due to the limited accessibility of gelatin bulk powder (*YouHerblt* 240 Bloom, €20.00 for 350 g of powder from Amazon), values within the middle range of the above were used, with a 15% weight percentage of gelatin and 8% weight percentage of NaCl, which yield Equation (5):

$$\begin{cases} \frac{BG}{BG+w+s} = 15.0\% \\ \frac{s}{BG+w+s} = 8\% \end{cases} \Rightarrow \begin{cases} BG = \frac{15}{77}w \\ s = \frac{8}{77}w \end{cases} \quad (5)$$

The same steps as for agar were followed to prepare the gelatin solution, with 140 g of gelatin powder mixed with 70 g of salt in 700 mL of hot (not boiling) water. Unlike agar, gelatin is better mixed at a lower temperature to prevent the formation of air pockets.

The resulting BG-based phantom is shown in Figure 21, right.

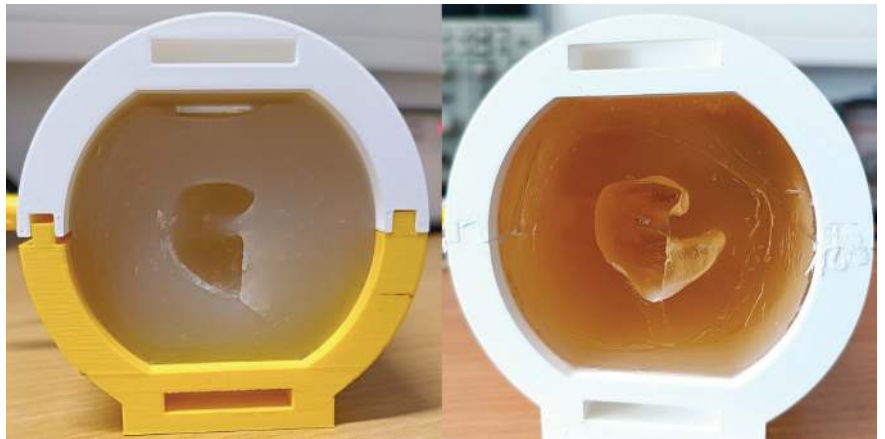


Figure 21. Ear-EEG phantoms made with agar (left) and ballistic gelatin (BG) (right).

Carbon Fiber-Doped Silicone (CF)

A non-perishable phantom that can maintain signal integrity indefinitely was also envisioned, as the aforementioned organic materials tend to deteriorate over time.

Conductive fillers have been tested to make silicone rubbers conductive, with carbon fibers found to be the most effective even at a 0.5% weight percentage measuring a 2 k Ω resistance. Increasing to 1% further lowered the resistance to 200 Ω , and higher percentages gave similar resistances, indicating that the percolation threshold (filler concentration at which a conductive network is fully established) was reached [29].

A step-by-step approach for doping the silicone was followed [30], which used the following materials:

- Chopped carbon fibers, 3 mm in length, €30.00 for 500 g from Amazon
- Two-part A/B system platinum-curable silicone, mixing ratio of 1:1, €23.00 for 630 mL from Amazon (two were required for the phantom)

The silicone was estimated to have a density of 1.11 g/cm³. Various carbon fiber weight percentages were tested according to Equation (6) to establish a percolation threshold:

$$\frac{CF}{CF + S} = X\% \text{ (where } X = 0.5, 1.0, 1.5, 2.0 \dots \text{)} \quad (6)$$

where CF is the weight of carbon fibers and S is the weight of silicone. It was observed that when the weight percentage of CF exceeded 1.0%, the mixture became too dense and unsuitable for molding with the current mixing methods. Therefore, only CF weight percentages of 0.5% and 1.0% were considered and prepared according to Equation (7):

$$\begin{cases} CF = \frac{S}{199}, \text{ for } X = 0.5\% \\ CF = \frac{S}{99}, \text{ for } X = 1.0\% \end{cases} \quad (7)$$

This non-perishable phantom was ultimately made for a CF weight percentage of 1.0%, with 8 g of carbon fibers (rounded to nearest unit) being used for a total of 700 mL of silicone, following the next steps [30]:

1. Measure 8 g of carbon fibers into a disposable cup (use a mask and gloves when handling carbon fibers)
2. Wet the carbon fibers with a small amount of rubbing alcohol, spread them around, and let it almost entirely evaporate (to release strands of hair that surround the carbon fibers)
3. Add the carbon fibers to 350 mL of part A silicone and mix thoroughly until the mix presents a grey/blueish tint (an electric mixer with a wider spatula attachment was used)
4. Add 350 mL of part B silicone and keep mixing for up to 25 min until it reaches the same tint
5. Pour into the phantom casing equally through each vent and let cure for 6 h.

To characterize the conductivity profiles of our chosen materials, fine strips of each material (30 $\times$ 10 $\times$ 2 mm) were prepared specifically for conductivity testing. Both ends of each sample were coated with conductive silver ink and the standard two-wire method was employed to measure the resistance across the sample. The conductivity of each sample was then calculated according to Equation (8):

$$\begin{cases} \rho = R \frac{A}{L} \\ \sigma = \frac{1}{\rho} \end{cases} \Rightarrow \sigma = \frac{L}{R \times A} \quad (8)$$

where ρ is the material resistivity, R is the measured resistance, A is the area of the cross-section of the sample, and L is the sample's length. σ is the material's conductivity as the inverse of resistivity. The measured conductivity values for each material are listed below in Table 1.

Table 1. Electrical conductivity (Siemens per meter) for samples of agar, BG, and CF-doped silicone (1%) as the proposed materials for the ear-EEG phantom.

	Conductivity [S/m]
Agar	0.309
BG	0.918
CF (1%)	14.035

While the sample preparation for the CF-doped silicone samples was successful, the mixing method employed was not suitable for a homogeneous distribution of carbon fibers in the larger phantom mold. This resulted in the creation of conductive (darker) and non-conductive (translucent, low to no carbon fibers) areas, as illustrated in Figure 22. Therefore this material was omitted from further testing.

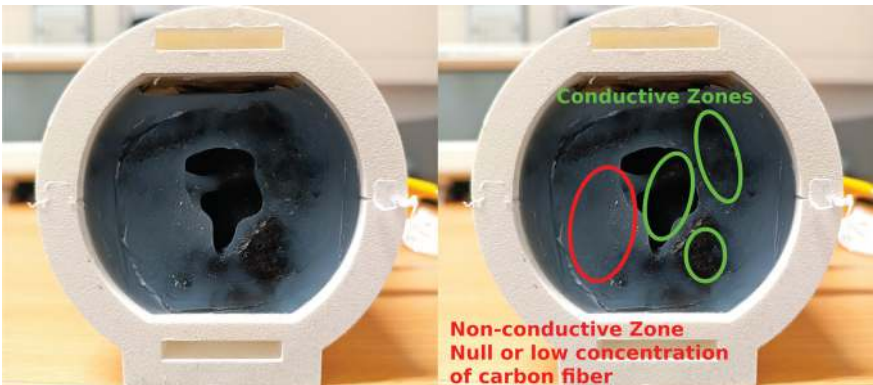


Figure 22. CF-doped silicone ear-EEG phantom—the lack of conductive homogeneity is highlighted on the right, with conductive and non-conductive zones visible.

2.2.2. Ear-EEG Phantom Testing Protocol and Setup

The agar and BG phantoms were tested over eight days, alternating daily for a total of four sessions each (i.e., “Day 4” is the last testing day per phantom), and refrigerated while not in use (4 °C). All phantoms were made using ear scans from the same subject such that the same pair of EEG earbuds could be used across all tests.

The phantom was first weighed without the side lids and with the AUX cables stacked on top. The antennas were connected to a signal generator, simulating a 10 Hz (alpha band) 100 mV square wave. The oscilloscope probe was jammed in the side of the phantom to measure the signal through the material alone and test for the material’s integrity. Then, both earpieces were inserted, and electrode contact impedances were measured using an analog impedance meter (D175, Digitimer), capable of measurement values up to 50 kΩ. The impedance meter’s reference electrode was connected to electrode Ex1. The oscilloscope probe was then connected directly to each electrode to visualize the simulated signal, which allowed us to test if any electrode was not functioning or had a diminished signal amplitude compared to the one measured through the material. To simulate an ear-EEG acquisition, the earbuds were then connected to an OpenBCI Cyton amplifier (8 channels, 250 Hz sampling rate, 24-bit ADC) and connected as per the recommended setup [31]. Ex1 was assigned to the bottom SRB2 pin (reference), and Ex2 was assigned to the bottom BIAS pin (noise-canceling) a priori. The remaining electrodes were connected in ascending order to the bottom N1P to N6P analog input pins (Ex3 is N1P and Ex8 is N6P on the board). Impedance values were also rechecked through the OpenBCI GUI. Then, 20 s recordings were obtained through the amplifier on each ear, with a gain of 24: a recording with the simulated wave and an acquisition with no signal to assess the recorded noise floor at each electrode, both offline bandpass-filtered between 0.3 and 100 Hz.

The protocol was repeated after applying conductive paste (Ten20) to each ear electrode. The testing setup is illustrated in Figure 23.

The RMS noise floor was calculated and the alpha simulation SNR was calculated (with 1% confidence) as per Equation (2).

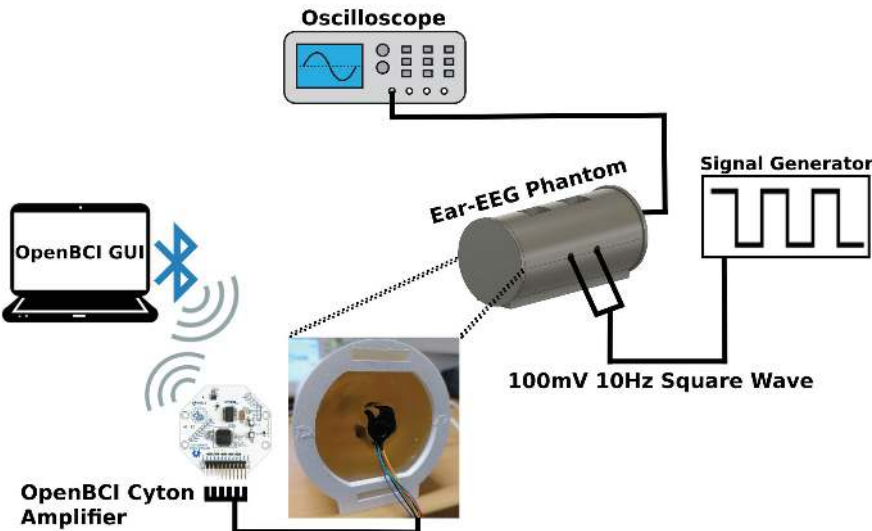


Figure 23. Schematic of testing setup of the proposed ear-EEG phantom.

2.2.3. Phantom Integrity and Durability

When measuring the signal through the phantom, it is important to track the usability of this tool over time. This was achieved by monitoring the net weight of each phantom and measuring the input signal through the medium over several days.

Table 2 shows that by Day 4, the agar and BG phantoms lost no more than 10 g in weight (with a higher percentage loss in the agar). The reason for this weight loss can be attributed to the evaporation of water in both materials.

Table 2. Measured mass (g) of the agar and BG ear-EEG phantoms over the testing days.

	Day 1	Day 2	Day 3	Day 4
Agar	855	851	850	845
BG	963	959	958	956

Table 3 shows the measures of signal integrity for each phantom over a one-week period. The agar phantom suffered the most significant amplitude loss, decreasing from an initial amplitude of 100 mV to an amplitude of 36 mV by day 4. The BG phantom signal reduced by 50% of its initial amplitude of 80 mV to 40 mV due to the drying of the material. Note, the phantoms endured extensive pilot testing before the first day of measurements and thus, were kept in sub-optimal refrigerating conditions, leading to the observed drying of the materials.

Table 3. Measured signal amplitude (mV) at the sides of each ear-EEG phantom directly through the material as a measure of signal integrity on agar and BG.

	Day 1	Day 2	Day 3	Day 4
Agar	44	40	40	36
BG	80	52	52	40

2.2.4. Electrode Impedance

Figure 24 shows an example of electrode contact impedances (measured via the OpenBCI Cyton board and validated via the analog impedance meter) for wet- and dry-electrode conditions taken during the agar phantom’s second day of recordings. Impedance measurements show a clear improvement across electrodes when applying conductive paste to the electrodes, reliably bringing the impedance under 10 kΩ, which is an acceptable level for EEG recordings (the same effect happened for BG measurements, with overall lower values than agar). From these data, we can see that channels ER3, EL3, EL7, and EL8 did not achieve suitable impedances in the dry-electrode condition, with values greater than 50 kΩ, indicative of no or very poor electrode contact. This effect was consistent across testing sessions (suggesting ER4 as a better channel than ER3 for dry recordings, at least on this specific subject). However, applying conductive paste to these electrodes reduced the impedances to a suitable level.

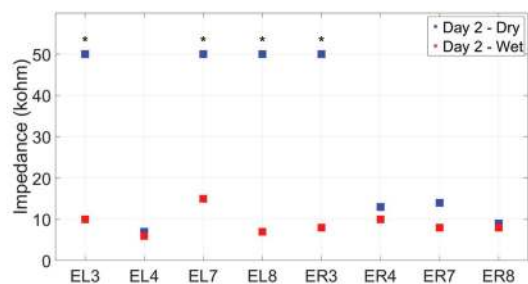


Figure 24. Contact impedance measures (kΩ) for the agar ear-EEG phantom for wet- and dry-electrode conditions (taken on Day 2 of testing). * indicates electrodes that surpassed an impedance of 50 kΩ in the dry condition.

2.2.5. Noise Floor Measurements

As with the impedance measurements, noise floor measurements (without the simulated input signal) greatly benefited from the use of conductive paste on each electrode. As seen in Figure 25, applying conductive paste to the electrodes lowered the noise floor close to or under 1 μVrms across all channels. These observations indicate that the electrode material and shape are adequate for wet-electrode recordings. Comparing the left- and right-ear results in the dry condition, we would predict that the left ear is more prone to noise (with RMS values on the order of hundreds of μV) when compared to the right-ear electrodes, suggesting better performance in the right earpiece of this particular subject.

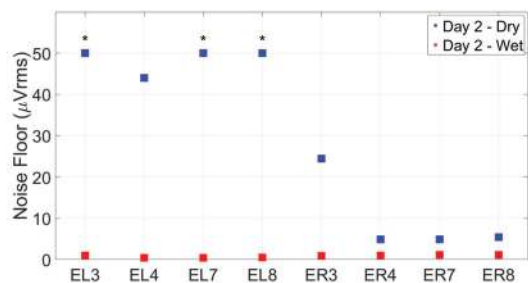


Figure 25. Noise floor measures (μVrms) for the agar ear-EEG phantom for wet- and dry-electrode conditions (taken on Day 2 of testing). * indicates electrodes that surpassed a noise floor of 50 μVrms in the dry condition.

2.2.6. Alpha-Wave Simulation

When simulating a known signal into the phantom through the two antennas, the 10 Hz input signal (to simulate alpha-band activity) was successfully recorded at all ear electrode sites on Day 1 of testing in agar and BG. As expected, the modulation SNR was higher in the ear canal electrodes (e.g., ER8), as they are further away from the within-ear reference (ER1) than, for example, ER3 or ER4.

In Figure 26, the resulting alpha modulation spectra at ER8 are present, with clear peaks at the 10 Hz and 30 Hz odd harmonic frequencies that compose the square wave. The noise floor of the spectra resembles what would be expected from an EEG recording, although that was not always the case in the following days of testing when the agar or BG phantom noise floor readings started to increase. Nonetheless, the phantom antenna signal delivery and materials proved effective in simulating a given neural signal, and for this purpose, provide a feasible validation framework.

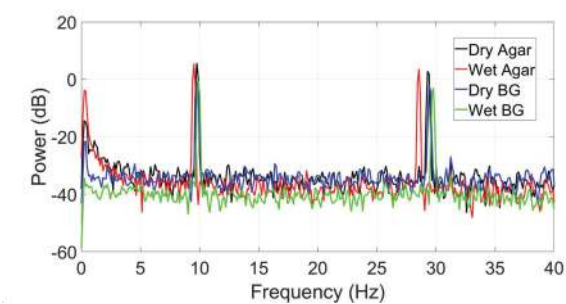


Figure 26. Power spectrum (dB) showing the synthetically generated alpha wave (10 Hz input signal) recorded using a custom ear-EEG device (electrode ER8) for the agar and BG phantoms in dry and wet-electrode conditions.

3. Results: Toolkit Use Case: Validation of an Ear-EEG Sensor

3.1. Ear-EEG Devices and Setup

The tested ear-EEG devices were developed by industry partner Segotia and fabricated by a third-party vendor (Figure 27). Each earbud was custom-fit to the participant’s ear (five subjects total, age avg. 36.2 ± 10.6 std.), based on a wax mold of their outer ear taken by a professional audiologist. This mold was scanned and digitized as an .stl file, digitally rendered into a discrete earbud form factor, and 3D-printed at a very high quality. On the external side, the earbud boasts an earphone insertion vent for sound delivery.



Figure 27. EEG earbuds developed by Segotia. Internal side of the tested earbuds (left). External side of the tested earbuds (right). Sensors are numbered in order of signal channel acquisition 1–8.

Each earbud contains eight 2×1 mm sintered Ag/AgCl disc electrodes by Easycap GmbH, Germany (Figure 27). One is located in the concha cymba area (1), four around the concha cavum (2–5), one near the antitragus (6), and two inside the ear canal (7 and 8) oriented in the posterior and anterior directions, respectively. Thin wires were welded to the electrodes, which exit the earbud near electrode 6. The electrode cabling was carefully color-coded for connecting to the amplifier (see Figure 27, left).

The ear-EEG data were acquired under wet- and dry-electrode conditions, simultaneously with conventional wet scalp-EEG using the same amplifier to provide a direct comparison with our scalp-EEG benchmark. However, since this amplifier only supports 24 recording channels outside the CMS and DRL, four scalp channels (F3, F4, P3, P4) from the control group were disconnected to accommodate all ear-EEG electrodes on the amplifier. The resulting ear- and scalp-EEG electrode configurations used are shown in Figure 28.

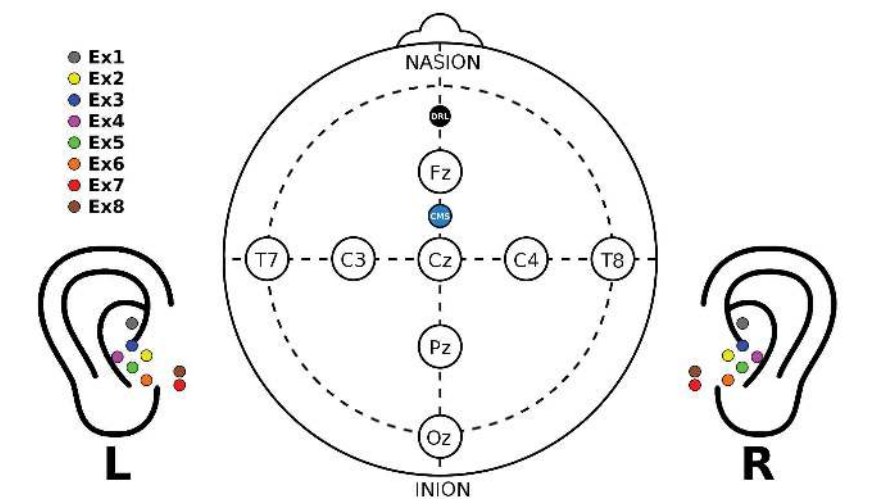


Figure 28. Ear- and scalp-EEG electrode configurations color-coded as in Figure 27. Electrode numbers are provided in the legend, where “x” is replaced by L or R to indicate the left or right ear, respectively.

The EEG earbuds were strapped to the participant’s back and shoulders. The ear-EEG cabling was taped to the cap with skin-safe tape and connected to the amplifier, as seen in Figure 29.

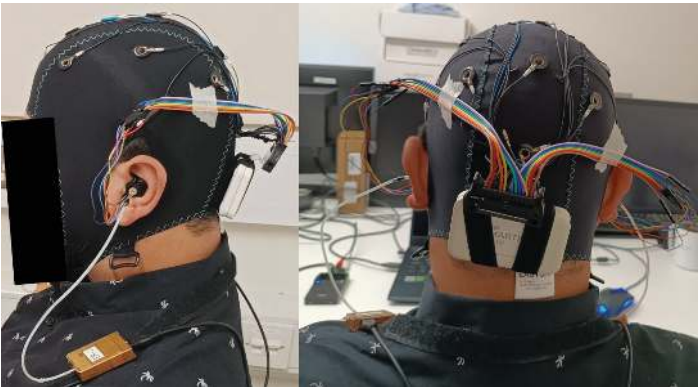


Figure 29. Ear- and scalp-EEG setup for side view (left) and posterior view (right).

A small amount of conductive electrode paste (Ten20) was used for wet-electrode ear recordings. Subjects inserted the earpieces themselves in a way that was snug and yet comfortable. Ear electrode impedance was generally too high or immeasurable by the Streamer software, and as such was not assessed. For the ear-EEG recordings, the lowest volume setting from the validation group was adopted for all subjects.

The same data preprocessing and statistical analysis were applied to the ear-EEG data as before. To identify the optimal ear-EEG re-referencing strategy, the following reference configurations were systematically tested:

- **Cz**: standard central scalp reference
- **T8**: scalp reference closer to the ear
- **ER3**: within-ear reference (e.g., relative to ER8) and between-ear reference (e.g., relative to EL8).

3.2. Ear-P Lab for Ear-EEG Validation

This section presents the ear-EEG data obtained for each of the EEG validation paradigms (using wet-electrode recordings), demonstrating how this tool can be used to assess novel ear-EEG devices.

3.2.1. Alpha Block

Figure 30 shows the grand average alpha power at ER8/EL8 for different ear-EEG reference configurations. Alpha modulation was 3.7 dB for the Cz reference, approximately 3 dB lower than the scalp-EEG benchmark taken at Oz. Re-referencing to T8, closer to the ear, alpha modulation decreases to 3.1 dB. Using a within-ear and between-ear reference resulted in lower modulation values of around 2 dB.

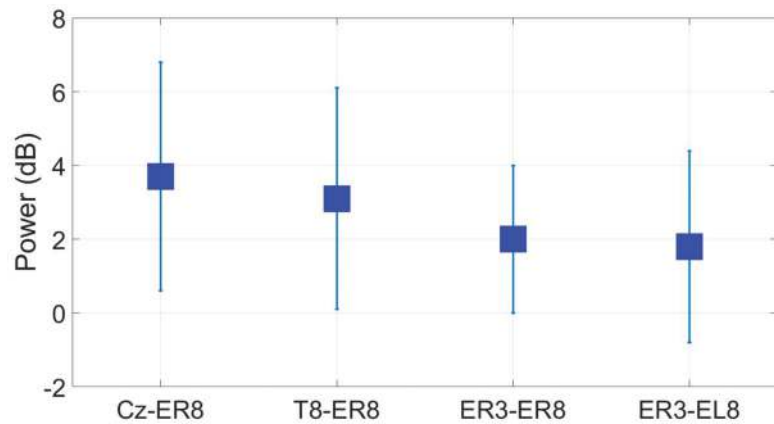


Figure 30. Grand average alpha modulation (wet ear-EEG) at ER8/EL8 for different referencing configurations. Omitted results are not significant based on a *t*-test ($p < 0.05$).

3.2.2. ASSR

The ASSR SNR at ER8 (Figure 31) for the Cz reference had an average value of 7.9 dB, a decrease of 1.5 dB relative to the scalp-EEG benchmark. Comparing the scalp and ear-EEG reference strategies, we see that the SNR is 1.2 dB higher for the within-ear ER3 reference over the temporal T8 site.

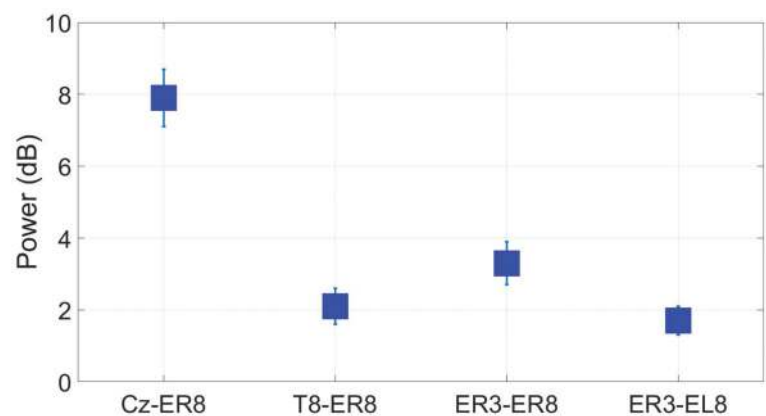


Figure 31. Grand average ASSR responses (wet ear-EEG) to a 40 Hz AM auditory stimulus at ER8/EL8 for different referencing configurations. Omitted results are not significant based on an *F*-test ($p < 0.05$).

3.2.3. SSVEP

The SSVEP measured at ER8 had an SNR of 6.4 dB and 5.9 dB when referenced to Cz and T8, respectively. This was much lower than the scalp-EEG benchmarks of 11 dB and 7.5 dB measured at Oz and T8, respectively. However, in-ear measures using the within-ear and between-ear referencing configurations were not significantly greater than zero, even for other in-ear electrodes (Figure 32).

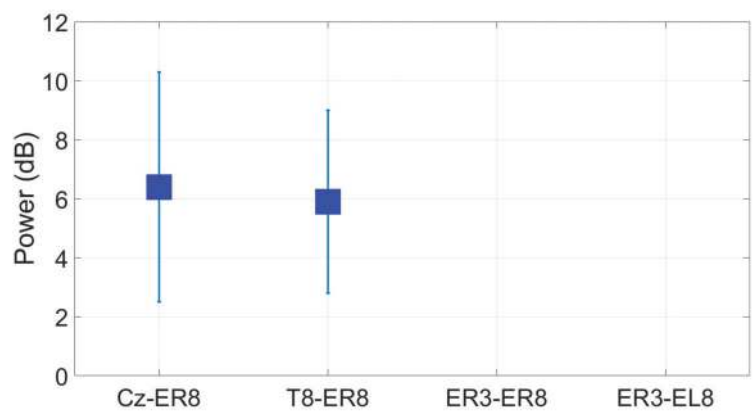


Figure 32. Grand average SSVEP responses (wet ear-EEG) to a 10 Hz visual stimulus at ER8/EL8 for different referencing configurations. Omitted results are not significant based on an *F*-test ($p < 0.05$).

3.2.4. AEP

When referenced to Cz, the AEP response at electrode ER8 displayed a significant N1 component (albeit inverted) with an amplitude of around 6 μ V for wet-electrode ear-EEG recordings. Referenced to T8, the N1 component was diminished but significant, with a peak amplitude of around 2 μ V (Figure 33).

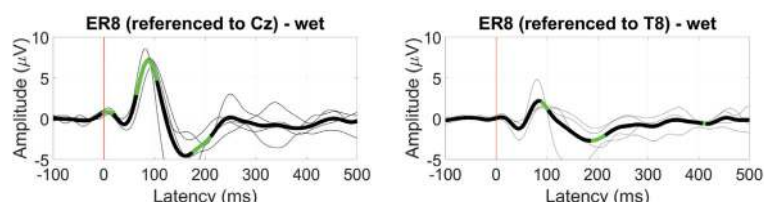


Figure 33. Grand average AEP waveform (black line, wet ear-EEG) at ER8 referenced to Cz (left) and T8 (right). Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons). Grey lines represent individual responses.

For the ear reference analysis (Figure 34), only the between-ear configuration measured at EL8 showed a significant N1 component with an amplitude of 1 μ V.

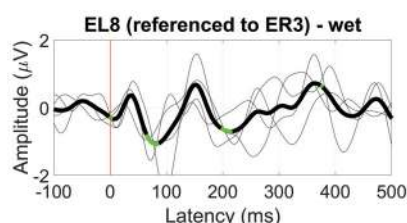


Figure 34. Grand average AEP waveform (black line, wet ear-EEG) at EL8 referenced to ER3. Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons). Grey lines represent individual responses.

3.2.5. VEP

The VEP waveform measured at Oz for the scalp-EEG benchmark was not evident near the ear at T8. For ear-EEG data, the VEP waveform was visible when using both the Cz and T8 referencing configuration, with amplitudes of around 4 and 2 μ V, respectively (Figure 35).

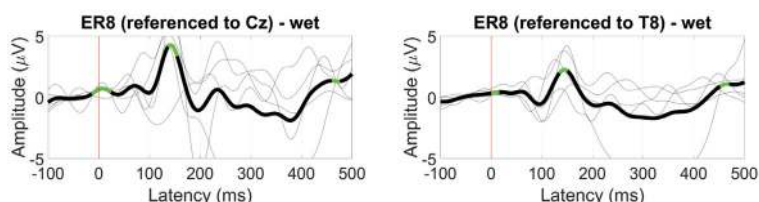


Figure 35. Grand average VEP waveform (black line, wet ear-EEG) at ER8 referenced to Cz (left) and T8 (right). Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons). Grey lines represent individual responses.

For the ER3 reference, only the between-ear configuration resulted in significant deflections in the VEP waveform, showing a 200 ms component with an amplitude of 1 μ V (Figure 36).

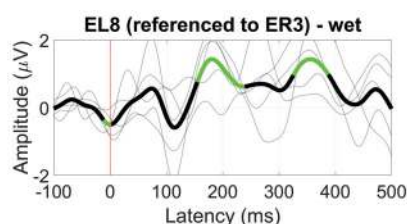


Figure 36. Grand average VEP waveform (black line, wet ear-EEG) at EL8 referenced to ER3. Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons). Grey lines represent individual responses.

3.2.6. AEP OddBall (MMN)

Significant MMN responses were observed at ER8 for both the Cz and T8 referencing configurations, with the deviant response being more prominent in the former (Figure 37).

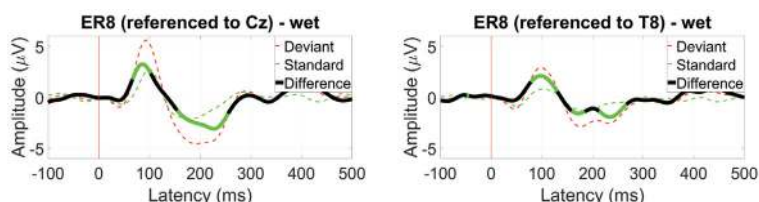


Figure 37. Grand average MMN waveform (wet ear-EEG) at ER8 referenced to Cz (left) and T8 (right). Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons).

For within-ear and between-ear referencing, no significant components resembling an MMN response were found.

3.2.7. VEP OddBall (P300)

The greatest observed difference between the target and standard VEP waveforms was a negative deflection at a latency of around 350 ms for Cz (5 μ V) and T8 (2 μ V) referencing, suggesting a later onset for this component than that observed on the scalp (Figure 38).

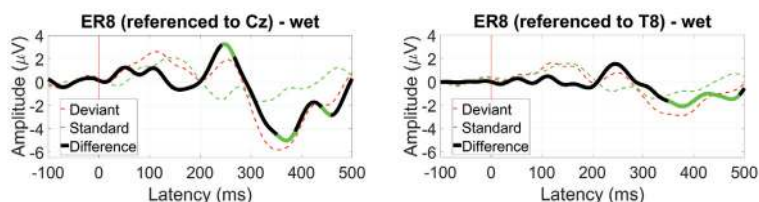


Figure 38. Grand average P300 waveform (wet ear-EEG) at ER8 referenced to Cz (left) and T8 (right). Statistically significant segments are highlighted in green, based on a t -test ($p < 0.05$, not corrected for multiple comparisons).

For within-ear and between-ear referencing, no significant components resembling a P300 response were found.

3.2.8. EOG (Blinks and Saccades)

Examining EOG blink ratios, the ER3 reference resulted in a blink ratio about 1.5 times higher than the scalp references near the ear, with diminished results for the between-ear configuration at EL8. The best ratio was obtained for the within-ear reference, with a blink ratio of 3 (Figure 39).

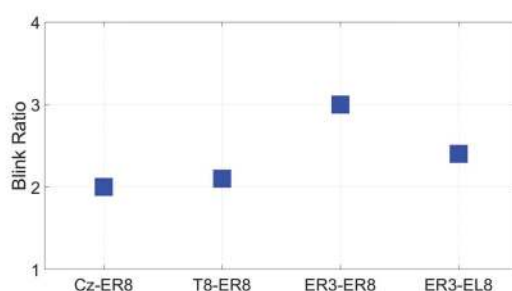


Figure 39. EOG amplitude ratio for soft and hard blinks (wet ear-EEG) for different reference configurations.

Saccade profiles at ER8 were maintained for each of the different eye directions, with greater amplitudes observed for the Cz reference, but better differentiation when referencing to T8 with a 10 μV difference between up and down saccades (Figure 40).

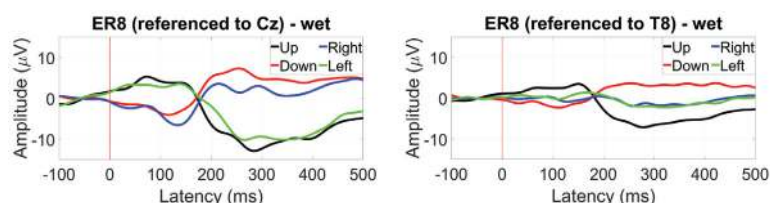


Figure 40. Grand average saccade profiles (wet ear-EEG) in the four cardinal directions at ER8 referenced to Cz (left) and T8 (right).

For within-ear referencing, the vertical saccade plane presents a better differentiation of 3 to 4 μV , while referencing between ears resulted in a greater amplitude difference between right and left saccades (the vertical saccades present the same profile, Figure 41).

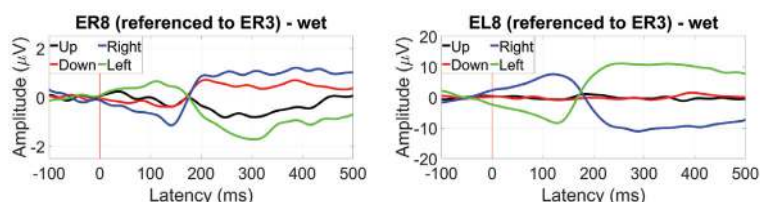


Figure 41. Grand average saccade profiles (wet ear-EEG) in the four cardinal directions at ER8 referenced within-ear (left) and between ears (right).

3.3. Reassessment of Dry Ear-EEG ASSR Data

The ear-EEG phantom results suggest that the ER4 electrode has better connectivity to the ear than ER3, which was previously selected as a reference, for the live ear-EEG recordings. Figure 42 shows a direct comparison of the ASSR dry data (most relevant auditory paradigm) reanalyzed with ER4 as a reference for this single subject that was used for molding the phantom. The suggested electrode ER4 from the phantom improved the SNR at ER8 (and the other electrodes) by about 2.3 dB, demonstrating how the phantom can be used to predict improved performance of ear-EEG devices without the need to conduct testing on a live subject.

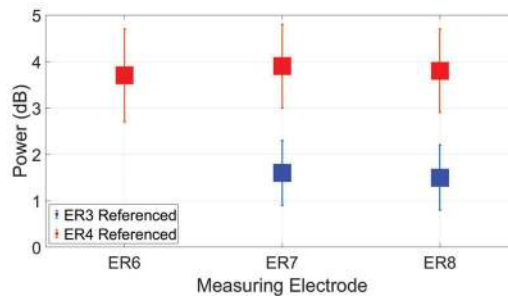


Figure 42. Reassessment of dry-electrode ear-EEG ASSR data for the subject used in the construction of the ear-EEG phantom. Data were re-referenced from ER3 (original reference) to ER4 (better electrode proposed by the phantom), resulting in an increase in SNR.

4. Discussion

Ear-EEG devices are an active field of research and development with potential to disrupt the way in which brain activity can be factored into daily life. We can expect an increase in new applications, form factors, and sensor technology as the field grows. Currently, there is scientific evidence to support the use of ear-EEG [20,21] and we set out to provide a set of tools that could benefit further testing and characterization of ear-EEG devices. Our approach was to develop a validation toolkit that would enable the characterization of ear-EEG devices from hardware to neural signal acquisition.

We developed an EEG acquisition application based on nine commonly-used EEG paradigms: EaR-P Lab. Our validation results indicate that EaR-P Lab can indeed be used to elicit the expected neural responses from scalp-EEG recordings. When evaluating these responses from the perspective of near-ear electrode locations such as T8, we observed that VEP responses were the most affected, indicating that ear-EEG may have a limited capability for more visually-driven applications.

We also developed an EEG phantom suitable for evaluating ear-EEG sensors by combining 3D scans of a participant's ear impressions and 3D printing a mold to be filled with conductive material. This allowed us to assess ear-EEG devices in terms of their electrode contact impedance, noise floor, and acquisition of a known (synthetically generated) signal. To the best of our knowledge, this is the first EEG phantom dedicated for ear-EEG devices, as traditional head phantoms neglect the structures of the ear needed for this type of detailed evaluation [7].

Our phantom development highlighted salt-doped agar as the best substrate material for this purpose, with conductivity values equivalent to those reported in the literature for whole-brain anatomy at 0.33 S/m [32,33]. While salt-doped BG and CF-doped silicone were also investigated as potential material substrates, there were notable limitations with these options.

BG showed a drastic decrease in signal transmission integrity across a week of testing. This could be related to the storage conditions, highlighting the susceptibility of the phantom to environmental conditions over time, which could hinder the repeatability of results when evaluating ear sensors. While agar is also an organic and perishable material, it showed a more stable performance despite being kept under the same storage conditions as the BG phantom. To address the susceptibility to degradation over time, we investigated the use of a synthetic non-perishable material composition for the phantom by doping platinum-cured silicone with carbon fibers. This approach, however, did not scale from the prepared samples for conductivity testing to the larger ear-EEG phantom.

The mechanism by which carbon fibers turn silicone into a conductive medium is different to that of salt-doped agar or BG. In the latter, salt fully dissolves in the medium, creating a homogeneous conductive substrate, while in the former, carbon fibers disperse in the silicone, creating conductive paths through heterogeneously mixed fibers. When the

samples were small strips, this dispersion of carbon fibers was effective throughout the volume; however, when a larger volume was used, the stirring method did not achieve dispersion of fibers over the full volume. This led to patches of non-conductive silicone, clearly visible in the final product. A more optimal mixing approach should be devised in the future to ensure that the CF phantom is conductive throughout. It should be noted, however, that the conductivity of this synthetic substrate is orders of magnitude larger than that of real anatomical structures and the agar/BG substrates. As an alternative to silicone, other commercially available synthetic ballistic gels could be considered (e.g., Perma-Gel, INC. or Humimic Medical, TM). However, as these are based on a combination of mineral oil and proprietary gellants, new methods to induce and control conductivity would need to be developed.

We validated these tools using custom-fit ear-EEG devices developed in-house at Segotia. Our results showed that the ear-EEG sensors were functional and EEG responses were successfully recorded when using a scalp reference (Cz and T8). However, in-ear and between-ear referencing configurations only showed significant EEG responses for steady-state paradigms and alpha blocking (although diminished AEP and VEP responses could also be seen). Further investigation is needed to identify the optimal referencing configuration for each of these ERP paradigms, as it has been shown that this is crucial for obtaining characteristic ERP responses [34]. However, this was not the scope of the current study. Importantly, we could see clear EOG deflections using both within-ear and between-ear reference configurations, which has promising implications for ear-based BCI technologies.

We also demonstrated how the data obtained with our ear-EEG phantom—without the need of a testing subject—could be used to select an optimal referencing location in the ear and when applied to the actual ear-EEG recordings of the same tested subject, it improved the SNR of the ASSR response.

Future work for our ear-EEG validation toolkit include improving the testing protocol and the assessment of a generic ear canal for comparing different earpiece designs. Our ear-EEG phantom methodology could also be improved in the future for studying how real-world factors, such as gait, impact ear-EEG [35].

The presented ear-EEG validation toolkit is available to the scientific community via the Open Science Framework (OSF) repository named “Brain Wearables: Validation toolkit for Ear Level EEG sensors” (<https://osf.io/2dxs4/>). All future modifications to the Ear-P Lab and the ear-EEG phantom developed by the authors will also be updated in the referenced repository.

5. Conclusions

In this study, we set out to address the need for a comprehensive characterization tool for ear-EEG sensors at the neural and hardware level. We presented a validation toolkit composed of a desktop application (EaR-P Lab) for acquiring EEG data to nine cardinal EEG paradigms relevant to the field of ear-EEG and a novel ear-EEG phantom for interrogation of the sensor and hardware signal chain quality. To facilitate better characterization of ear-EEG devices, our ear-EEG phantom is the first electrical phantom (designed based on 3D impressions of the human ear) to include the ear canal structures. These tools were validated with traditional scalp recordings and bench-top evaluations before being applied in a use-case study, which characterized the functionality of a custom-made ear-EEG device.

Through our use-case study, we demonstrated the utility of this toolkit for the characterization of ear-EEG sensors and relating hardware level observations, such as sensor placement and contact, to neural level outcomes, such as ASSR signal strength. To our knowledge, this is the first publication where controlled testing through an EEG phantom is applied to ear-EEG technology. By making this toolkit available through the Open Science Framework, we hope to facilitate and accelerate future research on and development of ear-EEG sensors and devices.

Author Contributions: Conceptualization, A.L.V. and M.J.C.; methodology, G.C. and A.L.V.; design, G.C. and M.J.C.; validation, G.C.; formal analysis, G.C.; investigation, G.C. and A.L.V.; writing—original draft preparation, G.C.; writing—review and editing, G.C., A.L.V. and M.J.C.; visualization, G.C., A.L.V. and M.J.C.; supervision, A.L.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval for this study was granted by the Research Ethics Committee of the School of Engineering, Trinity College Dublin, with reference No.: 3182.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The presented ear-EEG validation toolkit and the validation data are available to the scientific community via the Open Science Framework (OSF) repository named: “Brain Wearables: Validation toolkit for Ear Level EEG sensors” (<https://osf.io/2dxs4/>). All future modifications to the Ear-P Lab and the ear-EEG phantom generated by the authors will also be updated in the mentioned repository.

Acknowledgments: We would like to thank Jack Maughan, PhD student at Trinity College Dublin, for the help and work performed with the conductivity characterization of the tested ear-EEG phantom molding materials.

Conflicts of Interest: During the development of this project and at the time of submission of this manuscript, M.J.C. was employed at Segotia, the third-party company who provided the ear-EEG sensors for this study.

Abbreviations

The following abbreviations are used in this manuscript:

EEG	Electroencephalography
BCI	Brain–Computer Interface
ERP	Event-Related Potential
ASSR	Auditory Steady-State Response
SSVEP	Steady-State Visual Evoked Potential
AEP	Auditory Evoked Potential
VEP	Visual Evoked Potential
MMN	Mismatch Negativity
EOG	Electro-Oculography
ISI	Interstimulus Interval
BG	Ballistic Gelatin
CF	Carbon Fibers

Appendix A. EaR-P Lab Paradigms and Settings

This appendix outlines the testing sequences initialized by each button in the Main Menu of EaR-P Lab.

Resting State

This test paradigm displays a white cross centered on the screen (Figure A1) for the selected duration of time. This option was implemented to record resting-state EEG or any custom test not readily provided by default in EaR-P Lab.

Auditory Steady-State Response (ASSR)

This test paradigm displays the same focus as Figure A1, while playing an amplitude-modulated sound file for an ASSR test. The duration of the experiment is the same as the duration of the sound file.

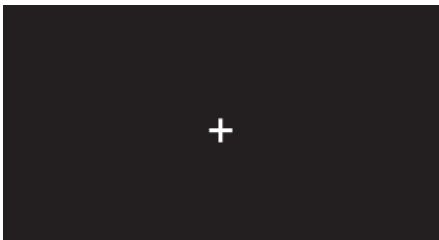


Figure A1. Focus cross utilized by the “Resting State”, “ASSR”, and “Alpha Block” paradigms.

Steady-State Visual Evoked Potential (SSVEP)

This paradigm consists of a flickering target like the one presented in Figure A2. The target consists of a radial black and white pattern centered on a black background with a diameter that is equal to the screen height. The flicker frequency and the duration of the test can be adjusted in the “Settings” menu.

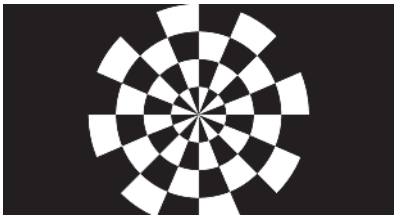


Figure A2. Target used for the “SSVEP” experiment in EaR-P Lab.

Alpha Block

This paradigm starts with onscreen instructions for the subject to press “SPACE” and close the eyes. After a specified amount of time, there is a transition phase where an artificially generated voice will warn the subject to open its eyes and look at the cross for the same time as before. These instructions repeat one more time, so in total, there are two phases with eyes closed and two with eyes open that alternate. At the beginning of each phase, the specific markers seen in Figure A3 are also sent through LSL.

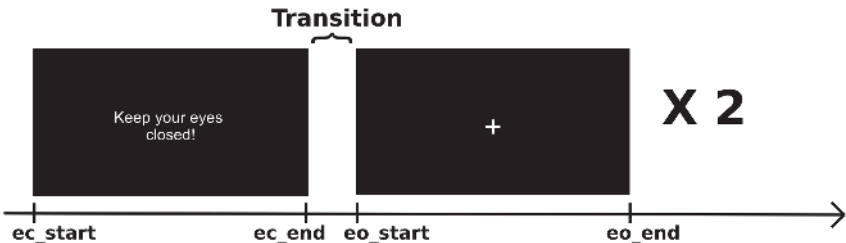


Figure A3. “Alpha Block” paradigm functioning, with the markers that are sent at the start and end of each phase.

Auditory Evoked Potential (AEP)

In this paradigm, a sequence of discreet auditory stimuli is played. The stimuli can be self-generated tones of specific frequency and duration via PsychoPy or custom-made sound files added to the EaR-P Lab source folder. The stimuli are separated by an adjustable ISI, and each time the stimuli start, the “aep” marker is sent to LabRecorder. In order to keep the subject engaged (to help avoid drowsiness while the test is running), the *Inscapes* animation is also played in the background, as shown in Figure A4. This animation was originally crafted to maintain children “engaged while minimizing the certain aspects of

cognitive processing” [36]. The original sound was removed from the animation, and it will loop back to the beginning until the test is finished.

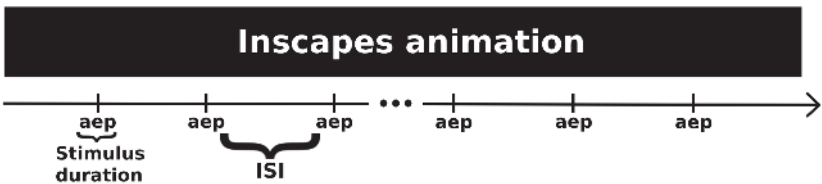


Figure A4. “AEP” sequence featured in EaR-P Lab.

Visual Evoked Potential (VEP)

This paradigm is similar to its auditory counterpart, with the default option of a shape being displayed on-screen for a specified amount of time and ISI. Different versions of VEP are possible to acquire through this paradigm. Flash VEP can be elicited where the screen is black and flashes white for a specified duration. Two patterned stimuli, a checkerboard and a dartboard (as shown in Figure A5), can be used for an onset–offset VEP or pattern–reversal VEP (if the chosen ISI is set to zero). A “vep” marker is sent every time the on-screen stimulus changes to a new one.

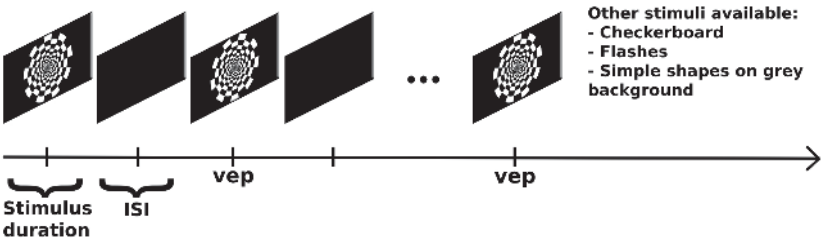


Figure A5. “VEP” button paradigm structure and options.

OddBall Paradigms

These paradigms are identical for visual and auditory modalities. The stimulus sequence starts with 20 standard stimuli followed by a randomly generated sequence of standard and target events, as depicted in Figure A6. The oddBall paradigm has a fixed, non-modifiable, 80–20 configuration (i.e., the target stimuli appears at random with a probability of 20%). For the auditory version of oddBall, similar to the regular “AEP” sequence, the *Inscapes* movie is played in the background. Additionally, two rules are also implemented during these paradigms, these being that there cannot be a succession of two targets in a row (i.e., there is always at least one standard stimulus before and after each oddBall) and that in the case there is a part of the sequence made up of eight standard events in a row (excluding the first twenty events), the next event is compulsorily a target stimulus. Markers for the “standard” and “target” as well as a marker for a “response” in case of an active task (e.g., P300) are sent accordingly.

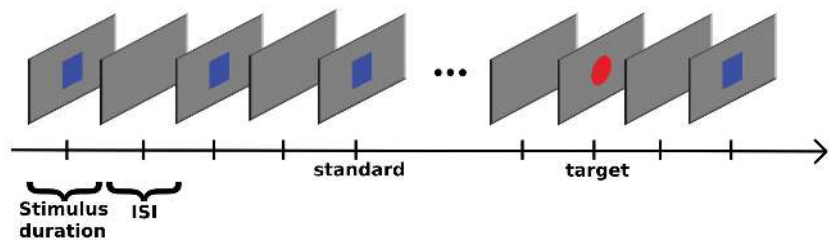


Figure A6. “OddBall”-type paradigm sequence—example of visual oddBalls.

Electro-Oculography (EOG)

This paradigm is split into two parts: eye blinks and saccades. During the eye-blinks phase, the subject is prompted to perform a total of four eye blinks, alternating between a normal or “soft” blink and an intentional or “hard” blink. For each blink, the subject has a three-second window that starts once “SPACE” is pressed, according to the instructions. As shown in Figure A3, the “soft_blink_start”, “soft_blink_end”, “hard_blink_start”, and “hard_blink_end” static markers are used to limit each blink window.

The next part of this paradigm is a “follow-the-dot” test, shown in Figure A7. Here, a red dot starting in the center of the screen disappears and reappears at one of the four main cardinal directions for the subject to follow to test for eye saccade amplitudes and profiles. The dot jumps a distance equal to half the screen height for each saccade, is held in that position for a specified duration, and returns to the middle during the ISI period. The appropriate marker stating “top”, “right”, “left”, or “bottom” is sent when the dot moves to the respective location. The static markers “follow_start” and “follow_end” also limit this part of the coded “EOG” paradigm.

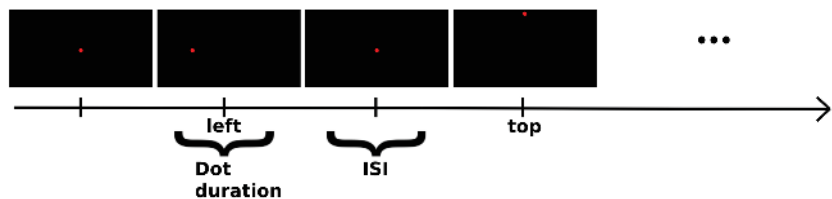


Figure A7. “Follow-the-dot” phase of the “EOG” paradigm in EaR-P Lab.

Settings Menu

The “Settings” button opens up a secondary window (Figure A8) with different boxes and options that control what happens and is shown in each paradigm mentioned above. These settings range from the type of stimuli to the frequency of the SSVEP, testing durations, total trial number for the AEP/VEP paradigms, how many oddBalls until the oddBall-type tasks are complete, and how many repetitions for each saccade in the EOG test. Additionally, there is some information about how to proceed to add more stimuli to EaR-P Lab, a disclaimer about the ISI for auditory paradigms, and a button to automatically measure the angle of the saccades in the EOG task, dependent on the dimensions of the stimuli presentation monitor used and the subject’s distance to said screen.

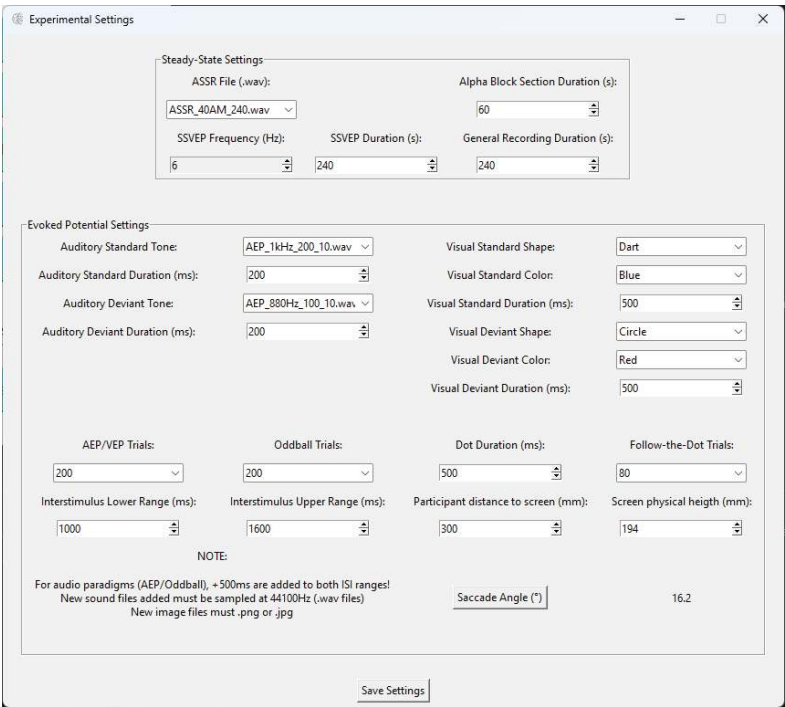


Figure A8. EaR-P Lab—Settings menu.

Markers Menu

Like the “Settings” menu, here a secondary window (Figure A9) is presented with text boxes about each paradigm so that the user can set its delimiting markers for each experiment. This feature is useful, for instance, to record the same type of paradigm within the same *.xdf* file and to help with accurate data parsing. At the bottom of the window, there is information about the static (meaning non-changeable in the GUI) markers used by EaR-P Lab, as shown in the above paragraphs.



Figure A9. EaR-P Lab—Markers menu.

Appendix B. Ear-EEG Phantom Dimensions

This appendix section will illustrate the detailed dimensions and design drawings of the ear-EEG phantom.

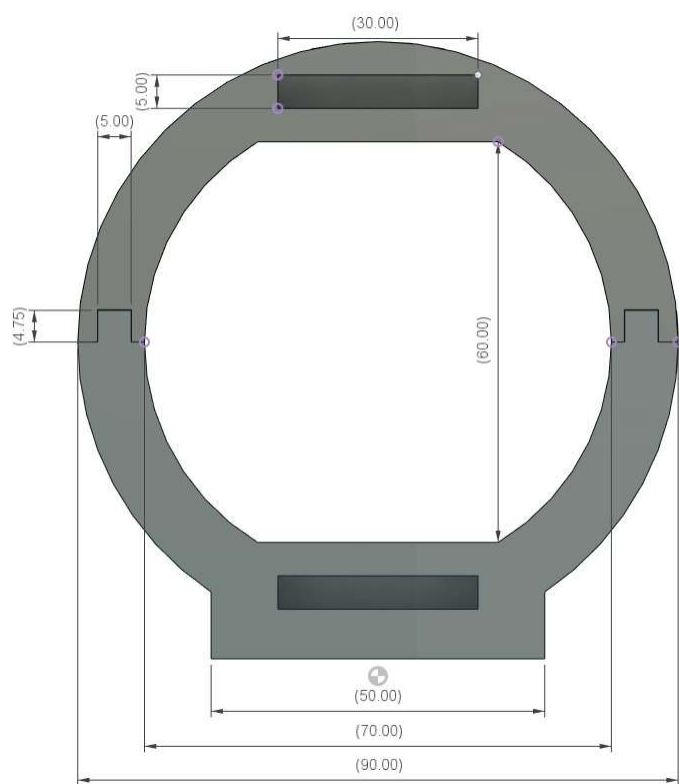


Figure A10. Proposed ear-EEG phantom prototype dimensions—side view.

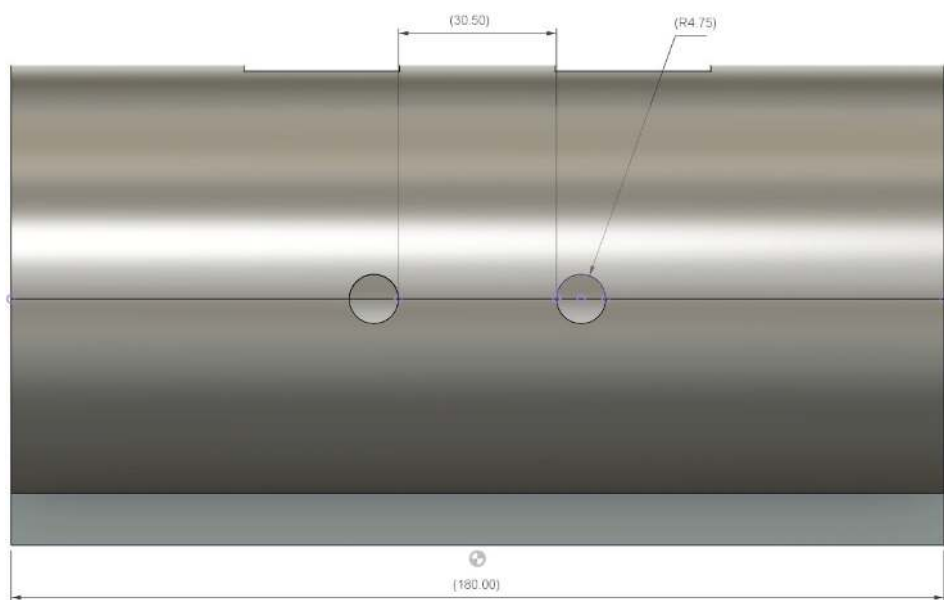


Figure A11. Proposed ear-EEG phantom prototype dimensions—front view.

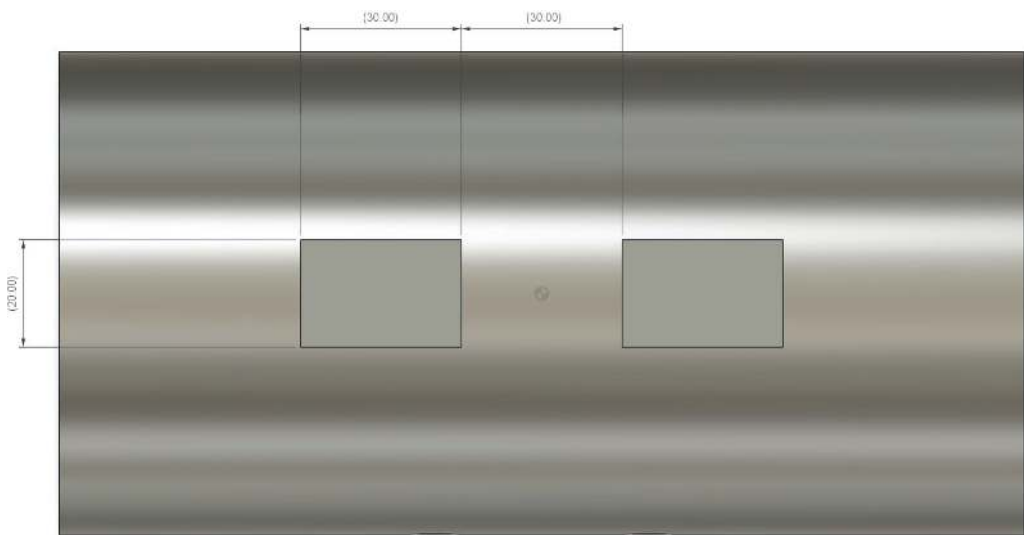


Figure A12. Proposed ear-EEG phantom prototype dimensions—top view.

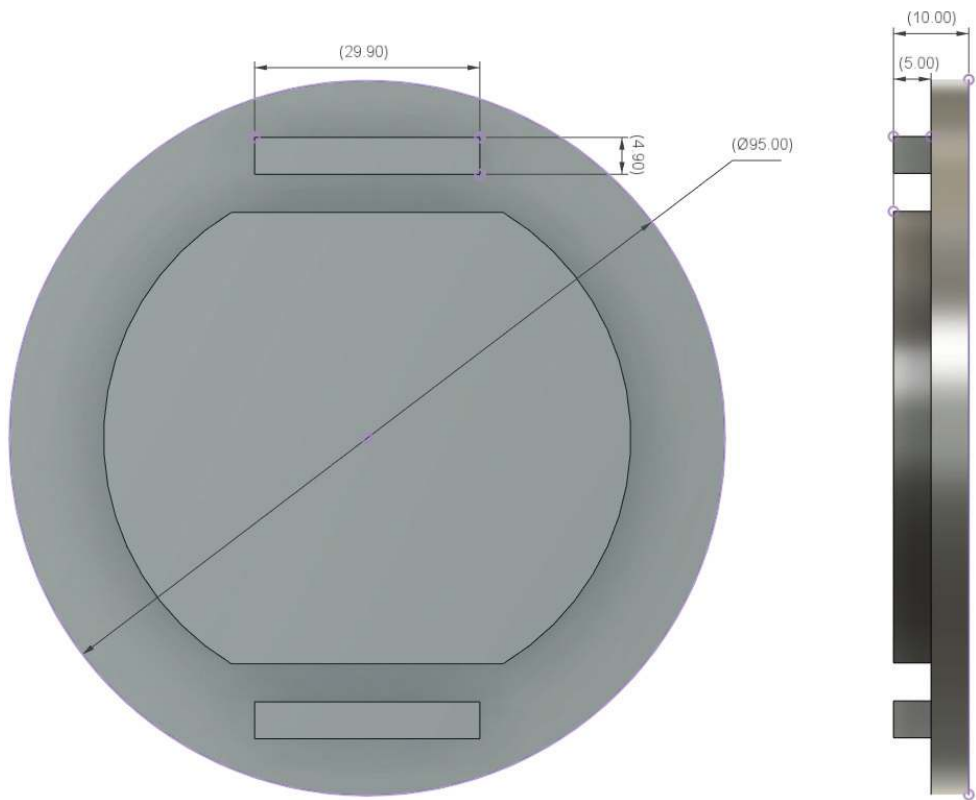


Figure A13. Proposed ear-EEG phantom prototype dimensions—lids.

References

1. Casson, A.; Yates, D.; Smith, S.; Duncan, J.; Rodriguez-Villegas, E. Wearable Electroencephalography. *IEEE Eng. Med. Biol. Mag.* **2010**, *29*, 44–56. [CrossRef] [PubMed]
2. Casson, A.J. Wearable EEG and beyond. *Biomed. Eng. Lett.* **2019**, *9*, 53–71. [CrossRef]
3. Soufneyestani, M.; Dowling, D.; Khan, A. Electroencephalography (EEG) Technology Applications and Available Devices. *Appl. Sci.* **2020**, *10*, 7453. [CrossRef]
4. Looney, D.; Kidmose, P.; Park, C.; Ungstrup, M.; Rank, M.; Rosenkranz, K.; Mandic, D. The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring. *IEEE Pulse* **2012**, *3*, 32–42. [CrossRef]
5. Wyckoff, S.N.; Sherlin, L.H.; Ford, N.L.; Dalke, D. Validation of a wireless dry electrode system for electroencephalography. *J. Neuroeng. Rehabil.* **2015**, *12*, 95. [CrossRef]
6. What Are Imaging Phantoms? Available online: <https://www.nist.gov/physics/what-are-imaging-phantoms> (accessed on 11 December 2023).
7. Hairston, W.D.; Slipher, G.A.; Yu, A.B. Ballistic gelatin as a putative substrate for EEG phantom devices. *arXiv* **2016**, arXiv:1609.07691. <https://doi.org/10.48550/ARXIV.1609.07691>.
8. Richer, N.; Downey, R.J.; Nordin, A.D.; Hairston, W.D.; Ferris, D.P. Adding neck muscle activity to a head phantom device to validate mobile EEG muscle and motion artifact removal. In Proceedings of the 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), San Francisco, CA, USA, 20–23 March 2019; pp. 275–278. [CrossRef]
9. Chowdhury, M.E.; Khandakar, A.; Hossain, B.; Alzoubi, K. Effects of the phantom shape on the gradient artefact of electroencephalography (EEG) data in simultaneous EEG-fMRI. *Appl. Sci.* **2018**, *8*, 1969. [CrossRef]
10. Audette, W.E.; Allen, L.V. *Design and Demonstration of a Head Phantom for Testing of Electroencephalography (EEG) Equipment Hearing Protection View Project DPOAE Level and Phase Mapping View Project*; Create, Inc: Hannover, Germany, 2020. [CrossRef]
11. Wood, S.; Martins, T.; Ibrahim, T.S. How to design and construct a 3D-printed human head phantom. *J. 3D Print. Med.* **2019**, *3*, 119–125. [CrossRef]
12. Marathe, A.R.; Ries, A.J.; McDowell, K. Sliding HDCA: Single-trial eeg classification to overcome and quantify temporal variability. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 201–211. [CrossRef]
13. Polite, D.; Prior, F.; Ponton, C.; Nolan, T.; Larson-Prior, L. Sources of non-physiologic noise in simultaneous EEG-fMRI data: A phantom study. In Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August–4 September 2010; pp. 5129–5132. [CrossRef]
14. Nordin, A.D.; Hairston, W.D.; Ferris, D.P. Dual-electrode motion artifact cancellation for mobile electroencephalography. *J. Neural Eng.* **2018**, *15*, 056024. [CrossRef]
15. Kuratko, D.; Lacik, J.; Koudelka, V.; Vejmla, C.; Wojcik, D.K.; Raida, Z. Forward Model of Rat Electroencephalogram: Comparative Study of Numerical Simulations With Measurements on Rat Head Phantoms. *IEEE Access* **2022**, *10*, 92023–92035. [CrossRef]
16. Owda, A.Y.; Casson, A.J. Investigating Gelatine Based Head Phantoms for Electroencephalography Compared to Electrical and Ex Vivo Porcine Skin Models. *IEEE Access* **2021**, *9*, 96722–96738. [CrossRef]
17. Velcescu, A.; Lindley, A.; Cursio, C.; Krachunov, S.; Beach, C.; Brown, C.A.; Jones, A.K.; Casson, A.J. Flexible 3D -printed EEG electrodes. *Sensors* **2019**, *19*, 1650. [CrossRef] [PubMed]
18. Kaveh, R.; Doong, J.; Zhou, A.; Schwendeman, C.; Gopalan, K.; Burghardt, F.L.; Arias, A.C.; Maharbiz, M.M.; Muller, R. Wireless User-Generic Ear EEG. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *14*, 727–737. [CrossRef] [PubMed]
19. Kappel, S.L.; Rank, M.L.; Toft, H.O.; Andersen, M.; Kidmose, P. Dry-Contact Electrode Ear-EEG. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 150–158. [CrossRef] [PubMed]
20. Kaongoen, N.; Choi, J.; Choi, J.W.; Kwon, H.; Hwang, C.; Hwang, G.; Kim, B.H.; Jo, S. The future of wearable EEG: A review of ear-EEG technology and its applications. *J. Neural Eng.* **2023**, *20*, 051002. [CrossRef] [PubMed]
21. Röddiger, T.; Clarke, C.; Breitling, P.; Schneegans, T.; Zhao, H.; Gellersen, H.; Beigl, M. Sensing with Earables: A Systematic Literature Review and Taxonomy of Phenomena. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 1–57. [CrossRef]
22. Peirce, J.; Gray, J.R.; Simpson, S.; ad Richard Höchenberger, M.M.; Sogo, H.; Kastman, E.; Lindeløv, J.K. PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods* **2019**, *51*, 195–203. [CrossRef]
23. Kothe, C.; Medine, D.; Boulay, C.; Grivich, M.; Stenner, T. Lab Streaming Layer Open Source Repository. Available online: <https://github.com/scn/labstreaminglayer> (accessed on 3 March 2023).
24. Razavi, M.; Yamauchi, T.; Janfaza, V.; Leontyev, A.; Longmire-Monford, S.; Orr, J. Multimodal-Multisensory Experiments. *Preprints* **2020**, 2020080614. [CrossRef]
25. Bridges, D.; Pitiot, A.; MacAskill, M.R.; Peirce, J.W. The timing mega-study: Comparing a range of experiment generators, both lab-based and online. *PeerJ* **2020**, *8*, e9414. [CrossRef]
26. Can PsychoPy Deliver Millisecond Precision. Available online: <https://www.psychopy.org/general/timing/millisecondPrecision> (accessed on 11 December 2023).
27. Bishop, D.V.M.; Hardiman, M.J. Measurement of mismatch negativity in individuals: A study using single-trial analysis. *Psychophysiology* **2010**, *47*, 697–705. [CrossRef] [PubMed]
28. Goverdovsky, V.; Looney, D.; Kidmose, P.; Mandic, D.P. In-Ear EEG From Viscoelastic Generic Earpieces: Robust and Unobtrusive 24/7 Monitoring. *IEEE Sens. J.* **2016**, *16*, 271–277. [CrossRef]

29. Saleem, A.; Frommann, L.; Soever, A. Fabrication of extrinsically conductive silicone rubbers with high elasticity and analysis of their mechanical and electrical characteristics. *Polymers* **2010**, *2*, 200–210. [CrossRef]
30. SILC Circuits: High Performance Conductive Silicone. Available online: <https://www.instructables.com/Silc-Circuits-High-Performance-Conductive-Silicone/> (accessed on 11 December 2023).
31. OpenBCI EEG Setup. Available online: <https://docs.openbci.com/GettingStarted/Biosensing-Setups/EEGSetup/> (accessed on 11 December 2023).
32. McCann, H.; Pisano, G.; Beltrachini, L. Variation in reported human head tissue electrical conductivity values. *Brain Topogr.* **2019**, *32*, 825–858. [CrossRef] [PubMed]
33. Koessler, L.; Colnat-Coulbois, S.; Cecchin, T.; Hofmanis, J.; Dmochowski, J.P.; Norcia, A.M.; Maillard, L.G. In-vivo measurements of human brain tissue conductivity using focal electrical current injection through intracerebral multicontact electrodes. *Hum. Brain Mapp.* **2016**, *38*, 974–986. [CrossRef] [PubMed]
34. Mikkelsen, K.B.; Kappel, S.L.; Mandic, D.P.; Kidmose, P. EEG Recorded from the Ear: Characterizing the Ear-EEG Method. *Front. Neurosci.* **2015**, *9*, 438. [CrossRef]
35. Oliveira, A.S.; Schlink, B.R.; Hairston, W.D.; König, P.; Ferris, D.P. Induction and separation of motion artifacts in EEG data using a mobile phantom head device. *J. Neural Eng.* **2016**, *13*, 036014. [CrossRef]
36. Vanderwal, T.; Kelly, C.; Eilbott, J.; Mayes, L.C.; Castellanos, F.X. Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *NeuroImage* **2015**, *122*, 222–232. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

A Combined Magnetolectric Sensor Array and MRI-Based Human Head Model for Biomagnetic FEM Simulation and Sensor Crosstalk Analysis

Mesut-Ömür Özden *, Giuseppe Barbieri and Martina Gerken *

Integrated Systems and Photonics, Department of Electrical and Information Engineering, Kiel University, Kaiserstraße 2, 24143 Kiel, Germany; giba@tf.uni-kiel.de

* Correspondence: omoz@tf.uni-kiel.de (M.-Ö.); mge@tf.uni-kiel.de (M.G.); Tel.: +49-431-880-6250 (M.G.)

Abstract: Magnetolectric (ME) magnetic field sensors are novel sensing devices of great interest in the field of biomagnetic measurements. We investigate the influence of magnetic crosstalk and the linearity of the response of ME sensors in different array and excitation configurations. To achieve this aim, we introduce a combined multiscale 3D finite-element method (FEM) model consisting of an array of 15 ME sensors and an MRI-based human head model with three approximated compartments of biological tissues for skin, skull, and white matter. A linearized material model at the small-signal working point is assumed. We apply homogeneous magnetic fields and perform inhomogeneous magnetic field excitation for the ME sensors by placing an electric point dipole source inside the head. Our findings indicate significant magnetic crosstalk between adjacent sensors leading down to a 15.6% lower magnetic response at a close distance of 5 mm and an increasing sensor response with diminishing crosstalk effects at increasing distances up to 5 cm. The outermost sensors in the array exhibit significantly less crosstalk than the sensors located in the center of the array, and the vertically adjacent sensors exhibit a stronger crosstalk effect than the horizontally adjacent ones. Furthermore, we calculate the ratio between the electric and magnetic sensor responses as the sensitivity value and find near-constant sensitivities for each sensor, confirming a linear relationship despite magnetic crosstalk and the potential to simulate excitation sources and sensor responses independently.

Keywords: biomagnetic sensor; crosstalk; finite-element method (FEM); human head model; magnetic fields; magnetolectric effect; MRI data; multiferroic device; multiscale model; sensor array

Citation: Özden, M.-Ö.; Barbieri, G.; Gerken, M. A Combined Magnetolectric Sensor Array and MRI-Based Human Head Model for Biomagnetic FEM Simulation and Sensor Crosstalk Analysis. *Sensors* **2024**, *24*, 1186. <https://doi.org/10.3390/s24041186>

Academic Editor: José Machado Da Silva

Received: 23 December 2023

Revised: 7 February 2024

Accepted: 7 February 2024

Published: 11 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the field of medical diagnostics, bioelectric measurements are commonly performed on patients to investigate possible pathological disorders. Well-known applications of this kind are the electrocardiogram (ECG) and the electroencephalogram (EEG), which measure and evaluate the electric activities of the human heart and brain, respectively. To perform them, electrodes are applied directly on the patient's body and the electric potential generated by the heart or brain activity is measured on the surface of the skin. While electrical measurements such as ECG and EEG are of the utmost importance and allow physicians to gather valuable vital information on a patient's health with cost-effective and widely available machines and with high temporal resolution, they do have drawbacks as well, leading to increasing interest in measurement applications based on the magnetic sensing of biological signals. As an alternative or complementary measurement technique to bioelectric signals [1], biomagnetic signals can also be measured with appropriate sensing devices such as fluxgate magnetometers [2,3], optically pumped magnetometers [4], SQUID systems [5], and magnetolectric (ME) magnetic field sensors [6–9].

In current clinical and research applications, SQUID sensors are used as the gold standard for magnetic measurements down to the femtotesla (fT) range, enabling their use in applications such as magnetoencephalography (MEG) [5,10]. Multi-channel SQUID

magnetometers with helmet-shaped dewars have the ability to cover several head regions and reduce measurement times simultaneously [5], and state-of-the-art devices achieve a noise spectral density of approximately $3 \text{ fT/Hz}^{1/2}$ [11], allowing their efficient and precise application in MEG.

While SQUIDs are still the sensing device of choice for many applications due to their extremely high sensitivity, ME sensors were recently shown to be promising in highly versatile and novel applications ranging from automatic real-time magnetic localization of an ultrasound probe [12], to magnetic detection of positions and orientations of deep brain stimulation (DBS) electrodes in patients [13], to magnetic motion analysis for swallowing detection in individuals suffering from dysphagia [14]. Crucially, measurements in the operation room or even wearable solutions could be possible with ME sensors for applications such as the detection and orientation of DBS electrodes or swallowing detection, which are currently not feasible with SQUID systems.

Even outside the scope of biomagnetic sensing, ME devices have the potential for a variety of applications, such as energy-efficient memory [15,16], antennas and energy harvesting [17,18], electric current sensing [19], and automotive applications [20]. Lastly, as opposed to many other magnetic sensing systems, ME sensors also offer the potential for room-temperature, passive and unshielded operation [9,21]. In this work, we demonstrate a combined multiscale 3D finite-element method (FEM) model including several ME sensors in different array configurations and a detailed anatomical human head model based on MRI data. Previous investigations report simplified spherical or realistic anatomical head models and magnetic field calculations for specific applications such as the EEG or MEG forward problem. Extremely detailed models on EEG and MEG applications exist, which even consider how the movement of the brain inside the cerebrospinal fluid relative to the inner skull due to subjects' changing body position can affect mesh generation [22,23]. However, the respective sensor systems for such applications are often not jointly evaluated at all, or simplified to point [24] or circular magnetometers [25]. However, ME sensors are not of a negligible size and their geometry plays a role in the sensitivity of measurements [26]. Therefore, the study framework presented in this work is necessary to simulate sensors and the head or throat in a single FEM model. In this case, we focus on the head, as the simulation framework is known from previous MEG cases, and we significantly extend this framework by integrating the sensors in the simulation. The key novelty of this approach is the complete integration of an MRI-based head model with a fully coupled ME sensor array model and its physical properties. This enables us to evaluate the mechanical, electrical, and magnetic behavior of the magnetoelectric sensing devices in a variable array configuration. Additionally, the inclusion of further components such as operation instruments could be added to the simulation framework.

With this extensive model, we aim to investigate the response of ME sensors to different excitation mechanisms such as homogeneously applied external magnetic fields and a dipole source inside the human head. We also evaluate the concept of magnetic crosstalk between adjacent ME sensors in the different array configurations, which is based on the high-permeability material utilized in the magnetostrictive layers of the devices. This crosstalk can lead to adverse sensing performance based on the location of each sensor inside the array, the distance to adjacent sensors and the excitation source, and the method of excitation (i.e., homogeneous vs. inhomogeneous excitation). Lastly, we calculate the ratio between electric sensor response and the magnetic flux density for each individual sensor in order to determine whether the separability of effects is visible in our combined FEM simulation, as expected for a linear model. While the inclusion of nonlinear effects due to magnetic material properties or secondary currents induced in the head requires a combined model, determining a simple linear relationship between the magnetic sensor excitation and the electric response justifies the investigation of excitation sources and sensor responses in separate models. This greatly reduces computational requirements, allows for higher resolution meshing, and enables arbitrary combinations of separately developed source and sensor models.

Figure 1 visualizes the concept of the simulations presented in this work. In (a), we illustrate a clipped view of the head model with its tissue regions of skin, skull, and white matter, as well as the point dipole located in the latter region. (b) demonstrates an electric current point dipole source inside the head inducing the propagation of an inhomogeneous magnetic field. This magnetic field propagates through the head, into the surrounding air environment, and becomes the excitation source for the ME sensor array. The individual sensors within the array can display vastly different responses to this excitation based on the previously mentioned parameters, which we aim to systematically investigate and discuss throughout this work.

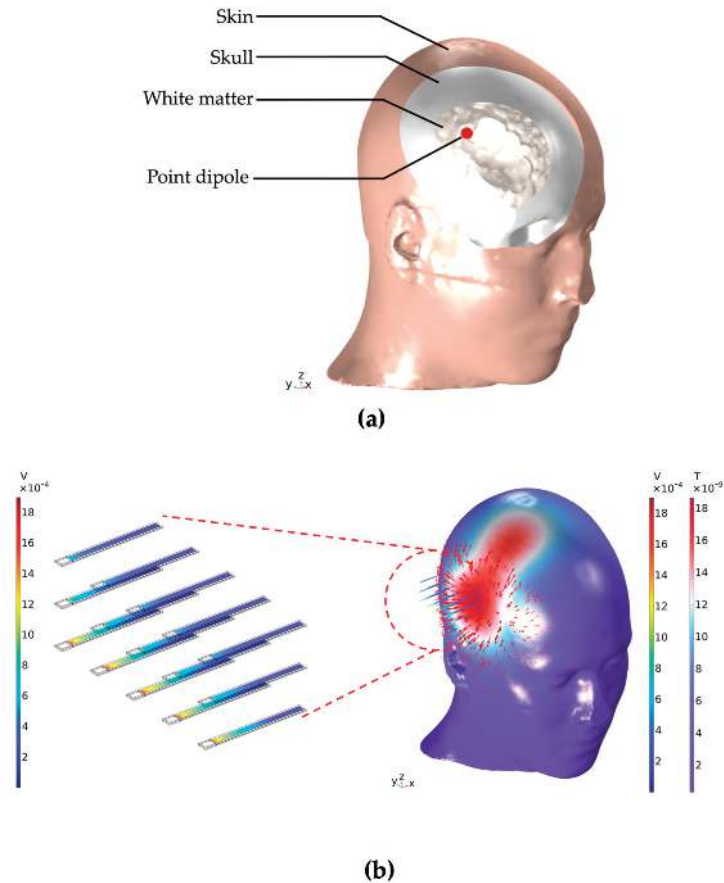


Figure 1. (a) Clipped view of the human head with its tissue regions of skin, skull, and white matter. The white matter region contains an electric current point dipole source—as would be obtained with a deep-brain stimulation electrode—which creates an electromagnetic field. This field propagates through the tissues of the head, through the air environment, and into the ME sensor array. Shown in (b) are the magnetic flux density norm on the head's surface with corresponding vector arrows and, based on this propagating magnetic field as the method of excitation, the electric response of an adjacent ME sensor array with 15 sensors. The sensors are located at a distance of 1.5 cm from the head and a distance of 1 cm between neighboring sensors. The resulting absolute potential on the surface of each individual piezoelectric layer is shown.

Our study is divided into the following sections. Section 2 offers insights into the setup and geometries of the separate ME sensor and human head models and their merging into a joint multiscale 3D FEM model. We also describe our simulation method and variations

within the model that are relevant to our investigation, as well as the process of obtaining the head model from raw MRI data. In Section 3, we showcase the results of magnetic field excitation and propagation, as well as the ME sensor array behavior for different excitation and array configurations. We highlight the magnetic and electric effects of the sensor response and investigate the magnetic crosstalk between adjacent sensors based on their position in the array and the relationship between electric response and magnetic excitation of the ME sensors. Lastly, Section 4 includes important discussion points of the obtained results and concludes our work with an emphasis on major insights gained by this study, as well as important implications for future research in the field of magnetoelectric sensors.

2. Models and Methods

The MRI-based human head model, the single ME sensor model, and the ME sensor array model were developed and combined using COMSOL Multiphysics 6.1 with its built-in *solid mechanics*, *magnetic fields*, and *electrostatics* interfaces. The software was used to set up and perform 3D FEM simulations in the frequency domain with magnetic excitation at the cantilever sensor's physical resonance frequency of 848 Hz. The software pipeline to process medical MRI data into the segmented 3D human head model, as well as the working principle, geometry, and physical properties of the ME sensors, will be explained in the following subsections. Lastly, we will describe the integration of both components into the combined multiscale head and sensor model before moving on to the results section of our study. We performed the simulations shown in this work using an Intel Xeon E5-2697A v4 CPU with 64 cores at 2.60 GHz, 503.8 GiB RAM and Ubuntu OS 22.04.3.

2.1. MRI-Based Human Head Model

In order to create an approximation of the head and improve the previous spherical model [27], MRI-based medical images were processed into a 3D FEM model of the human head for this investigation. In the literature, there are now many open- and closed-source databases from which to extract these types of medical raw data. The specific model considered in this study is derived from the NY Head Model constructed by the Parra Lab group at the City University of New York [28]. The segmentation data were obtained by averaging three different MRI sources for various tissues: the brain is acquired from the symmetric ICBM-152 v2009, non-brain tissues are extrapolated from the symmetric ICBM-152 v6, and the lower portion of the head from [28]. The general model was based and validated on four individual heads, whereby a precise FEM model was built for each of them [29]. The final segmentation files after the averaging operation involve the symmetric geometry of the entire head. This was taken into consideration due to the symmetric properties exhibited by objects imported into COMSOL Multiphysics, which can significantly reduce the computational cost of any simulations.

The provided dataset from [28] was processed and tissue regions for the skin, skull, and brain were extracted. To process the *.nifti* segmentation files, two software programs were used: *MATLAB* R2023b for creating closed and volumetric 3D objects, and *3D Slicer* 5.2.2 for refining geometry operations. In particular, the *iso2mesh* 3D library in *MATLAB*, which provides excellent computational capabilities for binary and grayscale volumetric images such as segmented MRI/CT scans, was utilized.

The *.nifti* files for each of the three tissues were loaded, opened, and converted into binary logical values to identify each grayscale level of the segmented geometry. Subsequently, a closing operation was performed on each geometry to obtain a final closed water-tight object, thus avoiding “holes” that could introduce discontinuities in the final mesh. Utilizing the *fillholes3d.m* function with a gap size of 55 for each of the three tissues resulted in the creation of the three objects. The gap size is a crucial parameter, representing the size of the hole to be filled in the geometry. In this case, a trade-off was sought, as large gap sizes would lead to a geometry significantly different from the original, especially for the brain, while very small values would introduce an almost negligible approximation in the geometry.

Subsequently, the 3D binary image was converted into an actual volume using the *imedge3d.m* function which extracts contour voxels from a binary image. Finally, the object was reconverted into the *.nifti* format and exported.

For the final processing, *3D Slicer* was used. The three objects were imported into the software after being processed in *MATLAB*. Due to the reduced complexity of the geometry after the previous closing operation, the 3D representation of the object was obtained using the Otsu thresholding method. Subsequently, cutting operations were performed in *3D Slicer*, allowing for the manual removal of small volumes from the object. These volumes are considered undesirable for the final purpose, as they would provide extra material for meshing without a specific purpose. Additionally, the lower part of the skull (including the first vertebrae of the spine) was partially removed, as this part of the head is negligible for the simulations.

After the cutting operation, some of the modified surfaces underwent shape changes. To address this, classic morphological opening and closing operations were applied to remove small extrusions remaining in the geometry and fill small residual holes. Finally, a smoothing operation was performed; more specifically, Gaussian smoothing was used.

Moreover, only for the skull, a “grow” operation was carried out using the *margin operation* tool to give it a thickness of about 3 mm. The uniform thickness of the skull adds a high degree of homogeneity to the final mesh but represents a strong approximation of the skull, while the real thickness of the skull is not uniform along the skullcap. The last step in *3D Slicer* was to export the geometry in *.stl* format, ready to be imported into *COMSOL Multiphysics 6.1*. Figure 2 illustrates the resulting geometry and mesh of the three tissue regions of the head.

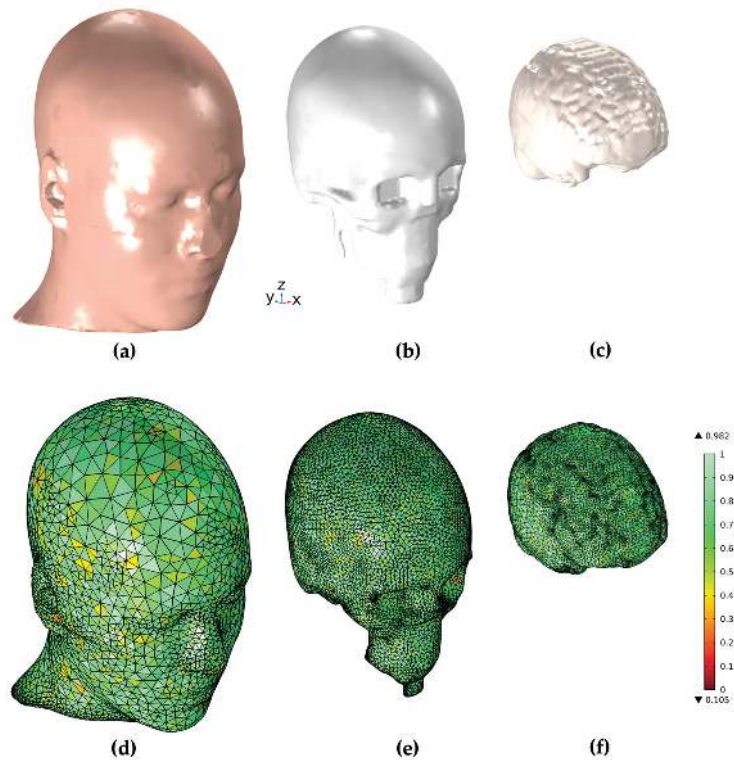


Figure 2. COMSOL Multiphysics geometry and mesh generation of the three tissue regions of the head. (a–c) show the geometry of the skin, skull, and white matter region, while (d–f) illustrate the meshed geometries for those regions in the stand-alone head model.

In order to perform simulations including the electromagnetic properties of the human head, each region was assigned a specific conductivity and relative permittivity value taken from the literature [30–32]. Table 1 gives an overview of the utilized values.

Table 1. The regions of the human head model with their respective specific conductivity and relative permittivity taken from literature.

Region	Specific Conductivity [30,31]	Relative Permittivity [32]
Skin	1.00 S/m	1,200,000
Skull	0.05 S/m	40,000–1,000,000
White Matter	0.43 S/m	30,000,000

2.2. ME Sensor Model

We designed and implemented the ME sensor based on the ME sensor models from our previous work [26,27]. Each sensor consisted of a substrate layer of silicon with 26.25 mm length, 2.45 mm width, and 300 μm thickness, a magnetostrictive layer of FeCoSiB with 22.90 mm length, 1.80 mm width, and 20 μm thickness, and a piezoelectric layer of aluminum nitride (AlN) with 25.60 mm length, 1.60 mm width, and 20 μm thickness. The AlN is polycrystalline, and both the magnetostrictive and piezoelectric layer can be produced for experimental measurements via an in-house magnetron sputtering process [33]. The magnetostrictive and piezoelectric layers were located on the opposite sides of the substrate layer. The thickness of the active layers was chosen at a factor of 10 times higher than that typically used in experimental sensors at Kiel University to reduce the computation time. The sensor operated in resonant bending mode and fixed-free configuration at a resonance frequency of 848 Hz. Figure 3 gives an overview of the ME sensor geometry.

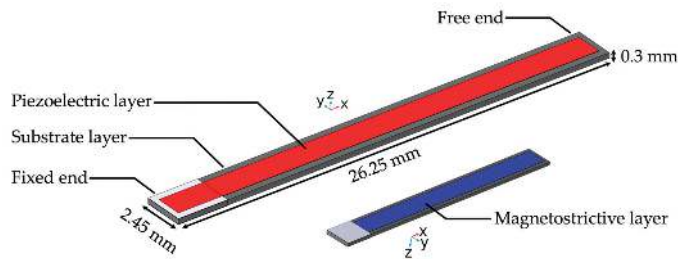


Figure 3. The ME sensor model. The cantilever sensor is shown with substrate layer in grey and the piezoelectric layer on the surface in red. The left end of the sensor is clamped, while the right end is free, resulting in the fixed-free bending mode operation. The smaller inset shows the opposite side of the substrate with the magnetostrictive layer on top in blue. The length, width, and thickness for the substrate layer are given in mm.

Table 2 gives an overview of the three layers forming the composite ME sensor. The magnetostrictive layer, piezoelectric layer, and substrate layer each have their unique material parameters which are given in Appendix A. A linearized material model was used at the sensor’s small-signal working point. Details of the applied boundary conditions and physical properties of the three layers are described in our previous work [26,27].

Table 2. The layers of the ME sensor model with their materials, length, width, and height.

Layer	Material	Length	Width	Height
Magnetostrictive	FeCoSiB	22.90 mm	1.80 mm	20.00 μm
Piezoelectric	AlN	25.60 mm	1.60 mm	20.00 μm
Substrate	Si	26.25 mm	2.45 mm	300.00 μm

The mathematical model and physical properties of cantilever ME sensors consisting of ideal and slipless composite layers are governed by systems of differential and constitutive equations which are characterized in detail in previous studies [30,34–38]. The equations that we discuss below define our 3D-FEM model and are utilized in three built-in physics interfaces in COMSOL Multiphysics 6.1: the *solid mechanics*, *magnetic fields*, and *electrostatics* interfaces. This set of equations was also utilized and described in our previous work containing ME sensors and a simplified human head model [12]. It is repeated here for easy access. Beginning with *solid mechanics*, this interface describes a set of equations that couples the mechanical, electrical, and magnetic properties of the ME sensor and contains specific terms for each layer.

$$-\rho\omega^2\mathbf{u} = \nabla \cdot \mathbf{S} \quad (1)$$

$$\mathbf{S} = \mathbf{C} : \boldsymbol{\varepsilon} \quad (2)$$

$$\boldsymbol{\varepsilon} = \frac{1}{2} [(\nabla\mathbf{u})^T + \nabla\mathbf{u}] \quad (3)$$

$$\mathbf{C} = \mathbf{C}(E, \nu). \quad (4)$$

The first set of Equations (1)–(4) corresponds to the *linear elastic material* node which covers the general mechanical properties of the model and includes $\mathbf{S}$ and $\boldsymbol{\varepsilon}$ for stress and strain, $\mathbf{u}$ as the displacement vector, the density ρ and the angular frequency ω . For the silicon substrate specifically, the coupling between stress and strain is a function of its Young's modulus and Poisson's ratio and is described by $\mathbf{C}$ in Equation (2).

The magnetostrictive and piezoelectric layers have separate physics nodes discussed below, starting with the *magnetostrictive material* node:

$$\mathbf{S} = c_H : \boldsymbol{\varepsilon} - \mathbf{H} \cdot e_{HS} \quad (5)$$

$$\mathbf{B} = \mu_0 \mu_{rS} \mathbf{H} + e_{HS} : \boldsymbol{\varepsilon}. \quad (6)$$

Equations (5) and (6) include the coupling matrix in Voigt notation e_{HS} and the elasticity matrix c_H . This node also governs the relation between the magnetic flux density vector $\mathbf{B}$, the magnetic field vector $\mathbf{H}$, and the relative permeability μ_{rS} of the material. Analogous to the *magnetostrictive material* node, the *piezoelectric material* node has a specific set of equations that describe its properties, including the coupling of the ME sensor's electric and elastic properties:

$$\nabla \cdot \mathbf{D} = \rho_v \quad (7)$$

$$\mathbf{S} = c_E : \boldsymbol{\varepsilon} - \mathbf{E} \cdot e_{ES} \quad (8)$$

$$\mathbf{D} = \varepsilon_0 \varepsilon_{rS} \mathbf{E} + e_{ES} : \boldsymbol{\varepsilon}. \quad (9)$$

In Equation (7), Gauss's law for the relation between the electric displacement field $\mathbf{D}$ and the volume charge density is applied to the model. Equations (8) and (9) include the coupling matrix in Voigt notation e_{ES} and the elasticity matrix c_E , as well as $\mathbf{E}$ for the electric field vector and ε_{rS} for the relative permittivity. We used the stress-magnetization and the stress-charge form for the magnetostrictive and piezoelectric material nodes, respectively. The Equations (1)–(9) that we described so far are applied in the *solid mechanics* physics node of our model and include elastic material properties, as well as coupling between the sensor layers. The electric and magnetic behavior of our model is governed in the *magnetic fields* and *electrostatics physics* nodes of our model given in Equations (10)–(12):

$$\nabla \times \mathbf{H} = \sigma \mathbf{E} + j\omega \mathbf{D} \quad (10)$$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (11)$$

$$\mathbf{E} = -\nabla V - j\omega \mathbf{A} \quad (12)$$

$$\mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E} \quad (13)$$

$$\mathbf{B} = \mu_0 \mu_r \mathbf{H}. \quad (14)$$

With Equations (10)–(12) we utilized Maxwell's equation and related electric and magnetic fields to the magnetic vector potential $\mathbf{A}$. The equation for the electric field strength used in the model is given in (12). Equations (13) and (14) apply via the boundary condition *Ampère's law* in the *magnetic fields* physics interface to the silicon substrate. Additionally, this boundary condition defines the general constitutive relations of $\mathbf{D}$ and $\mathbf{E}$ for the magnetostrictive material, as well as $\mathbf{B}$ and $\mathbf{H}$ for the piezoelectric material, in instances where no more specific physical properties are assigned, i.e., no magnetostriction for the piezoelectric material or piezoelectricity for the magnetostrictive material. With the set of equations given in (1)–(14), our multiscale combined ME sensor and human head model can be simulated with different electric and magnetic excitation methods, and full magnetoelectric coupling of the sensor layers.

2.3. ME Sensor Array Model

The aim of this work is to combine 15 individual ME sensors into a sensor array and investigate the response of each individual sensor based on its relative position in the array and distance from the neighboring sensors for different types of magnetic excitation. To achieve this goal, we organized the sensors into a 3×5 grid to form an array with three rows and five columns of adjacent sensors. The long axis of the sensors was parallel to the x -axis and the piezoelectric layer faced upwards in z -direction. The distances between the sensors in the array varied from 5 mm up to 5 cm in the vertical and horizontal directions simultaneously. Figure 4 illustrates the array configuration with the distance between neighboring sensors set to 5 mm. An array of these or similar dimensions was chosen due to in-house fabrication approaches at Kiel University and the possibility of using the array in an operation room, as a smaller array with variable positioning would potentially allow surgical procedures without covering the entire head.

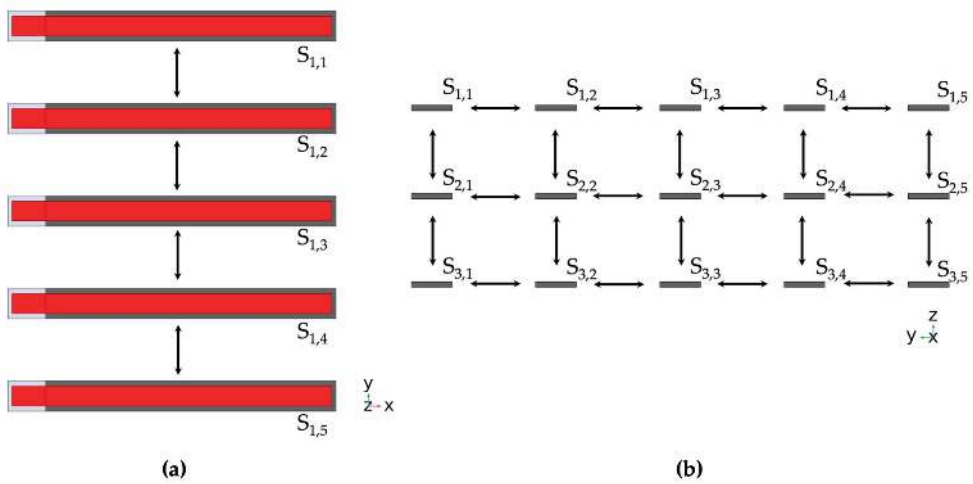


Figure 4. (a) Top-down view and (b) frontal view (on fixed end) of the 15-sensor array. Between each adjacent sensor is a variable distance between 1 mm and 5 cm. This figure shows equidistant sensor placement with 5 mm between the neighbors in both vertical and horizontal direction. The variable distance allowed us to investigate the magnetic crosstalk between sensors and analyze the influence of the magnetostrictive layers on nearby ME sensors.

2.4. Combined MRI-Based Human Head and ME Sensor Array Model

After establishing the human head and ME sensor array models separately, we combined both parts into the joint multiscale model. The sensors in the array can be located at arbitrary positions in space and distances both from the head and adjacent sensors. The challenge to be overcome was to combine both the anatomical head model, as well as the ME sensor array model with 15 sensors, into one combined multiscale and multi-physics model. A core consideration for this was the differences in dimensions between the different model domains. While the sensors' PE and MS layers had a thickness of only 20 μm , the diameter of the head was approximately 20 cm. This translated to a vastly different size and resolution of mesh elements for the differently sized domains, because the mesh is designed such that extremely thin layers are of a sufficiently small size, while larger domains are modeled with large element sizes to reduce the computational load. The number of degrees of freedom (DOF) solved for in a combined model with three ME sensors demonstrated in previous work [27] is approximately 25 million, while the number increases to approximately 70 million DOF when including 15 ME sensors in the array. The number of DOF is determined by the amount of mesh elements, as well as the utilized physics in the model, and is an important figure of merit for the computational requirements to solve a given FEM model. Details on the mesh parameters for all structures in a model with three ME sensors are given in Appendix B, with sizes referring to the longest edge of tetrahedral mesh elements.

Increasing the number of sensors from three to 15 posed a challenge in terms of computation time and hardware requirements for the computer running our simulations, since the number of DOF almost tripled in the combined model with a larger array and full mechanical, magnetic and electric coupling. Using the mesh parameters given in Appendix B, we were not able to achieve a successful simulation run with a converged solution. In order to achieve a solution, we iteratively adjusted some parts of the mesh to be coarser than in the previous study [27], while keeping the mesh resolution as high as possible for the sensor geometry. The adjusted mesh parameters for this investigation step included the head geometry, the air environment, and the substrate layer of the ME sensors. The magnetostrictive and piezoelectric layers remained unchanged. The mesh

parameters for the combined model with 15 sensors are given in Appendix C. Inspecting the mesh yielded some elements of low quality for the figure of merit skewness, but no significant number of elements of poor quality (defined by skewness under 0.1 according to COMSOL's guidelines). The coarser mesh might have adversely impacted the accuracy of the solutions provided in the results section, but was necessary to enable our simulations to finish successfully.

Due to challenges with the numerical stability of our solutions, two simplifying conditions had to be applied to our model. Firstly, based on communications with COMSOL employees, calculating 3D FEM models with very high differences in material parameters such as the specific conductivity may result in a failure to find a converged solution [39]. In this case, this affects the near-zero specific conductivity of air which fills most of the modelling space. A recommended solution for this is to artificially increase the specific conductivity of the material in question sufficiently; thus, we set the conductivity of our air domain to 1×10^{-6} S/m. Secondly, we increased the numerical stability of our simulations with dipole excitation by positioning a second dipole with the same orientation and a magnitude $1 \mu\text{A}\cdot\text{m}$ at a distance of 2 mm directly below the main excitation dipole of magnitude $1 \text{ mA}\cdot\text{m}$. Due to the factor of a thousand between dipole magnitudes, the contribution of the secondary dipole to the overall electromagnetic field is considered negligible, while empirically improving the convergence of the utilized indirect solver we used. The application of similar conditions to improve numerical stability was also discussed in our previous work [27].

3. Results

The results in this section are categorized into two different types of magnetic field excitation. First, we looked at the array's response in a constant, homogeneous magnetic field, which is applied to the entire model volume. Following that, we replaced the homogeneous magnetic field excitation with a single electric current point dipole source inside the white matter region of the head model. For the investigation of the crosstalk effect, the distance between adjacent sensors in both the horizontal and vertical directions was varied in four steps within an interval between 5 mm and 5 cm. As the distance between neighboring sensors changes, the magnetic flux between them changes direction and is guided inside the highly permeable magnetostrictive layers of the sensors. For each sensor distance, a simulation with either homogeneous or dipole excitation was performed and the response for each sensor in the array was evaluated. Different distances between sensors result in different sensor responses depending on the position inside the array, as the following sections will demonstrate.

3.1. Homogeneous Excitation

We performed magnetic excitation of the ME sensor array using a homogeneous magnetic field strength of 1 A/m in x -direction applied to the entire model environment. The aim of this study step was to establish and validate the basic sensor response to a simple excitation field and investigate potential crosstalk effects independently of influences of inhomogeneous field effects. Before each array simulation, validation steps were conducted for homogeneous and dipole excitation, i.e., the sensor material parameters were set to those of the air environment in order to eliminate geometric or numeric inconsistencies and validate the magnetic flux density inside the model domain without the presence of high-permeability sensor material. Figure 5a shows a schematic of the full array and highlights the five sensors of the middle row, namely sensors $S_{2,1}$, $S_{2,2}$, $S_{2,3}$, $S_{2,4}$, and $S_{2,5}$. We further highlight the exemplary behavior of these five sensors in Figure 5b, where we display the sensors in a homogeneous excitation magnetic field in x -direction and the corresponding magnetic flux density inside the magnetostrictive layer for each sensor. The distance between adjacent sensors was 1 cm. The outermost sensors $S_{2,1}$ and $S_{2,5}$ exhibited the highest magnetic flux density, followed by sensors $S_{2,2}$ and $S_{2,4}$. The central sensor, $S_{2,3}$, exhibited the lowest magnetic flux density out of all sensors in the array due to its central

position and the resulting crosstalk with its adjacent sensors. Finally, Figure 5c gives a plot of the magnetic flux density norm inside every sensor of the full array with 15 sensors for different distances between adjacent sensors. The previously observed behavior of high crosstalk between adjacent sensors at low distances is clearly visible, while the closer a sensor is to the center of the array, the stronger the effect. At a distance of 5 cm between the sensors, the crosstalk effect is negligible, and all sensors exhibit approximately the same response. At each individual distance between adjacent sensors, the sensor with the lowest magnetic flux density was the middle sensor ($S_{2,3}$), while the highest flux density was shared between the four corner sensors ($S_{1,1}$, $S_{1,5}$, $S_{3,1}$, $S_{3,5}$). At the lowest distance of 5 mm, the flux density in the corner sensors was 7.9% higher than in the middle sensor. For the central sensor ($S_{2,3}$), the magnetic flux density was 15.6% lower at a distance of 5 mm compared to a distance of 5 cm to its neighbors. Notably, the crosstalk effect was significantly stronger with up to 11% for sensors with closely vertically adjacent sensors, compared to horizontally adjacent sensors, based on simulations considering exclusively horizontally or vertically adjacent sensors. For the electric potential, Figure 5d offers similar general behavior for the crosstalk effect based on the electric behavior. Here, while the values converge for all sensors at a distance of 5 cm between neighbors for the magnetic flux density, the electric potential still sees a difference of approximately 4% between the highest (S_{33}) and lowest (S_{12}) at that distance. We expect numerical error to this degree based on the calculation of the fully coupled magnetoelectric effect between the layers as a possibility for the slightly diverging behavior of the electric response of the sensors.

3.2. Dipole Excitation

After evaluation of the behavior of the ME sensor array in homogeneous magnetic field excitation, we moved towards a specific inhomogeneous excitation mechanism. An electric current point dipole was placed at coordinates $x = 30$ mm, $y = -20$ mm, and $z = 30$ mm inside the white matter compartment of the head geometry, within the approximated right cerebral hemisphere. This configuration serves as a representative of a deep-brain stimulation scenario. The dipole moment, combined with the electric properties of the head, results in an induced magnetic field propagating through the head and the air environment into the ME sensor array, enabling us to evaluate its response and gain further insights into the behavior of the individual sensors in inhomogeneous excitation.

Homogeneous excitation

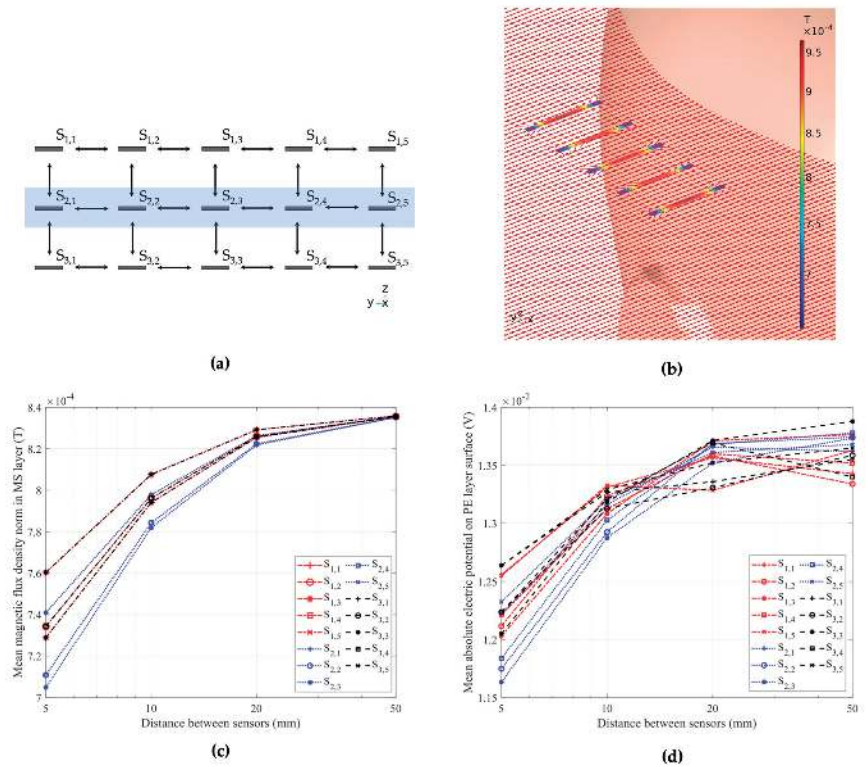


Figure 5. (a) Schematic of the 15-sensor array in top-down and frontal view. The blue rectangle marks the second row of five sensors, which we used to visualize the crosstalk effects between sensors. The distance between sensors in this schematic is 5 mm. (b) The 15-sensor array in a homogeneous magnetic field. The magnetic field strength is 1 A/m and is applied in x -direction. The highlighted middle row of five sensors within the array, namely the sensors $S_{2,1}$, $S_{2,2}$, $S_{2,3}$, $S_{2,4}$, and $S_{2,5}$, is shown, with clear crosstalk effects between sensors. The distance between the sensors is 1 cm in this exemplary position. (c) The magnetic flux density inside the MS layers and (d) the electric potential on the surface of the PE layer of each of the 15 sensors of the array at different distances from adjacent sensors, showcasing crosstalk at small distances between the sensors and increased effects in the central sensors.

Based on the orientation of the dipole and the ME sensor, the sensor response can be vastly different due to its directional sensitivity, as we showed in our previous work with an array of orthogonally oriented ME sensors and different dipole orientations [27]. For this work, we exemplarily investigated only one dipole orientation (y -direction) and one sensor orientation for each sensor in the array (x -direction), but modifications to this model with arbitrary configurations for arrays and sources are possible for further analysis. The dipole can be configured with an arbitrary dipole moment direction, amplitude, and location in the head. We chose an exemplary dipole moment of 1 mA·m in y -direction. The chosen dipole amplitude is in agreement with studies on deep brain stimulation (DBS) and head models from the literature [32,40]. As seen in previous investigations, it is not trivial to predict the behavior of the sensor array when excited by an inhomogeneous magnetic field. As seen in Figure 6, the individual sensors' magnetic and electric behavior does not follow specific patterns with increasing distance to neighbors. Some indicators can partly predict the behavior; for example, the fact that in the vertical array configuration, the bottom row of

sensors exhibits a higher magnetic flux density based on proximity to the human head and the magnetic field propagation through the tissue. In this case, all five sensors in the bottom row of sensors ($S_{3,1}$ – $S_{3,5}$) exhibited the highest magnetic and electric response at distances of between 5 mm and 1 cm between adjacent sensors. Similarly, four out of the five sensors ($S_{3,1}$ – $S_{3,4}$) and three out of the four sensors ($S_{3,2}$ – $S_{3,4}$) with the highest magnetic flux density and electric potential at distances of 2 cm and 5 cm from their neighbors, respectively, are sensors from the bottom row of the array.

Dipole excitation

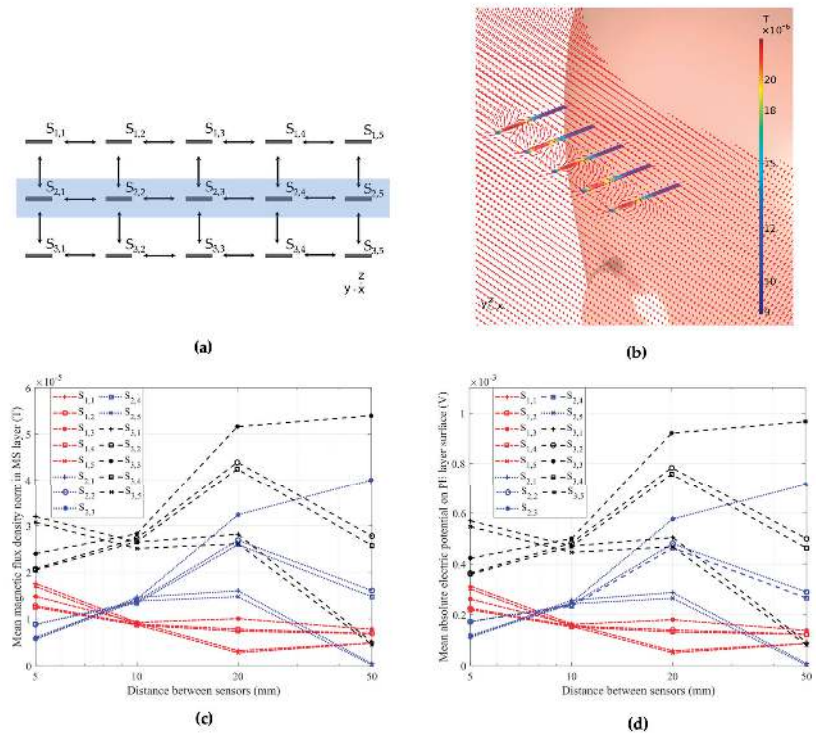


Figure 6. (a) Schematic of the 15-sensor array in top-down and frontal view. The blue rectangle marks the second row of five sensors, which we used to visualize the crosstalk effects between sensors. The distance between sensors in this schematic is 5 mm. (b) The 15-sensor array in a dipole magnetic field. The dipole moment is 1 mA·m in magnitude and oriented in y-direction. The highlighted middle row of five sensors within the array, namely the sensors $S_{2,1}$, $S_{2,2}$, $S_{2,3}$, $S_{2,4}$, and $S_{2,5}$, is shown, with possible crosstalk and flux concentration effects between sensors. The distance between the sensors is 1 cm in these exemplary positions. (c) The magnetic flux density inside the MS layers and (d) the electric potential on the surface of the PE layer of each of the 15 sensors of the array at different distances from adjacent sensors, showcasing an increased sensor response in the bottom row of sensors in the array ($S_{3,1}$ – $S_{3,5}$), but inconsistent behavior with increasing sensor distance.

4. Discussion

Magnetic crosstalk effects between the adjacent ME sensors are clearly observed for the homogeneous excitation case in Figure 5; they are more challenging to visualize for the inhomogeneous case in Figure 6 with a strong spatial variation of the magnetic field strength and direction. To gain a better understanding of the observed results, we investigated the sensor sensitivity. As the ME sensor has a previously demonstrated directional sensitivity [27], we investigated the effect of the x -component of the magnetic field, which is parallel to the cantilever's long axis. We calculated two ratios between

important model parameters in order to discuss the presented sensor behavior in both homogeneous and inhomogeneous excitation cases. The first ratio was between the mean absolute B_x component of the magnetic flux density inside the MS layer of the sensor, and the same layer with its material parameters set to those of air. The second ratio was between the mean absolute electric potential as the electric sensor response and the B_x component of the magnetic flux density inside the magnetostrictive layer of each sensor. Figure 7a,b demonstrate the results for the ratios between sensor and air magnetic flux densities, while Figure 7c,d visualize the ratios between electric potential and sensor magnetic flux density for both excitation cases and each individual sensor inside the array. Here, the horizontal axis determines the sensor number with horizontally adjacent sensor columns ($S_{1,n}$ – $S_{1,n}$), while the markers differentiate between the vertically adjacent rows of sensors ($S_{m,1}$ – $S_{m,3}$), as illustrated in Figures 5a and 6a. The four different colors represent the different distances between adjacent sensors, with distances of 5 mm, 1 cm, 2 cm, and 5 cm in the horizontal and vertical directions, respectively. The results show that, while we observe magnetic flux densities and piezoelectric voltages that are highly dependent on the sensor position and distance to its neighbors in both homogeneous and inhomogeneous excitation, the ratio between electric response and excitation field was near constant for each excitation case.

As seen in Figure 7a,b, the ratio between the magnetic flux densities inside the sensor geometry and the corresponding air volumes were not constant with up to two orders of magnitude in the homogeneous and four orders of magnitude in the inhomogeneous case between the ratios at different sensor positions. This large spread showcases the field concentration effect of the magnetostrictive material and highly position-dependent behavior of the sensors. Contrary to the highly variable ratios between the magnetic flux densities, Figure 7c illustrates a narrow range of values between 16.0 and 16.6 for the ratio between electric potential and magnetic flux density for the homogeneous excitation case and any sensor at the investigated array positions and distances. For the inhomogeneous case, Figure 7d shows different behavior between the middle row of sensors (Row 2, $S_{2,1}$ – $S_{2,5}$) at a distance of 5 mm between adjacent sensors and every other configuration. A factor between 17.4 and 18.0 was calculated for every sensor position, with outliers for the first and last sensor in the second row ($S_{2,1}$ and $S_{2,5}$) at a distance of 5 cm exhibiting a factor of approximately 19.0.

The entire second row of sensors ($S_{2,1}$ – $S_{2,5}$) at the minimum distance of 5 mm between sensors exhibited a factor between 19.5 and 19.7. Despite the outlying row of sensors in the inhomogeneous case, all 15 sensors in either homogeneous or dipole excitation fields exhibited similar, near-constant ratios between the electric response and the magnetic excitation. Considering the clear crosstalk effects demonstrated for the homogeneous case in Figure 5 and the seemingly inconsistent behavior in dipole excitation seen in Figure 6, with up to two orders of magnitude difference in magnetic flux density between different sensor positions in Figure 6c, these near-constant ratios demonstrated a highly linear relation between the excitation magnetic field and the electric sensor response, including potential crosstalk and flux concentration effects. Therefore, the change in the electric sensor response may be attributed to the change in the magnetic field in the sensor.

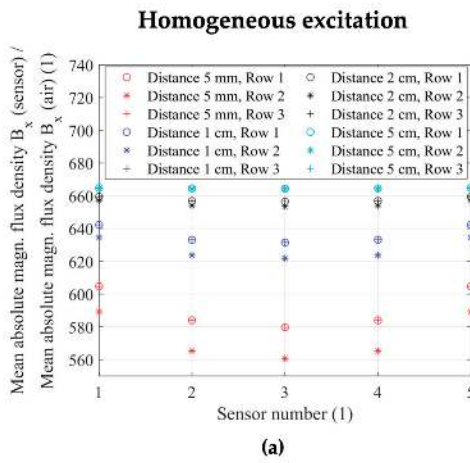
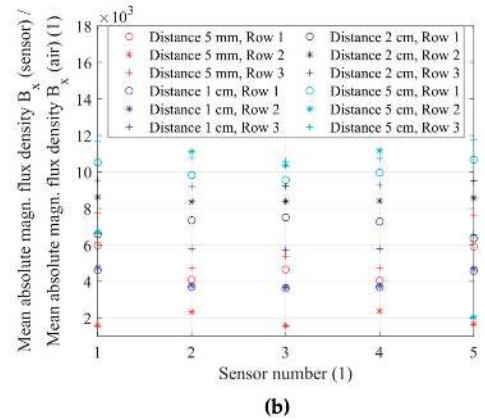
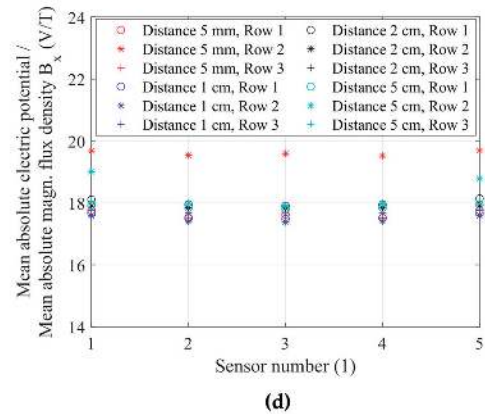
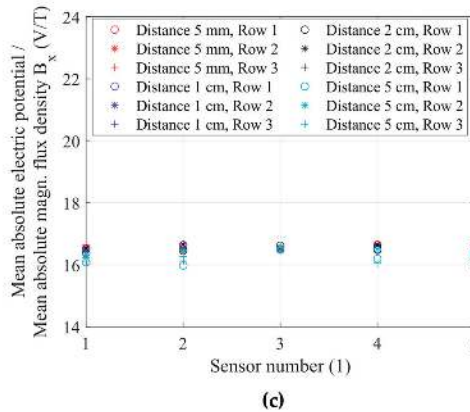
Ratio of magnetic flux density
in sensor and air**Dipole excitation**Ratio of electric potential
and magnetic flux density in sensor

Figure 7. (a,b) The ratio between the mean absolute x-component of the magnetic flux density inside the MS layer geometry for sensor material parameters and sensor material parameters set to air in (a) homogeneous and (b) dipole excitation cases. This ratio is a measure of the field concentration and depends on the sensor position in the array. (c,d) The ratio between the mean absolute electric potential over the PE layer and the mean absolute magnetic flux density component in x-direction in each sensor of the array and for different distances between adjacent sensors. (c) illustrates the near-constant factors between 16.0 and 16.6 for the ratio for each sensor and each position within the array in homogeneous excitation. (d) shows a factor between 17.4 and 18.0 for most sensors, with the outlying second row of sensors ($S_{2,1}$ – $S_{2,5}$) at the minimum distance of 5 mm between adjacent sensors exhibiting a factor between 19.5 and 19.7.

5. Conclusions

We have shown a combined ME sensor array and MRI-based human head model for joint biomagnetic field simulations and analysis of ME sensor behavior. The combined model allowed us to place an electric dipole source inside the head and simulate the excitation and propagation of an electromagnetic field through the head's tissue regions, the air environment, and into the array of ME sensors. As an alternative study step to the dipole excitation, a homogeneous excitation field was also applied to the entire simulation environment and the sensor responses were evaluated.

The results that we presented offer insights on the response of individual ME sensors within an array configuration to different excitation mechanisms. For an array of 15 ME

sensors in a homogeneous magnetic excitation field, a magnetic crosstalk effect between adjacent sensors is clearly visible in Figure 5. The magnitude of this effect decreases with increasing distance between adjacent sensors and becomes negligible at a distance of 5 cm. In the case of inhomogeneous excitation, a strong change in the sensor response is obtained as seen in Figure 6 due to the position-dependent magnetic field strength and direction. The sensor response, in this case, depends on the position and orientation of the array relative to the source inside the head, as well as the propagation of the electromagnetic field through the head based on the electric tissue properties and geometry of the head [27]. In excitation fields that have large vector components in directions other than the sensitive direction of the ME sensor, differences between sensor positions may be more pronounced, as opposed to a homogeneous excitation field in a sensitive direction, which is demonstrated in this study. Another effect that has to be considered is the inverse magnetostrictive effect, which is caused by strain in the magnetostrictive layer and can lead to superimposing local fields on top of the excitation field [34]. This effect could have a varying degree of influence on the behavior of adjacent sensors, particularly in inhomogeneous magnetic fields.

To investigate whether the linear relation between the excitation magnetic field and the electric sensor response is maintained, we calculated the ratio of these quantities for each individual ME sensor in both excitation setups and four different distances between adjacent sensors inside of the array. We demonstrated that even with large differences up to factors of two orders of magnitude in the magnetic and electric sensor responses between sensors in various combinations of array geometry and magnetic field sources, the ratio between the electric potential and magnetic flux density inside the sensors was near-constant in homogeneous (16.0–16.6) and inhomogeneous (17.4–19.7) excitation. This result is shown in Figure 7 and confirms that the behavior of each ME sensor was linear as expected for the linear model, even though a highly position-dependent field concentration is demonstrated. Therefore, for the linear approximation, a separation of the excitation source and sensor model in both homogeneous and inhomogeneous excitation configurations is possible. In the next step, the FEM model can be extended to investigate nonlinear effects due to nonlinear material properties as well as effects due to nonlinear secondary currents induced inside the head.

In summary, this work contains insights into the response of ME sensors within different array configurations for homogeneous and inhomogeneous dipole magnetic field excitation for the small-signal linearized case. The sensor response strongly depends on the excitation vector field and confirms the influence of magnetic crosstalk between sensors. Further research may include additional simulations with different array configurations in inhomogeneous fields to evaluate the near-constant sensitivity of the sensors. Additionally, results could be compared between separately evaluated source and sensor simulations and the sensor response in combined models such as the one presented in this work. Future excitation models could also evolve the head geometry demonstrated in this work to contain more tissue regions in higher resolution, as well as different biological or artificial excitation sources such as implanted DBS electrodes, giving rise to potential localization and orientation investigations with patient- or application-specific head and ME sensor models.

Author Contributions: Conceptualization, M.-Ö.Ö. and M.G.; methodology, M.-Ö.Ö., G.B. and M.G.; validation, M.-Ö.Ö.; investigation, M.-Ö.Ö.; writing—original draft preparation, M.-Ö.Ö.; writing—review and editing, M.-Ö.Ö., G.B. and M.G.; visualization, M.-Ö.Ö. and G.B.; supervision, M.G.; funding acquisition, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) via the collaborative research center CRC 1261 “Magnetolectric Sensors: From Composite Materials to Biomagnetic Diagnostics”.

Data Availability Statement: The data that the findings of this work are based on are available from the corresponding authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A. Material Parameters

The material parameters of AlN, FeCoSiB, and silicon at the working point used for the ME sensor are based on literature and in-house characterization and are given in this section [35,41–43].

AlN:

$$c_{E,\text{AlN}} = \begin{pmatrix} 41 & 14.9 & 9.9 & 0 & 0 & 0 \\ 14.9 & 41 & 9.9 & 0 & 0 & 0 \\ 9.9 & 9.9 & 38.9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 12.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 12.5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 12.5 \end{pmatrix} \times 10^{10} \text{ Pa} \quad (\text{A1})$$

$$e_{ES,\text{AlN}} = \begin{pmatrix} 0 & 0 & 0 & 0 & -0.48 & 0 \\ 0 & 0 & 0 & -0.48 & 0 & 0 \\ 9.9 & 9.9 & 38.9 & 0 & 0 & 0 \\ -0.58 & -0.58 & 1.55 & 0 & 0 & 0 \end{pmatrix} \times \text{C/m}^2 \quad (\text{A2})$$

$$\rho_{\text{AlN}} = 3300 \text{ kg/m}^3 \quad (\text{A3})$$

$$\varepsilon_{\text{AlN}} = 80 \times 10^{-12} \text{ F/m} \quad (\text{A4})$$

$$\mu_{\text{AlN}} = 0.4\pi \times 10^{-6} \text{ H/m} \quad (\text{A5})$$

FeCoSiB:

$$c_{H,\text{FeCoSiB}} = \begin{pmatrix} 150 & 45 & 45 & 0 & 0 & 0 \\ 45 & 150 & 45 & 0 & 0 & 0 \\ 45 & 45 & 150 & 0 & 0 & 0 \\ 0 & 0 & 0 & 40 & 0 & 0 \\ 0 & 0 & 0 & 0 & 40 & 0 \\ 0 & 0 & 0 & 0 & 0 & 40 \end{pmatrix} \times 10^{10} \text{ Pa} \quad (\text{A6})$$

$$e_{HS,\text{FeCoSiB}} = \begin{pmatrix} 8500 & -2833.3 & -2833.3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \times \text{N/(Am)} \quad (\text{A7})$$

$$\rho_{\text{FeCoSiB}} = 7250 \text{ kg/m}^3 \quad (\text{A8})$$

$$\varepsilon_{\text{FeCoSiB}} = 8.85 \times 10^{-12} \text{ F/m} \quad (\text{A9})$$

$$\mu_{\text{FeCoSiB}} = 1.13 \times 10^{-3} \text{ H/m} \quad (\text{A10})$$

Silicon:

$$E_{\text{Si}} = 170 \times 10^9 \text{ Pa} \quad (\text{A11})$$

$$\nu_{\text{Si}} = 0.28 \quad (\text{A12})$$

$$\rho_{\text{Si}} = 2329 \text{ kg/m}^3 \quad (\text{A13})$$

$$\epsilon_{Si} = 103.59 \times 10^{-12} \text{ F/m}$$

(A14)

$$\mu_{Si} = 0.4\pi \times 10^{-6} \text{ H/m}$$

(A15)

Appendix B. Mesh Parameters for Original 3-Sensor Array Model

Head and sensor model geometry parameters for a model with 3 ME sensors [27]. The mesh is designed so that extremely thin layers of PE and MS material are of a sufficiently small size, while larger structures such as the head and airbox are allowed larger element sizes to decrease the total number of degrees of freedom and therefore the computational load for the simulations. Values given are for the longest edges of tetrahedral mesh elements.

Mesh Parameter	PE and MS Layer	Substrate Layer	Skin	Skull	White Matter	Airbox
Max. element size	$5.00 \times 10^{-4} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$3.61 \times 10^{-2} \text{ m}$
Min. element size	$9.00 \times 10^{-5} \text{ m}$	$1.00 \times 10^{-4} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$
Max. element growth rate	3	3	1.45	1.45	1.45	1.4
Curvature factor	0.3	0.3	0.5	0.5	0.5	0.4
Resolution of narrow regions	3	3	0.6	0.6	0.6	0.7

Appendix C. Mesh Parameters for Modified 15-Sensor Array Model

In order to be able to run simulations with the computationally highly demanding 15-sensor array model on our hardware, modifications to the mesh parameters given in Appendix B had to be made. The following table includes the modified minimum element sizes for the substrate layer, the head geometry, and the surrounding air volume. The sensors' piezoelectric and magnetostrictive layers remain unchanged.

Mesh Parameter	PE and MS Layer	Substrate Layer	Skin	Skull	White Matter	Airbox
Max. element size	$5.00 \times 10^{-4} \text{ m}$	$1.00 \times 10^{-3} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$5.25 \times 10^{-2} \text{ m}$	$3.61 \times 10^{-2} \text{ m}$
Min. element size	$9.00 \times 10^{-5} \text{ m}$	$3.00 \times 10^{-4} \text{ m}$	$6.56 \times 10^{-3} \text{ m}$	$6.56 \times 10^{-3} \text{ m}$	$6.56 \times 10^{-3} \text{ m}$	$2.63 \times 10^{-3} \text{ m}$
Max. element growth rate	3	3	1.45	1.45	1.45	1.4
Curvature factor	0.3	0.3	0.5	0.5	0.5	0.4
Resolution of narrow regions	3	3	0.6	0.6	0.6	0.7

References

1. Muthuraman, M.; Moliadze, V.; Mideksa, K.G.; Anwar, A.R.; Stephani, U.; Deuschl, G.; Freitag, C.M.; Siniatchkin, M. EEG-MEG Integration Enhances the Characterization of Functional and Effective Connectivity in the Resting State Network. *PLoS ONE* **2015**, *10*, e0140832. [CrossRef]

2. Primdahl, F. The fluxgate magnetometer. *J. Phys. E* **1979**, *12*, 241–253. [CrossRef]

3. Auster, H.U.; Glassmeier, K.H.; Magnes, W.; Aydogar, O.; Baumjohann, W.; Constantinescu, D.; Fischer, D.; Fornacon, K.H.; Georgescu, E.; Harvey, P.; et al. The THEMIS Fluxgate Magnetometer. *Space Sci. Rev.* **2008**, *141*, 235–264. [CrossRef]

4. Jodko-Władzińska, A.; Wildner, K.; Pałko, T.; Władziński, M. Compensation System for Biomagnetic Measurements with Optically Pumped Magnetometers inside a Magnetically Shielded Room. *Sensors* **2020**, *20*, 4563. [CrossRef] [PubMed]

5. Wikswo, J.P. SQUID magnetometers for biomagnetism and nondestructive testing: Important questions and initial answers. *IEEE Trans. Appl. Supercond.* **1995**, *5*, 74–120. [CrossRef]

6. Galopin, N.; Mininger, X.; Bouillault, F.; Daniel, L. Finite Element Modeling of Magnetoelectric Sensors. *IEEE Trans. Magn.* **2008**, *44*, 834–837. [CrossRef]

7. Reermann, J.; Durdaut, P.; Salzer, S.; Demming, T.; Piorra, A.; Quandt, E.; Frey, N.; Höft, M.; Schmidt, G. Evaluation of magnetoelectric sensor systems for cardiological applications. *Measurement* **2018**, *116*, 230–238. [CrossRef]

8. Zuo, S.; Schmalz, J.; Ozden, M.-O.; Gerken, M.; Su, J.; Niekie, F.; Lofink, F.; Nazarpour, K.; Heidari, H. Ultrasensitive Magneto-electric Sensing System for Pico-Tesla MagnetoMyoGraphy. *IEEE Trans. Biomed. Circuits Syst.* **2020**, *14*, 971–984. [CrossRef]
9. Elzenheimer, E.; Bald, C.; Engelhardt, E.; Hoffmann, J.; Hayes, P.; Arbustini, J.; Bahr, A.; Quandt, E.; Höft, M.; Schmidt, G. Quantitative Evaluation for Magnetolectric Sensor Systems in Biomagnetic Diagnostics. *Sensors* **2022**, *22*, 1018. [CrossRef]
10. Mandal, P.K.; Banerjee, A.; Tripathi, M.; Sharma, A. A Comprehensive Review of Magnetoencephalography (MEG) Studies for Brain Functionality in Healthy Aging and Alzheimer’s Disease (AD). *Front. Comput. Neurosci.* **2018**, *12*, 60. [CrossRef]
11. Yalaz, M.; Noor, M.S.; McIntyre, C.C.; Butz, M.; Schnitzler, A.; Deuschl, G.; Höft, M. DBS electrode localization and rotational orientation detection using SQUID-based magnetoencephalography. *J. Neural Eng.* **2021**, *18*, 026021. [CrossRef]
12. Bald, C.; Bergholz, R.; Schmidt, G. Automatic Localization of an Ultrasound Probe with the Help of Magnetic Sensors. *Curr. Dir. Biomed. Eng.* **2022**, *8*, 317–320. [CrossRef]
13. Yalaz, M.; Maling, N.; Deuschl, G.; Juárez-Paz, L.M.; Butz, M.; Schnitzler, A.; Helmers, A.-K.; Höft, M. MaDoPO: Magnetic Detection of Positions and Orientations of Segmented Deep Brain Stimulation Electrodes: A Radiation-Free Method Based on Magnetoencephalography. *Brain Sci.* **2022**, *12*, 86. [CrossRef] [PubMed]
14. Hoffmann, J.; Roldan-Vasco, S.; Krüger, K.; Niekie, F.; Hansen, C.; Maetzler, W.; Orozco-Arroyave, J.R.; Schmidt, G. Pilot Study: Magnetic Motion Analysis for Swallowing Detection Using MEMS Cantilever Actuators. *Sensors* **2023**, *23*, 3594. [CrossRef]
15. Bibes, M.; Barthélémy, A. Towards a magnetolectric memory. *Nat. Mater.* **2008**, *7*, 425–426. [CrossRef] [PubMed]
16. Garten, L.M.; Staruch, M.L.; Bussmann, K.; Wollmershauser, J.; Finkel, P. Enhancing Converse Magnetolectric Coupling Through Strain Engineering in Artificial Multiferroic Heterostructures. *ACS Appl. Mater. Interfaces* **2022**, *14*, 25701–25709. [CrossRef]
17. Liang, X.; Chen, H.; Sun, N.; Luo, B.; Golubeva, E.; Müller, C.; Mahat, S.; Wei, Y.; Dong, C.; Zaeimbashi, M.; et al. Mechanically Driven Solidly Mounted Resonator-Based Nanoelectromechanical Systems Magnetolectric Antennas. *Adv. Eng. Mater.* **2023**, *25*, 2300425. [CrossRef]
18. Zaeimbashi, M.; Nasrollahpour, M.; Khalifa, A.; Romano, A.; Liang, X.; Chen, H.; Sun, N.; Matyushov, A.; Lin, H.; Dong, C.; et al. Ultra-compact dual-band smart NEMS magnetolectric antennas for simultaneous wireless energy harvesting and magnetic field sensing. *Nat. Commun.* **2021**, *12*, 3141. [CrossRef]
19. Bichurin, M.; Petrov, R.; Leontiev, V.; Semenov, G.; Sokolov, O. Magnetolectric Current Sensors. *Sensors* **2017**, *17*, 1271. [CrossRef]
20. Petrov, R.; Leontiev, V.; Sokolov, O.; Bichurin, M.; Bozhkov, S.; Milenov, I.; Bozhkov, P. A Magnetolectric Automotive Crankshaft Position Sensor. *Sensors* **2020**, *20*, 5494. [CrossRef]
21. Viehland, D.; Wuttig, M.; McCord, J.; Quandt, E. Magnetolectric magnetic field sensors. *MRS Bull.* **2018**, *43*, 834–840. [CrossRef]
22. Rice, J.K.; Rorden, C.; Little, J.S.; Parra, L.C. Subject position affects EEG magnitudes. *NeuroImage* **2013**, *64*, 476–484. [CrossRef] [PubMed]
23. Vorwerk, J.; Oostenveld, R.; Piastra, M.C.; Magyari, L.; Wolters, C.H. The FieldTrip-SimBio pipeline for EEG forward solutions. *Biomed. Eng. OnLine* **2018**, *17*, 37. [CrossRef]
24. Piastra, M.C.; Nüßing, A.; Vorwerk, J.; Bornfleth, H.; Oostenveld, R.; Engwer, C.; Wolters, C.H. The Discontinuous Galerkin Finite Element Method for Solving the MEG and the Combined MEG/EEG Forward Problem. *Front. Neurosci.* **2018**, *12*, 30. [CrossRef]
25. Huang, M.X.; Mosher, J.C.; Leahy, R.M. A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Phys. Med. Biol.* **1999**, *44*, 423–440. [CrossRef]
26. Özden, M.-Ö.; Teplyuk, A.; Gümüş, Ö.; Meyners, D.; Höft, M.; Gerken, M. Magnetolectric cantilever sensors under inhomogeneous magnetic field excitation. *AIP Adv.* **2020**, *10*, 025132. [CrossRef]
27. Özden, M.-Ö.; Schmalz, J.; Gerken, M. A Combined Magnetolectric Sensor and Human Head Model for Biomagnetic FEM Simulations. *IEEE Sens. J.* **2023**, *23*, 30259–30270. [CrossRef]
28. Huang, Y.; Parra, L.C.; Haufe, S. The New York Head—A precise standardized volume conductor model for EEG source localization and tES targeting. *NeuroImage* **2016**, *140*, 150–162. [CrossRef]
29. Haufe, S.; Huang, Y.; Parra, L.C. A highly detailed FEM volume conductor model based on the ICBM152 average head template for EEG source imaging and TCS targeting. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Milan, Italy, 25–29 August 2015; pp. 5744–5747.
30. Hasanyan, D.; Gao, J.; Wang, Y.; Viswan, R.; Li, M.; Shen, Y.; Li, J.; Viehland, D. Theoretical and experimental investigation of magnetolectric effect for bending-tension coupled modes in magnetostrictive-piezoelectric layered composites. *J. Appl. Phys.* **2012**, *112*, 013908. [CrossRef]
31. Peters, M.J.; de Munck, J.C. The Influence of Model Parameters on the Inverse Solution Based on MEGs and EEGs. *Acta Otolaryngol.* **1991**, *111*, 61–69. [CrossRef]
32. Van Uiter, R.; Weinstein, D.; Johnson, C. Volume Currents in Forward and Inverse Magnetoencephalographic Simulations Using Realistic Head Models. *Ann. Biomed. Eng.* **2003**, *31*, 21–31. [CrossRef]
33. Greve, H.; Woltermann, E.; Quenzer, H.-J.; Wagner, B.; Quandt, E. Giant magnetolectric coefficients in (Fe₉₀Co₁₀)₇₈Si₁₂B₁₀-AlN thin film composites. *Appl. Phys. Lett.* **2010**, *96*, 182501. [CrossRef]
34. Schmalz, J.; Spetzler, E.; McCord, J.; Gerken, M. Investigation of Unwanted Oscillations of Electrically Modulated Magnetolectric Cantilever Sensors. *Sensors* **2023**, *23*, 5012. [CrossRef]
35. Gugat, J.L.; Schmalz, J.; Krantz, M.C.; Gerken, M. Magnetic Flux Concentration Effects in Cantilever Magnetolectric Sensors. *IEEE Trans. Magn.* **2016**, *52*, 7403008. [CrossRef]

36. Blackburn, J.F.; Vopsaroiu, M.; Cain, M.G. Verified finite element simulation of multiferroic structures: Solutions for conducting and insulating systems. *J. Appl. Phys.* **2008**, *104*, 074104. [CrossRef]
37. Guo, M.; Dong, S. A resonance-bending mode magnetoelectric-coupling equivalent circuit. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2009**, *56*, 2578–2586. [CrossRef] [PubMed]
38. Petrov, V.M.; Srinivasan, G.; Bichurin, M.I.; Galkina, T.A. Theory of magnetoelectric effect for bending modes in magnetostrictive-piezoelectric bilayers. *J. Appl. Phys.* **2009**, *105*, 063911. [CrossRef]
39. COMSOL Multiphysics Support. Private Communication, 14 June 2023. Available online: <https://www.comsol.de/support> (accessed on 6 February 2024).
40. Volkmann, J.; Herzog, J.; Kopper, F.; Deuschl, G. Introduction to the programming of deep brain stimulators. *Mov. Disord.* **2002**, *17*, S181–S187. [CrossRef]
41. Piorra, A.; Jahns, R.; Teliban, I.; Gugat, J.L.; Gerken, M.; Knöchel, R.; Quandt, E. Magnetoelectric thin film composites with interdigital electrodes. *Appl. Phys. Lett.* **2013**, *103*, 032902. [CrossRef]
42. Gugat, J.L.; Krantz, M.C.; Gerken, M. Two-Dimensional Versus Three-Dimensional Finite-Element Method Simulations of Cantilever Magnetoelectric Sensors. *IEEE Trans. Magn.* **2013**, *49*, 5287–5293. [CrossRef]
43. Ludwig, A.; Quandt, E. Optimization of the ΔE -effect in thin films and multilayers by magnetic field annealing. In Proceedings of the IEEE International Digest of Technical Papers on Magnetism Conference, Amsterdam, The Netherlands, 28 April–2 May 2002; p. AE2. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

Intensity Thresholds for External Workload Demands in Basketball: Is Individualization Based on Playing Positions Necessary?

Sergio J. Ibáñez ¹, Carlos D. Gómez-Carmona ^{1,2,*}, Pablo López-Sierra ¹ and Sebastián Feu ¹

¹ Research Group in Optimization of Training and Sports Performance (GOERD), Department of Didactics of Music Plastic and Body Expression, Faculty of Sport Science, University of Extremadura, 10003 Cáceres, Extremadura, Spain; sibanez@unex.es (S.J.I.); plopezp@alumnos.unex.es (P.L.-S.); sfeu@unex.es (S.F.)

² BioVetMed & SportSci Research Group, International Excellence Campus “Mare Nostrum”, Faculty of Sport Sciences, University of Murcia, 30720 San Javier, Murcia, Spain

* Correspondence: carlosdavid.gomez@um.es

Abstract: Currently, basketball teams use inertial devices for monitoring external and internal workload demands during training and competitions. However, the intensity thresholds preset by device manufacturers are generic and not adapted for specific sports (e.g., basketball) and players’ positions (e.g., guards, forwards, and centers). Using universal intensity thresholds may lead to failure in accurately capturing the true external load faced by players in different positions. Therefore, the present study aimed to identify external load demands based on playing positions and establish different intensity thresholds based on match demands in order to have specific reference values for teams belonging to the highest competitive level of Spanish basketball. Professional male players ($n = 68$) from the Spanish ACB league were monitored during preseason official games. Three specific positions were used to group the players: guards, forwards, and centers. Speed, accelerations, decelerations, impacts/min, and player load/min were collected via inertial devices. Two-step clustering and k-means clustering categorized load metrics into intensity zones for guards, forwards, and centers. Guards covered more distance at high speeds (12.72–17.50 km/h) than forwards and centers ($p < 0.001$). Centers experienced the most impacts/min ($p < 0.001$). Guards exhibited greater accelerations/decelerations, albeit mostly low magnitude ($p < 0.001$). K-means clustering allowed the setting of five zones revealing additional thresholds. All positions showed differences in threshold values ($p < 0.001$). The findings provide insights into potential disparities in the external load during competition and help establish position-specific intensity thresholds for optimal monitoring in basketball. These data are highly applicable to the design of training tasks at the highest competitive level.

Citation: Ibáñez, S.J.; Gómez-Carmona, C.D.; López-Sierra, P.; Feu, S. Intensity Thresholds for External Workload Demands in Basketball: Is Individualization Based on Playing Positions Necessary? *Sensors* **2024**, *24*, 1146. <https://doi.org/10.3390/s24041146>

Academic Editor: José Machado Da Silva

Received: 23 December 2023

Revised: 29 January 2024

Accepted: 1 February 2024

Published: 9 February 2024

Keywords: team sports; kinematics; impacts; player load; inertial devices

1. Introduction

Basketball is an intermittent high-intensity sport characterized by repeated explosive movements including sprints, jumps, and shuffles, as well as skill-related actions with the ball [1]. The physical and physiological demands during competitions can be substantial, given the fast transitions between offense and defense [2,3]. Players take on specific positions based on their roles during gameplay. The most common position classification includes three (guards, forwards, and centers) or five roles (point guard, shooting guard, small forward, power forward, and center) [1,4]. With increasing specialization, some systems differentiate up to eight or thirteen positions [5,6]. Various methods have been utilized for playing position classification, such as game statistics analysis [7,8] or principal component analysis [6,9].



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

The increasing specialization of players into positions is closely tied to the gameplay dynamics and rules that characterize high-level basketball [4]. International governing body FIBA dictates regulations that shape spacing, timeouts, substitutions, and other structural elements that compel teams to strategize roles [10]. At the elite professional level, tailored tactics maximizing each position's efficacy based on their court movement and actions are paramount. In this sense, players in different basketball positions present different characteristics: (a) guards are shorter, lighter and rely more on speed, agility and shooting skills; (b) forwards are versatile players who operate closer to the basket and are involved in more rebounding plays; and (c) centers are the tallest players, focusing on post play near the basket and defense around the paint [8,11]. These distinct roles result in variability in movement patterns and physiological responses [1,12]. Guards and forwards covered more distance, changes in direction, high-intensity sprints, and dribbles than centers [3,13]. The greater distances covered and accelerations performed by guards contribute to a higher external workload [3,14]. Otherwise, centers received more impacts, collisions, and contact with opponents [15,16]. The differences extend to internal workload indicators, such as heart rate responses, with higher values for guards than forwards and centers suggesting greater cardiovascular intensity [1,17]. Identifying position-specific demands allows the individualization of training programs to optimize performance during games [18].

Player tracking systems and microtechnology are often used to monitor external workload variables across playing positions, including the distance covered, accelerations/decelerations, impacts, and player load [19,20]. To classify the intensity of movements, commonly, five ranges have been utilized (e.g., covered distance: standing, walking, jogging, running, and sprinting) [1,3]. These intensity thresholds are generic for the device manufacturer (e.g., covered distance: standing >6 km/h, walking 6–12 km/h, jogging: 12–18 km/h, running 18–24 km/h, and sprinting >24 km/h) [17,21]. However, these predetermined thresholds may not accurately reflect the real demands of experienced basketball players due to the maximum speed registered in competition being between 18 and 22 km/h [3,22]. In this sense, it is necessary to improve the methodological aspects of using individualized thresholds in team sports [23,24].

To address the abovementioned issue, different methods have been proposed to calculate intensity thresholds in team sports based on maximum values, using Gaussian distributions, k-means clustering, or spectral clustering [25]. In basketball, k-means clustering and two-step clustering have been utilized to individualize intensity thresholds of distance covered, accelerations, decelerations, impacts, and player load in male and female basketball players [26,27]. The advantage of two-step clustering is the robust mathematical method that automatically provides a number of thresholds, while k-means clustering allows a set of five thresholds based on expert judgment and the previous literature [27,28]. However, these studies have established external load intensity ranges considering the whole team.

Since previous research has found differences in the external load performed based on the specific playing position of the players, as they have different roles during the game [1,3,12], an improved understanding of position-specific intensity thresholds can optimize training and game management. Therefore, the present study aims to describe and establish intensity thresholds based on mathematical models (k-means clustering and two-step clustering) of external load variables, considering the specific playing position of basketball players, as well as to compare the results obtained to determine which mathematical model is more useful in classifying external demands.

2. Materials and Methods

2.1. Study Design

The present study aims to establish position-specific intensity thresholds and determine differences in external workload demands during basketball games across playing positions. To test this, an observational quantitative study design was utilized, given the inability to actively manipulate or intervene in official basketball competitions [29]. The

independent variable was playing position, with three levels: guards, forwards, and centers. The dependent variables were external workload metrics monitored via inertial devices during live games, including speed, accelerations, decelerations, impacts per minute, and player load per minute. These specific external workload variables were selected as key indicators of match physical demands based on the previous basketball literature [1,3]. The continuous tracking of these variables allows the quantification of external workload profiles by positions. Furthermore, established mathematical clustering techniques like two-step clustering and k-means clustering can categorize the external workload data into zones to derive intensity thresholds fitted to each playing position [25,26]. The findings can elucidate disparities between playing positions and establish more targeted, individualized intensity thresholds for optimal training and workload management in basketball.

2.2. Participants and Sample

We included 68 professional male basketball players from 6 different teams competing in the top-tier Spanish ACB league during the 2022–2023 preseason. Data were collected during six official preseason games involving ACB teams. It is important to mention that the Spanish basketball first division regulations prohibit inertial devices during official in-season competitions, so data had to be registered in two official preseason tournaments. To be included in the study, players had to meet the following criteria: (a) belong to the official roster of their respective first team in the ACB league; (b) play at least 5 min during the preseason tournaments analyzed; (c) have no musculoskeletal injuries in the 15 days prior to the games that would have limited their maximum performance; and (d) undergo a 10-day adaptation period with the inertial devices used for monitoring before data collection.

The study obtained approval from the University Bioethics Committee (233/2019) and adhered to the ethical guidelines outlined in the Declaration of Helsinki (2013). Participation was voluntary, with written informed consent obtained from all players. Coaching staff and team managers were informed about the study's purpose, procedures, and potential risks and benefits prior to providing consent for their teams' participation. Players were also briefed about the study details and provided consent before data collection. All participant data were anonymized, and confidentiality was maintained.

2.3. Variables and Equipment

To evaluate the external workload performed across playing positions during matches, five kinematic and neuromuscular variables were chosen that are commonly used by basketball teams according to previous research [19]:

- **Velocity:** the speed at which a player moves across the court, measured in kilometers per hour (km/h).
- **Acceleration:** the rate at which a player increases his velocity, measured in meters per second squared (m/s^2).
- **Deceleration:** the rate at which a player decreases his velocity, measured in meters per second squared (m/s^2).
- **Impacts/min:** the number of times a player makes contact with another player or with the ground with a g-force higher than 1 g per minute. It is measured in counts per minute (n/min).
- **Player load/min:** measurement derived from the accelerometer of the total body load in its 3 axes of movement (vertical, anteroposterior, and mediolateral), calculated as the square root of the sum of the accelerations divided by sampling frequency [30]. It is a sum of distance covered, accelerations and decelerations, and impacts and is measured in arbitrary units per minute (a.u./min).

To obtain the external workload variables, all basketball players were equipped with WIMU PRO™ inertial devices (RealTrack Systems, Almería, Spain), which were located on the player's back (between T2 and T4 at intra-scapular level) and fitted using an anatomical harness. To record kinematical variables (speed, accelerations, and decelerations), the

inertial devices are equipped with ultra-wideband technology (UWB) for recording in indoor environments at a frequency of 33 Hz. The UWB system consisted of eight antennae that were placed around the court following the recommendations provided by Pino-Ortega et al. [31] for switching on and calibrating to guarantee the reliability and validity of the measurements. On the other hand, to record neuromuscular variables (impacts and player load), the inertial devices were equipped with different microsensors (four tri-axial accelerometers, three gyroscopes, and one magnetometer) that were set at 100 Hz [30].

2.4. Procedures

Firstly, the clubs were contacted to provide a clear understanding of the research objectives, along with the potential benefits and risks for the athletes involved. Upon obtaining consent from the clubs, further engagement was undertaken with the tournament organizers to seek authorization for the placement of UWB equipment on the playing fields. Once the proposal was approved, informed consent forms were signed by the coaches and players of the participating clubs.

Then, data collection was performed during two preparatory tournaments (six official games) of the ACB League, Spain's top professional basketball league. To gather the necessary data, all players were equipped with WIMU PROTM inertial devices (RealTrack Systems, Almeria, Spain) using an anatomical harness. To minimize disruption to the players' pre-match preparations, the equipment was fitted upon their arrival at the venue. The placement and calibration of the UWB system, as well as microsensors of inertial devices, took place 90 min before the start of each match, ensuring no interference with team preparations.

Following the conclusion of each match, data were extracted from the inertial devices using specialized software, SPROTM (v. 990, RealTrack Systems, Almeria, Spain). Subsequent analysis involved selecting relevant time sequences during which players were actively involved in the matches. The raw data were then exported to an Excel spreadsheet to create the database. Finally, data analysis was performed with statistical software, utilizing two methodologies: (a) unconditional analysis employing two-step clustering; and (b) widely adapted five-group k-means clustering algorithm. These approaches have been commonly employed in the existing literature to date, allowing for meaningful grouping and interpretation of the results [25,27,28].

2.5. Statistical Analysis

Raw data for the speed (km/h), positive and negative change in speed (accelerations/decelerations, m/s²), the impacts (count/min), and player load per minute (a.u./min) generated by all players during the matches were imported into the statistical package. These data generated a total sample of 3,345,703 cases for the speed variable, 22,943,386 cases for the acceleration variable, 26,451,233 cases for the deceleration variable, and 257 cases for the impacts and player load variables.

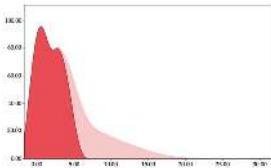
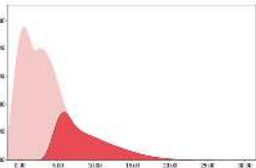
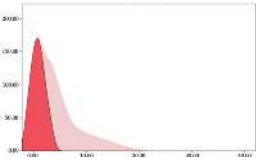
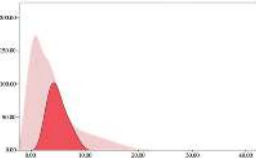
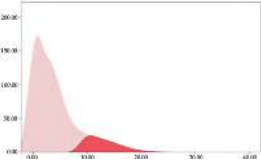
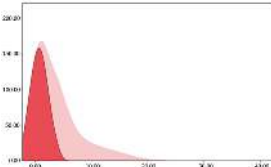
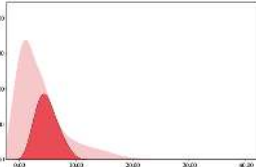
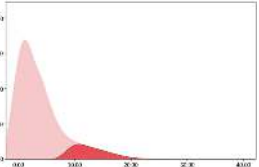
Then, two cluster analyses were performed with each of the five variables selected in this study: (a) two-step clustering to identify the load zones automatically; and (b) five-range k-means clustering. Previous research has employed these mathematical methods to classify intensity thresholds in basketball [26,27]. After identifying the centers of each cluster, a new variable identifying the membership of each case was generated to identify the lower and upper values and thus determine the working area of each variable. An ANOVA was performed to identify the existence of differences between each of the identified clusters. Data analysis was performed using the Statistical Package for the Social Sciences (SPSS, IBM, SPSS Statistics, v.25.0, Armonk, NY, USA). Statistical differences were considered if $p < 0.05$.

3. Results

3.1. Velocity by Playing Positions

Tables 1 and 2 show the two-step clustering and k-means clustering analysis of velocity by playing positions during matches. Two-step clustering indicates two ranges in guards (<4.75 and >4.76 km/h) and three ranges in forwards (<2.74, 2.75 to 8.81, >8.82 km/h) and centers (<2.70, 2.71 to 8.69, >8.70 km/h). The highest distance in all positions was covered at low intensity (guards: 68.8%, forwards: 49.1%, and centers: 50.4%).

Table 1. Two-step clustering of velocity by playing positions.

Role	Speed (km/h)	Low/Walking			Moderate/Jogging	High/Sprinting	
Guard	Cluster Centers	1.75			8.82		
	Ranges	<4.75			>4.76		
	%	68.8%			31.2%		
	Distribution						
Forward	Centers	0.84			4.92	12.87	
	Ranges	<2.74			2.75 to 8.81	>8.82	
	%	49.1%			37.1%	13.9%	
	Distribution						
Center	Centers	0.80			4.86	12.74	
	Ranges	<2.70			2.71 to 8.69	>8.70	
	%	50.4%			36.8%	12.7%	
	Distribution						

Otherwise, k-means clustering presents two zones over than the ranges provided by two-step clustering (running: 12.72 to 17.50 km/h; and sprinting: >17.51 km/h). Regarding k-means clustering, a higher percentage of standing and walking demands were performed by forwards and centers, while a higher percentage of jogging, running, and sprinting demands were performed by guards. ANOVA identified statistical differences in five-range k-means clustering in the total of cases ($F = 10,027,814.10$; $p < 0.001$) and per playing positions (guards: $F = 2,493,424.24$, $p < 0.001$; forwards: $F = 4,578,381.41$, $p < 0.001$; centers: $F = 3,789,342.25$, $p < 0.001$).

Table 2. Five-range k-means clustering of speed by playing positions.

Role	Speed (km/h)	Very Low/ Standing	Low/ Walking	Moderate/ Jogging	High/ Running	Very High/ Sprinting
Guard	Centers	0.70	4.12	8.26	12.61	17.33
	Ranges	<2.46	2.47 to 6.34	6.35 to 10.67	10.68 to 15.22	>15.23
	%	43.77%	35.35%	12.47%	6.39%	2.01%
Forward	Centers	0.81	4.40	9.09	13.81	18.57
	Ranges	<2.67	2.68 to 6.93	6.94 to 11.74	11.75 to 16.50	>16.51
	%	48.33%	32.56%	11.26%	6.02%	1.83%
Centers	Centers	0.87	4.67	9.88	15.17	20.06
	Ranges	<2.86	2.87 to 7.51	7.52 to 12.97	12.98 to 18.42	>18.43
	%	52.27%	31.98%	10.52%	4.54%	0.69%
Total	Centers	0.92	4.73	9.85	14.70	19.35
	Ranges	<2.95	2.96 to 7.58	7.59 to 12.71	12.72 to 17.50	>17.51
	%	51.7%	32.0%	10.60%	4.60%	1.00%

3.2. Changes in Speed by Playing Positions

The two-step clustering and k-means clustering analysis of speed changes (accelerations and decelerations) by playing positions during matches are shown in Tables 3 and 4. Two-step clustering of accelerations indicates five ranges in guards (<0.31, 0.32 to 0.68, 0.69 to 1.23, 1.24 to 2.44, >2.45 m/s²), four ranges in forwards (<0.50, 0.51 to 1.10, 1.11 to 2.45, >2.46 m/s²), and three ranges in centers (<0.55, 0.56 to 1.34, >1.35 m/s²). Only guards and forwards presented a very high acceleration zone, with 1.3 and 1.1% of total accelerations. The two-step clustering of decelerations indicates three ranges in guards (>−0.48, −0.49 to −1.28, <−1.29 m/s²), and four ranges in forwards (>−0.35, −0.36 to −0.79, −0.80 to −1.64, <−1.65 m/s²) and centers (>−0.29, −0.30 to −0.67, −0.68 to −1.26, <−1.27 m/s²).

Table 3. Two-step clustering of changes in speed (accelerations and decelerations) by playing positions.

Accelerations						
Role	Acc (m/s ²)	Very Low	Low	Moderate	High	Very High
Guard	Centers	0.12	0.49	0.88	1.63	3.57
	Ranges	<0.31	0.32 to 0.68	0.69 to 1.23	1.24 to 2.44	>2.45
	%	34.5%	23.1%	34.8%	6.4%	1.3%
Distribution						
						
Forward	Centers	0.20	0.78	1.5	3.59	
	Ranges	<0.50	0.51 to 1.10	1.11 to 2.45	>2.46	
	%	43.5%	47.0%	8.4%	1.1%	
Distribution						
						
Centers	Centers	0.21	0.85	2.03		
	Ranges	<0.55	0.56 to 1.34	>1.35		
	%	47.3%	46.6%	6.1%		
Distribution						
						

Table 3. Cont.

Decelerations							
Role	Dec (m/s ²)						
Guard	Centers						
	Ranges						
	%						
		Very Low	Low	Moderate	High	Very High	
		−0.22 −0.48 to −0.00 58.2%	−0.55 −0.79 to −0.36 37.0%	−1.01 −1.64 to −0.80 18.2%	−1.86 < −1.29 4.0%		
		Distribution					
Forward	Centers						
	Ranges						
	%						
		−0.16 −0.35 to −0.00 42.7%	−0.55 −0.79 to −0.36 37.0%	−1.01 −1.64 to −0.80 18.2%	−2.20 < −1.65 2.1%		
		Distribution					
Centers	Centers						
	Ranges						
	%						
		−0.13 −0.29 to −0.00 37.4%	−0.46 −0.67 to −0.30 36.4%	−0.88 −1.26 to −0.68 22.5%	−1.72 < −1.27 3.7%		
		Distribution					

Table 4. Five-range k-means clustering of speed by playing positions.

Accelerations						
Role	Acc (m/s ²)	Very Low	Low	Moderate	High	Very High
Guard	Centers	0.45	1.36	3.49	6.80	13.84
	Ranges	<0.96	0.94 to 2.64	2.65 to 5.53	5.54 to 12.00	>12.01
	%	85.31%	13.66%	0.92%	0.10%	0.01%
Forward	Centers	0.39	1.09	2.86	5.53	9.19
	Ranges	<0.84	0.85 to 2.18	2.19 to 4.51	4.52 to 7.86	>7.87
	%	72.30%	26.20%	1.31%	0.18%	0.01%
Centers	Centers	0.44	1.24	3.40	6.84	25.33
	Ranges	<0.93	0.94 to 2.56	2.57 to 5.67	5.68 to 14.61	>14.61
	%	80.40%	18.63%	0.90%	0.06%	0.01%
Total	Centers	0.46	1.29	3.34	6.49	14.23
	Ranges	<0.95	0.96 to 2.53	2.54 to 5.31	5.32 to 12.25	>12.26
	%	83.49%	15.49%	0.93%	0.11%	0.06%
Decelerations						
Role	Dec (m/s ²)	Very Low	Low	Moderate	High	Very High
Guard	Centers	−0.27	−0.86	−2.27	−4.39	−9.21
	Ranges	−0.59 to −0.00	−1.55 to −0.60	−3.51 to −1.59	−7.40 to −3.52	−18.99 to −7.41
	%	67.56%	30.01%	2.30%	0.12%	0.01%
Forward	Centers	−0.26	−0.86	−1.96	−4.27	−9.17
	Ranges	−0.59 to −0.00	−1.51 to −0.60	−3.44 to −1.52	−7.37 to −3.45	−19.45 to −7.38
	%	66.58%	30.76%	2.55%	0.10%	0.01%
Centers	Centers	−0.27	−0.87	−2.05	−5.62	−13.84
	Ranges	−0.59 to −0.00	−1.58 to −0.60	−4.32 to −1.59	−11.50 to −4.33	−26.03 to −11.51
	%	68.28%	30.04%	1.64%	0.03%	0.01%
Total	Centers	−0.26	−0.82	−1.78	−3.72	−7.91
	Ranges	−0.56 to 0.0	−1.37 to −0.57	−2.98 to −1.38	−6.27 to −2.99	−14.55 to −6.28
	%	64.34%	32.41%	3.07%	0.17%	0.01%

K-means clustering presents two zones over the ranges provided by two-step clustering in accelerations (high, 5.32–12.25 m/s²; very high, >12.26 m/s²) and decelerations (high, −2.99 to −6.27 m/s²; very high, <−6.28 m/s²), but these zones did not represent more than 1%. The majority of accelerations and decelerations are between the very-low and low categories (>96%). ANOVA identified statistical differences in five-range k-means clustering in the total of cases ($F = 10,460,669.16$; $p < 0.001$) and per playing position (guards: $F = 3,066,120.22$, $p < 0.001$; forwards: $F = 5,612,608.43$, $p < 0.001$; centers: $F = 4,106,991.93$, $p < 0.001$).

3.3. Impacts per Minute by Playing Position

Tables 5 and 6 show the two-step clustering and k-means clustering analysis of impacts per minute by playing positions during matches. Two-step clustering indicates three ranges in guards (<57.65, 88.63 to 139.27, >143.11 n/min) and forwards (<125.52, 128.10 to 158.09, >161.83 n/min), and two ranges in centers (<142.00, >143.45 n/min).

Table 5. Two-step clustering of impacts per minute by playing position.

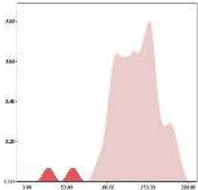
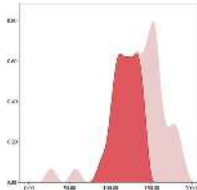
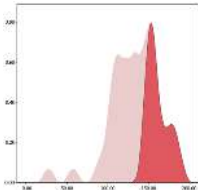
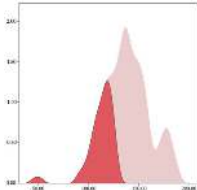
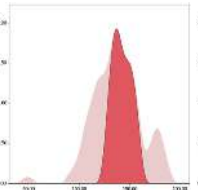
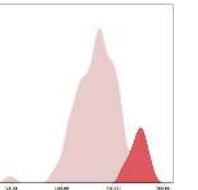
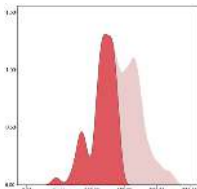
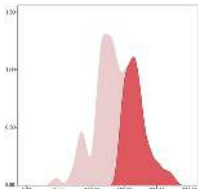
Role	Imp (n/min)	Low	Moderate	High	Very High
Guard	Centers	42.58	119.60		160.63
	Ranges	<57.65	88.63 to 139.27		>143.11
	%	3.8%	54.7%		41.5%
Distribution					
Forwards	Centers		111.78	142.22	174.85
	Ranges		<125.52	128.10 to 158.09	>161.83
	%		31.3%	53.5%	15.2%
Distribution					
Centers	Centers		114.81		168.64
	Ranges		<142.00		>143.45
	%		55.2%		44.8%
Distribution					

Table 6. Five-range k-means clustering of impacts per minute by playing position.

Role	Imp (n/min)	Very Low	Low	Moderate	High	Very High
Guard	Centers	27.50	57.65	88.63	139.27	188.10
	Ranges	<27.50	57.65 to 88.62	88.63 to 127.14	129.04 to 157.44	160.31 to 188.10
	%	1.89%	1.89%	12.47%	41.51%	16.98%
Forward	Centers	49.27	93.75	125.52	163.78	186
	Ranges	<49.27	93.75 to 119.23	120.56 to 139.67	140.63 to 163.78	171.75 to 186.00
	%	1.01%	21.21%	33.33%	32.32%	12.12%
Centers	Centers	45.33	78	123.09	188.1	223.47
	Ranges	<45.33	68.77 to 104.44	108.33 to 142.00	143.45 to 177.67	181.41 to 223.47
	%	0.95%	11.43%	42.86%	34.29%	10.48%
Total	Centers	57.45	107.4	132.01	155.1	182.69
	Ranges	<78.00	83.22 to 119.60	120.53 to 143.60	143.81 to 169.14	169.43 to 223.47
	%	2.72%	23.74%	33.07%	26.46%	14.01%

K-means clustering presents a very low zone that was not detected by two-step clustering. In addition, a higher percentage of the moderate numbers of impacts per minute was performed by forwards and centers while a higher percentage of impacts by centers were performed in high and very-high zones. ANOVA identified statistical differences in

five-range k-means clustering in the total cases ($F = 593.048$; $p < 0.001$) and per playing position (guards: $F = 114.19$, $p < 0.001$; forwards: $F = 298.24$, $p < 0.001$; centers: $F = 246.51$, $p < 0.001$).

3.4. Player Load per Minute by Playing Position

The two-step clustering and k-means clustering analysis of player load per minute by playing position during matches are shown in Tables 7 and 8. The two-step clustering of accelerations indicates two ranges in all playing positions (guards: <1.10 , >1.11 a.u./min; forwards: <1.25 , >1.26 a.u./min; centers: <1.11 , >1.12 a.u./min). K-means clustering presents three zones (very low, low, and high), more than the ranges provided by two-step clustering. The majority of player load per minute accumulated by centers and forwards is with low and moderate magnitude, while guards have high and very-high magnitude. ANOVA identified statistical differences in five-range k-means clustering in the total cases ($F = 600.956$; $p < 0.001$) and per playing position (guards: $F = 94.23$, $p < 0.001$; forwards: $F = 204.46$, $p < 0.001$; centers: $F = 253.43$, $p < 0.001$).

Table 7. Two-step clustering of player load per minute by playing position.

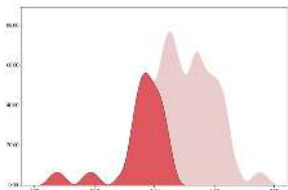
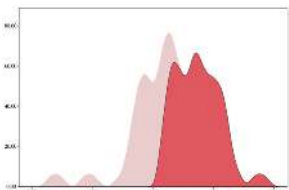
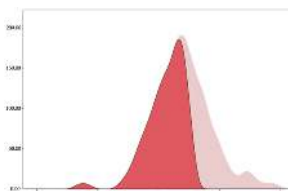
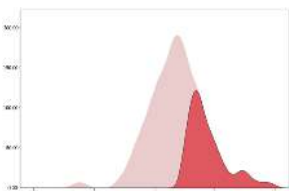
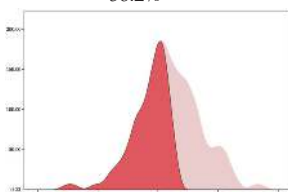
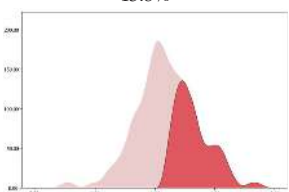
Role	PL (a.u./min)	Moderate	High
Guards	Centers	0.89	1.37
	Ranges	0.19 to 1.10	1.11 to 1.88
	%	35.8%	64.2%
Forwards	Centers	1.06	1.44
	Ranges	0.38 to 1.25	1.26 to 1.93
	%	64.6%	35.4%
Centers	Centers	0.92	1.33
	Ranges	0.27 to 1.11	1.12 to 1.83
	%	56.2%	43.8%
	Distribution		
	Distribution		
	Distribution		

Table 8. Five-range k-means clustering of player load per minute by playing positions.

Role	Imp (n/min)	Very Low	Low	Moderate	High	Very High
Guard	Centers	0.19	0.47	0.93	1.21	1.52
	Ranges	0.19 to 0.47	0.48 to 1.06	1.07 to 1.33	1.34 to 1.64	1.65 to 1.88
	%	1.89%	1.89%	28.30%	35.85%	32.08%
Forward	Centers	0.38	0.96	1.21	1.43	1.77
	Ranges	0.38 to 0.38	0.39 to 1.03	1.04 to 1.25	1.26 to 1.48	1.49 to 1.93
	%	1.01%	31.31%	43.43%	19.19%	5.05%
Centers	Centers	0.56	0.93	1.20	1.49	1.83
	Ranges	0.27 to 0.50	0.51 to 0.91	0.92 to 1.14	1.15 to 1.40	1.41 to 1.83
	%	5.71%	40.95%	39.05%	13.33%	0.95%
Total	Centers	0.36	0.89	1.14	1.40	1.72
	Ranges	0.19 to 0.50	0.51 to 1.01	1.02 to 1.27	1.28 to 1.56	1.57 to 1.93
	%	1.95%	27.24%	39.30%	26.85%	4.67%

4. Discussion

This research aimed to find out the different kinematic and neuromuscular load zones according to specific positions in basketball players. The present study revealed clear differences in external workload demands during basketball games based on playing positions. Guards covered a greater total distance, especially at high intensities like jogging, running, and sprinting. This aligns with previous research showing guards performing more high-intensity running and changes in direction [3]. The need to frequently transition from offense to defense places extensive movement demands on guards. In contrast, centers and forwards exhibited lower speed profiles, concentrating movement close to the basket. Centers, in particular, operated predominantly in the key area, as evidenced by the highest impacts per minute. Their role revolves around screening, boxing out, and defending the paint [1,32]. These disparities in game demands likely stimulate position-specific physiological responses. For example, guards demonstrate higher heart rate intensity, reflecting greater cardiovascular demands from constantly moving up and down the court [17]. Identifying differences between positions is crucial for optimizing training and workload management, as this makes it possible to personalize training according to the specificity required for each variable.

The personalization of training is increasingly vital in team sports like basketball. While the training regimen has a collective objective requiring cooperation, each player must develop position-specific skills in order for the team to withstand diverse competitive challenges. Kozina et al. determined that training individualization is necessary, and systems should facilitate coaches’ work [33]. After a systematic review, Reina et al. [12] recommended personalized training based on the specific demands of each basketball position, supported by research on position-specific loads [11,34]. Therefore, identifying demands by position is key since roles require predominant technical/tactical actions that must be considered by coaches for the personalized preparation of players for games.

To establish protocols or reference values to individualize training, various statistical methods have been employed that allow the grouping of variables [11], verify whether statistical differences exist between specific positions based on load variables [3,12], or divide the load of different variables into correlative ranges to establish the optimal work zones [26]. One of the most-used analyses is utilized to determine reference values and to divide the variables into statistically different groups in their different modalities: hierarchical [35], k-means [26], and two steps [36]. The two-step classification method selects the number of groups based on the statistical differences that exist, without the researcher being able to intervene. This classification method sometimes has a low applicability, since the discrimination of the loads that coaches require does not occur [27]. The use of k-means analysis allows the pre-establishing of groups according to scientific evidence, allowing greater contextualization and applicability to established thresholds and thus improving the training process.

The results of the two-step classification method for the speed and acceleration variables show different groupings based on specific positions. The number of groups identified and the thresholds of each group are not directly applicable or comparable to the thresholds presented in the literature [27], but it is useful to know the descriptive behavior of the different specific positions of basketball players. The identification of two speed groups for the guards, and three groups for the forwards and pivots, shows that the guards have less variation in speed. During games, guards set the pace of the game and the construction of the play, trying to play fast or slow since they usually work under predetermined plays before the game [16]. On the other hand, forwards and pivots have greater oscillations in speed because they are more reactive groups that have to interact with the environment to make decisions to try to surprise the rival [10,37]. Regarding acceleration and deceleration, this classification is inverted. The guards have a greater number of zones because they have a greater variety of different movements by having more space on the field and participate more in the play [14], while pivots have less space and a narrower range of motions, executing more defined movements, focused essentially on enhancing their speed and agility for gestures such as jumping to shoot or rebound [37].

In the comparison between specific positions with the results of the k-means cluster, differences are found between specific positions. The guards have lower values in the lower speed zones and higher values in the faster speed zones. This is also seen with the percentage distribution of speed values; the distribution is lower as the pace increases, but it is always higher in the guards than in the pivots. Guards are players who run and accelerate less than pivots but maintain these high values for longer. These results align with those found by Petway et al. [3], who detected that guards performed high-speed values during displacements and also achieved them in acceleration. The high-intensity load is distributed at a very low percentage in high-performance settings across various specific positions. These observations align with Scanlan et al. [38], who emphasize the intermittent nature of movements made in basketball, with an average of 1750 changes in speed between the different intensity zones throughout a game. In basketball, not only it is important to know the external load's diverse intensity zones but it is crucial to be aware of the speed distribution in each of the zones and the behaviors that lead athletes to different intensities in order to plan training and prepare players to withstand the demands of competition [18,26,39].

Homogeneity characterizes the neuromuscular load classifications for player load and impacts, with comparable groupings within variables across positions. Previous research found that loads were higher in competitions than training, with pivots incurring greater impact loads than forwards and guards [3,15]. This trend manifests in the present study's results, revealing the two-step classification method's impact distribution with higher values in pivots regarding intensity values. Moreover, the k-means cluster analysis evidences this difference, as pivots have elevated values in zones 4 and 5, indicating that they sustain more impacts than guards, with up to 45% of the impacts per minute relative to the total time-weighted impacts. Conversely, guards exhibit higher player load values than pivots both in men's [2,3,27], and women's basketball [12,15,26]. The findings again are related to the k-means cluster analysis, where guards display greater intensity across player load zones, accompanied by higher values in these zones. This leads to guards moving with higher absolute acceleration and more frequently than pivots [11,16]. These results represent the court functions performed by specific positions, influenced by tactics and players' physical capacities [4]. While pivots' loading arises from constant opponent contact and struggling to score near the basket, guards generate substantial physical loading by repeating highly explosive motions over time. Therefore, these aspects should prompt conditioning coaches to develop two targeted strategies for personalized work for training and recovery processes.

Although the present study is the first approach to the classification of kinematic and neuromuscular demands in professional basketball by specific positions for the individualization of training loads and recovery processes, different limitations must be considered.

First, the sample included only male professionals from a single league in preseason games. Evaluating female athletes and youth players would discern whether thresholds differ by sex and skill level. Similarly, analyzing data across full regular seasons with higher game intensity could refine and validate the reported thresholds. Second, the study derived thresholds based solely on external load metrics. A more comprehensive understanding should examine internal workload responses relative to the external classifications. Connecting physiological indicators like heart rate to the intensity zones would inform the true physical demands. Third, tactical and contextual factors like the game pace, match status, or attack–defense phase were not considered but may modulate workload outputs. Finally, the customized ranges provide objective classifications, but some overlap between positions still exists. Future approaches could apply more advanced machine learning techniques to improve the sensitivity and individual customization of thresholds. Overall, position-specific thresholds allow more effective workload monitoring to improve player health, performance, and training design.

5. Conclusions

The present study aimed to describe and establish intensity thresholds for common external workload variables in basketball based on playing positions. The results demonstrate clear differences in external load profiles between guards, forwards, and centers during competitions. Guards covered greater distance in high-intensity zones including jogging, running, and sprinting compared to forwards and centers. Guards also exhibited a higher number of accelerations and decelerations, though the magnitudes were predominantly low. Forwards and centers performed the majority of moderate speed activity. Centers experienced the highest frequency of impacts and collisions with other players.

The findings highlight the disparities in game demands across playing positions arising from distinct roles during gameplay. Guards cover more of the court and execute quicker multidirectional movements, leading to greater acceleration/deceleration loads. The versatile roles of forwards position them closer to the basket, with more rebounding tasks. Centers operate predominantly around the key area with frequent box-outs, screens, and other contact actions.

Establishing intensity zones using mathematical techniques like k-means clustering and two-step clustering allows the individualization of thresholds fitted to each playing position. Though some overlap exists between positions, the tailored thresholds provide stronger practical utility compared to universal thresholds from device manufacturers. Teams can implement these findings to improve the monitoring of the external workload in basketball due to individualizing thresholds and can also enhance their analysis of competition performance and refine training programs tailored to game demands. On a practical level, coaches could apply these findings by:

1. Using the thresholds as guidelines in training drills to expose players to competition intensity by position. During training, tasks will be designed to address specific demands per position, accounting for work thresholds and action quantities within each work range. For example, guards would perform more running and sprinting tasks, while large amounts of screening contact are implemented in center drills.
2. Considering the thresholds when interpreting external loads from monitoring devices. A certain volume of impacts may signal high intensity for a guard but a normal range for centers during games due to different standards. Targeted thresholds facilitate more sensitive alert systems to prompt interventions around excessive loads and guide return-to-play protocols.
3. Individualizing post-game recovery programming by prescribing active rest for positions accruing heavy accelerations/decelerations versus more passive modalities for those incurring extensive impacts. In this sense, tailored fitness programs will be created to enable players to recover from competition demands, compensating for produced imbalances.

Author Contributions: Conceptualization, S.J.I. and C.D.G.-C.; methodology, S.F. and P.L.-S.; software, S.F. and C.D.G.-C.; formal analysis, S.J.I. and P.L.-S.; investigation, C.D.G.-C. and P.L.-S.; data curation, S.F. and C.D.G.-C.; writing—original draft preparation, P.L.-S. and C.D.G.-C.; writing—review and editing, C.D.G.-C., S.F. and S.J.I.; visualization, C.D.G.-C., S.F. and S.J.I.; supervision, S.F. and S.J.I.; funding acquisition, S.F. and S.J.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Spanish National Agency of Investigation through the project “Scientific and Technological Support to analyze the Training Workload of Basketball teams according to sex, level of the players, and season period” (PID2019-106614GB-I00) MCIN/AEI/10.13039/501100011033.

Institutional Review Board Statement: This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Universidad de Extremadura (no. 233/2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the Organic Law 3/2018, of 5 December, on the Protection of Personal Data and Guarantee of Digital Rights of the Government of Spain, which requires that this information must be in custody.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Stojanović, E.; Stojiljković, N.; Scanlan, A.T.; Dalbo, V.J.; Berkelmans, D.M.; Milanović, Z. The Activity Demands and Physiological Responses Encountered During Basketball Match-Play: A Systematic Review. *Sports Med.* **2018**, *48*, 111–135. [CrossRef]
2. García, F.; Vázquez-Guerrero, J.; Castellano, J.; Casals, M.; Schelling, X. Differences in Physical Demands between Game Quarters and Playing Positions on Professional Basketball Players during Official Competition. *J. Sports Sci. Med.* **2020**, *19*, 256–263.
3. Petway, A.J.; Freitas, T.T.; Calleja-González, J.; Leal, D.M.; Alcaraz, P.E. Training load and match-play demands in basketball based on competition level: A systematic review. *PLoS ONE* **2020**, *15*, e0229212. [CrossRef]
4. Bianchi, F.; Facchinetti, T.; Zuccolotto, P. Role Revolution: Towards a New Meaning of Positions in Basketball. *Electron. J. Appl. Stat. Anal.* **2017**, *10*, 712–734.
5. Alagappan, M. *From 5 to 13 Redefining the Positions in Basketball*; Ayasdi Inc.: Boston, MA, USA, 2012.
6. Cheng, A. Using Machine Learning to Find the 8 Types of Players in the NBA. Available online: <https://medium.com/fastbreak-data/classifying-the-modern-nba-player-with-machine-learning-539da03bb824> (accessed on 19 April 2021).
7. Fernández-Cortés, J.A.; Mandly, M.G.; García-Rubio, J.; Ibáñez, S.J. Contribution of Professional Basketball Players According to the Specific Position and the Competition Phase. *E-Balónmano Com* **2021**, *17*, 223–232.
8. Sampaio, J.; Janeira, M.; Ibáñez, S.; Lorenzo, A. Discriminant analysis of game-related statistics between basketball guards, forwards and centres in three professional leagues. *Eur. J. Sport Sci.* **2006**, *6*, 173–178. [CrossRef]
9. Gómez-Carmona, C.D.; Mancha-Triguero, D.; Pino-Ortega, J.; Ibáñez, S.J. Exploring Physical Fitness Profile of Male and Female Semiprofessional Basketball Players through Principal Component Analysis—A Case Study. *J. Funct. Morphol. Kinesiol.* **2021**, *6*, 67. [CrossRef]
10. Zhang, S.; Lorenzo, A.; Gómez, M.-A.; Liu, H.; Gonçalves, B.; Sampaio, J. Players’ technical and physical performance profiles and game-to-game variation in NBA. *Int. J. Perform. Anal. Sport* **2017**, *17*, 466–483. [CrossRef]
11. Salazar, H.; Castellano, J.; Svilar, L. Differences in External Load Variables between Playing Positions in Elite Basketball Match-Play. *J. Hum. Kinet.* **2020**, *75*, 257–266. [CrossRef]
12. Reina, M.; García-Rubio, J.; Ibáñez, S.J. Training and Competition Load in Female Basketball: A Systematic Review. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2639. [CrossRef]
13. Figueira, B.; Mateus, N.; Esteves, P.; Dadeliené, R.; Paulauskas, R. Physiological Responses and Technical-Tactical Performance of Youth Basketball Players: A Brief Comparison between 3 × 3 and 5 × 5 Basketball. *J. Sports Sci. Med.* **2022**, *21*, 332–340. [CrossRef]
14. Castagna, C.; Chaouachi, A.; Rampinini, E.; Chamari, K.; Impellizzeri, F. Aerobic and Explosive Power Performance of Elite Italian Regional-Level Basketball Players. *J. Strength Cond. Res.* **2009**, *23*, 1982–1987. [CrossRef]
15. Staunton, C.; Wundersitz, D.; Gordon, B.; Kingsley, M. Accelerometry-Derived Relative Exercise Intensities in Elite Women’s Basketball. *Int. J. Sports Med.* **2018**, *39*, 822–827. [CrossRef]
16. Delextrat, A.; Badiella, A.; Saavedra, V.; Matthew, D.; Schelling, X.; Torres-Ronda, L. Match activity demands of elite Spanish female basketball players by playing position. *Int. J. Perform. Anal. Sport* **2015**, *15*, 687–703. [CrossRef]

17. Puente, C.; Abián-Vicén, J.; Areces, F.; López, R.; Del Coso, J. Physical and Physiological Demands of Experienced Male Basketball Players During a Competitive Game. *J. Strength Cond. Res.* **2017**, *31*, 956–962. [CrossRef]
18. Edwards, T.; Spiteri, T.; Piggott, B.; Bonhotal, J.; Haff, G.G.; Joyce, C. Monitoring and Managing Fatigue in Basketball. *Sports* **2018**, *6*, 19. [CrossRef]
19. Russell, J.L.; McLean, B.D.; Impellizzeri, F.M.; Strack, D.S.; Coutts, A.J. Measuring Physical Demands in Basketball: An Explorative Systematic Review of Practices. *Sports Med.* **2021**, *51*, 81–112. [CrossRef]
20. Fox, J.L.; Scanlan, A.T.; Stanton, R. A Review of Player Monitoring Approaches in Basketball: Current Trends and Future Directions. *J. Strength Cond. Res.* **2017**, *31*, 2021–2029. [CrossRef]
21. Vázquez-Guerrero, J.; Jones, B.; Fernández-Valdés, B.; Moras, G.; Reche, X.; Sampaio, J. Physical demands of elite basketball during an official U18 international tournament. *J. Sports Sci.* **2019**, *37*, 2530–2537. [CrossRef]
22. Caparrós, T.; Casals, M.; Solana, A.; Peña, J. Low External Workloads Are Related to Higher Injury Risk in Professional Male Basketball Games. *J. Sports Sci. Med.* **2018**, *17*, 289–297.
23. Clemente, F.; Ramirez-Campillo, R.; Beato, M.; Moran, J.; Kawczynski, A.; Makar, P.; Sarmiento, H.; Afonso, J. Arbitrary absolute vs. individualized running speed thresholds in team sports: A scoping review with evidence gap map. *Biol. Sport* **2023**, *40*, 919–943. [CrossRef]
24. Polglaze, T.; Hogan, C.; Dawson, B.; Buttfield, A.; Osgnach, C.; Lester, L.; Peeling, P. Classification of Intensity in Team Sport Activity. *Med. Sci. Sports Exerc.* **2018**, *50*, 1487–1494. [CrossRef]
25. Sweeting, A.J.; Cormack, S.J.; Morgan, S.; Aughey, R.J. When Is a Sprint a Sprint? A Review of the Analysis of Team-Sport Athlete Activity Profile. *Front. Physiol.* **2017**, *8*, 432. [CrossRef]
26. Ibáñez, S.J.; Gómez-Carmona, C.D.; Mancha-Triguero, D. Individualization of Intensity Thresholds on External Workload Demands in Women's Basketball by K-Means Clustering: Differences Based on the Competitive Level. *Sensors* **2022**, *22*, 324. [CrossRef]
27. Ibáñez, S.J.; López-Sierra, P.; Lorenzo, A.; Feu, S. Kinematic and Neuromuscular Ranges of External Loading in Professional Basketball Players during Competition. *Appl. Sci.* **2023**, *13*, 11936. [CrossRef]
28. Leroy, A.; Marc, A.; Dupas, O.; Rey, J.L.; Gey, S. Functional Data Analysis in Sport Science: Example of Swimmers' Progression Curves Clustering. *Appl. Sci.* **2018**, *8*, 1766. [CrossRef]
29. Ato, M.; López-García, J.J.; Benavente, A. Un sistema de clasificación de los diseños de investigación en psicología. *Anal. Psicol.* **2013**, *29*, 1038–1059. [CrossRef]
30. Gómez-Carmona, C.D.; Bastida-Castillo, A.; García-Rubio, J.; Ibáñez, S.J.; Pino-Ortega, J. Static and Dynamic Reliability of WIMU PROTM Accelerometers according to Anatomical Placement. *Proc. Inst. Mech. Eng. Part P J. Sports Eng. Technol.* **2019**, *233*, 238–248.
31. Pino-Ortega, J.; Bastida-Castillo, A.; Gómez-Carmona, C.D.; Rico-González, M. Validity and reliability of an eight antennae ultra-wideband local positioning system to measure performance in an indoor environment. *Sports Biomech.* **2024**, *23*, 145–155. [CrossRef]
32. Fox, J.L.; Stanton, R.; Scanlan, A.T. A Comparison of Training and Competition Demands in Semiprofessional Male Basketball Players. *Res. Q. Exerc. Sport* **2018**, *89*, 103–111. [CrossRef]
33. Kozina, Z.L.; Prusik, K.; Prusik, K. The Concept of Individual Approach in Sport. *Pedagog. Psychol. Med.-Biol. Probl. Phys. Train. Sports* **2015**, *19*, 28–38. [CrossRef]
34. Pion, J.; Segers, V.; Stautemas, J.; Boone, J.; Lenoir, M.; Bourgois, J.G. Position-specific performance profiles, using predictive classification models in senior basketball. *Int. J. Sports Sci. Coach.* **2018**, *13*, 1072–1080. [CrossRef]
35. Duman, E.A.; Sennaroğlu, B.; Tuzkaya, G. A cluster analysis of basketball players for each of the five traditionally defined positions. *Proc. Inst. Mech. Eng. Part P J. Sports Eng. Technol.* **2021**, ahead of print. [CrossRef]
36. Mateus, N.; Esteves, P.; Gonçalves, B.; Torres, I.; Gomez, M.A.; Arede, J.; Leite, N. Clustering performance in the European Basketball according to players' characteristics and contextual variables. *Int. J. Sports Sci. Coach.* **2020**, *15*, 405–411. [CrossRef]
37. Cui, Y.; Liu, F.; Bao, D.; Liu, H.; Zhang, S.; Gómez, M. Key Anthropometric and Physical Determinants for Different Playing Positions during National Basketball Association Draft Combine Test. *Front. Psychol.* **2019**, *10*, 2359. [CrossRef]
38. Scanlan, A.T.; Dascombe, B.J.; Reaburn, P.; Dalbo, V.J. The physiological and activity demands experienced by Australian female basketball players during competition. *J. Sci. Med. Sport* **2012**, *15*, 341–347. [CrossRef]
39. Fox, J.L.; Stanton, R.; Sargent, C.; Wintour, S.-A.; Scanlan, A.T. The Association between Training Load and Performance in Team Sports: A Systematic Review. *Sports Med.* **2018**, *48*, 2743–2774. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

BackMov: Individualized Motion Capture-Based Test to Assess Low Back Pain Mobility Recovery after Treatment

Fernando Villalba-Meneses ^{1,2,3,*}, Cesar Guevara ⁴, Paolo A. Velásquez-López ², Isaac Arias-Serrano ², Stephanie A. Guerrero-Ligüa ², Camila M. Valencia-Cevallos ², Diego Almeida-Galárraga ², Carolina Cadena-Morejón ⁵, Javier Marín ^{1,3} and José J. Marín ^{1,3}

¹ IDERGO (Research and Development in Ergonomics), I3A (Instituto de Investigación en Ingeniería de Aragón), University of Zaragoza, C/Mariano Esquillor s/n, 50018 Zaragoza, Spain; jmarinbone@unizar.es (J.M.); jjmarin@unizar.es (J.J.M.)

² School of Biological Sciences and Engineering, Yachay Tech University, Hacienda San José s/n, San Miguel de Urcoquí 100119, Ecuador; paolo.velasquez@yachaytech.edu.ec (P.A.V.-L.); aarias@yachaytech.edu.ec (I.A.-S.); stephanie.guerrero@yachaytech.edu.ec (S.A.G.-L.); camila.valencia@yachaytech.edu.ec (C.M.V.-C.); dalmeida@yachaytech.edu.ec (D.A.-G.)

³ Department of Design and Manufacturing Engineering, University of Zaragoza, C/Mariano Esquillor s/n, 50018 Zaragoza, Spain

⁴ Centro de Investigación en Mecatrónica y Sistemas Interactivos—MIST, Universidad Tecnológica Indoamérica, Quito 170103, Ecuador; cesarguevara@uti.edu.ec

⁵ School of Mathematical and Computational Sciences, Yachay Tech University, Hacienda San José s/n, San Miguel de Urcoquí 100119, Ecuador; ccadena@yachaytech.edu.ec

* Correspondence: 762301@unizar.es

Citation: Villalba-Meneses, F.; Guevara, C.; Velásquez-López, P.A.; Arias-Serrano, I.; Guerrero-Ligüa, S.A.; Valencia-Cevallos, C.M.; Almeida-Galárraga, D.; Cadena-Morejón, C.; Marín, J.; Marín, J.J. BackMov: Individualized Motion Capture-Based Test to Assess Low Back Pain Mobility Recovery after Treatment. *Sensors* **2024**, *24*, 913. <https://doi.org/10.3390/s24030913>

Academic Editor: José Machado Da Silva

Received: 21 November 2023

Revised: 24 January 2024

Accepted: 27 January 2024

Published: 31 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Low back pain (LBP) is a common issue that negatively affects a person's quality of life and imposes substantial healthcare expenses. In this study, we introduce the (Back-pain Movement) BackMov test, using inertial motion capture (MoCap) to assess lumbar movement changes in LBP patients. The test includes flexion–extension, rotation, and lateralization movements focused on the lumbar spine. To validate its reproducibility, we conducted a test-retest involving 37 healthy volunteers, yielding results to build a minimal detectable change (MDC) graph map that would allow us to see if changes in certain variables of LBP patients are significant in relation to their recovery. Subsequently, we evaluated its applicability by having 30 LBP patients perform the movement's test before and after treatment (15 received deep oscillation therapy; 15 underwent conventional therapy) and compared the outcomes with a specialist's evaluations. The test-retest results demonstrated high reproducibility, especially in variables such as range of motion, flexion and extension ranges, as well as velocities of lumbar movements, which stand as the more important variables that are correlated with LBP disability, thus changes in them may be important for patient recovery. Among the 30 patients, the specialist's evaluations were confirmed using a low-back-specific Short Form (SF)-36 Physical Functioning scale, and agreement was observed, in which all patients improved their well-being after both treatments. The results from the specialist analysis coincided with changes exceeding MDC values in the expected variables. In conclusion, the BackMov test offers sensitive variables for tracking mobility recovery from LBP, enabling objective assessments of improvement. This test has the potential to enhance decision-making and personalized patient monitoring in LBP management.

Keywords: low back pain; minimal detectable change; deep oscillation therapy; inertial measurement unit; range of motion

1. Introduction

Low back pain (LBP) is the main cause of absenteeism and disability in industrialized societies. Prolonged duration of LBP can significantly affect quality of life due to

biomechanical alterations, such as movement alteration, muscular compensation, postural change, joint degeneration or core dysfunction, reducing the ability to perform daily activities [1]. LBP is an extremely common symptom worldwide and occurs in all age groups, from children to the elderly population [2]. Approximately 10–20% of patients develop chronic LBP, defined as pain and disability persisting for more than 12 weeks [3]. Therefore, there is a clear necessity of reducing the disability caused by this problem; hence, efficient methodologies are needed to improve the function of people suffering these problems.

Traditional physiotherapy treatments have proven to be effective in enhancing function and reducing disability in patients with chronic LBP. Therefore, graded activity or exercise programs that focus on improving function and preventing disability are recommended as primary treatment strategies [4]. Consequently, there is a heightened focus on physical treatments that allow self-management, with less emphasis on pharmacological and surgical interventions [4].

Analyzing how a patient responds to a therapeutic approach by administering pre-treatment and post-treatment tests can help determine whether this should be modified, substituted or discontinued [5,6]. In this regard, some researchers propose the use of tools and indicators to assess the minimal detectable change (MDC) in order to evaluate and ensure actual response to treatments in clinical practice [5,6]. The use of these kinds of tools could reduce or facilitate the monitoring of the effect of treatments without the need to constantly test patients with questionnaire-based measures [7].

The MDC is the minimum quantity of change that can be detected in order to be considered “real” instead of that resulting from potential measurement error [5]. This index represents the variability of the measures of each variable. If a change of one variable is detected and it is lower than its MDC value, it would not be considered statistically significant, since it is lower than the variability of the test [8]. The MDC is important for clinical decision-making because it can provide a threshold value for therapists, clinical therapists, and clinical researchers to determine whether the results represent a real change or reflect intrinsic variability of measurement [9].

The integration of MDC into therapeutic strategies such as DOT or traditional physiotherapy treatment facilitates the evaluation of patient progress. Consequently, physicians are able to monitor patient improvement to ensure that the administered treatments or therapies are yielding satisfactory outcomes [6]. Currently, assessing improvements is normally conducted through qualitative techniques after treatment, either by observation of body movements or through interviews with the patient, as is used for assessing the state of patients before treatment [10]. As exemplified by Marín et al. [10,11], the integration of inertial measurement units (IMUs) for movement analysis via full-body motion capture (MoCap) into rehabilitation assessment based on medical examination is feasible. This is mainly because IMU-based technologies are amongst the most prevalent methodologies employed for MoCap.

MoCap provides information about spatio-temporal and kinematic variables [10–12]. These variables are particularly useful for monitoring the progress of patients with musculoskeletal disorders and can offer many opportunities in the field of rehabilitation to aid decision-making using measurements before and after treatment, intervention, or therapy [5,10,11,13,14]. IMUs are electronic devices that capture motion through signal processing of output data from various embedded sensors (accelerometers, gyroscopes, and magnetometers) [15–17]. In addition, IMUs have become particularly important because they do not require external cameras and can be embedded in wearable technology [6,18].

Despite the ostensibly broad applicability of clinical movement analysis [19,20], its comprehensive integration into routine clinical practice encounters certain hurdles, most notably the complexity of data analysis stemming from measurement processes. This necessitates the development of strategies for the automatic and consistent processing of the spatiotemporal and kinematic variables information that is generated [10,12]. A particular requirement is the standardization of methods to facilitate a comparative analysis of

variables generated from two distinct measurement sessions, for instance, those conducted pre-and post-treatment, or at different junctures during the rehabilitation process [11]. Furthermore, the incorporation of MoCap-based tests needs to address the inherent heterogeneity among patients in routine clinical rehabilitation practice. This entails surmounting the challenge of carrying out individual patient assessments, tracking intra-patient session data over time, individually managing recovery trajectories, and comparing the efficacy of treatments across diverse patient profiles [11].

This study introduces the BackMov test, leveraging inertial MoCap sensor technology for comprehensive lumbar movement analysis. Crafted to objectively track patient progress during rehabilitation, the BackMov test involves two pivotal measurement sessions: pre- and post-treatment. The primary goal is to precisely quantify advancements or setbacks in patients with low back pain (LBP) post-rehabilitation, establishing a correlation between positive changes in kinematic variables of lumbar movements and recovery after therapy. To validate the test's reliability, a test-retest involving three lumbar movements was conducted on 37 healthy volunteers, utilizing the minimal detectable change (MDC) as a statistical benchmark.

To illustrate its clinical relevance, the BackMov test was administered to 30 LBP-diagnosed patients, capturing the same three-segment movement before and after therapy. The resulting data were then compared to evaluations by a specialist physician. Considered as a novel tool, the BackMov test aims to provide a quantifiable assessment of movement recovery in LBP patients post-therapy. By doing so, it can support specialists in determining the efficacy of treatments in restoring mobility, thereby reducing the disability caused by LBP. The BackMov test enhances decision-making and introduces a systematic approach to the management and treatment of LBP, offering tailored patient monitoring to ensure more effective rehabilitation strategies.

2. Materials and Methods

2.1. Participants and Protocol

In this study, 37 healthy volunteers, both male and female, were recruited to determine the reproducibility of the test using the minimum detectable change (MDC) criterion. A call for volunteers between 18 and 65 years of age was made through social media. The inclusion criteria for participation in this study included having a diagnosis of low back pain (LBP) made by a doctor or specialist and the experiencing of a reduction in mobility or increased difficulty in movement. The exclusion criteria included individuals presenting any disease or disability that may have hindered their movement, as well as those who were high-level athletes. Additionally, individuals engaging in dangerous activities during the study period were excluded, as well as individuals undergoing any specific drug treatment to alleviate pain. To be eligible for participation, volunteers must not have received physiotherapeutic treatment during the previous 6 months. All participants were required to sign an informed consent form and to attend treatment sessions in person. The final consent was signed on 12 March 2023. Following recruitment, patients underwent the MoCap movement test to evaluate flexion–extension, lateralization, and rotation movements. Among these patients, fifteen (7 men and 8 women) received traditional treatment, including massage, exercise, heat therapy, and cold therapy. The remaining fifteen (7 men and 8 women) were treated with the deep oscillation method to analyze their response to an additional treatment different from the conventional one. Moreover, it was important to know how the clinician evaluated the patient. First, a clinical review was performed, including collecting information about the type of work the patient performed, if the patient practiced any sport, if the patient had experienced any impactful events in the last few weeks or months, and if any type of chronic disease was present. Secondly, the clinician observed the entire posterior trunk, detecting anomalies in the whole spine, accompanied by a palpatory examination of the lumbar spine. Third, the treatments were applied. For the conventional treatment, the clinician applied a combination of different treatments that have yielded positive results. First, a heat treatment was applied for 10 min [21], then a magneto machine was applied for another 20 min [22]. Finally, a massage series was performed [23], which

was coupled with three William’s exercises (the first three exercise describe in William’s program) [24] consisting of two series of 10 repetitions of each exercise for five minutes. These treatments were used in order to consider their effects in combination, producing a multi-factorial therapy program that could guarantee yielding results in both the short and long run of the experiment [25].

A deep oscillation treatment was also performed as a novel treatment based on research findings on electrostatic fields used in chronic pain treatment [26]. The clinicians asked the patient to lie down on the table and to remove their T-shirt in order to provide a visible lumbar zone for performance of the treatment. Then, heat was applied to the lumbar region for 10 min. Next, talcum powder was applied and deep oscillation treatment was immediately performed for 15 min. Finally, the patient was massaged on the treated area for 5 min. These massages were also coupled with the William’s exercises as described for the previous treatment. Patients had to perform three sessions per week, i.e., in total, to complete 12 treatment sessions.

To evaluate the patient’s physical improvements, the clinicians used the Short Form-36 Physical Functioning scale specific to low back pain pathologies [27].

Patient-related information, such as age, BMI, diagnosis, pain level were not included in this document because these variables were not considered when conducting this study. However, information regarding the patients and the procedures performed on each patient is summarized in Table 1. Further details on the status of the patients prior to the treatments can be found in the Supplementary Materials where details of BMI, pain level, age, and a brief description of the situation of the patients is provided. A flowchart of the methodology for ROM classification is presented in Figure 1.

Table 1. Participants.

Group	Male	Female	Total
DOT	7	8	15
TT	7	8	15
Total	14	16	30

Deep oscillation treatment (DOT); Traditional treatment (TT).

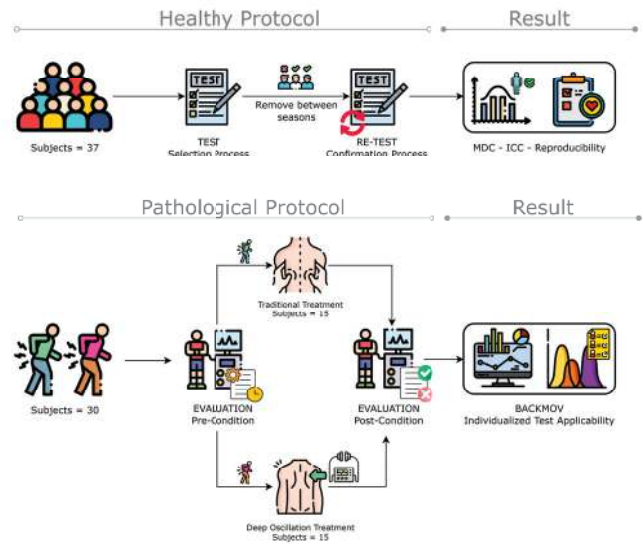


Figure 1. Methodology flowchart of the data acquisition and algorithm implementation for ROM classification (figure designed from Freepik illustrations).

2.2. Ethical Statement

The study was conducted in accordance with the Declaration of Helsinki and the protocol was approved by the Ethics Committee for Research on Human Beings of the Pontificia Universidad Católica del Ecuador (PUCE), Ecuador (N° EO-146-2022). Written informed consent was obtained from each participant.

2.3. Technology and Instrumentation

We employed the Move Human (MH) Sensors MoCap system which was developed by IDERGO (Investigación y Desarrollo en Ergonomía Research Group, V19-07.011, University of Zaragoza, Zaragoza, Spain) using NGIMU (x-io technologies, Bristol, UK). This system relies on inertial measurement units (IMUs) securely positioned on three key areas of the body: sensor 1 is collocated in the superior head region (forehead), sensor 2 is collocated in the cervical region (specifically at C7), and sensor 3 is collocated in the sacral region (at the iliac crest level) (see Figure 2) for comprehensive analysis. This system provides accurate information on the rotations and displacements of each body segment at a frequency of 60 Hz.

The MH-Sensors system enables the visualization and real-time monitoring of movement on a digital representation of a human body, or avatar, which is adjusted to the subject's body dimensions. In addition, the inertial sensor system houses three types of sensors: accelerometers, gyroscopes, and magnetometers. These signals are combined, enabling the clinician to obtain three rotation angles in the axes of space (as well as the angular velocities and accelerations).

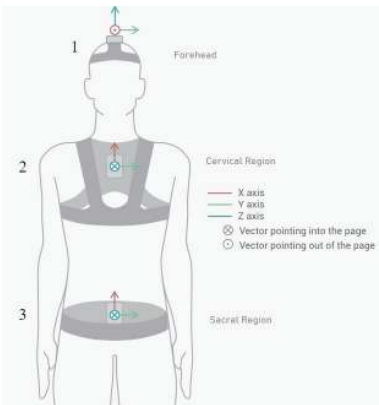


Figure 2. Inertial sensor placements and orientations on patient (sensor 1 in the forehead region, sensor 2 in the cervical region, and sensor 3 in the sacral region).

2.4. Variable

The dorsal–lumbar movements analyzed to obtain data were flexion–extension (Flex), right–left rotation (Rot), and right–left lateralization (Lat) (See Figure 3). As a result, we acquired information relating to spatiotemporal and kinematic variables. Each variable was calculated for each dorsal–lumbar movement (Flex, Rot, and Lat). Information for the variables considered in the study is summarized in Table 2.

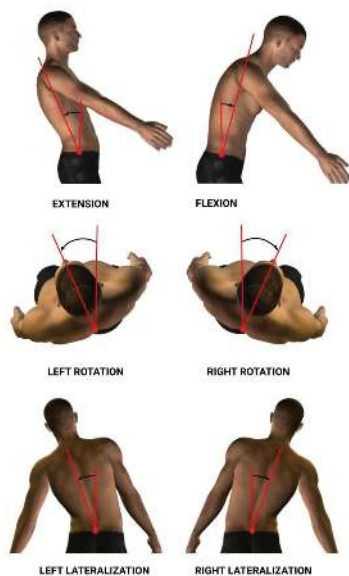


Figure 3. Flexion–extension, right–left rotation, and right–left lateralization lumbar movements.

Table 2. Summary of spatiotemporal and kinematic variables for each dorsal–lumbar movement.

Variable	Description
Max.Range	Overall range of motion in the dorsal–lumbar movement (Maximum extension range plus maximum flexion range).
Max (Min)	Highest value achieved within the range of motion in the dorsal–lumbar movement (Max is for flexion and Min for extension).
Max.Mean (Min.Mean)	Average range of motion in the dorsal–lumbar movement (Max is for flexion and Min for extension).
Coefficient of variation (%) (CV)	Indicator of variability in motion data within the dorsal–lumbar movement, expressed as a percentage.
Average speed (°/s) (Speed)	Average speed of motion in degrees per second in the dorsal–lumbar movement.
Average SpeedUp (°/s ²) (SpeedUp)	Average acceleration of motion in degrees per second squared in the dorsal–lumbar movement.
Mobility area (AreaMean) *	Displays the total area under the motion curve in the dorsal–lumbar movement, which can provide information about the number of movements performed within a specific timeframe. A larger area indicates greater patient mobility when performing movements.
Variability area (AreaStd) *	Indicates the total area under the variability curve in the dorsal–lumbar region, providing insight into the overall variability of movement within a specific timeframe. It depends on the similarity or resemblance of the different cycles of the resulting graph. A larger area reflects greater variability in the cycles of dorsal–lumbar movement by the patient, indicating less uniformity or regularity in those movements.

Table 2. Cont.

Variable	Description
Relative variability (coef%) *	Represents a percentage-based measure of relative variability in the dorsal–lumbar region, offering a standardized assessment of variation in motion data.
Harmony *	Evaluates the overall smoothness and coordination of movements in the dorsal–lumbar region. The variable used to calculate the harmony of motion was the Pearson correlation coefficient between the angle and the angular acceleration. A value of “−1” would indicate a maximum correlation, that is, a straight line in the relationship between these two variables.
Offset (°) *	Measures the time delay or phase difference between different movements or segments in the dorsal–lumbar region, expressed in degrees. In this case, these waves are the signals of angle and acceleration of dorsal–lumbar movement.

Variables with * are justified based on the work [28].

2.5. Magnitude-Based Decision (MBD) to Monitor Individuals with LBP

To measure the effects of the treatment on patients, it is necessary to use the statistical approach, magnitude-based decision MBD, in order to compare the effect size with a predetermined threshold. When discussing individual monitoring, it is important to consider that each patient performed six complete cycles during each session (one session prior to the treatments and one session after the treatments; these two sessions were conducted for both types of treatment, traditional treatment and deep oscillation treatment), attempting to reach their maximum range in each exercise in order to obtain the necessary information to evaluate the respective variables needed for assessment. Thus, for individual monitoring of each variable, it was possible to compare two sets of measurements: one from the pre-treatment session (n_1 samples, X_1 mean, and SD_1 standard deviation) and another from the post-treatment session (n_2 , X_2 mean, and SD_2 standard deviation). For this purpose, we used the same discerning method as described in Marin et al. [11], the magnitude-based decision method. This method provides the probability that a change (which is defined by the confidence interval of the difference, CIdiff) exceeds a specific threshold ($-\delta$, $+\delta$) [19], in this case, the MDC [29] (1).

MDC values at 95% confidence were calculated using the following expression:

$$MDC_{95} = 1.96\sqrt{2SEM}; \quad SEM = SD_{pool}\sqrt{1 - ICC} \tag{1}$$

In this equation, SD stands for the weighted mean of the standard deviation between test and retest, ICC is the intraclass correlation coefficient, and SEM is the measurement standard error.

In this way, a change is only considered substantial if it overcomes the test’s inherent faults, so that it can be stated that the change observed is real and is not the product of a measurement error. This can be especially helpful to physicians when used in conjunction with clinical data.

By calculating the CIdiff, we can create a graph for each variable that depicts the threshold ($-\delta$, $+\delta$) and the t-distribution of the change between the pre- and post-series, like the one presented in Figure 4. Analysis of where changes occur with respect to the threshold is made clearer using this depiction. To conduct this analysis statistically, we determined specific domains of probability denoted negative change (N), trivial change (T), and positive change (P). These domains are defined by the proportion of the t area that falls inside; as such, we have a “negative” region ($-infy$, $-\delta$), a “trivial” region ($-\delta$, $+\delta$), and a “positive” region ($+\delta$, $+infy$).

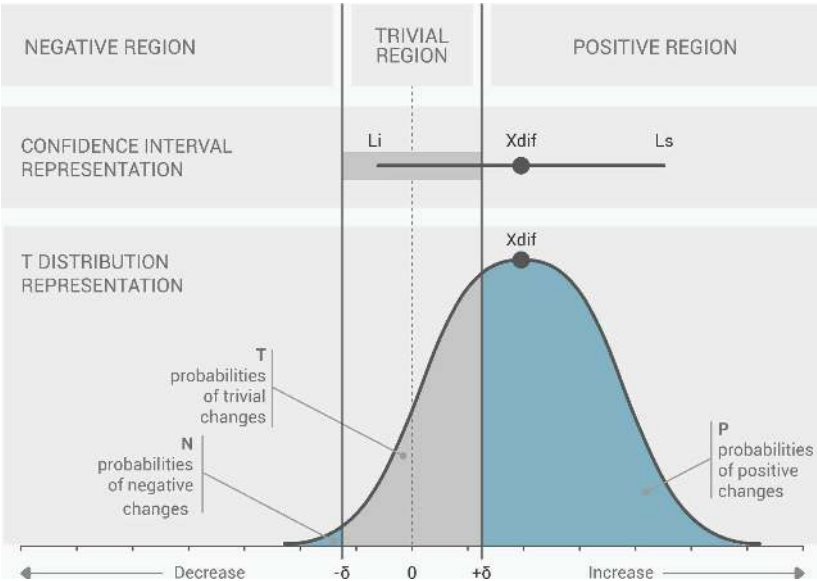


Figure 4. Magnitude-based choices (MBD) threshold: representation of the difference between the pre- and post-series of a single variable. Change within a topic (difference in means between pre- and post-series) is referred to as X_{dif} . Li : the change’s lower limit; Ls : the change’s upper limit, according to J. Marin, J. J. Marin, T. Blanco, J. de la Torre, I. Salcedo, E. Martitegui [11].

In this case, a change is regarded as null or insignificant if it does not surpass the threshold in any direction where the percentages of P and N are both less than 5% ($N < 5\%$ and $P < 5\%$). A change is classified as uncertain if both P and N surpass 5% ($P > 5\%$ and $N > 5\%$), because it occurs concurrently in both directions. With a predetermined likelihood of change, any further measure of $CIdiff$ can be classified as either positive (increment) or negative (decrement). The probability of change is P when there is an increase in the change, and N when there is a decrease in the change. According to the classification of the likelihood of change, 5 to 25% is considered “unlikely”, 25 to 75% is “possible”, 75 to 95% is “likely”, 95 to 99% is “very likely”, and greater than 99% is considered “extremely likely”.

Apart from the application of the MDC, another important point to observe in the individual analysis is that most changes in the variables are neither necessarily beneficial nor harmful. Increasing or decreasing the magnitude of a particular variable may be beneficial to one patient but harmful to another because LBP affects patients differently. Therefore, the results must be interpreted individually for each patient.

3. Results

3.1. MDC Index Tables Results from Test-Retest Lumbar Movement Analysis

Tables 3–5 are used to show the results of the calculated MDC values from the healthy subjects (absolute value of the MDC at 95% and dimensionless value of the effect size MDC_{es} at 95%). The tables also summarize the values of the means (μ), standard deviations (SD), and results of the variability through ICC of each of the analyzed variables. Each table summarizes the results for a specific lumbar movement: flexion–extension, rotation, and lateralization, respectively.

Table 3. Test-retest results from flexion–extension movement tests. Minimal detectable changes index.

	Test μ $\bar{x}$ (SD)	Retest μ $\bar{x}$ (SD)	ICC	MDC95	MDC95 _{es}
Flex.MaxRange (°)	93.03 (15.59)	91.35 (15.98)	0.94	10.79	0.68
Flex.Max (°)	64.78 (10.32)	64.60 (10.60)	0.94	7.09	0.68
Flex.MaxMean (°)	63.06 (10.67)	62.74 (10.42)	0.95	6.77	0.64
Flex.MaxCV (%)	2.16 (1.36)	2.12 (1.26)	0.086	1.37	1.05
Flex.Min (°)	−28.48 (8.87)	−27.12 (9.16)	0.89	8.23	0.91
Flex.MinMean (°)	−25.37 (8.20)	−24.19 (8.87)	0.85	9.09	1.06
Flex.MinCV (%)	−9.44 (4.26)	−9.66 (5.42)	0.57	8.82	1.81
Flex.Speed.MaxMean (°/s) *	110.38 (17.32)	110.67 (22.59)	0.85	21.30	1.06
Flex.Speed.MinMean (°/s)	−89.47 (18.65)	−88.76 (17.60)	0.88	17.66	0.97
Flex.Speed.AreaMean (°/s)	216.42 (54.87)	217.98 (75.79)	0.87	66.49	1.00
Flex.Speed.AreaStd (°/s)	122.64 (82.01)	106.47 (46.29)	0.55	123.92	1.86
Flex.Speed.AreaCoef (%)	55.05 (30.81)	50.66 (22.88)	−0.11	79.41	2.93
Flex.SpeedUp.MaxMean (°/s)	482.02 (144.35)	506.80 (122.37)	0.91	109.47	0.82
Flex.SpeedUp.MinMean (°/s)	−494.68 (157.85)	−564.66 (163.66)	0.60	282.67	1.76
Flex.SpeedUp.Harmony (°/s)	−0.52 (0.12)	−0.51 (0.08)	0.59	0.18	1.79
Flex.SpeedUp.Offset (°/s)	115.74 (6.26)	115.23 (5.92)	0.75	8.48	1.39

Flex: flexion–extension movement; μ : mean; SD: standard deviation; ICC: intraclass correlation coefficient; MDC95_{es}: minimal detectable change in dimensionless value effect size at 95%; MDC95: minimal detectable change in absolute value at 95%. Parameters that showed significant differences between groups were marked with an asterisk (*).

Table 4. Test-retest results from rotation movement tests. Minimal detectable changes index.

	Test μ $\bar{x}$ (SD)	Retest μ $\bar{x}$ (SD)	ICC	MDC95	MDC95 _{es}
Rot.MaxRange (°)	82.03 (11.24)	81.35 (12.88)	0.93	8.65	0.72
Rot.Max (°)	41.03 (5.73)	40.05 (6.96)	0.87	6.26	0.98
Rot.MaxMean (°)	38.90 (5.65)	38.14 (6.69)	0.89	5.73	0.93
Rot.MaxCV (%)	4.12 (2.31)	3.80 (1.87)	0.42	4.44	2.11
Rot.Min (°)	−41.93 (6.93)	−41.60 (7.17)	0.88	6.69	0.95
Rot.MinMean (°)	−39.92 (6.64)	−39.10 (6.86)	0.94	4.46	0.66
Rot.MinCV (%)	−3.73 (2.40)	−4.50 (4.28)	0.58	6.20	1.79
Rot.Speed.MaxMean (°/s)	85.37 (20.82)	83.58 (21.51)	0.91	17.16	0.81
Rot.Speed.MinMean (°/s)	−82.53 (20.00)	−81.81 (21.24)	0.94	14.14	0.69
Rot.Speed.AreaMean (°/s)	166.40 (59.04)	161.89 (64.08)	0.94	40.76	0.66
Rot.Speed.AreaStd (°/s)	91.87 (42.67)	88.44 (41.59)	0.07	112.75	2.68
Rot.Speed.AreaCoef (%)	59.38 (30.81)	59.16 (28.92)	0.36	64.72	2.21
Rot.SpeedUp.MaxMean (°/s)	435.34 (105.80)	467.90 (107.38)	0.33	241.20	2.26

Table 4. Cont.

	Test μ $\bar{x}$ (SD)	Retest μ $\bar{x}$ (SD)	ICC	MDC95	MDC95 _{es}
Rot.SpeedUp.MinMean (°/s)	−474.21 (142.97)	−447.03 (121.93)	0.19	330.98	2.49
Rot.SpeedUp.Harmony (°/s)	−0.52 (0.13)	−0.50 (0.13)	0.78	0.16	1.29
Rot.SpeedUp.Offset (°/s)	121.49 (9.03)	119.88 (8.47)	0.78	11.50	1.31

Rot: rotation movement; μ : mean; SD: standard deviation; ICC: intraclass correlation coefficient; MDC95_{es}: minimal detectable change in dimensionless value effect size at 95%; MDC95: minimal detectable change in absolute value at 95%.

Table 5. Test-retest results from lateralization movement tests. Minimal detectable changes index.

	Test μ $\bar{x}$ (SD)	Retest μ $\bar{x}$ (SD)	ICC	MDC95	MDC95 _{es}
Lat.MaxRange (°)	82.20 (9.97)	84.88 (9.29)	0.93	7.00	0.73
Lat.Max (°)	42.62 (6.06)	44.32 (5.30)	0.92	4.43	0.78
Lat.MaxMean (°)	41.16 (5.83)	42.16 (5.83)	0.94	3.70	0.67
Lat.MaxCV (%)	2.93 (1.62)	3.86 (2.82)	0.25	5.53	2.40
Lat.Min (°)	−39.58 (5.12)	−40.56 (5.50)	0.87	5.32	1.00
Lat.MinMean (°)	−37.90 (5.10)	−38.64 (5.09)	0.95	3.00	0.59
Lat.MinCV (%)	−3.48 (2.64)	−3.53 (3.10)	−0.03	8.09	0.59
Lat.Speed.MaxMean (°/s)	70.91 (17.15)	71.98 (16.46)	0.91	14.09	0.84
Lat.Speed.MinMean (°/s)	−73.19 (17.34)	−71.98 (16.46)	0.91	14.16	0.83
Lat.Speed.AreaMean (°/s)	150.35 (55.01)	156.26 (53.89)	0.95	32.90	0.60
Lat.Speed.AreaStd (°/s)	73.14 (31.86)	74.77 (35.00)	0.44	69.57	2.08
Lat.Speed.AreaCoef (%)	53.18 (27.17)	51.76 (24.60)	0.53	49.05	1.89
Lat.SpeedUp.MaxMean (°/s)	356.25 (91.40)	360.27 (100.21)	0.35	213.94	2.23
Lat.SpeedUp.MinMean (°/s)	−384.04 (106.50)	−355.99 (98.21)	0.38	223.91	2.19
Lat.SpeedUp.Harmony (°/s)	−0.55 (0.15)	−0.53 (0.15)	0.68	0.23	1.56
Lat.SpeedUp.Offset (°/s)	123.58 (10.18)	122.24 (10.44)	0.67	16.31	1.58

Lat: lateralization movement; μ : mean; SD: standard deviation; ICC: intraclass correlation coefficient; MDC95_{es}: minimal detectable change in dimensionless value effect size at 95%; MDC95: minimal detectable change in absolute value at 95%.

3.2. Results of the Patient-Level Study

The results of this study involved 30 patient-level sets of data summarized into two group analyses corresponding to the two therapy treatments (15 patients each group). In this section, we present the summarized data for the two group analysis. The first group corresponds to the deep oscillation treatment therapy group; its results are summarized in Tables 6–8 and Figures 5–7; each table and figure is presented for one of the three movements (flexion–extension, rotation, and lateralization, respectively). Tables 9–11 and Figures 8–10 present data for the second group subject to traditional treatment therapy. These tables include the change between the pre- and post-series of the analyzed variables, the threshold MDC, and the MBD numerical results (i.e., the N, T, and P values; the variables that underwent a significant change are marked with a *). The figures use the confidence interval representation to show the information included in the Tables. The numerical results and the biomechanical interpretation of the single patient analysis are presented as Supplementary Materials in an Excel file (Supplementary Materials S2 and S3).

Table 6. Results of study for flexion–extension movement with deep oscillation treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Flex.DO.MaxRange (°)	64.26	80.32	16.06	2.79	0/0/100
Flex.DO.Max (°) *	51.46	61.72	10.25	1.83	0/0/100
Flex.DO.MaxMean (°) *	49.19	59.88	10.69	1.75	0/0/100
Flex.DO.MaxCV (%) *	3.51	2.36	−1.16	0.35	90/09/01
Flex.DO.Min (°) *	−12.79	−18.60	−5.81	2.12	95/05/0
Flex.DO.MinMean (°)	−10.08	−16.17	−6.08	2.35	97/03/0
Flex.DO.MinCV (%)	−22.16	−17.89	4.28	2.28	34/11/55
Flex.DO.Speed.MaxMean (°/s) *	77.41	103.41	26.01	5.50	0/0/100
Flex.DO.Speed.MinMean (°/s) *	−68.59	−89.53	−20.95	4.56	100/0/0
Flex.DO.Speed.AreaMean (°/s) *	116.34	195.13	78.79	17.17	0/0/100
Flex.DO.Speed.AreaStd (°/s)	53.62	69.27	15.65	32.00	0/100/0
Flex.DO.Speed.AreaCoef (%)	54.35	37.61	−16.75	20.50	31/69/0
Flex.DO.SpeedUp.MaxMean (°/s) *	360.73	437.13	76.40	28.27	0/06/94
Flex.DO.SpeedUp.MinMean (°/s)	−387.66	−472.13	−84.46	72.99	61/39/0
Flex.DO.SpeedUp.Harmony (°/s)	−0.57	−0.60	−0.03	0.05	37/60/03
Flex.DO.SpeedUp.Offset (°/s)	114.59	117.19	2.60	2.19	04/40/56

Flex: flexion–extension movement; DO: deep oscillation; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes.. Parameters that showed significant differences between groups were marked with an asterisk (*).

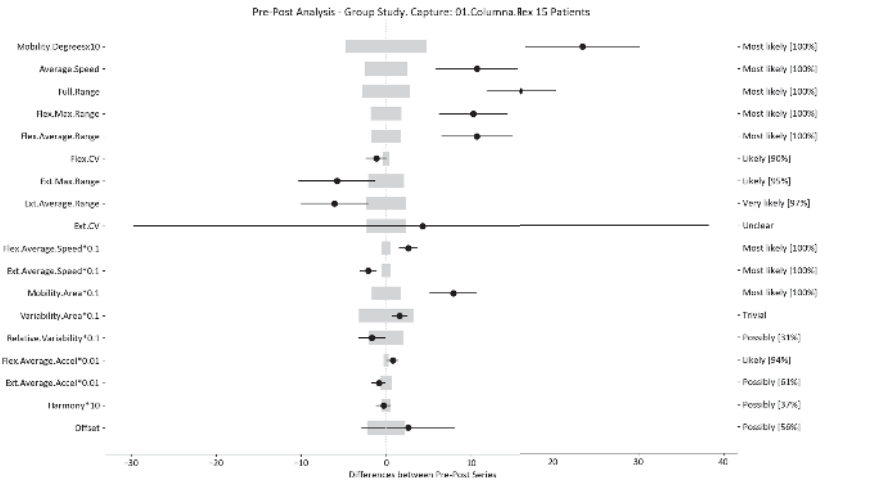


Figure 5. Flexion–extension analysis and confidence interval for DO group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

Table 7. Results of study for rotation movement with deep oscillation treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Rot.DO.MaxRange (°) *	63.64	75.20	11.56	2.23	0/0/100
Rot.DO.Max (°) *	30.59	37.59	7.00	1.62	0/0/100
Rot.DO.MaxMean (°) *	28.58	35.94	7.37	1.48	0/0/100
Rot.DO.MaxCV (%)	6.25	3.98	−2.27	1.15	93/07/0
Rot.DO.Min (°) *	−33.06	−37.61	−4.55	1.73	98/02/0
Rot.DO.MinMean (°) *	−31.10	−35.71	−4.61	1.15	99/01/0
Rot.DO.MinCV (%)	−5.10	−4.37	0.72	1.60	0/84/16
Rot.DO.Speed.MaxMean (°/s) *	70.08	87.03	16.95	4.43	0/0/100
Rot.DO.Speed.MinMean (°/s) *	−69.55	−94.09	−24.55	3.65	100/0/0
Rot.DO.Speed.AreaMean (°/s) *	115.47	181.37	65.90	10.52	0/0/100
Rot.DO.Speed.AreaStd (°/s)	51.55	68.28	16.73	29.11	0/99/01
Rot.DO.Speed.AreaCoef (%)	48.39	41.32	−7.07	16.71	0/100/0
Rot.DO.SpeedUp.MaxMean (°/s) *	351.89	472.20	120.31	62.28	0/08/92
Rot.DO.SpeedUp.MinMean (°/s)	−371.03	−447.01	−75.98	85.46	41/59/0
Rot.DO.SpeedUp.Harmony (°/s)	−0.57	−0.64	−0.06	0.04	76/24/0
Rot.DO.SpeedUp.Offset (°/s)	124.96	129.07	4.11	2.97	0/30/70

Rot: rotation movement; DO: deep oscillation; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes. Parameters that showed significant differences between groups were marked with an asterisk (*).

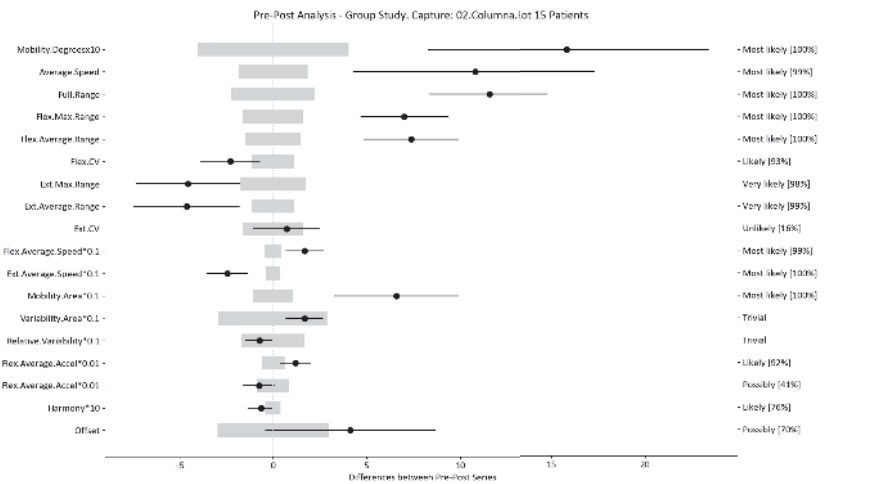


Figure 6. Rotation analysis and confidence interval for DO group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

Table 8. Results of study for lateralization movement with deep oscillation treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Lat.DO.MaxRange (°) *	61.69	74.96	13.27	1.81	0/0/100
Lat.DO.Max (°) *	31.98	37.97	5.99	1.14	0/0/100
Lat.DO.MaxMean (°) *	30.70	36.80	6.10	0.96	0/0/100
Lat.DO.MaxCV (%)	3.75	2.79	−0.97	1.43	19/81/0
Lat.DO.Min (°) *	−29.71	−36.99	−7.28	1.37	100/0/0
Lat.DO.MinMean (°) *	−28.59	−36.00	−7.41	0.77	100/0/0
Lat.DO.MinCV (%)	−3.81	−2.34	1.46	2.09	0/70/30
Lat.DO.Speed.MaxMean (°/s) *	58.93	77.67	18.74	3.64	0/0/100
Lat.DO.Speed.MinMean (°/s) *	−61.35	−82.58	−21.23	3.66	100/0/0
Lat.DO.Speed.AreaMean (°/s) *	109.71	170.13	60.42	8.49	0/0/100
Lat.DO.Speed.AreaStd (°/s)	38.45	51.01	12.56	17.96	0/90/10
Lat.DO.Speed.AreaCoef (%)	43.31	35.54	−7.77	12.66	20/80/0
Lat.DO.SpeedUp.MaxMean (°/s) *	269.18	355.10	85.92	55.24	0/18/82
Lat.DO.SpeedUp.MinMean (°/s)	−278.67	−362.77	−84.10	57.81	82/18/0
Lat.DO.SpeedUp.Harmony (°/s)	−0.67	−0.71	−0.03	0.06	14/85/01
Lat.DO.SpeedUp.Offset (°/s)	132.43	135.13	2.70	4.21	01/76/23

Lat: lateralization movement; DO: deep oscillation; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes. Parameters that showed significant differences between groups were marked with an asterisk (*).

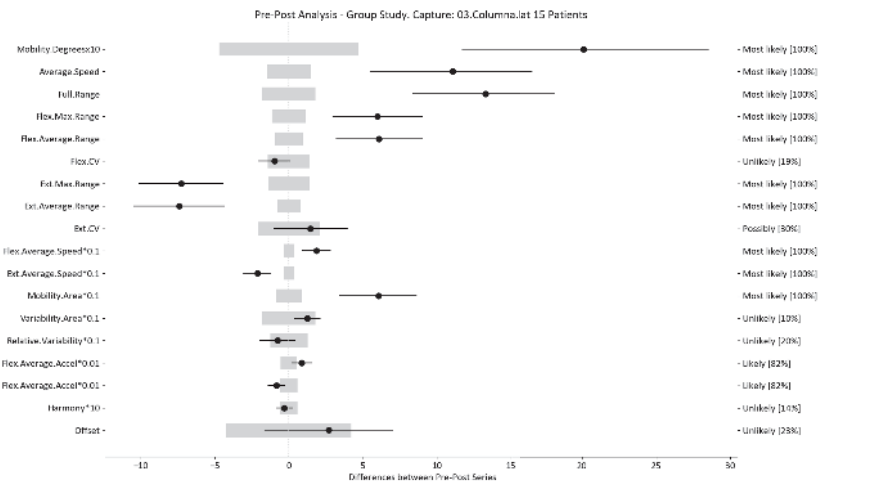


Figure 7. Lateralization analysis and confidence interval for DO group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

Table 9. Results of study for flexion–extension movement with traditional treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Flex.TT.MaxRange (°) *	72.31	85.05	12.74	2.79	0/0/100
Flex.TT.Max (°) *	54.00	66.15	12.15	1.83	0/0/100
Flex.TT.MaxMean (°) *	51.09	64.20	13.11	1.75	0/0/100
Flex.TT.MaxCV (%) *	4.74	2.63	−2.11	0.35	100/0/0
Flex.TT.Min (°)	−18.30	−18.90	−0.60	2.12	23/67/10
Flex.TT.MinMean (°)	−16.11	−15.75	0.37	2.35	11/69/20
Flex.TT.MinCV (%) *	−10.61	−18.75	−8.13	2.28	62/8/30
Flex.TT.Speed.MaxMean (°/s) *	76.57	101.12	24.55	5.50	0/0/100
Flex.TT.Speed.MinMean (°/s) *	−71.44	−91.20	−19.76	4.56	100/0/0
Flex.TT.Speed.AreaMean (°/s) *	130.87	195.99	65.12	17.17	0/0/100
Flex.TT.Speed.AreaStd (°/s)	53.74	777.77	24.04	32.00	0/87/13
Flex.TT.Speed.AreaCoef (%)	59.27	52.44	−6.82	20.50	8/92/0
Flex.TT.SpeedUp.MaxMean (°/s) *	320.41	447.90	127.49	28.27	0/3/97
Flex.TT.SpeedUp.MinMean (°/s)	−324.74	−459.96	−135.23	72.99	92/8/0
Flex.TT.SpeedUp.Harmony (°/s)	−0.58	−0.62	−0.04	0.05	33/67/0
Flex.TT.SpeedUp.Offset (°/s)	117.97	118.46	0.48	2.19	0/90/10

Flex: flexion–extension movement; TT: traditional treatment; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes. Parameters that showed significant differences between groups were marked with an asterisk (*).

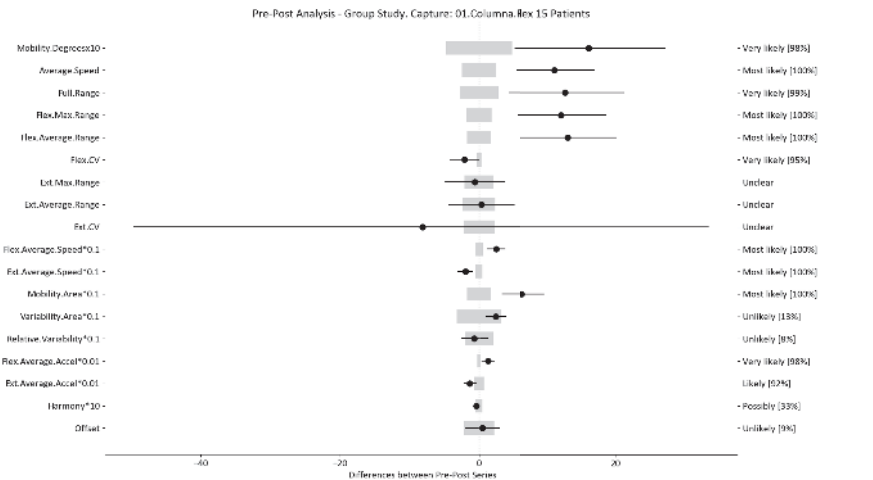


Figure 8. Flexion–extension analysis and confidence interval for TT group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

Table 10. Results of study for rotation movement with traditional treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Rot.TT.MaxRange (°) *	69.42	77.84	8.42	2.23	0/0/100
Rot.TT.Max (°)	33.78	38.51	4.73	1.62	0/27/73
Rot.TT.MaxMean (°)	31.94	37.21	5.27	1.48	0/14/86
Rot.TT.MaxCV (%)	4.46	2.93	−1.52	1.15	75/25/0
Rot.TT.Min (°) *	−35.64	−39.33	−3.69	1.73	87/13/0
Rot.TT.MinMean (°) *	−33.65	−37.42	−3.77	1.15	92/8/0
Rot.TT.MinCV (%)	−4.95	−3.76	1.18	1.60	0/70/30
Rot.TT.Speed.MaxMean (°/s) *	73.33	90.45	17.12	4.43	0/0/100
Rot.TT.Speed.MinMean (°/s) *	−70.86	−90.90	−20.04	3.65	100/0/0
Rot.TT.Speed.AreaMean (°/s) *	126.99	185.66	58.67	10.52	0/0/100
Rot.TT.Speed.AreaStd (°/s)	52.05	68.55	16.50	29.11	0/100/0
Rot.TT.Speed.AreaCoef (%)	43.21	37.21	−6.00	16.71	0/100/0
Rot.TT.SpeedUp.MaxMean (°/s)	369.28	460.78	91.50	62.28	0/20/80
Rot.TT.SpeedUp.MinMean (°/s)	−349.55	−440.21	−90.66	85.46	60/40/0
Rot.TT.SpeedUp.Harmony (°/s)	−0.59	−0.67	−0.08	0.04	93/7/0
Rot.TT.SpeedUp.Offset (°/s)	126.58	132.75	6.17	2.97	0/5/95

Rot: rotation movement; TT: traditional treatment; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes. Parameters that showed significant differences between groups were marked with an asterisk (*).

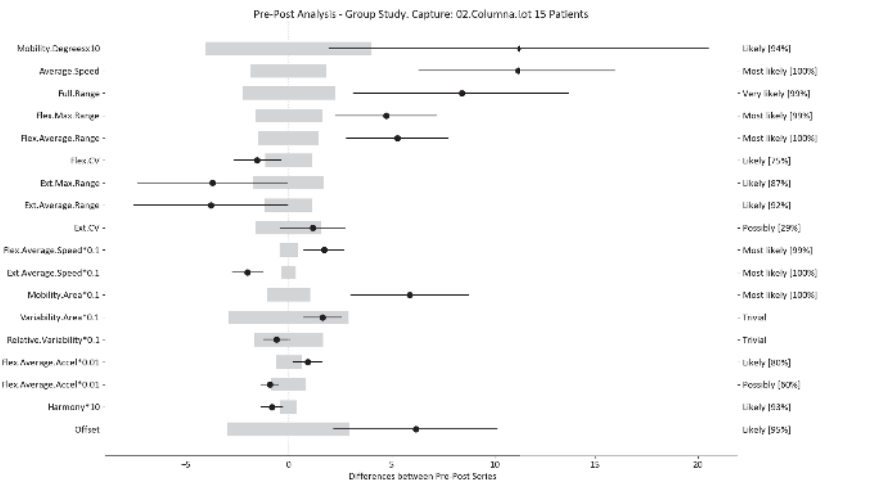


Figure 9. Rotation analysis and confidence interval for TT group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

Table 11. Results of study for lateralization movement with traditional treatment.

	Value Pre	Value Post	Mean dif	±MDC	N/U/P
Lat.TT.MaxRange (°) *	60.25	68.09	7.84	1.81	0/5/95
Lat.TT.Max (°) *	30.74	35.48	4.74	1.14	0/3/97
Lat.TT.MaxMean (°) *	29.46	34.08	4.61	0.96	0/3/97
Lat.TT.MaxCV (%)	3.97	3.40	−0.57	1.43	4/96/0
Lat.TT.Min (°)	−29.51	−32.62	−3.11	1.37	84/16/0
Lat.TT.MinMean (°) *	−27.73	−31.38	−3.65	0.77	95/4/1
Lat.TT.MinCV (%)	−4.97	−3.30	1.68	2.09	0/74/26
Lat.TT.Speed.MaxMean (°/s) *	55.73	67.91	12.18	3.64	0/0/100
Lat.TT.Speed.MinMean (°/s) *	−54.56	−68.37	−13.81	3.66	100/0/0
Lat.TT.Speed.AreaMean (°/s) *	88.84	127.80	38.96	8.49	0/0/100
Lat.TT.Speed.AreaStd (°/s)	30.67	36.87	6.20	17.96	0/100/0
Lat.TT.Speed.AreaCoef (%)	35.39	29.76	−5.62	12.66	0/100/0
Lat.TT.SpeedUp.MaxMean (°/s)	244.55	291.50	46.95	55.24	0/68/32
Lat.TT.SpeedUp.MinMean (°/s)	−248.87	−282.96	−34.09	57.81	3/97/0
Lat.TT.SpeedUp.Harmony (°/s)	−0.67	−0.73	−0.06	0.06	54/46/0
Lat.TT.SpeedUp.Offset (°/s)	132.91	137.55	4.65	4.21	0/40/60

Lat: lateralization movement; TT: traditional treatment; MDC: minimal detectable change threshold; N: probability of negative changes; U: probability of unknown/trivial changes; P: probability of positive changes. Parameters that showed significant differences between groups were marked with an asterisk (*).

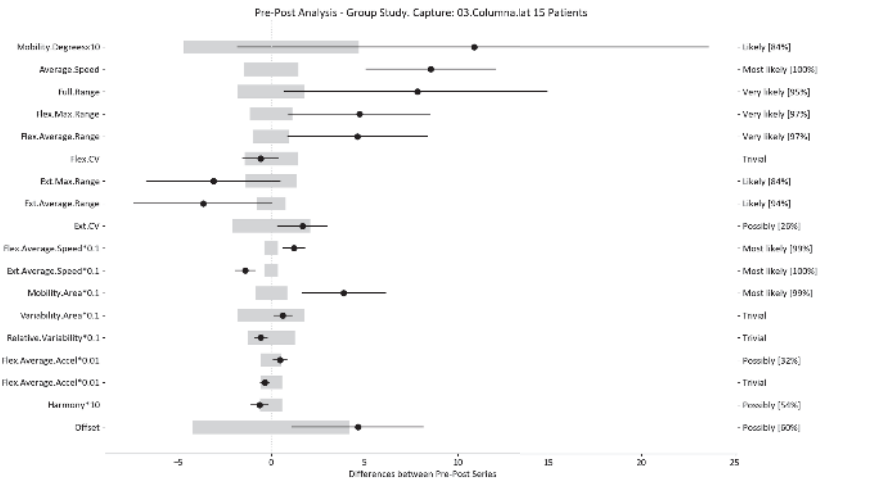


Figure 10. Lateralization analysis and confidence interval for TT group. Wide light gray bars: MDC threshold. Thin black bars: confidence intervals of differences.

In relation to the individual data of all patients for both treatments, this information is summarized in the Supplementary Materials. When deciding how to interpret the results, we only considered the changes that exceeded the threshold δ at a level interpreted as ‘very likely’ (above >95% probability); we denominated these as ‘real’ changes.

3.3. Results of the Low-Back-Specific Version of the SF-36 Physical Functioning Scale

As established before, this study involved 30 separate patients, each of whom was assessed by a physician specialist using a low-back-specific version of the SF-36 Physical Functioning scale two times; once before the treatment sessions and once after such treatments. In this section, we present the scores determined for the same patients referred to in the previous section (see Table 12). All the scores, as well as the more detailed answers of the patients to the questions from the rest of the patients, are presented in the Supplementary Materials.

Table 12. Low-back-specific SF-36 Physical Functioning scale average scores and differences in both treatments.

	DOT Group	TT Group
Pre-Treatment	60.00	65.78
Post-treatment	74.67	75.67
Difference	14.67	9.89

4. Discussion

To assess the reproducibility in our test-retest results, we evaluated the ICC values, which exceeded 0.7. These results are satisfactory in comparison to those reported in a similar study [30], where use was made of an optical motion capture system to analyze the same movements but in a seated position; in this study only moderate ICC values (0.6–0.7) for most of their variables were obtained. In another study by Megan O’Grady et al. [31], use was similarly made of an IMU system to analyze lumbar movement (flexion–extension, lateral flexion left to right, and rotation). When calculating the reproducibility, these authors obtained values ($0.95 < ICC < 1.00$) for the kinematics they were evaluating. The values obtained in our study with respect to the same variables were very similar to those reported in Megan O’Grady’s study [31] ($ICC > 0.85$). Both of these findings indicate the reliability of IMU systems and their potential to assess the kinematics of lumbar movement. Most of the kinetics variables also exhibited high reproducibility ($ICC > 0.85$), with the exception of the accelerations for the lateralization and rotation movement. When compared to other studies that also dealt with MDC, we note that our results for the ROM were comparable with those obtained in similar studies [32,33]. However, we were unable to find details of previous MDC studies focusing on the velocities of these movements, so we are not able to fully support our findings with respect to these variables.

The utility of the test is determined by the MDC value of the variables, with Tables 3–5 (marked with a *) indicating the variables with the lowest MDC ranges and the greatest ICC in our study. Furthermore, based on the statistical application of the MDC technique in a sample of 30 patients, these are the variables most relevant for measuring change at the individual level and, hence, the efficacy of an intervention/treatment. The findings were compared to those of various studies of individuals with LBP, where the most reliable variables are those in which the biggest change was identified in relation to improvement [34–39]. As a result, the BackMov test could give clinicians reliable and easy to understand information about the change in a particular characteristic seen in a patient, enabling them to clearly see the improvement or impairment of a patient. The clinician, who can now be aware of these specific characteristics in the patient’s circumstances may find it easier to decide whether or not to continue with a treatment or whether to change it in favor of a treatment to target the specific change that they may have noticed.

For example, we draw attention to the graphs depicted in Figures 5–10 with regard to the findings of this study. These graphs were created to graphically and intuitively display the changes observed in the set of variables, making it easier for the physician to recognize the factors that need further consideration and analysis. This information can be useful, as a significant number of guidelines primarily focus on assessing the

disability that the LBP may be causing the patient [40]. In relation to this, we saw that the improvement in ROM and velocities could be directly related to the abilities of individuals to perform physical activities. This is supported by the observation that for patients that exhibited a significant change in these variables when responding to the questionnaire conducted by the physicians (see Supplementary Materials S5), their answers showed that most of them were better when performing activities that involved a high degree of movement from their point of view, such as going up the stairs or walking. Furthermore, as the velocities of the patients also seemed to have improved, this could be related to the therapies helping patients to overcome their fear of movement so that they were less hesitant to move, therefore decreasing the overall time it took them to perform an activity [41,42]. Overcoming the fear of movement would correlate with the improvement in these variables with overall improvement in performing daily activities. Furthermore, as movement increases this also helps to reduce muscle tension and stiffness that may have been caused as a result of the individuals trying to move as little as possible to avoid pain [43–45]. The results of our MDC comparisons are in line with evaluation of progress in the patients conducted by the specialist physicians using the low-back-specific form of the SF-36 Physical Functioning scale. The results were that both groups showed real changes in the variables that can be associated with disability and thus may influence recovery of LBP patients. In this way, the graphs are intended to make it easier to identify a change in a variable related to disability, thus facilitating monitoring of treatment without the need to keep completing questionnaires to assess progress. Furthermore, if the clinician needs to pursue a more extensive and rigorous analysis, the tables (Tables 6–11) contain more detailed information on the variables and the changes in the patients kinematics.

Personalized medicine has become an important goal for health professionals [46]. The response to therapy for conditions like LBP problems can be objectively described through individual patient assessment, so these kinds of tools that reduce the time to assess progression can be good for the future of therapy. Although there seems to be benefit in using the BackMov test for the purposes described, there are certain issues that need to be addressed as they may cause the results to not be reliable if they are not taken into account. A clear issue is the fact that IMUs suffer from integration drift, which involves small errors appearing during the measurement of acceleration and angular velocity that can be progressively combined into larger errors in velocity and angle, which are compounded into still greater errors in position [47]. Another issue that may arise from this test that must be taken into consideration is the effect of noise associated with other variables, including pathology, treatment used between pre- and post-testing, illness progression, and even private personal events [48–52]. Despite the promising results and the potential utility of the BackMov test in assessing lumbar movement recovery in LBP patients post-therapy, it is crucial to acknowledge certain limitations that may impact the interpretation of our findings. One notable limitation is the exclusion criteria employed in participant selection. Excluding individuals with pre-existing diseases or disabilities that might hinder their movement was essential for maintaining homogeneity in the study cohort. However, it is important to recognize that the applicability of the BackMov test to a broader patient population, including those with comorbidities or disabilities, remains to be explored. Another limitation stems from the exclusion of high-level athletes from the study cohort. While this decision aimed to control for the potential influence of exceptional physical fitness on lumbar movement, it has inadvertently limited the generalizability of our findings to this specific subgroup. The BackMov test's effectiveness in assessing movement recovery in athletes with LBP warrants further investigation to ascertain its relevance and applicability across diverse patient profiles.

Additionally, the exclusion of individuals engaged in dangerous activities during the study period raises questions about the test's external validity in real-world scenarios where patients may need to resume such activities post-therapy. The BackMov test's capacity to capture improvements in lumbar movement relevant to occupational or recreational activities involving increased risk remains an avenue for future research. Moreover, the exclusion of individuals undergoing drug-specific treatments to alleviate pain introduces

another layer of complexity. Many LBP patients often rely on pharmacological interventions for pain management. The impact of these treatments on lumbar movement and the BackMov test's sensitivity to changes in patients undergoing drug-specific interventions were not explored in this study. Future investigations should consider incorporating these variables to enhance the test's applicability in a broader clinical context.

As an important note, while the BackMov test appears to possess great potential to detect when a patient experiences a relevant change in a particular variable (high likelihood of change), the physician should still not only rely on these results for determining whether or not this change is important for treatment. This is significant because, despite the fact that this kind of metric is meant to offer unbiased data in support of a diagnosis, the MDC itself is not a definitive diagnosis and should not be used as the only tool to determine a clinical decision as it is based on merely an association of certain variables with the disability produced by the LBP [53–55]. Some of these associations are still in need of further evaluation to be considered directly proportional to the effects of LBP. Furthermore, these values do not account for the psychological aspect that must be addressed when dealing with these kinds of problems [56,57]. Regarding the study's implications for research and clinical practice, we believe that the use of the BackMov test as a 'logical' guideline can be the foundation for future patient studies for the evaluation of treatments. However, it is necessary to further develop the rules that allow for more effective qualification of the change detected in the variables. This would further support use of the method in the evaluation of treatments or the monitoring of a patient's overall recovery based on the premise of disability recovery from improved kinematics of lumbar movements.

5. Future Work

To enhance the understanding and potential of this methodology, a more comprehensive and rigorous approach is proposed. In the newly suggested methodology, a control group should be included to observe the progress of LBP patients in the absence of any treatment. Additionally, other therapeutic treatments for the pathology should be incorporated. In future studies, greater attention must be given to the psychological and psychosocial aspects of the disease. Furthermore, the utilization of alternative statistical tools, such as the minimal important difference (MID), should be considered to complement the minimum detectable change (MDC). Determining MID thresholds, alongside MDC, would offer a more valuable means to assess therapies, provide evidence for diagnoses, and facilitate the monitoring of patient recoveries. Lastly, unless more evidence emerges regarding the relationship between disability and kinematic variables or the association of fear of movement with kinematics, a dedicated study focusing on establishing these connections should be planned to support the use of this methodology.

6. Conclusions

In conclusion, we can say that the test-retest of the BackMov test that was conducted on a sample of healthy young subjects was able to show excellent reproducibility and validity for variables such as the Max.Range, the Flex.Average.Range, the Ext.Average.Range, the Flex.Average.Speed and the Ext.Average.Speed in all three assessed lumbar movements, thus making the use of IMUs a reliable tool for assessment of range of motion (ROM) evaluation recovery in LBP patients. The assessments made by the expert physician were mainly validated using the objective pre- and post-test outcomes for 30 patients with LBP difficulties. In this approach, the results were demonstrated to be applicable to the evaluated patient, as certain factors were sensitive enough to detect significant changes associated with the progression (improvement, worsening, or no changes) in LBP. From this, we draw the conclusion that the application of clinical lumbar movements analysis based on IMU technology in rehabilitation could be advantageous in this branch of the medical field as an extra tool that may help specialists to assess the extent of recovery of any lost mobility caused by LBP. The results of this study also may contribute to improving medical decision-making and the individualized follow-up of patients; however, more

extensive and rigorous studies must be undertaken in order to fully validate the potential for use of these tools.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s24030913/s1>. This article contains spreadsheet of the results of data collection (BMI, Ages, gender, pain intensity, and brief examination from both groups), Supplementary material S1; variables data collection and complete statistics tables from both groups, Supplementary material S2 and supplementary material S3; collection and complete statistics tables from the Healthy Group, supplementary material S4; and the results from the low-back-specific Short Form (SF)-36 Physical Functioning scale from each patient, supplementary material S5.

Author Contributions: Conceptualization, J.M., J.J.M., F.V.-M., C.G. and P.A.V.-L.; methodology, I.A.-S., J.M., J.J.M., C.G., F.V.-M., S.A.G.-L., C.M.V.-C. and P.A.V.-L.; software, F.V.-M., C.G. and P.A.V.-L.; validation, C.G., F.V.-M. and P.A.V.-L.; formal analysis, C.G., F.V.-M. and P.A.V.-L.; investigation, D.A.-G., C.C.-M., F.V.-M. and P.A.V.-L.; resources, J.M., J.J.M., F.V.-M., C.C.-M. and P.A.V.-L.; data curation, C.G., F.V.-M. and P.A.V.-L.; writing—original draft preparation, J.M., C.G., C.C.-M., J.J.M., I.A.-S., P.A.V.-L., S.A.G.-L., C.M.V.-C. and F.V.-M.; writing—review and editing, I.A.-S., J.M., J.J.M., C.G., S.A.G.-L., C.M.V.-C., C.C.-M., F.V.-M. and P.A.V.-L.; visualization, J.M., J.J.M., C.G., D.A.-G., C.C.-M., F.V.-M. and P.A.V.-L.; supervision, D.A.-G., C.C.-M., J.M., J.J.M. and F.V.-M.; project administration, D.A.-G., J.M., C.G., J.J.M. and F.V.-M.; funding acquisition, D.A.-G., C.G. and F.V.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad Indoamérica project number INV-0014-05-019.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Bioethics Committee of the Pontificia Universidad Católica del Ecuador (N° EO-146-2022, approved on 22 March 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data are contained within the article.

Acknowledgments: We would like to thank all the people who voluntarily participated in this study, the I3A—University Institute of Engineering Research of Aragon, University of Zaragoza, Zaragoza, Spain, as well as the physiotherapy clinic “FisioMedics”—Rehabilitation Center Physics from Ibarra-Ecuador, who collaborated with us unconditionally and their representative, Karen Mora.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zahra, N.A.I.; Sheha, E.A.A.E.; Elsayed, H.A. Low Back Pain, Disability and Quality of Life among Health Care Workers. *Int. J. Pharm. Res. Allied Sci.* **2020**, *9*, 34–44.
2. Hartvigsen, J.; Hancock, M.J.; Kongsted, A.; Louw, Q.; Ferreira, M.L.; Genevay, S.; Hoy, D.; Karppinen, J.; Pransky, G.; Sieper, J.; et al. What low back pain is and why we need to pay attention. *Lancet* **2018**, *391*, 2356–2367. [CrossRef] [PubMed]
3. Maher, C.G. Effective physical treatment for chronic low back pain. *Orthop. Clin. N. Am.* **2004**, *35*, 57–64. [CrossRef]
4. Foster, N.E.; Anema, J.R.; Cherkin, D.; Chou, R.; Cohen, S.P.; Gross, D.P.; Ferreira, P.H.; Fritz, J.M.; Koes, B.W.; Peul, W.; et al. Prevention and treatment of low back pain: Evidence, challenges, and promising directions. *Lancet* **2018**, *391*, 2368–2383. [CrossRef] [PubMed]
5. De la Torre, J.; Marin, J.; Polo, M.; Marín, J.J. Applying the minimal detectable change of a static and dynamic balance test using a portable stabilometric platform to individually assess patients with balance disorders. *Healthcare* **2020**, *8*, 402. [CrossRef]
6. Marín, J.; Blanco, T.; de la Torre, J.; Marín, J.J. Gait analysis in a box: A system based on magnetometer-free IMUs or clusters of optical markers with automatic event detection. *Sensors* **2020**, *20*, 3338. [CrossRef]
7. Bailly, F.; Trouvin, A.P.; Bercier, S.; Dadoun, S.; Deneuville, J.P.; Faguer, R.; Fassier, J.B.; Koleck, M.; Lassalle, L.; Le Vraux, T.; et al. Clinical guidelines and care pathway for management of low back pain with or without radicular pain. *Jt. Bone Spine* **2021**, *88*, 105227. [CrossRef]
8. Deane, J.A.; Papi, E.; Phillips, A.T.; McGregor, A.H. Reliability and minimal detectable change of the ‘Imperial Spine’ marker set for the evaluation of spinal and lower limb kinematics in adults. *BMC Res. Notes* **2020**, *13*, 495. [CrossRef] [PubMed]
9. López-de Uralde-Villanueva, I.; Sarría Visa, T.; Moscardó Marichalar, P.; del Corral, T. Minimal detectable change in six-minute walk test in children and adolescents with cystic fibrosis. *Disabil. Rehabil.* **2021**, *43*, 1594–1599. [CrossRef]
10. Marín, J.; Blanco, T.; Marín, J.J.; Moreno, A.; Martitegui, E.; Aragüés, J.C. Integrating a gait analysis test in hospital rehabilitation: A service design approach. *PLoS ONE* **2019**, *14*, e0224409. [CrossRef]

11. Marin, J.; Marin, J.J.; Blanco, T.; de la Torre, J.; Salcedo, I.; Martitegui, E. Is my patient improving? Individualized gait analysis in rehabilitation. *Appl. Sci.* **2020**, *10*, 8558. [CrossRef]
12. Prakash, C.; Kumar, R.; Mittal, N. Recent developments in human gait research: Parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artif. Intell. Rev.* **2018**, *49*, 1–40. [CrossRef]
13. Zhou, H.; Hu, H. Human motion tracking for rehabilitation—A survey. *Biomed. Signal Process. Control* **2008**, *3*, 1–18. [CrossRef]
14. Simon, S.R. Quantification of human motion: Gait analysis—Benefits and limitations to its application to clinical problems. *J. Biomech.* **2004**, *37*, 1869–1880. [CrossRef]
15. Cloete, T.; Scheffer, C. Benchmarking of a full-body inertial motion capture system for clinical gait analysis. In Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08—Personalized Healthcare through Technology, Vancouver, BC, Canada, 20–25 August 2008; pp. 4579–4582. [CrossRef]
16. Cooper, G.; Sheret, I.; McMillian, L.; Siliverdis, K.; Sha, N.; Hodgins, D.; Kenney, L.; Howard, D. Inertial sensor-based knee flexion/extension angle estimation. *J. Biomech.* **2009**, *42*, 2678–2685. [CrossRef] [PubMed]
17. Mooney, R.; Corley, G.; Godfrey, A.; Quinlan, L.R.; ÓLaighin, G. Inertial sensor technology for elite swimming performance analysis: A systematic review. *Sensors* **2016**, *16*, 18. [CrossRef]
18. Marin, J.; Blanco, T.; Marin, J.J. Octopus: A design methodology for motion capture wearables. *Sensors* **2017**, *17*, 1875. [CrossRef] [PubMed]
19. Marxreiter, F.; Gaßner, H.; Borozdina, O.; Barth, J.; Kohl, Z.; Schlachetzki, J.C.; Thun-Hohenstein, C.; Volc, D.; Eskofier, B.M.; Winkler, J.; et al. Sensor-based gait analysis of individualized improvement during apomorphine titration in Parkinson's disease. *J. Neurol.* **2018**, *265*, 2656–2665. [CrossRef]
20. Baker, R. Gait analysis methods in rehabilitation. *J. Neuroeng. Rehabil.* **2006**, *3*, 4. [CrossRef]
21. Freiwald, J.; Magni, A.; Fanlo-Mazas, P.; Paulino, E.; Sequeira de Medeiros, L.; Moretti, B.; Schleip, R.; Solarino, G. A role for superficial heat therapy in the management of non-specific, mild-to-moderate low back pain in current clinical practice: A narrative review. *Life* **2021**, *11*, 780. [CrossRef]
22. Abdulla, F.A.; Alsaadi, S.; Sadat-Ali, M.; Alkhamis, F.; Alkawaja, H.; Lo, S. Effects of pulsed low-frequency magnetic field therapy on pain intensity in patients with musculoskeletal chronic low back pain: Study protocol for a randomised double-blind placebo-controlled trial. *BMJ Open* **2019**, *9*, 6561. [CrossRef]
23. Furlan, A.D.; Giraldo, M.; Baskwill, A.; Irvin, E.; Imamura, M. Massage for low-back pain. *Cochrane Database Syst. Rev.* **2015**, *9* [CrossRef] [PubMed]
24. Voinea, A.; Iacobini, A. Williams' program for low back pain. *Marathon* **2014**, *6*, 210–214.
25. Mayer, J.M.; Ralph, L.; Look, M.; Erasala, G.N.; Verna, J.L.; Matheson, L.N.; Mooney, V. Treating acute low back pain with continuous low-level heat wrap therapy and/or exercise: A randomized controlled trial. *Spine J.* **2005**, *5*, 395–403. [CrossRef]
26. Steinhäuser, S.; Ganter, M.T.; Stadelmann, V.; Hofer, C.K. Whole-Body Electrostatic Pain Treatment in Adults with Chronic Pain: A Prospective Multicentric Observational Clinical Trial. *Pain Ther.* **2023**, *13*, 69–85. [CrossRef] [PubMed]
27. Davidson, M.; Keating, J.L.; Eyres, S. A low back-specific version of the SF-36 Physical Functioning scale. *Spine* **2004**, *29*, 586–594. [CrossRef]
28. Baydal, M.; Pozo, D.; Moya, F.; Fuentes, V. *Cinematografía del Raquis Cervical. Definición de Patrones de Movimiento para la Valoración Funcional en el Síndrome del Latigazo Cervical*; Universitat Politècnica de València: Valencia, Spain, 2012; pp. 35–36.
29. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **1999**, *8*, 135–160. [CrossRef] [PubMed]
30. Wildenbeest, M.H.; Kiers, H.; Tuijt, M.; van Dieën, J.H. Reliability of measures to characterize lumbar movement patterns, in repeated seated reaching, in a mixed group of participants with and without low-back pain: A test-retest, within- and between session. *J. Biomech.* **2021**, *121*, 110435. [CrossRef]
31. O'Grady, M.; O'Dwyer, T.; Connolly, J.; Condell, J.; Esquivel, K.M.; O'Shea, F.D.; Gardiner, P.; Wilson, F. Measuring spinal mobility using an inertial measurement unit system: A reliability study in axial spondyloarthritis. *Diagnostics* **2021**, *11*, 490. [CrossRef]
32. Bauer, C.M.; Heimgartner, M.; Rast, F.M.; Ernst, M.J.; Oetiker, S.; Kool, J. Reliability of lumbar movement dysfunction tests for chronic low back pain patients. *Man. Ther.* **2016**, *24*, 81–84. [CrossRef]
33. Kolber, M.J.; Pizzini, M.; Robinson, A.; Yanez, D.; Hanney, W.J. The reliability and concurrent validity of measurements used to quantify lumbar spine mobility: An analysis of an iPhone® application and gravity based inclinometry. *Int. J. Sport. Phys. Ther.* **2013**, *8*, 129–137.
34. Barrett, C.J.; Singer, K.P.; Day, R. Assessment of combined movements of the lumbar spine in asymptomatic and low back pain subjects using a three-dimensional electromagnetic tracking system. *Man. Ther.* **1999**, *4*, 94–99. [CrossRef]
35. Elnaggar, I.; Nordin, M.; Sheikhzadeh, A.; Parnianpour, M.; Kahanovitz, N. Effects of Spinal Flexion and Extension Exercises on Low-Back Pain and Spinal Mobility in Chronic Mechanical Low Back Pain Patients. *Spine* **1991**, *16*, 967–972. [CrossRef]
36. Konstantinou, K.; Foster, N.; Rushton, A.; Baxter, D.; Wright, C.; Breen, A. Flexion Mobilizations With Movement Techniques: The Immediate Effects on Range of Movement and Pain in Subjects with Low Back Pain. *J. Manip. Physiol. Ther.* **2007**, *30*, 178–185. [CrossRef]
37. Laird, R.A.; Kent, P.; Keating, J.L. How consistent are lordosis, range of movement and lumbo-pelvic rhythm in people with and without back pain? *BMC Musculoskelet. Disord.* **2016**, *17*, 403. [CrossRef]

38. Rose-Dulcina, K.; Genevay, S.; Dominguez, D.; Armand, S.; Vuillerme, N. Flexion-Relaxation Ratio Asymmetry and Its Relation with Trunk Lateral ROM in Individuals with and Without Chronic Nonspecific Low Back Pain. *Spine* **2020**, *45*, E1–E9. [CrossRef] [PubMed]
39. Williams, J.M.; Haq, I.; Lee, R.Y. A novel approach to the clinical evaluation of differential kinematics of the lumbar spine. *Man. Ther.* **2013**, *18*, 130–135. [CrossRef] [PubMed]
40. Chiarotto, A.; Boers, M.; Deyo, R.A.; Buchbinder, R.; Corbin, T.P.; Costa, L.O.; Foster, N.E.; Grotle, M.; Koes, B.W.; Kovacs, F.M.; et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. *Pain* **2018**, *159*, 481–495. [CrossRef] [PubMed]
41. Naqvi, W.M.; Vaidya, L.; Kumar, K. Impact of low back pain on fear of movement and functional activities. *Int. J. Res. Pharm. Sci.* **2020**, *11*, 4830–4835. [CrossRef]
42. Ippersiel, P.; Teoli, A.; Wideman, T.H.; Preuss, R.A.; Robbins, S.M. The relationship between pain-related threat and motor behavior in nonspecific low back pain: A systematic review and meta-analysis. *Phys. Ther.* **2022**, *102*, pzab274. [CrossRef]
43. Koppenhaver, S.; Gaffney, E.; Oates, A.; Eberle, L.; Young, B.; Hebert, J.; Proulx, L.; Shinohara, M. Lumbar muscle stiffness is different in individuals with low back pain than asymptomatic controls and is associated with pain and disability, but not common physical examination findings. *Musculoskelet. Sci. Pract.* **2020**, *45*, 102078. [CrossRef]
44. Russo, M.; Deckers, K.; Eldabe, S.; Kiesel, K.; Gilligan, C.; Viece, J.; Crosby, P. Muscle control and non-specific chronic low back pain. *Neuromodulation Technol. Neural Interface* **2018**, *21*, 1–9. [CrossRef]
45. Kett, A.R.; Sighting, F. Sedentary behaviour at work increases muscle stiffness of the back: Why roller massage has potential as an active break intervention. *Appl. Ergon.* **2020**, *82*, 102947. [CrossRef]
46. Hamburg, M.; Collins, F. The Path to Personalized Medicine. *Path Pers. Med.* **2010**, *363*, 301–304. [CrossRef]
47. Zhao, J. A Review of Wearable IMU (Inertial-Measurement-Unit)-based Pose Estimation and Drift Reduction Technologies. *J. Phys. Conf. Ser.* **2018**, *1087*, 042003. [CrossRef]
48. Deyo, R.A.; Tsui-Wu, Y.J. Descriptive Epidemiology of Low-Back Pain and Its Related Medical Care in the United States. *Spine* **1987**, *12*, 264–268. [CrossRef]
49. Ha, J.Y. Evaluation of Metabolic Syndrome in Patients with Chronic Low Back Pain: Using the Fourth Korea National Health and Nutrition Examination Survey Data. *Chonnam Med. J.* **2011**, *47*, 160. [CrossRef]
50. Kawi, J. Chronic low back pain patients' perceptions on self-management, self-management support, and functional ability. *Pain Manag. Nurs.* **2014**, *15*, 258–264. [CrossRef] [PubMed]
51. Noonan, A.; Brown, S. Paraspinal muscle pathophysiology associated with low back pain and spine degenerative disorders. *JOR Spine* **2021**, *4*, e1171. [CrossRef] [PubMed]
52. Tsuji, T.; Matsudaira, K.; Sato, H.; Vietri, J. The impact of depression among chronic low back pain patients in Japan. *BMC Musculoskelet. Disord.* **2016**, *17*, 447. [CrossRef] [PubMed]
53. La Touche, R.; Perez-Fernandez, M.; Barrera-Marchessi, I.; Lopez-De-Uralde-Villanueva, I.; Villafañe, J.H.; Prieto-Aldana, M.; Suso-Marti, L.; Paris-Aleman, A. Psychological and physical factors related to disability in chronic low back pain. *J. Back Musculoskelet. Rehabil.* **2019**, *32*, 603–611. [CrossRef]
54. Nzamba, J.; Van Damme, S.; Favre, J.; Christe, G. The relationships between spinal amplitude of movement, pain and disability in low back pain: A systematic review and meta-analysis. *Eur. J. Pain* **2023**, *28*, 37–53. [CrossRef]
55. Nordstoga, A.L.; Meisingset, I.; Vasseljen, O.; Nilsen, T.I.; Unsgaard-Tøndel, M. Longitudinal associations of kinematics and fear-avoidance beliefs with disability, work ability and pain intensity in persons with low back pain. *Musculoskelet. Sci. Pract.* **2019**, *41*, 49–54. [CrossRef]
56. Knoop, J.; Rutten, G.; Lever, C.; Leemeijer, J.; de Jong, L.J.; Verhagen, A.P.; van Lankveld, W.; Staal, J.B. *Lack of Consensus Across Clinical Guidelines Regarding the Role of Psychosocial Factors Within Low Back Pain Care: A Systematic Review*; Elsevier Inc.: Amsterdam, The Netherlands, 2021; Volume 22, pp. 1545–1559. [CrossRef]
57. Buruck, G.; Tomaschek, A.; Wendsche, J.; Ochsmann, E.; Dörfel, D. Psychosocial areas of worklife and chronic low back pain: A systematic review and meta-analysis. *BMC Musculoskelet. Disord.* **2019**, *20*, 480. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Article

Analytical Performance of the Factory-Calibrated Flash Glucose Monitoring System FreeStyle Libre2™ in Healthy Women

Zhuoxiu Jin ¹, Alice E. Thackray ^{1,2}, James A. King ^{1,2}, Kevin Deighton ³, Melanie J. Davies ^{2,4} and David J. Stensel ^{1,2,5,6,*}

¹ National Centre for Sport and Exercise Medicine, School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough LE11 3TU, UK; z.jin@lboro.ac.uk (Z.J.); a.e.thackray@lboro.ac.uk (A.E.T.); j.a.king@lboro.ac.uk (J.A.K.)

² National Institute for Health and Care Research (NIHR) Leicester Biomedical Research Centre, University Hospitals of Leicester National Health Service (NHS) Trust and the University of Leicester, Leicester LE1 5WW, UK; melanie.davies@uhl-tr.nhs.uk

³ Nuffield Health, Epsom, Surrey KT18 5AL, UK; kevin.deighton@nuffieldhealth.com

⁴ Diabetes Research Centre, University of Leicester, Leicester LE5 4PW, UK

⁵ Faculty of Sport Sciences, Waseda University, Tokorozawa 359-1192, Japan

⁶ Department of Sports Science and Physical Education, The Chinese University of Hong Kong, Ma Liu Shui, Hong Kong 999077, China

* Correspondence: d.j.stensel@lboro.ac.uk

Abstract: Continuous glucose monitoring (CGM) is used clinically and for research purposes to capture glycaemic profiles. The accuracy of CGM among healthy populations has not been widely assessed. This study assessed agreement between glucose concentrations obtained from venous plasma and from CGM (FreeStyle Libre2™, Abbott Diabetes Care, Witney, UK) in healthy women. Glucose concentrations were assessed after fasting and every 15 min after a standardized breakfast over a 4-h lab period. Accuracy of CGM was determined by Bland–Altman plot, 15/15% sensor agreement analysis, Clarke error grid analysis (EGA) and mean absolute relative difference (MARD). In all, 429 valid CGM readings with paired venous plasma glucose (VPG) values were obtained from 29 healthy women. Mean CGM readings were 1.14 mmol/L (95% CI: 0.97 to 1.30 mmol/L, $p < 0.001$) higher than VPG concentrations. Ratio 95% limits of agreement were from 0.68 to 2.20, and a proportional bias (slope: 0.22) was reported. Additionally, 45% of the CGM readings were within ± 0.83 mmol/L (± 15 mg/dL) or $\pm 15\%$ of VPG, while 85.3% were within EGA Zones A + B (clinically acceptable). MARD was 27.5% (95% CI: 20.8, 34.2%), with higher MARD values in the hypoglycaemia range and when VPG concentrations were falling. The FreeStyle Libre2™ CGM system tends to overestimate glucose concentrations compared to venous plasma samples in healthy women, especially during hypoglycaemia and during glycaemic swings.

Keywords: continuous glucose monitoring; blood glucose; agreement; glycaemic dynamics

Citation: Jin, Z.; Thackray, A.E.; King, J.A.; Deighton, K.; Davies, M.J.; Stensel, D.J. Analytical Performance of the Factory-Calibrated Flash Glucose Monitoring System FreeStyle Libre2™ in Healthy Women. *Sensors* **2023**, *23*, 7417. <https://doi.org/10.3390/s23177417>

Academic Editor: José Machado Da Silva

Received: 4 August 2023

Revised: 21 August 2023

Accepted: 22 August 2023

Published: 25 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Continuous glucose monitoring (CGM) devices automatically track interstitial glucose concentrations via a small sensor attached to the upper arm or abdomen. CGM is commonly used in clinical practice to help people with diabetes maintain healthy glycaemic management and prevent hypoglycaemic episodes. Studies have demonstrated high accuracy of CGM devices in individuals with diabetes [1,2], supporting the safety of CGM systems for informing treatment decisions in adults with type 1 diabetes or individuals with type 2 diabetes who are on insulin [3–5].

Along with clinical applications, CGM systems are also widely used for research [6–10] to provide comprehensive glycaemic profiles, including glucose nadir [7], glucose peak [9] and area under the curve [10]. These devices are often utilised in studies under free-living conditions, as they remove the barrier of collecting blood samples in the laboratory. However, given the limited studies conducted in healthy individuals [11,12] and the various types of CGM systems on the market [13], we are unable to draw firm conclusions on the accuracy of specific CGM systems. A study led by Akintola and colleagues [11] identified good agreement of the Medtronic ENLiTE CGM system with venous serum glucose concentrations in normoglycaemic participants. Another study also showed overall high accuracy of the Dexcom G6 CGM system compared with capillary blood glucose concentrations in healthy individuals [12].

The FreeStyle Libre™ CGM system is also commonly used for both clinical and research applications [8,14–16]. The FreeStyle Libre™ sensors are factory-calibrated, which reduces the likelihood of inaccuracies compared to user-calibrated devices such as the Medtronic ENLiTE and Dexcom. Some studies have used the first-generation FreeStyle Libre™ CGM system to calculate daily glycaemic variables in healthy adults [8,16] but its accuracy compared with venous blood measurements has not been established in this population. Since the second generation of the sensor (FreeStyle Libre2™) received CE mark clearance in the EU in 2018 and FDA approval in 2020, the accuracy of FreeStyle Libre2™ sensors has been tested in people with type 1 diabetes or type 2 diabetes only [17], but not in healthy individuals. At least one study has observed differences in glycaemic variables assessed by CGM between normoglycemic individuals and people with impaired glucose tolerance [18], showing higher daytime average interstitial glucose concentrations and higher postprandial peaks in a group with abnormal glucose tolerance. In addition, studies examining the accuracy of CGM systems compared with glucose measurements from venous blood have provided inconsistent findings [1,15]. Therefore, the present study compared glucose responses assessed using the FreeStyle Libre2™ CGM sensor with those obtained from venous plasma samples to investigate the accuracy of the sensor and its suitability for research purposes in healthy women.

2. Methods

2.1. Participants

Thirty healthy, normal-weight females provided written informed consent to participate in this study. One participant did not have sensor data; thus, data from 29 participants were subsequently integrated into the study analysis. Eligibility criteria were female, 18 to 35 years old; body mass index between 18.5 and 24.9 kg/m²; non-smokers (vaping was considered smoking); no known medical conditions; not taking any medications that might influence the study outcomes including those that may influence the CGM readings during the lab study period such as acetaminophen (paracetamol), acetylsalicylic acid (aspirin) or ascorbic acid (vitamin C); habitually consumed three meals a day; no clinically diagnosed eating disorders; not dieting and stable weight for 3 months before the study (<3 kg change in weight); no severe dislike or allergy to any of the study foods; and regular menstrual cycle in the past 6 months (those taking the oral contraceptive pill were not eligible for the study). Study procedures were approved by the Loughborough University Ethics Committee and complied with the Helsinki Declaration guidelines. Participants were informed of the purpose, procedures and potential risks in the study before providing written informed consent.

2.2. Study Design

The study was a single-arm trial consisting of a pre-assessment visit and a main trial. On the pre-assessment visit day, anthropometric parameters were measured using standard techniques. Body mass and height were measured on an integrated scale and stadiometer in light clothes and without shoes (Seca 285, Hamburg, Germany). Waist and hip circumferences were measured with a nonelastic flexible tape while standing (Seca

201, Hamburg, Germany). A CGM sensor (FreeStyle Libre 2™, Abbott Diabetes Care, Witney, UK) was then fitted. According to the manufacturer’s recommendations, the sensor was placed over the participant’s posterior upper arm. The arm on which the sensor was positioned was determined based on the participant’s preference. The pre-assessment visit was scheduled two days before the main trial. One day before the main trial, participants were required to refrain from any strenuous exercise, alcohol and caffeine. Participants were provided with the same Margherita pizza (Goodfella, Green Isle Foods Ltd., Co Kildare, Ireland, 3524 kJ, carbohydrate 90 g, protein 40 g, fat 34 g) and were instructed to consume as much of this as they desired in the evening and to avoid eating or drinking anything else except plain water from 10:00 pm. On the main trial day, participants arrived at the laboratory at 8:00 am after fasting overnight for at least 10 h. A venous cannula (Venflon 20 G/32 mm, BOC Ohmeda, Sweden) was inserted into an antecubital vein on the opposite arm to the CGM sensor. Participants rested for 30 min and then consumed a fixed breakfast within 15 min, consisting of two slices of white bread (100 g), a bowl of corn flakes (15 g), bananas (150 g, weighed with the skin), strawberry jam (15 g) and semi-skimmed milk (200 g) (2121 kJ, carbohydrate 95.2 g, protein 18.2 g, fat 5.6 g). A timer was started once participants finished the breakfast (t = 0 min). Throughout the testing period, participants remained sedentary within the laboratory, leaving them free to read, watch videos or use a computer. The study protocol is shown in Figure 1.

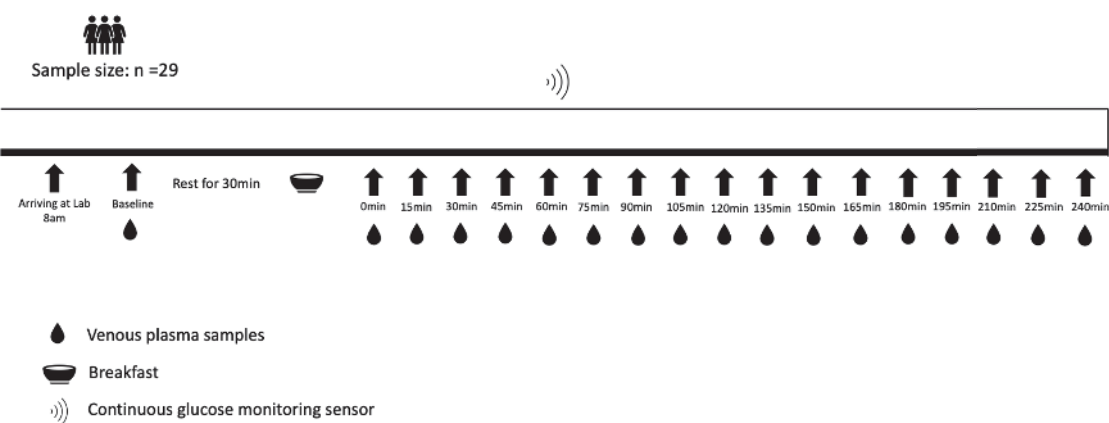


Figure 1. Study protocol.

2.3. FreeStyle Libre2™ CGM Sensor

The current study used a sensor-based flash glucose monitoring system (FreeStyle Libre2™; Abbott Diabetes Care, Witney, UK). This factory-calibrated sensor continuously monitors interstitial glucose concentration utilising wired enzyme technology (osmium mediator and glucose oxidase enzyme co-immobilized on an electrochemical sensor). Real-time glucose concentrations were obtained by scanning the sensor every 15 min via the Librelink application on participants’ smartphones. A glucose trend arrow (indicating rate and direction of change in glucose concentration) was also displayed on the screen. The trend arrows are categorised into rising quickly (increasing > 0.111 mmol/L/min), rising (increasing 0.056–0.111 mmol/L/min), changing slowly (not increasing/decreasing > 0.056 mmol/L/min), falling (decreasing 0.056–0.111 mmol/L/min) and falling quickly (decreasing > 0.111 mmol/L/min) [19].

2.4. Blood Samples

Venous blood samples were collected via cannula at baseline (before breakfast), immediately after breakfast ($t = 0$ min) and every 15 min from $t = 0$ min until 240 min for venous plasma glucose (VPG) measurement. Venous blood was drawn into pre-chilled 4.9 mL K3 EDTA tubes using the Sarstedt S-Monovette® system (Sarstedt, Nümbrecht, Germany). These blood samples were then centrifuged at 3500 rpm ($2054 \times g$) for 10 min (Heraeus Labofuge 400R, Thermo Fisher Scientific, Waltham, MA, USA). The separated plasma was stored in a -80 °C freezer (TwinGuard, Panasonic, Kadoma, Japan) until biochemical analysis. VPG concentrations were assayed using the ABX Pentra Glucose PAP CP kit by colorimetry on a Pentra C400 clinical chemistry analyser (HORIBA Medical, Montpellier, France), with an intraassay coefficient of variation of 0.5%.

2.5. Statistical Analyses

A linear mixed model was used to compare the mean differences between venous plasma glucose concentrations and CGM readings, and Brysbaert and Stevens's method [20] was used to calculate the effect size. Analyses were conducted using the 'lme4' package for linear mixed-effects models in R (version 4.0.5). The agreement between glucose measurements derived from CGM and venous plasma samples was assessed using Bland–Altman analysis [21]. The proportional bias was calculated using linear mixed models.

We also applied the 15/15% sensor agreement analysis (the proportion of CGM readings within 15% of the reference VPG values ≥ 5.6 mmol/L or within 15 mg/dL (0.83 mmol/L) of the reference VPG values < 5.6 mmol/L) and the Clarke error grid analysis (EGA) to measure the analytical accuracy of CGM systems according to the International Organization for Standardization criteria (ISO 15197:2013) [22]. The EGA was performed using the 'ega' package in R (version 4.0.5). In EGA, the diagonal represents perfect agreement between the CGM and reference VPG pairs, whereas the points below and above the line indicate under- and over-estimation of the reference VPG concentrations, respectively. The data points in zone A represent the CGM values that differ by less than 20% from the VPG reference pairs or are < 3.5 mmol/L, given that the reference VPG values are also < 3.5 mmol/L (indicating hypoglycaemia). The current study involved a healthy population; therefore, we customised the EGA hypoglycaemia cut-off value at 3.5 mmol/L (rather than 3.9 mmol/L) as recommended for people who do not have diabetes [23]. Readings in Zone A are typically considered safe for clinical decision-making. Zone B represents CGM values that deviate from the reference VPG values by more than 20% but would lead to benign or no treatment, thus considered acceptable but necessitating closer monitoring. Values falling within zone C may lead to overcorrect acceptable VPG values, potentially causing the actual VPG values to deviate from the target range of 3.5–10 mmol/L. Data points in zone D would result in failure to detect either hypoglycaemia (< 3.5 mmol/L) or hyperglycaemia (> 10 mmol/L). Zone E represents readings that are physiologically implausible and beyond clinical plausibility.

Additionally, a linear mixed model was used to calculate the mean absolute relative difference (MARD) for all paired measurements to determine the accuracy of the CGM sensor. The MARD was calculated using the following formula: mean [absolute value (CGM readings–VPG concentrations)/VPG concentrations], as previously performed in other studies [21,22,24]. We further investigated the MARD in different glycaemic ranges and during different rates of change based on the VPG concentrations, as high accuracy at rapid rates of blood glucose change is critical to capture important glycaemic variabilities. The rates of change in our study were calculated according to a previously used formula [25]:

$$rate = (VPG_i - VPG_{i-1}) / (T_i - T_{i-1})$$

Above, T_i and T_{i-1} are the timepoints of the i th and $(i-1)$ th venous blood sample, respectively, and VPG_i and VPG_{i-1} are the VPG concentrations corresponding to their related timestamps. We calculated Pearson's correlation coefficients (r) between CGM and VPG using a multilevel approach and calculated each participant's correlation coefficient to determine individual differences in the agreement of CGM readings and VPG levels. Data are presented as mean (standard deviation) or mean (confidence intervals), as appropriate. Statistical analysis and graphs were performed using R version 4.0.5 software. The level of significance was set at $p < 0.05$.

3. Results

Sensor readings from the CGM could not be recorded for one participant due to a technical issue, so results are presented for 29 individuals. The characteristics of the participants are summarised in Table 1. Our dataset comprised 429 valid CGM readings with paired VPG measurements.

Table 1. Characteristics of participants ($n = 29$ women).

	Mean $\pm$ SD
Age (years)	26 $\pm$ 4
Height (m)	1.66 $\pm$ 0.07
Body mass (kg)	60.0 $\pm$ 7.2
Body mass index (kg/m ²)	21.5 $\pm$ 2.0
Waist circumference (cm)	69.9 $\pm$ 5.3
Hip circumference (cm)	95.7 $\pm$ 4.1
Ethnicity	
Asian	22
White	4
Arab	2
Latino	1

Figure 2 shows meal-related changes in glucose concentrations measured by CGM and VPG. The mean CGM sensor readings were 1.14 mmol/L higher than VPG concentrations ($p < 0.001$), with the 95% confidence interval ranging from 0.97 to 1.30 mmol/L and a large effect size $d = 0.83$.

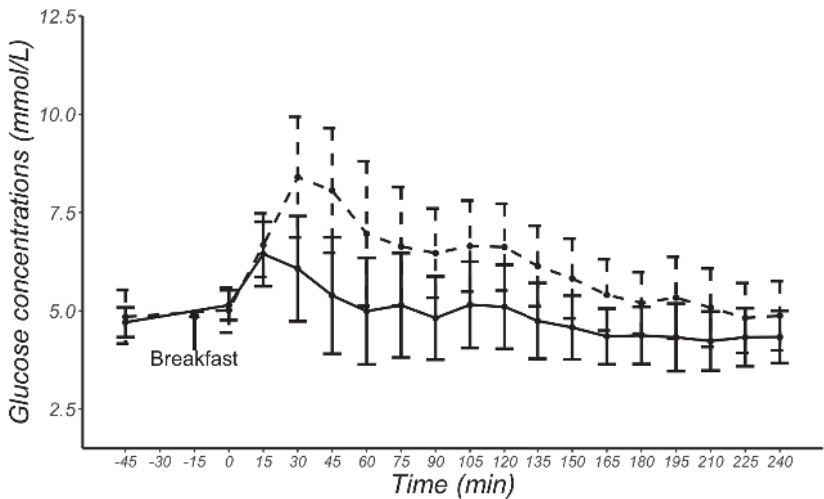


Figure 2. Mean and standard deviation glucose concentrations at each time point obtained from continuous glucose monitoring sensors (dashed line) and from venous plasma glucose samples (solid line) in 29 female participants.

A Bland–Altman plot of the raw individual glucose measurements is presented in Figure 3a, with a systematic bias of 1.14 mmol/L (95% CIs: 1.01, 1.26 mmol/L) and the 95% LoA ranging from -1.44 to 3.72 mmol/L. According to previous published statistical methods [21,26], we also used a natural logarithmic transformation to mitigate the heteroscedasticity. The transformed data are presented in Figure 3b, which shows a systematic bias of 0.20 (95% CIs: 0.17 , 0.23), and the 95% LoA are between -0.38 to 0.79 on a log scale. The ratio 95% limits of agreement range from 0.68 to 2.20 . This means that, in 95% of cases, the CGM readings are between 0.68 and 2.20 times the VPG concentrations. The plot displayed in Figure 3b demonstrates a proportional bias with increasing mean differences between the glucose concentrations from CGM and venous plasma as the mean of the two measurements increases.

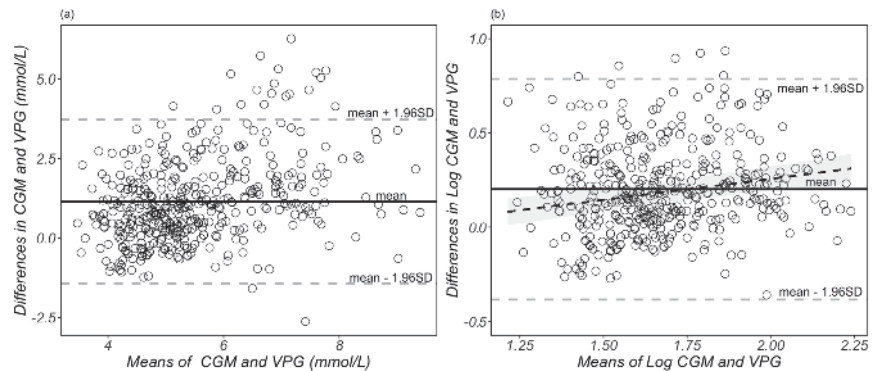


Figure 3. Bland–Altman plots of glucose measurements from raw (a) and natural log-transformed (b) data. Each dot represents a paired (continuous glucose monitoring and venous plasma) glucose measurement ($n = 429$ data points derived from 29 participants). The bias of the measurements and limits of agreement are represented as the solid black lines and grey dotted lines, respectively. The black dotted line and shaded grey area in (b) represents the proportional bias (slope = 0.22) and the 95% confidential intervals. CGM, continuous glucose monitoring; VPG, venous blood plasma glucose concentrations.

According to the 15/15% agreement analysis, 45% of CGM readings were within ± 0.83 mmol/L or $\pm 15\%$ of reference VPG concentrations. Furthermore, 56% of CGM readings were within ± 20 mg/dL (± 1.11 mmol/L) or $\pm 20\%$ of VPG concentrations, while 74% of CGM readings were within ± 30 mg/dL (± 1.67 mmol/L) or $\pm 30\%$ of the VPG concentrations (Figure 4).

The results obtained by EGA were as follows: 48.0% in zone A, 45.7% in zone B, 6.3% in zone D and no values in zones C and E (Figure 5). All the data points that fell into zone D suggest that CGM failed to detect blood glucose values < 3.5 mmol/L. This study recruited healthy individuals only; in this case, the data that fell into zone D indicate that CGM tends to overestimate the VPG concentrations. This could result in instances of hypoglycaemia going undetected in healthy individuals monitoring their blood glucose by CGM.

The overall MARD value for the FreeStyle Libre2™ CGM system was 27.5% (95% CI: 20.8, 34.2%). The MARD values across different glycaemic zones varied, with the MARD value being elevated for those in the hypoglycaemic range (< 3.5 mmol/L). Larger MARD values were also apparent when the rate of change in plasma glucose concentrations was categorised as falling and falling quickly (Table 2).

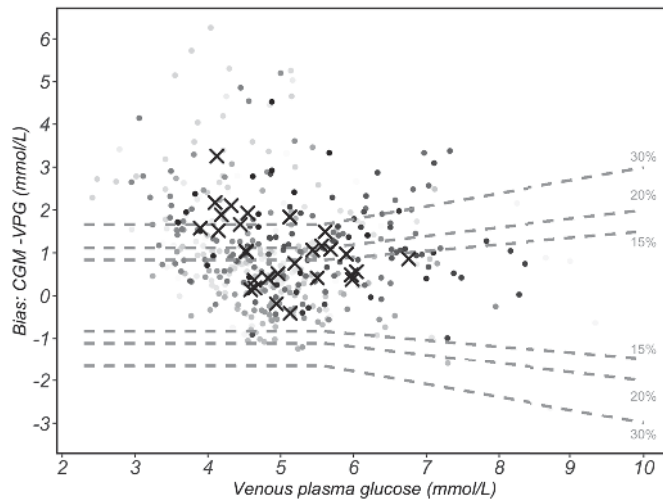


Figure 4. The 15/15% sensor agreement analysis. The 15% grey dotted lines form the area that continuous glucose monitoring (CGM) values differ by less than 15% of the paired and venous plasma glucose (VPG) reference values when VPG values are ≥ 5.6 mmol/L or differ by less than 15 mg/dL (0.83 mmol/L) of VPG values when VPG are < 5.6 mmol/L. These lines also indicate the 20%/20 mg/dL (1.1 mmol/L) and 30%/30 mg/dL (1.67 mmol/L) agreement rates. Circles with varied grey shading represent each of the 429 paired measurements. Circles with same grey shade belong to the same participant. Black crosses represent the mean value of each participant ($n = 29$).

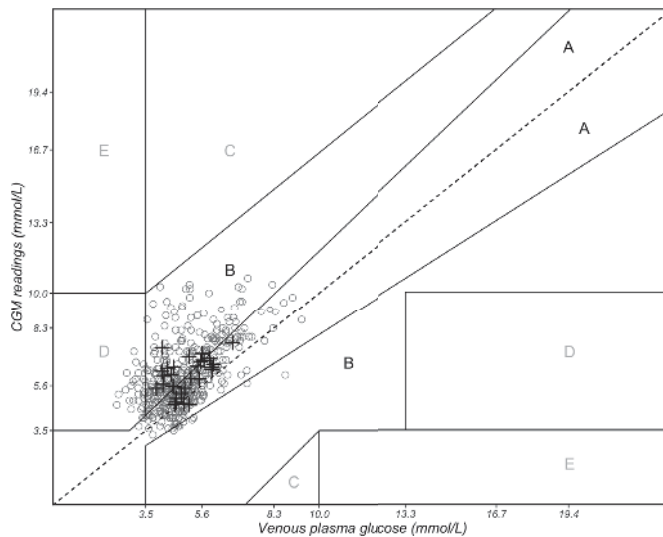


Figure 5. Clarke error grid analysis of paired continuous glucose monitoring (CGM) and venous plasma glucose (VPG) measurements. Zone A represents CGM values that differ by less than 20% of the VPG reference values or are < 3.5 mmol/L if the reference values are also < 3.5 mmol/L (indicating hypoglycaemia). Zone B represents CGM values that deviate from the reference VPG concentrations by $> 20\%$ but would result in benign or no treatment. Zones A and B are clinically acceptable, while values in zones C, D and E are potentially dangerous and, therefore, clinically significant. The solid circles represent the 429 paired measurements, and the black crosses represent the mean value of each of the 29 participants.

Table 2. Mean absolute relative difference (MARD) for 429 paired CGM and venous plasma glucose measurements obtained in 29 women.

	MARD (95% CI)
Overall (<i>n</i> = 429)	27.5% (20.8, 34.2%)
Glycaemic zones	
Hypoglycaemia (<3.5 mmol/L; <i>n</i> = 29)	69.5% (51.4, 82.9%)
Euglycaemia (3.5–10.0 mmol/L; <i>n</i> = 400)	25.9% (19.9, 31.8%)
Rate of changes	
Rising quickly (>0.111 mmol/L/min; <i>n</i> = 12)	14.2% (9.6, 18.8%)
Rising (0.056–0.111 mmol/L/min; <i>n</i> = 34)	13.8% (9.9, 17.8%)
Changing slowly (<0.056 mmol/L/min; <i>n</i> = 313)	28.4% (21.2, 35.6%)
Falling (0.056–0.111 mmol/L/min; <i>n</i> = 26)	63.7% (45.7, 81.7%)
Falling quickly (>0.111 mmol/L/min; <i>n</i> = 8)	67.3% (33.9, 100%)

Abbreviations: MARD, mean absolute relative difference; CIs, confidence intervals.

There was an average correlation coefficient (*r* value) of 0.63 (95% CI: 0.57, 0.68) between CGM readings and VPG concentrations using the multilevel method (*p* < 0.001). The individual correlation coefficients between CGM readings and VPG concentrations were positive for 28 participants, ranging from 0.20 to 0.96, whilst a negative correlation coefficient was identified for one participant (*r* = −0.24) (Figure 6). The individual glycaemic patterns exhibited a time lag between the two measurements, with the glucose peak appearing earlier with VPG than CGM in 25 participants (Figure 7).

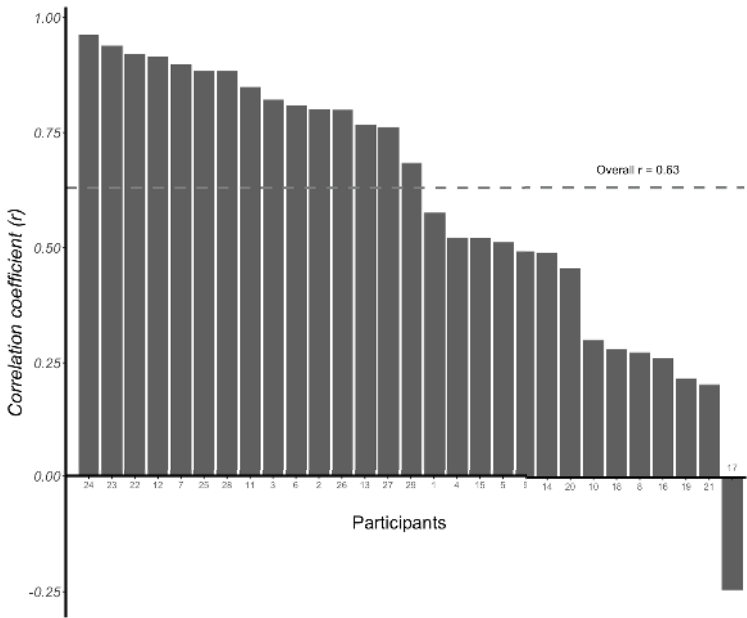


Figure 6. Individual correlation coefficients (*r* values) between venous plasma glucose concentrations and continuous glucose monitoring (CGM) readings determined for each of the 29 participants. Overall correlation coefficient *r* = 0.63 (95% CI: 0.57, 0.68).

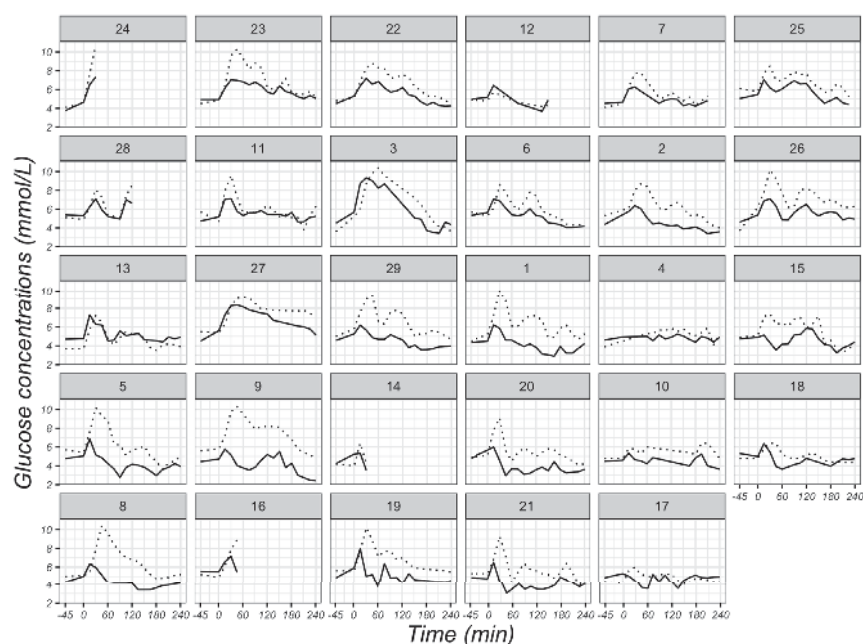


Figure 7. Individual glucose concentrations at different time points derived from venous plasma (solid lines) and CGM sensors (dotted lines) in 29 participants. Glucose concentrations were measured fasted at baseline (−45 min), immediately after breakfast consumption (0 min) and at 15-min intervals thereafter in the postprandial state until 240 min after breakfast consumption. The number of the individual plot represents the ID of each participant, and the order of plots is dependent on the magnitude of the r value from each participant in Figure 6.

4. Discussion

In this group of young, healthy, female participants, postprandial glucose concentrations determined by CGM were higher than VPG concentrations. There was good agreement between CGM and VPG concentrations when participants exhibited steady glucose concentrations, but agreement was poor during hypoglycaemia (<3.5 mmol/L) and during decreases in glucose.

CGM measures glucose concentrations from subcutaneous interstitial fluid, not blood, for which values are determined by the rate of glucose diffusion from plasma into the interstitial fluid and the rate of glucose uptake by subcutaneous tissue cells [27]. Thus, factors affecting cellular metabolism, glucose delivery and capillary permeability may alter interstitial glucose concentrations. In this assessment of 29 healthy individuals, glucose concentrations obtained from CGMs were higher than those measured in venous plasma, which is in line with findings reported previously [11]. Specifically, Akintola and colleagues investigated the accuracy of the Medtronic ENLiTE CGM system in participants with normoglycaemia and reported higher concentrations of glucose obtained via CGM than those obtained from venous serum samples during the day. Conversely, this study demonstrated lower CGM readings compared with paired VPG concentrations during night-time. Glucose concentrations in our study were assessed using CGM and venous sampling from the cubital vein, rather than arterial or capillary sources. After meal ingestion, blood glucose is absorbed by the small intestine, subsequently entering the bloodstream, and then diffusing from capillaries to interstitial fluid. After this exchange, the blood, which still carries some remaining glucose, flows into the veins. Therefore, in individuals with healthy cellular function, the CGM would be expected to elicit higher values than VPG as glucose enters the interstitial fluid prior to its journey towards the veins. However, this paradigm can vary

notably in individuals with diabetes. Studies recruiting people with diabetes or prediabetes report higher blood glucose concentrations derived from venous blood than those obtained from CGM devices [28,29]. For those with diabetes, insulin levels which are insufficient to facilitate the uptake of glucose into cells and/or insulin resistance may result in higher values in VPG. This suggests that there might be differences in glucose metabolism between people with and without diabetes. For healthy individuals, the overestimation of CGM may lead to an erroneous perception of having hyperglycaemia.

We used the Bland–Altman method to assess the agreement between glucose concentrations obtained from a CGM sensor and VPG samples. Our study demonstrated a proportional bias between the mean and mean difference values of CGM and VPG, consistent with previous results in healthy adults [30], suggesting a greater degree of inaccuracy at higher circulating glucose concentrations.

There are no established guidelines for evaluating the measurement performance of CGM systems. The accuracy criteria currently used for CGM systems are identical to those for self-monitoring blood glucose systems (SMBG), including statistics calculated from paired references and sensor glucose determinations [31]. The International Organization for Standardization (ISO) is a global association of national standardisation bodies, and the ISO 15197:2013 document has specified two requirements for SMBG [22]: (1) at least 95% of SMBG results must be within 0.83 mmol/L of comparators of <5.6 mmol/L or within 15% of comparator glucose values ≥ 5.6 mmol/L (15/15% agreement analysis), compared to a traceable laboratory method; (2) at least 99% of results must fall within zones A and B in an EGA. The current study found that only 45% of the total CGM readings were within ± 15 mg/dL (± 0.83 mmol/L) or $\pm 15\%$ of venous reference values, 56% were within ± 20 mg/dL (± 1.11 mmol/L) or $\pm 20\%$ of venous reference values and 74% were within ± 30 mg/dL (± 1.67 mmol/L) or $\pm 30\%$ of venous reference values. The EGA shows that 85% of the CGM readings were in zones A and B. Altogether, these data indicate suboptimal accuracy of CGM sensors compared to glucose concentrations measured from venous samples.

The MARD parameter is also commonly used to characterise the analytical performance in CGM systems. A round table conference held in 2013 relating to CGM reached a consensus that MARD values should be less than 14% and values higher than 18% represent poor accuracy [32]. Our study identified an overall 27.5% (95% CIs: 20.8, 34.2%) MARD value of the FreeStyle Libre2TM sensor, indicative of poor accuracy. A pooled analysis of data from previous studies using the first generation of the FreeStyle LibreTM CGM devices identified an overall median absolute relative difference value (MedARD) of 12.7% (IQR 5.9–23.5%) in individuals with type 1 diabetes [15].

When calculating the MARD values in different glucose ranges, we found that MARD values were much higher in the hypoglycaemia range (<3.5 mmol/L) compared with the euglycaemia range (3.5–10 mmol/L). Due to the limited sample size and the recruitment of healthy individuals, we did not detect any hyperglycaemic values. One study has reported lower MedARD values in hyperglycaemia zones in people with type 1 diabetes [15]. Therefore, our results combined with findings from previous work, suggest that the MARD produces more relative errors during hypoglycaemia. The present study revealed that the MARD was elevated when VPG concentrations were falling. This is consistent with the findings of a study observing higher MedARD values when sensor glucose concentrations were increasing or decreasing [25]. This indicates that CGM devices may not provide accurate glucose concentrations during rapid glucose swings such as in response to food intake.

The collective findings of our study suggest a discrepancy of glucose concentrations between CGM and VPG, especially in the hypoglycaemic zone (<3.5 mmol/L) and when blood glucose levels are falling. An explanation for this discrepancy could be the physiological differences between blood and interstitial fluid. Generally, glucose concentrations in blood and interstitial fluid are similar under steady-state conditions [25,33], which is in line with our results when participants fasted. However, in the case of swings in glycaemia

(defined as a rate of change > 0.055 mmol/L/min), there is a physiological time delay, with previous work suggesting that the FreeStyle Libre2™ sensor exhibits a 2-min time lag [17]. This time delay is also displayed in Figure 7, which shows that peaks in VPG were generally lower and occurred sooner than peaks identified by CGM. The higher MARD value while VPG is falling could partly be explained by the time lag.

A small sample size and high variability between individuals are limitations in our study. The data were constrained practically and financially on the number of participants that could be recruited. There are known physiological differences between the sexes in glucose homeostasis, with women tending to have higher levels of insulin sensitivity compared to men [34,35]. Therefore, we restricted our recruitment to a single sex to avoid any confounding influence of sex on the study outcomes. For all the women included in the study, we calculated each individual's correlation coefficient, which indicates the variance in the alignment of the participant's intrinsic physiology with the CGM calibration algorithm. The extent of this inter-individual variability depends on many factors. We standardised metabolism by discouraging vigorous physical activity and prescribing a standardised evening meal the evening before the main trial. However, we could not standardise glucose concentration in different tissues in individuals to minimise the physiological time lag during rapid changes in blood glucose, as glucose is transferred from the capillary endothelium to the interstitial fluid by simple diffusion down a concentration gradient. In addition, insertion of the sensor could cause trauma at the site, disrupting tissue structure and causing an inflammatory reaction that consumes glucose. Several participants reported discomfort when the sensor was inserted, and several participants removed the sensor immediately at the end of the trial. Moreover, the implanted glucose sensor could have been placed far from a blood vessel, which may cause an extended delay in the interchange between interstitial fluid and venous blood [36]. Future studies may be improved by employing a larger sample size of healthy participants, assessing glucose in different situations (e.g., fasted, fed, exercising) and measuring glucose over multiple days to help confirm and improve the accuracy of CGM systems.

5. Conclusions

Our study reports suboptimal accuracy of the FreeStyle Libre 2™ CGM sensor for measuring glucose concentrations compared to values obtained from venous plasma samples in young, healthy women, especially during hypoglycaemia and during glycaemic swings. A correction may need to be applied by the manufacturer and/or researcher(s) if glucose data from the CGM are compared to other compartments such as arterial, venous or capillary glucose concentrations. For research purposes, if a study design contains a non-euglycemic range or a rapid change in blood glucose concentrations, CGM devices may not accurately capture blood glucose concentrations during such periods. Since there are irremediable differences between CGM and VPG, as they present glucose concentrations in two different compartments, it would be worthwhile to consider which one is more related to pertinent physiological activities.

Author Contributions: Conceptualisation, Z.J., A.E.T., M.J.D. and D.J.S.; methodology, Z.J., A.E.T., J.A.K. and D.J.S.; formal analysis, Z.J. and K.D.; investigation, Z.J. and D.J.S.; data curation, Z.J.; writing—original draft preparation, Z.J.; writing—review and editing, Z.J., A.E.T., J.A.K., K.D., M.J.D. and D.J.S.; visualisation, Z.J. and K.D.; supervision, A.E.T., J.A.K. and D.J.S.; project administration, D.J.S.; funding acquisition, D.J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Institute for Health and Care Research (NIHR) Leicester Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. Zhuoxiu Jin is funded by the China Scholarship Council.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Loughborough University (reference number: 2021-5828-5026 and date of approval: 10 September 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are accessible through the Loughborough University Research Repository. https://repository.lboro.ac.uk/articles/dataset/Continuous_glucose_monitor_CGM_and_venous_plasma_glucose_VPG_data_for_Exploration_of_the_correlation_between_glucose_dynamics_and_energy_intake_in_women/23814873.

Acknowledgments: The authors thank the participants for their participation, and Tristan Boetti, Abi Woodliffe-Thomas and Josh Hart for their assistance with data collection.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CGM	Continuous Glucose Monitoring
VPG	Venous Plasma Glucose
EGA	Error Grid Analysis
MARD	Mean Absolute Relative Difference
ISO	International Organization for Standardization

References

- Ji, L.; Guo, L.; Zhang, J.; Li, Y.; Chen, Z. Multicenter Evaluation Study Comparing a New Factory-Calibrated Real-Time Continuous Glucose Monitoring System to Existing Flash Glucose Monitoring System. *J. Diabetes Sci. Technol.* **2021**, *17*, 208–213. [CrossRef] [PubMed]
- Moser, O.; Mader, J.K.; Tschakert, G.; Mueller, A.; Groeschl, W.; Pieber, T.R.; Koehler, G.; Messerschmidt, J.; Hofmann, P. Accuracy of continuous glucose monitoring (CGM) during continuous and high-intensity interval exercise in patients with type 1 diabetes mellitus. *Nutrients* **2016**, *8*, 489. [CrossRef]
- Beck, S.E.; Kelly, C.; Price, D.A.; Aronoff, S.; Bao, S.; Bhargava, A.; Biggs, W.; Billings, L.; Blevins, T.; Bode, B.W.; et al. Non-adjunctive continuous glucose monitoring for control of hypoglycaemia (COACH): Results of a post-approval observational study. *Diabet. Med.* **2022**, *39*, e14739. [CrossRef] [PubMed]
- Elbalsky, M.; Haszard, J.; Smith, H.; Kuroko, S.; Galland, B.; Oliver, N.; Shah, V.; de Bock, M.I.; Wheeler, B.J. Effect of divergent continuous glucose monitoring technologies on glycaemic control in type 1 diabetes mellitus: A systematic review and meta-analysis of randomised controlled trials. *Diabet. Med.* **2022**, *39*, e14854. [CrossRef] [PubMed]
- Forlenza, G.P.; Argento, N.B.; Laffel, L.M. Practical considerations on the use of continuous glucose monitoring in pediatrics and older adults and nonadjunctive use. *Diabet. Technol. Ther.* **2017**, *19*, S-13–S-20. [CrossRef] [PubMed]
- Kanamori, K.; Ihana-Sugiyama, N.; Yamamoto-Honda, R.; Nakamura, T.; Sobe, C.; Kamiya, S.; Kishimoto, M.; Kajio, H.; Kawano, K.; Noda, M. Postprandial glucose surges after extremely low carbohydrate diet in healthy adults. *Tohoku J. Exp. Med.* **2017**, *243*, 35–39. [CrossRef] [PubMed]
- Kim, J.; Lam, W.; Wang, Q.; Parikh, L.; Elshafie, A.; Sanchez-Rangel, E.; Schmidt, C.; Li, F.; Hwang, J.; Belfort-DeAguiar, R. In a Free-Living Setting, Obesity Is Associated with Greater Food Intake in Response to a Similar Premeal Glucose Nadir. *J. Clin. Endocr.* **2019**, *104*, 3911–3919. [CrossRef]
- Wyatt, P.; Berry, S.E.; Finlayson, G.; O’driscoll, R.; Hadjigeorgiou, G.; Drew, D.A.; Al Khatib, H.; Nguyen, L.H.; Linenberg, I.; Chan, A.T.; et al. Postprandial glycaemic dips predict appetite and energy intake in healthy individuals. *Nat. Metab.* **2021**, *3*, 523–529. [CrossRef]
- Camps, S.G.; Kaur, B.; Lim, J.; Loo, Y.T.; Pang, E.; Ng, T.; Henry, C.J. Improved Glycemic Control and Variability: Application of Healthy Ingredients in Asian Staples. *Nutrients* **2021**, *13*, 3102. [CrossRef]
- Lagerpusch, M.; Enderle, J.; Later, W.; Eggeling, B.; Pape, D.; Müller, M.J.; Bosy-Westphal, A. Impact of glycaemic index and dietary fibre on insulin sensitivity during the refeeding phase of a weight cycle in young healthy men. *Br. J. Nutr.* **2013**, *109*, 1606–1616. [CrossRef]
- Akintola, A.A.; Noordam, R.; Jansen, S.W.; de Craen, A.J.; Ballieux, B.E.; Cobbaert, C.M.; Mooijaart, S.P.; Pijl, H.; Westendorp, R.G.; van Heemst, D. Accuracy of continuous glucose monitoring measurements in normo-glycemic individuals. *PLoS ONE* **2015**, *10*, e0139973. [CrossRef] [PubMed]

12. Schierbauer, J.R.; Günther, S.; Haupt, S.; Zimmer, R.T.; Zunner, B.E.M.; Zimmermann, P.; Wachsmuth, N.B.; Eckstein, M.L.; Aberer, F.; Sourij, H.; et al. Accuracy of Real Time Continuous Glucose Monitoring during Different Liquid Solution Challenges in Healthy Adults: A Randomized Controlled Cross-Over Trial. *Sensors* **2022**, *22*, 3104. [CrossRef] [PubMed]
13. Heinemann, L.; Schoemaker, M.; Schmelzeisen-Redecker, G.; Hinzmann, R.; Kassab, A.; Freckmann, G.; Reiterer, F.; Del Re, L. Benefits and limitations of MARD as a performance parameter for continuous glucose monitoring in the interstitial space. *J. Diabetes Sci. Technol.* **2020**, *14*, 135–150. [CrossRef] [PubMed]
14. Price, C.; Ditton, G.; Russell, G.B.; Aloï, J. Reliability of inpatient CGM: Comparison to standard of care. *J. Diabetes Sci. Technol.* **2023**, *17*, 329–335. [CrossRef]
15. Moser, O.; Sternad, C.; Eckstein, M.L.; Szadkowska, A.; Michalak, A.; Mader, J.K.; Ziko, H.; Elsayed, H.; Aberer, F.; Sola-Gazagnes, A.; et al. Performance of intermittently scanned continuous glucose monitoring systems in people with type 1 diabetes: A pooled analysis. *Diabetes Obes. Metab.* **2022**, *24*, 522–529. [CrossRef]
16. Tsereteli, N.; Vallat, R.; Fernandez-Tajes, J.; Delahanty, L.M.; Ordoas, J.M.; Drew, D.A.; Valdes, A.M.; Segata, N.; Chan, A.T.; Wolf, J.; et al. Impact of insufficient sleep on dysregulated blood glucose control under standardised meal conditions. *Diabetologia* **2022**, *65*, 356–365. [CrossRef]
17. Alva, S.; Bailey, T.; Brazg, R.; Budiman, E.S.; Castorino, K.; Christiansen, M.P.; Forlenza, G.; Kipnes, M.; Liljenquist, D.R.; Liu, H. Accuracy of a 14-day factory calibrated continuous glucose monitoring system with advanced algorithm in pediatric and adult population with diabetes. *J. Diabetes Sci. Technol.* **2022**, *16*, 70–77. [CrossRef]
18. Hanefeld, M.; Sulk, S.; Helbig, M.; Thomas, A.; Köhler, C. Differences in glycemic variability between normoglycemic and prediabetic subjects. *J. Diabetes Sci. Technol.* **2014**, *8*, 286–290. [CrossRef]
19. Kudva, Y.C.; Ahmann, A.J.; Bergenstal, R.M.; Gavin, J.R.; Kruger, D.F.; Midyett, L.K.; Miller, E.; Harris, D.R. Approach to using trend arrows in the FreeStyle Libre flash glucose monitoring systems in adults. *J. Endocr. Soc.* **2018**, *2*, 1320–1337. [CrossRef]
20. Brysbaert, M.; Stevens, M. Power analysis and effect size in mixed effects models: A tutorial. *J. Cogn.* **2018**, *1*, 9. [CrossRef]
21. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [CrossRef]
22. Jendrike, N.; Baumstark, A.; Kamecke, U.; Haug, C.; Freckmann, G. ISO 15197: 2013 Evaluation of a Blood Glucose Monitoring System's Measurement Accuracy. *J. Diabetes Sci. Technol.* **2017**, *11*, 1275–1276. [CrossRef] [PubMed]
23. Frier, B.M. Defining hypoglycaemia: What level has clinical relevance? *Diabetologia* **2009**, *52*, 31–34. [CrossRef]
24. Obermaier, K.; Schmelzeisen-Redecker, G.; Schoemaker, M.; Klötzer, H.-M.; Kirchsteiger, H.; Eikmeier, H.; del Re, L. Performance evaluations of continuous glucose monitoring systems: Precision absolute relative deviation is part of the assessment. *J. Diabetes Sci. Technol.* **2013**, *7*, 824–832. [CrossRef]
25. Pleus, S.; Schoemaker, M.; Morgenstern, K.; Schmelzeisen-Redecker, G.; Haug, C.; Link, M.; Zschornack, E.; Freckmann, G. Rate-of-change dependence of the performance of two CGM systems during induced glucose swings. *J. Diabetes Sci. Technol.* **2015**, *9*, 801–807. [CrossRef] [PubMed]
26. Nevill, A.M.; Atkinson, G. Assessing agreement between measurements recorded on a ratio scale in sports medicine and sports science. *Br. J. Sports Med.* **1997**, *31*, 314–318. [CrossRef] [PubMed]
27. Cengiz, E.; Tamborlane, W.V. A tale of two compartments: Interstitial versus blood glucose monitoring. *Diabetes Technol. Ther.* **2009**, *11*, S-11–S-16. [CrossRef]
28. Sato, T.; Oshima, H.; Nakata, K.; Kimura, Y.; Yano, T.; Furuhashi, M.; Tannno, M.; Miki, T.; Miura, T. Accuracy of flash glucose monitoring in insulin-treated patients with type 2 diabetes. *J. Diabetes Investig.* **2019**, *10*, 846–850. [CrossRef]
29. Faerch, K.; Amadid, H.; Bruhn, L.; Clemmensen, K.K.B.; Hulman, A.; Ried-Larsen, M.; Blond, M.B.; Jørgensen, M.E.; Vistisen, D. Discordance between Glucose Levels Measured in Interstitial Fluid vs in Venous Plasma after Oral Glucose Administration: A Post-Hoc Analysis from the Randomised Controlled PRED Trial. *Front. Endocrinol.* **2021**, *12*, 1–8. [CrossRef]
30. Dye, L.; Mansfield, M.; Lasikiewicz, N.; Mahawish, L.; Schnell, R.; Talbot, D.; Chauhan, H.; Croden, F.; Lawton, C. Correspondence of continuous interstitial glucose measurement against arterialed and capillary glucose following an oral glucose tolerance test in healthy volunteers. *Br. J. Nutr.* **2010**, *103*, 134–140. [CrossRef]
31. Clarke, W.L.; Kovatchev, B. Continuous glucose sensors: Continuing questions about clinical accuracy. *J. Diabetes Sci. Technol.* **2007**, *1*, 669–675. [CrossRef] [PubMed]
32. Wernerman, J.; Desai, T.; Finfer, S.; Foubert, L.; Furnary, A.; Holzinger, U.; Hovorka, R.; Joseph, J.; Kosiborod, M.; Krinsley, J.; et al. Continuous glucose control in the ICU: Report of a 2013 round table meeting. *Crit. Care* **2014**, *18*, 226. [CrossRef] [PubMed]
33. Szadkowska, A.; Michalak, A.; Łosiewicz, A.; Kuśmierczyk, H.; Krawczyk-Rusiecka, K.; Chrzanowski, J.; Gawrecki, A.; Zozulińska-Ziółkiewicz, D.; Fendler, W. Impact of factory-calibrated Freestyle Libre System with new glucose algorithm measurement accuracy and clinical performance in children with type 1 diabetes during summer camp. *Pediatr. Diabetes* **2021**, *22*, 261–270. [CrossRef]
34. Tramunt, B.; Smati, S.; Grandgeorge, N.; Lenfant, F.; Arnal, J.F.; Montagner, A.; Gourdy, P. Sex differences in metabolic regulation and diabetes susceptibility. *Diabetologia* **2020**, *63*, 453–461. [CrossRef] [PubMed]

35. Mauvais-Jarvis, F. Gender differences in glucose homeostasis and diabetes. *Physiol. Behav.* **2018**, *187*, 20–23. [CrossRef]
36. Klueh, U. Analysis: On the path to overcoming glucose-sensor-induced foreign body reactions. *J. Diabetes Sci. Technol.* **2013**, *7*, 452–454. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Article

Comprehensive Understanding of Foot Development in Children Using Capacitive Textile Sensors

Sarah De Guzman ¹, Andrew Lowe ¹, Cylie Williams ², Anubha Kalra ^{1,*} and Gautam Anand ¹¹ School of Engineering, Computing and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand² Faculty of Medicine, Nursing and Health Sciences, Monash University, Melbourne, VIC 3800, Australia

* Correspondence: anubha.kalra@aut.ac.nz

Abstract: Knowledge of foot growth can provide information on the occurrence of children's growth spurts and an indication of the time to buy new shoes. Podiatrists still do not have enough evidence as to whether footwear influences the structural development of the feet and associated locomotor behaviours. Parents are only willing to buy an inexpensive brand, because children's shoes are deemed expendable due to their rapid foot growth. Consumers are not fully aware of footwear literacy; thus, views of consumers on children's shoes are left unchallenged. This study aims to embed knitted smart textile sensors in children's shoes to sense the growth and development of a child's feet—specifically foot length. Two prototype configurations were evaluated on 30 children, who each inserted their feet for ten seconds inside the instrumented shoes. Capacitance readings were related to the proximity of their toes to the sensor and validated against foot length and shoe size. A linear regression model of capacitance readings and foot length was developed. This regression model was found to be statistically significant (p -value = 0.01, standard error = 0.08). Results of this study indicate that knitted textile sensors can be implemented inside shoes to get a comprehensive understanding of foot development in children.

Citation: De Guzman, S.; Lowe, A.; Williams, C.; Kalra, A.; Anand, G. Comprehensive Understanding of Foot Development in Children Using Capacitive Textile Sensors. *Sensors* **2022**, *22*, 9499. <https://doi.org/10.3390/s22239499>

Academic Editor: José Machado Da Silva

Received: 26 September 2022

Accepted: 26 November 2022

Published: 5 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: foot length development; textile sensors; proximity capacitive sensing

1. Introduction

Literature on foot morphology indicates that in the first 3 years of life, the foot grows approximately 2 mm in length per month. More detailed studies indicate different foot-growth rates for children ages one to five years old. Both genders grow at similar rates; however, boys' feet are likely to grow one size longer and one size wider [1,2]. It is also indicated that the development of the child's foot is strongly influenced by the point at which the child starts to stand and start walking. Most manufacturers create shoes that are miniature adult shoes. However, children and specifically toddlers do not have the same foot anatomy and characteristics as adult feet [3–7]. During infancy, there is no need for a shoe as they will not be walking. From the phase when they start to walk, from ages 1 to 2 years, the sole purpose of the shoe is protection from weather and the environment [8]. Therefore, at this age, shoes must be very soft and flexible to allow for freedom to move as though one were barefooted. Thick soles would impinge foot development as they add strain to the foot and ankle joints [9]. Podiatrists still do not have enough evidence as to whether footwear influences the structural development of the feet and associated locomotor behaviour [3,8,10]. Information on foot growth can provide important insights on topics such as the occurrence of a child's growth spurt and an indication of when to buy new shoes [11]. Although foot-growth research is known to be one of the best indicators of a growth spurt and possibly various diseases, information regarding when parents buy new shoes and how their children's feet grow inside these shoes is lacking [11]. This work describes the development and validation of in-shoe monitoring tools to observe children's

foot growth and development, including foot shape and structure (i.e., anthropometric parameters), and specifically foot length. This paper aims to study the technical feasibility of textile sensors in measuring foot length by creating and validating shoe foot-length measurement systems to alleviate the pain points of both parents and podiatrists. This was undertaken by developing two prototypes, grounded heel and non-grounded heel, with embedded textile sensors for capacitive proximity sensing of the foot. Capacitive sensing is based on measuring electrical output through coupling between conductive and dielectric mediums. The change in capacitance detected through these sensors is directly proportional to the changes in pressure applied to the foot. Hysteresis is one of the critical drawbacks of capacitive sensors employing polymers, as they tend to lose their water-retention properties over time; thus, in this study, we made use of textile sensors to avoid this problem.

Textile sensors are made of intertwining fibres (also known as yarns or thread) formed by knitting, weaving, felting, and crocheting yarns together [10,12]. The current trend of using smart textiles shows various potential applications along with several challenges. Many smart textile technologies have been proven in laboratory studies, but there are still uncertainties in mass production. An initial investigation by Harms, Amft, and Troster [13] showed positive trends in simulations predicting the behaviour of loose and partially fitting garments.

2. Materials and Methods

In this study, the development and feasibility of textile sensors to better indicate the foot length of a child while the foot is still inside the shoe are discussed. There are two prototypes developed to test out its feasibility:

- Prototype 1 explains the process of selecting suitable sensing materials, sensing configurations, and electrical components. The materials and sensing configuration are evaluated quantitatively and qualitatively.
- Prototype 2 is a design iteration of the previous prototype. It uses the selected materials from Prototype 1 and refined design configuration. The effectiveness, reliability, and feasibility of this prototype are tested and evaluated with children using a wireless electronics module.

Both prototypes follow the proximity-capacitive-sensing principle. The desired sensing electrodes are made of conductive textiles located at the toe cap of the shoe. These are connected by coaxial cables to the Arduino 101 and MPR121 capacitive sensing chips. These two chips make up the embedded electronics unit to enable reading the fabric sensor's output. The output is then passed on to a data logger for further analysis and evaluation. In Sensor Prototype 1, the Arduino 101 is connected to a PC to enable data logging. In Sensor Prototype 2, the Arduino 101 is replaced by a smaller electronic Pycom unit to enable wireless communication from the electronics unit to an Android mobile phone. Wireless communication is implemented as it created a safer testing environment during the evaluation of the prototype with children.

2.1. Sensor Prototype 1

The Sensor Prototype 1 design requirements are broken up into two categories: mechanical aspects and electrical/electronic components and software. For the mechanical aspects, the sensor layout and materials and knitting techniques are included. Physical connections from the knitted conductive textiles sensors and other e-textile components and the interfacing of MPR121 (capacitive sensing chip) to Arduino constitute the electrical/electronic components and software. These prototype components and test procedures are explained in the following subsections.

2.1.1. Sensor Placement, Size, and Layer Layout

To create Sensor Prototype 1, conductive textile materials from Auckland University of Technology (AUT) Textile and Design Lab were obtained, and samples were knitted by the knitting technician for a specific size of toe cap. The toe-cap dimensions were obtained from

the Bobux design database (Bobux is a New Zealand children's shoe company that started creating soft sole shoes for infants in 1991). Two different fabric sensors were acquired, and these were installed inside two different shoes of the same size. The fabric sensors are connected with a coaxial cable to reduce outside interference/noise in the readings. The sensors were read initially by the Agilent E4980A LCR meter (Keysight Technologies, Santa Rosa, CA, USA, Agilent E4980A) and again by the Arduino 101 with an MPR121 capacitive sensing chip. Arduino 101 was connected to a PC and data were logged through Arduino's built-in serial monitor. The system layout is depicted in Figure 1.

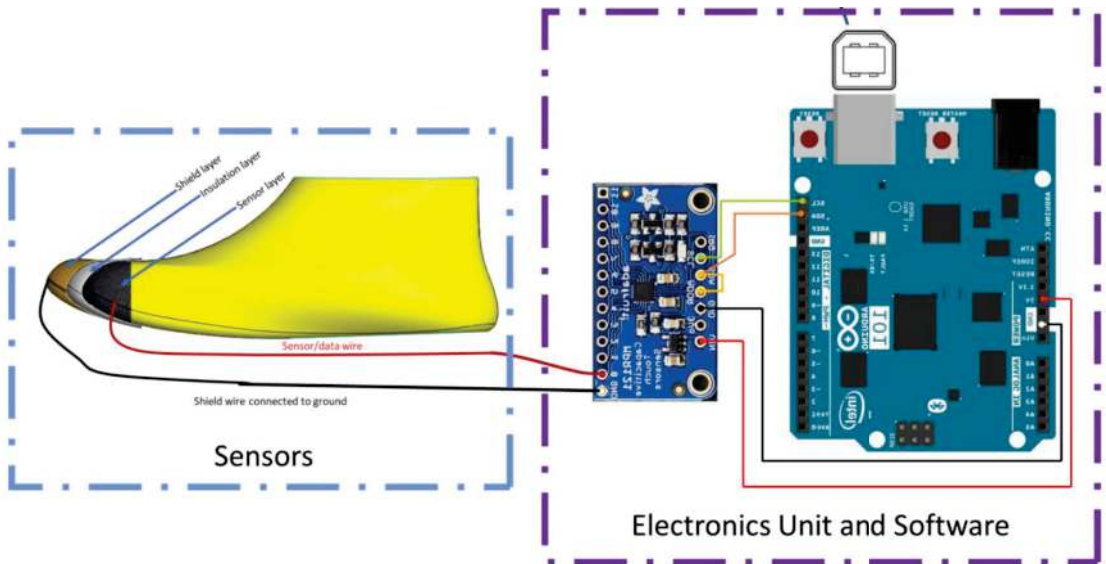


Figure 1. System layout for Sensor Prototype 1.

The evaluation of conductive textiles to select suitable sensor materials was carried out by inserting different-sized phantom feet. The readings were obtained using an LCR meter and these results were compared and graphed. The criterion for choosing a suitable sensor material is the linearity of the sensor results.

Normally, foot length is measured outside the shoes. In typical foot-measuring devices, the heel of the child needs to coincide with the curved part of the device, and a horizontal bar is moved until it touches the longest part of the toe. The number that this bar stops at indicates the size of the child's foot. Ideally, the heel sits at the heel counter. This is designed so that it can hold the heel in position, giving the foot more stability during variable movements, together with the whole shoe upper, which firmly holds the foot inside the shoe. This creates a basis for where to place the sensors. The heel will not experience much change, because the heel will always stay in the heel counter. Therefore, the toe area is an ideal place to put the sensors, as this would see the most change in length. This area is also ideal as layers can be easily fitted and concealed. In shoe manufacturing, polyester material is sandwiched between the inner layer of the shoe and the protective outer layer of the shoe as depicted in Figure 2. This polyester material (called a toe puff) hardens the toe area, which keeps the shoe in shape and adds more protection from scuffing during child's play.

Since the sensing technology is based on proximity sensing, there needs to be a conductive area that would serve as the electrode/sensing area. This is achieved using knitted conductive threads placed in the toe-cap area of an existing Bobux shoe. The sensor shape is based on the toe-cap template provided by Bobux shown in Figure 3. The print area indicates the sizing limit of the sensor shape.

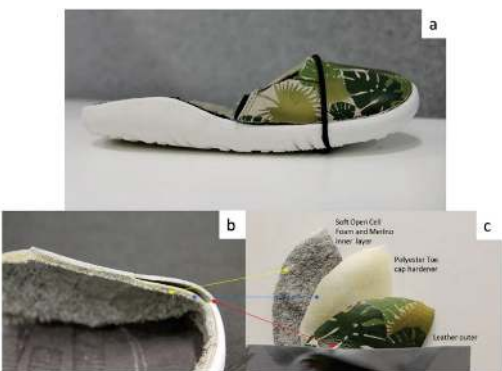


Figure 2. (a) Example of Bobux shoe with midsole removed; (b,c) Upper of the toe area showing the cross-section and layers of the shoe.

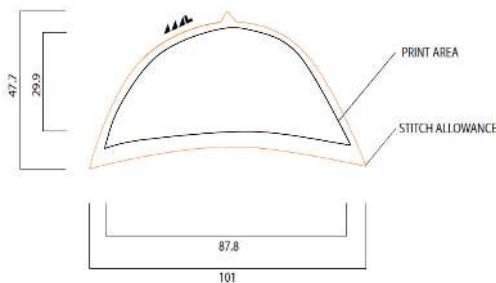


Figure 3. Toe-cap template of a Bobux shoe.

Apart from the sensor, there must also be an insulation layer, to support the sensor in place, and a shield layer to protect the sensor from false touches that can occur outside the shoe. This is an unwanted signal, which comes from an external influence, not the shoe wearer. The detection of the foot should only happen from the inside of the foot. The shield layer is another conductive fabric that helps decrease the interference from the outside of the shoe. The sensor, insulation, and shield make up the three layers of the toe-cap area of the shoe. The textile sensor transmits data to the electronics unit using the black and red wires shown in Figure 4. The black wire is connected between the shield and the ground (GND) pin of the electronics unit; the red wire is connected between the sensor and the signal pin of the electronics unit.

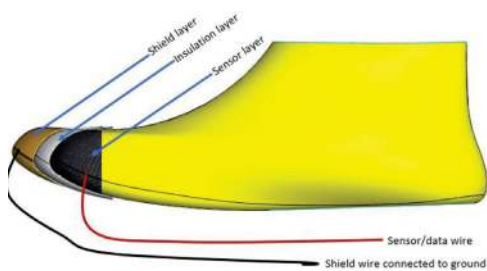


Figure 4. Sensor layer and wire layout.

Additionally, the red wire (sensor/data wire) needs to also be shielded so as not to transmit false touches. This is achieved using a coaxial cable, where the innermost wire of

the coaxial cable is connected to the sensor layer, and the outermost loose-braided copper wires are connected to the shield layer.

2.1.2. Sensor Materials

There were two conductive threads available in AUT Textile and Design Lab with different conductivities, as shown in Figure 5:

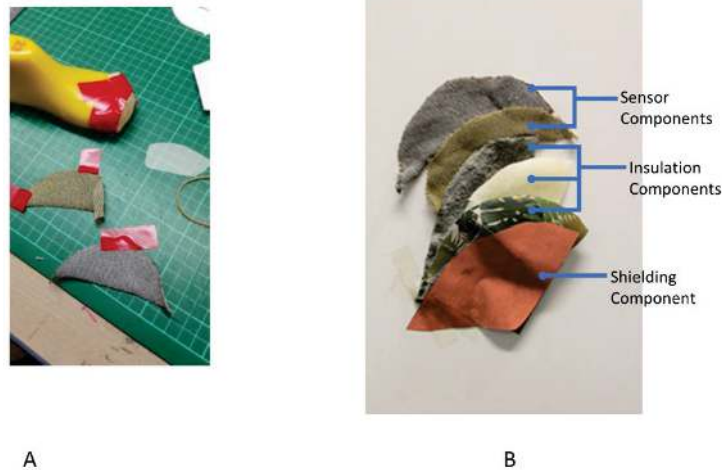


Figure 5. (A): Samples cut to shape using toe-cap template with (i) 99% silver thread sample, (ii) 80/20% polyester/stainless steel sample, (B): different layers of Sensor Prototype 1, showing the sensor, insulation, and shielding components.

(a) Silver thread—2/117 dtex 99% silver-plated nylon, linear resistance 3 k Ω /m (Statex Productions and Vertriebs GMBH, 2010)

(b) Stainless steel thread—2/50 Nm, 80% polyester and 20% stainless steel, linear resistance 530 k Ω /m (Ehrmann, Heimlich, Brucken, Weber, and Haug, 2014)

Both threads were plain knitted as a sheet with ribbed edges using a SIG 123 SV 14 (Shima Seiki, Castle Donington, UK) gauge knitting machine. The sheets were laser cut with the toe-cap template. This is to prevent fraying around the sensor edges. The laser-cut sensors were shaped around the toe-cap area of the shoe last using masking tape. This is undertaken so that the knit would create a solid shape, rendering it easy to attach inside the shoe as shown in Figure 6.

Figure 5B shows the different layers of sensor, insulation, and shielding components of Sensor Prototype 1. Sensor components shown include stainless steel and silver knitted sensors. These two materials were evaluated for their effectiveness, and the results presented in Section 3. Insulation components are the layers of the shoe, which includes (top to bottom) the sock liner, toe hardener/toe puff, and synthetic leather. The shielding component shown is copper taffeta. The sensor and shielding components are separated by the insulation components.

Sensor Prototype 1 uses conventional braided shielded copper wires. This type of wire contains a conductive inner core, usually copper insulated with PVC. This is surrounded by another set of copper wires braided around it, used for shielding. The innermost wire is shielded so that less software filtering is required to isolate the data needing to be read. This removes the interference that can occur when it is installed inside the shoes. This interference can occur when the skin from other parts of the foot touches the wire, such as when a child inserts their foot inside the shoe. This sliding motion along the wire can cause unwanted readings.



Figure 6. Example of the fabric sensors installed using 80/20 PES/SS sample. (a–c) low tack masking tape is used to shape the fabric sensor insulating the other layer. (d) completed sensor to go inside the toe area of the shoe.

2.1.3. Electrical/Electronic Components and Software

The electrical/electronic components will read the capacitance signals from the textile sensor and connections. There are two components used for the capacitance reading and for the acquisition of data from the sensor to the PC: the MPR121 and Arduino 101, as shown in Figure 7.

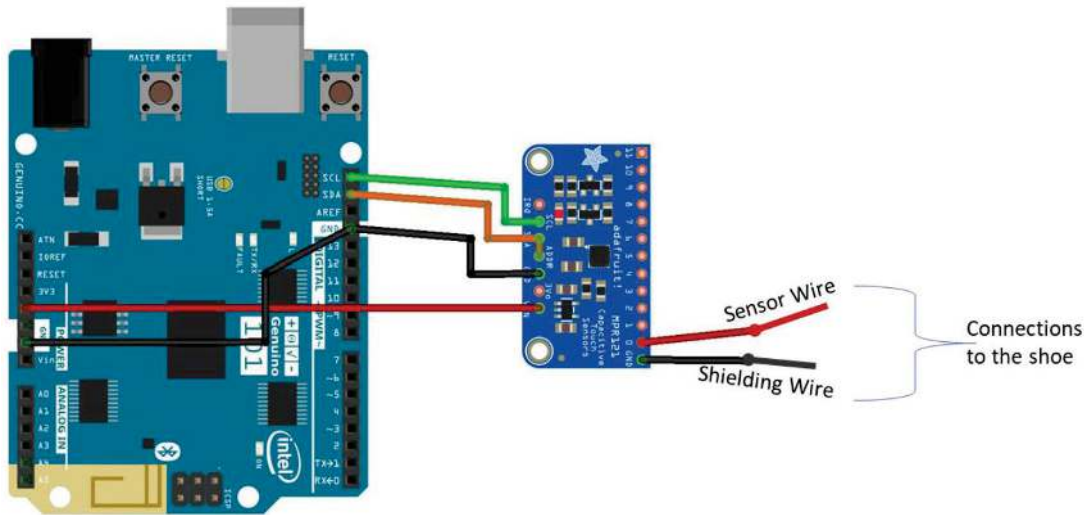


Figure 7. Connections between Arduino 101 and Adafruit MPR121.

The breakout board from Adafruit Industries includes an NXP semiconductor capacitive sensing chip called the MPR121. This can have up to 12 electrode inputs. One input electrode is used in foot-length measurement. This board was chosen due to its small form factor and capacitive sensing range from 10 pF to 2000 pF with a resolution of up to 0.01 pF. This range is desirable due to the unknowns of the capacitance of the fabric sensors.

This capacitive sensing module uses a CTMU to determine capacitance changes. This involves the measurement of voltage, as the system assigns a known current on the sensor.

The circuit then examines the shift in the voltage while constant current is being applied over a specified amount of time.

MPR121 can be set up as a touch and proximity sensor. In this case, it is set up as a proximity sensor. This microcontroller built into the Adafruit board is already wired so that it can step down higher voltages for up to 5V, as it only needs 3.3 V. MPR121 is also an I2C compliant device; it has a configurable I2C address that can be used to change register values to program the microcontroller as wanted. MPR121 is interfaced with an Arduino 101 module, this allows communication and programming in Arduino IDE without having to change the source code. Communication is via I2C, and the information from the sensing module is accessed through registers via serial data.

Code Implementation

For MPR121 to communicate and read the correct capacitance values, an algorithm is followed as recommended by the MPR121 datasheet. Open-source code and libraries that are readily available online were used with alterations to fit the paper's objective. Pseudocode is described in Figure 8, detailing the functions used in the program.

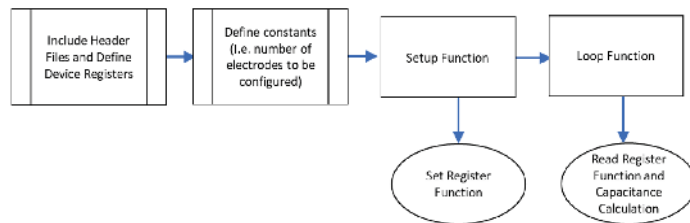


Figure 8. Pseudocode of the Arduino algorithm.

The algorithm components are briefly explained as follows:

- Include header files and define device registers: As MPR121 is an external sensing chip, libraries and header files (i.e., `#include "Adafruit_MPR121.h"`) are included so that variables inside this file can be easily accessed. After introducing the header files, the device addresses must be defined so that Arduino knows which information to access.
- Define constant, reading, and configuration registers: Once the device is addressed and found, configuration registers are defined so that their values can be changed in the setup functions. There are corresponding values for these configuration registers and only some registers are selected, as not all the electrodes are being used. Some registers are not configured. However, they are introduced here as they will be read later, in the loop function. These registers change based on the changes made to the configuration register. The registers that need to be read are *ELEC_Current*, for reading the electrode's current; *ELEC_Time*, the register for the electrode's charge time; and ADC Registers, *ADCLSB*, and *ADCMSB*, for reading the electrode's voltage.
- Setup function: the baud rate is defined to know how quickly the microcontroller samples and how often it obtains information from the serial port. It is set to read at the fastest rate, at 115,200 baud. The introduced configuration registers are changed to their corresponding values to perform as it needs. To set these registers to their new value, the device itself must be soft reset (*SOFT RESET*). This resets the registers into their default values. After this, the MPR121 is at stop mode (*ELEC_CFG* bit assignment is 0b10000000), so the remaining registers, such as the auto configuration registers (*AUTOCONFIG0*) and voltage operation limits (*UPLIMIT*, *LO LIMIT*, *TARGETLIMIT*), may be changed. Once registers are set, the MPR121 is changed to run mode so that these settings can be applied and used. To enable run mode, the electrode-configuration register-bit assignment changes to enable electrode detection and proximity sensing at electrode 0.

- Loop function: In this function, certain registers are read to enable the calculation of capacitance. As mentioned earlier, MPR121 operates using CTMU, which means that it needs current, voltage, and time to calculate capacitance. These can be read from reading registers defined earlier (*ELEC_Current*, *ELEC_Time*, *ADCMSB*, and *ADCLSB*). The current register can be read as is and does not need any other conversions. The time register contains two values, each corresponding to the charge time for that electrode. The electrode used for Sensor Prototype 1 is Electrode 0, the time for this register is accessed on the *ELEC_Time* register; however, this contains time values for electrodes 0 and 1.

2.1.4. Sensor-Material Evaluation and Code-Verification Testing

This is undertaken to evaluate whether the capacitance reading truly increases when the toe area gets closer to the sensor, which is laid out in a non-planar pattern. Additionally, this is to evaluate if the programmed capacitive sensing circuit displays a similar output as the LCR meter, which is a more reliable source of readings. In addition, an evaluation is needed to compare the silver knitted threads and stainless steel threads in terms of sensing proximity.

To test whether capacitance truly increases when the toe gets closer to the sensor, 3D-printed shoe lasts were covered in aluminium conductive tape. This is to replicate the conductive characteristics of the human foot and will be referred to as the foot phantom. Although aluminium may not have the same conductivity and electrical behaviour as the human foot, these were sufficient to undertake the tests carried out for Sensor Prototype 1. These lasts are of varied sizes; the range used was sizes 20–23 with equivalent lengths between 13.5 cm (Size 20)–15.5 cm (Size 23) as shown in Figure 9A. The test setup is shown in Figure 9B.

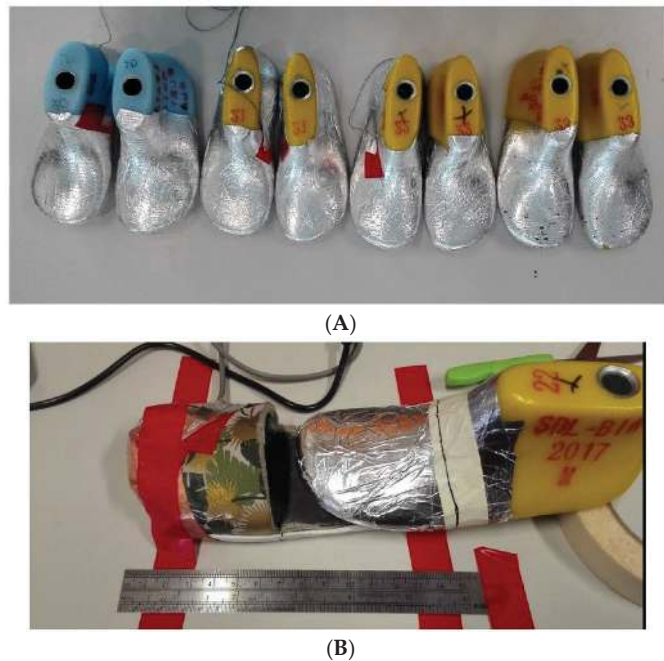


Figure 9. (A) Foot phantom, showing different sizes (showing sizes 20–23, left to right), (B) Test setup.

To assess whether the Arduino was programmed correctly, the textile sensors inside the shoes were connected to both the Agilent E4980A LCR meter and the Arduino and the readings were observed using the Arduino serial data window and recorded.

Finally, the results recorded from both materials were then evaluated. A criterion used to select material is that the capacitance should increase linearly as the foot approaches the textile sensor.

The results and discussion of Prototype 1 are found in Section 3.

2.2. Sensor Prototype 2

The objective of this prototype is to create a full textile and connection system so that it can conform to any shape and be more robust to flexing. An enclosure is created to protect the electronics unit during testing with real human feet. In addition, the prototype attempts to increase the sensitivity of the fabric sensors, so that they can read a larger range of values over a small distance.

Based on Prototype 1, the maximum sensing range corresponds to a foot phantom 2 cm away from the sensor. This corresponded to the insole information provided in Figure 10a. This refers to the 2 cm mark, which matches a foot from size 20 (13.5 cm) to size 23 (15.5 cm). From this information, a new striped pattern, Figure 10b, was devised to create a sensor that can sense a larger range, while also improving the structure of the knit integration with other threads to be implemented as a sock liner.

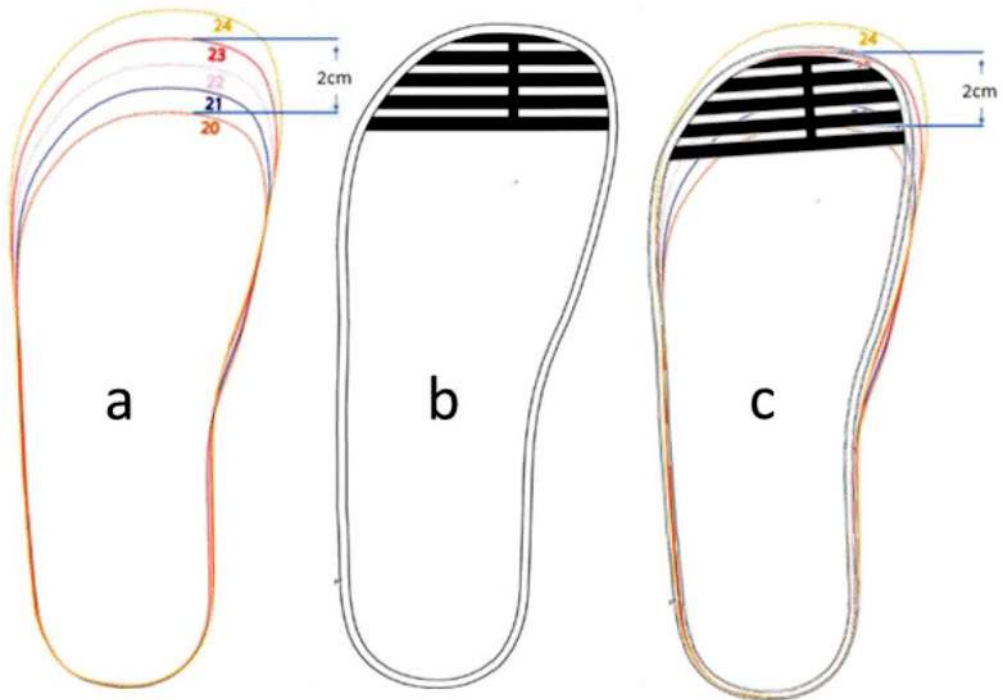


Figure 10. (a) Insole information from Bobux showing the 2 cm mark. (b) Sensor pattern segmented but connected in the middle. (c) Sensor pattern overlapped with insole information.

The striped pattern was drawn in Solidworks and corresponds to a size-23 shoe insole and toe cap. The dimensions of the pattern are shown in Figure 11. The thickness of the conductive knits and gaps are determined by the minimum number of courses (rows) the knitting machine can achieve. A SIG 123 SV 14-gauge knitting machine was used, which can do a minimum of two courses.

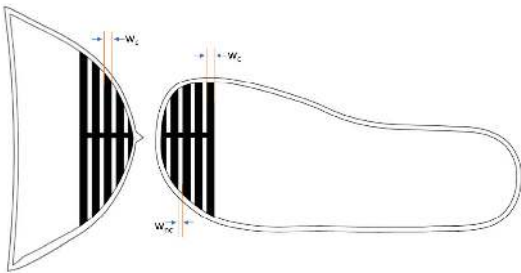


Figure 11. Striped pattern drawn on size-23 toe cap (**right**) and insole (**left**); w_c = width of conductive knit, and w_{nc} = width of non-conductive knit.

Two courses are equivalent to 2 mm thickness; the conductive traces are knitted with four courses (w_c) and the non-conductive thread gaps (w_{nc}) are knitted with two courses.

Before implementing this design with the knitting technician, the striped pattern was first implemented as a cut-out of copper stripes. The template was printed on a one-to-one scale, and the copper tape was cut out to the same scale. The stripe running across the conductive widths was replicated in the mock-up by connecting a pure stainless steel conductive thread with a single header pin. Figure 12 contains an amber-coloured tape, called polyimide tape, which provides insulation and electrical isolation from the surrounding electronics and human touch.

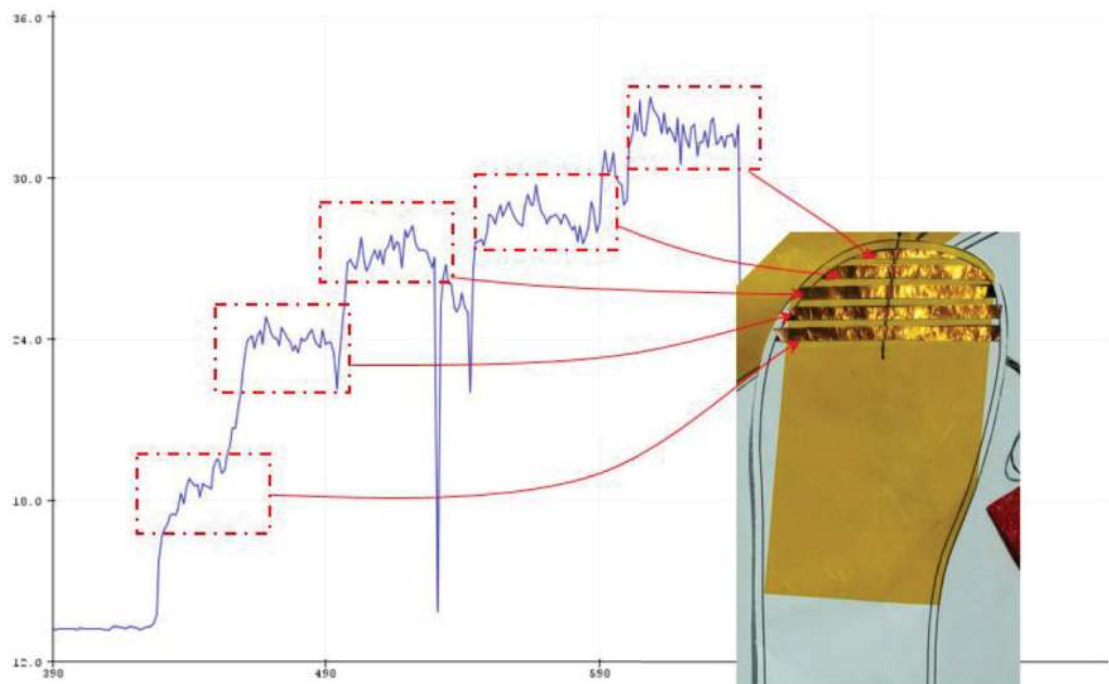


Figure 12. Incremental increase when lasts covered each copper stripe.

The mock-up was plugged into the Arduino with an MPR121 unit, to see whether capacitance increased whenever a copper stripe is covered. This was undertaken by covering each stripe with each size of the foot phantom (size 20–23) and observing whether there was an incremental change whenever a copper stripe is covered. The results were read and logged using the Arduino serial monitor built into the PC.

The graph depicted in Figure 12 shows an incremental increase when each stripe is covered by a size 20–23 foot phantom. Although this proved that the striped design works, this test is not an overall indication of the final design, as the final prototype must be made of fabric and not paper and copper tape.

2.3. Sensor Manufacturing—Sensor Fitment Iterations

Different iterations of sensors, mainly three versions, are assembled and visually evaluated. Version 1 is similar to the copper stripes but replaced with knitted 80 PES/20 SS conductive thread. For Version 2, the textile sensor was made and backed up with a polyester thread (white), to hold up the conductive stripes evenly. The stripes are also more complete, as it does not have to be hand-sewn together.

Version 3 was knitted as a sock. This eliminates the handwork of joining the upper and the sole-fabric sensor. The type of knitting used for the samples is tubular knitting using Shima Seiki Whole Garment technology. This technique allows for two different materials to be knitted together. It incorporates non-conductive polyester threads in between the conductive threads, allowing for a better mechanical structure and preventing the conductive rows from touching each other. This also allows for a seamless connection along the stripes, ensuring each row is connected as one round stripe. The main connection from the fabric sensor is performed the same way in Version 2, where an insulated stainless-steel thread was stripped and sewn onto the sensor. The knitted sensor sock was coated in PVA glue while it was inside a last. The glue was cured and dried inside an oven at 100 °C until dry to touch. This is to keep the sensor in shape and the insulated conductive thread connection in place, for easy installation inside the shoes.

2.4. Final Sensor Assemblies

Version 3 was chosen for manufacturing the final assemblies of the sock sensors because it achieved a full textile connection and the seamless integration of the knitted sensor from upper to sole. There were two knitted sock-sensor samples created for the test, characterized as grounded heel (GH) and non-grounded heel (NGH), as shown in Figure 13. Both samples were knitted with a combination of non-conductive polyester thread and 80/20 PET/SS conductive thread. GH contained a conductive knit patch on the heel and NGH did not. GH and NGH samples had the same knit layout on the toe area.

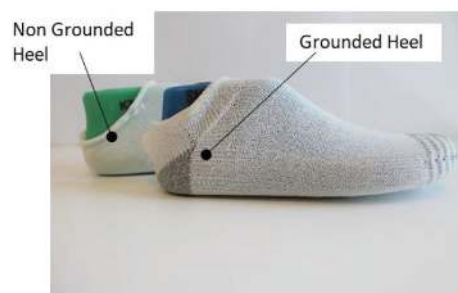


Figure 13. Two samples showing non-grounded heel (NGH) and grounded heel (GH) samples.

The sensors were installed inside Bobux shoes. The GH sensor contained an additional ground plane at the heel. This was connected to the shield layer on the outer upper of the shoe (copper taffeta). For both GH and NGH, insulated stainless steel thread was used as the sensor wire, carrying data to the electronics unit. A non-insulated stainless steel conductive thread is wrapped around this thread, to replicate the shielding effect of a coaxial cable, preventing noise from the environment.

Figure 14 shows the physical connections of the shoe with the GH sensor. The black cloth contains connections from the copper taffeta and the extra ground plane on the heel. The connections shown in Figure 14(right) are connected and crimped to a blue female FCI clincher. This was used as it gives the best contact for the traces while ensuring reliable contact between the conductive threads and the electronics unit.

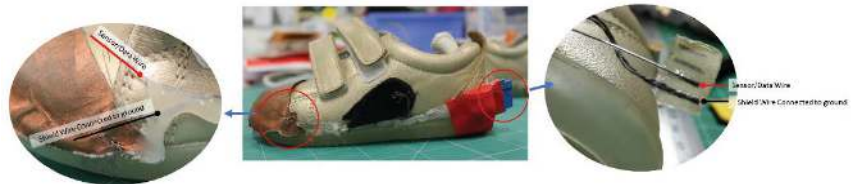


Figure 14. (Left) Connections from the sensor and shield layer on the toe cap; (Middle) Overview of the grounded heel shoe; (Right) Connections of wires attached to a Z-axis tape.

2.5. Pycom Expansion Board and WiPy 2.0

A wireless system was implemented so that the device can be tested on children's feet. This is undertaken to avoid any hazard that may occur during testing, such as a broken connection between the Arduino and the PC due to accidental disconnection while the child is walking.

The Arduino is replaced by two modules, a Pycom Expansion Board and WiPy 2.0. The Pycom Expansion Board is used because of its easy plug-and-play configuration, which allows for external microcontroller boards to be plugged in. In this situation, the MPR121 is still used and is plugged into the exposed female pin headers using a custom-made PCB.

The Pycom expansion board includes a mini-USB plug so that it can be easily programmed and deliver power using an external power source such as readily available power banks. The expansion board contains a step-down voltage regulator, so that when a higher voltage is plugged in, it will only take what is needed to prevent overvoltage (Pycom Ltd., Guildford, UK, 2017).

The WiPy 2.0 is a separate microcontroller that is attached by inserting the headers on the inner part of the Pycom Expansion Board. This microcontroller features Bluetooth low energy (BLE), Bluetooth classic, and wi-fi. This is advantageous, as it can be integrated with older Android devices that do not support BLE. It can also communicate over wi-fi, therefore eliminating the need for a wired connection to log and acquire data. This is compatible with MPR121, as it can communicate with the I2C bus protocol (Pycom Ltd., 2017).

An external custom-made PCB shield was manufactured to decrease the number of wires connecting the MPR121 to the Pycom Expansion Board headers, as shown in Figure 15.



Figure 15. PCB shield assembled on top of the Pycom Expansion Board.

The Arduino 101 is replaced by the Pycom Expansion Board and WiPy 2.0. Although both WiPy 2.0 and Arduino 101 support a Bluetooth low energy module, the WiPy 2.0 also had Bluetooth classic. This version had more software support in terms of open-source code. It was also more compatible with older Android phones, as they only work with Bluetooth classic.

An Android app was made using Android Studio to allow for seamless data collection, creating a wireless data-logging system. This meant that the electronics unit was capable of operating while it was strapped on a child, without worrying about physical disconnections. The code of MPR121 implemented in Arduino in Prototype 1 was also implemented in WiPy 2.0; this provided the same reading conditions with a different electronics unit.

A hardware enclosure as shown in Figures 16 and 17 was created to prevent physical disconnections between the circuit and the shoe. The hardware enclosure also prevents possible electric shocks that may come from low-voltage electronics while they are being tested on children. The design of the enclosure was based on an original Pycom enclosure provided on purchase.



Figure 16. 3D-printed enclosure.



Figure 17. 3D-printed enclosure with elastic straps to fit child’s ankle. (a) Front, (b) Back.

2.5.1. Overall Connection and Assembly of Shoes and to Electronics Unit

In Prototype 1, the overall system layout considered the use of Arduino. In this section, Pycom with MPR121 is implemented. The overall physical connections with pin assignments are laid out in Figure 18.

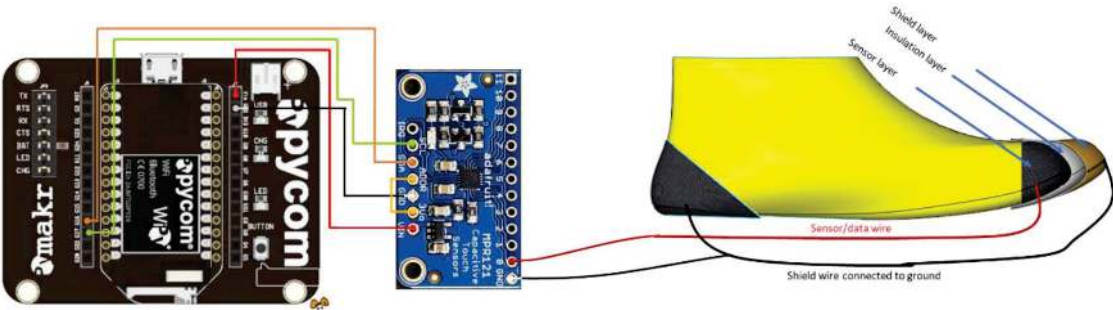


Figure 18. Physical connection from the electronics unit to the shoe sensor unit.

The grounded heel sensor and non-grounded heel sensor connected to Pycom and MPR121 are shown in Figure 19.

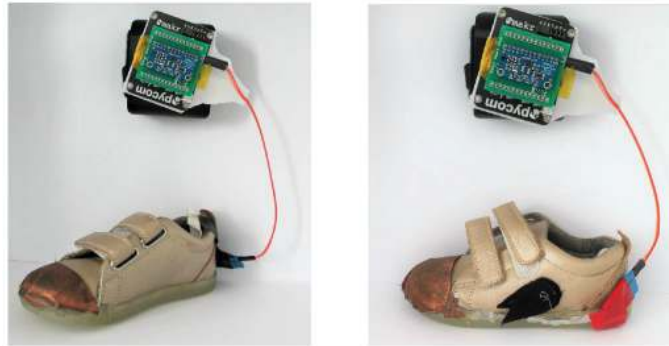


Figure 19. (Left) Non-grounded heel sensor connected to Pycom and MPR121, (Right) Grounded heel sensor connected to Pycom.

2.5.2. Evaluation with Children

A study on children was undertaken to prove the effectiveness and reliability of the textile sensors. Ethics approval was granted by the Auckland University of Technology Ethics Committee (AUTC). All children in this study were recruited from the general community using online platforms and email recruitment from the AUT Early Childhood Centre and the existing Bobux customer database. In-person recruitment also occurred using posters put up in the Bobux outlet store. Parents and/or guardians gave written consent and children assented to participate. During the test held in the Bobux retail store, the parents and caregivers needed to be present to give them familiarity with the environment. The shoe prototypes were tested on 30 children, and 14 datasets were taken from that.

The inclusion criteria put in place for this work were that only developing children aged 11–36 months old would be included, and parents would be asked if their children had any known health conditions that could have an impact on walking. These include genetic conditions that change walking, neuromuscular conditions that are exhibited in tiptoe walking, or orthopaedic conditions that result in foot-structure change.

Necessary exclusion criteria were: only children that can stand and walk properly could participate, and children's foot length needed to be under 150 mm, as the sensors are instrumented inside a size 23 Bobux shoe, which has a foot length of 150 mm.

The test procedure called for the participants/children to be in a static (standing) position, with their heels at the back of the shoe, and with socks worn on both feet. During the test, the children's guardians were asked to count from 1 to 10 after physical measurements—i.e., after foot length and foot circumference were measured. During the test, other parameters were recorded, such as Bobux shoe size (according to the Bobux shoe ruler), recorded by placing the child's heel at the heel curve of the Bobux shoe ruler, and the slider on the opposite end being slid until it touched the longest point of the foot, as shown in Figure 20.



Figure 20. Bobux shoe-size measurement.

Foot length was also measured using a tape measure, in which a tape measure was placed at the back of the Bobux shoe ruler. This indicated the foot length in centimetres. Foot circumference was measured using a tape measure (as shown in Figure 21), in which the tape measure was wrapped around the widest part of the foot while the participant was standing up.



Figure 21. Tape measure wrapped around the widest part of the foot.

These were recorded to create a comparison between capacitance readings, foot length, indicative Bobux shoe size, and possibly foot volume.

3. Results and Discussion

This section includes the results of both Sensor Prototype 1 and Prototype 2.

3.1. Sensor Prototype 1 Results and Evaluation

Tests were performed to see whether the programmed MPR121 interfaced with Arduino 101 read a similar output as the LCR meter (Keysight Technologies, Santa Rosa, CA, USA, Agilent E4980A), which is a more reliable source of measuring capacitance. Table 1 displays the capacitance readings from both knitted textile sensors, where the readings were derived from an LCR meter and the Arduino.

Table 1. Capacitance readings from both knitted textile sensors. Readings derived from LCR meter and Arduino.

Shoe Last Inserted	Distance of the Last to the Fabric Sensor (cm)	Capacitance Readings (pF)			
		Silver Samples		Stainless Steel Samples	
		Arduino	LCR Meter	Arduino	LCR Meter
No last	6	53.19	99.92	26.3	37
Size 23	2	54	101.26	29	39
Size 22	1.5	77	107.45	35.5	40
Size 21	1	75.5	105.4	42.5	42
Size 20	0	80.5	108	76	48

Figure 22 compares the capacitance readings of the silver samples, and Figure 23 compares the capacitance readings of the stainless-steel samples.

Both samples in Figure 22 behaved the same way, regardless of which measuring system was used, as the Arduino readings followed the same pattern as those from the LCR meter. However, the scaling between the two readings is different, because the readings did not have the same starting point and there was a large offset between two readings. There is also an irregularity, where the reading of Arduino at 1.5 cm is higher compared to the readings at 1 cm.

Both samples in Figure 23 also behaved similarly regardless of the measuring system. As the phantom foot approached closer to the sensor, the capacitance increased. The trends

for both Arduino and LCR meter are similar; however, between the 2 cm and 1 cm mark, they have two different slopes. The Arduino detected a much larger change on the last centimetre compared to the LCR meter, resulting in a much steeper gradient and dramatic change as the foot came much closer to the sensor.

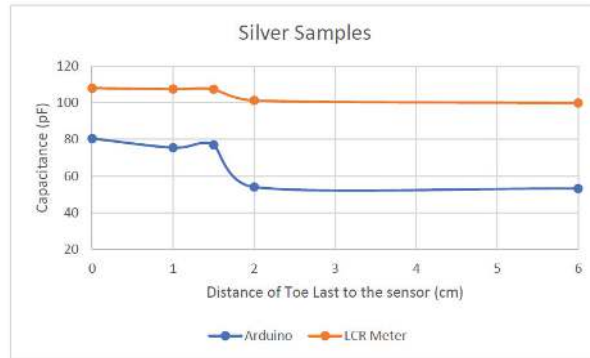


Figure 22. Silver sample capacitance readings comparison between Arduino and LCR meter.

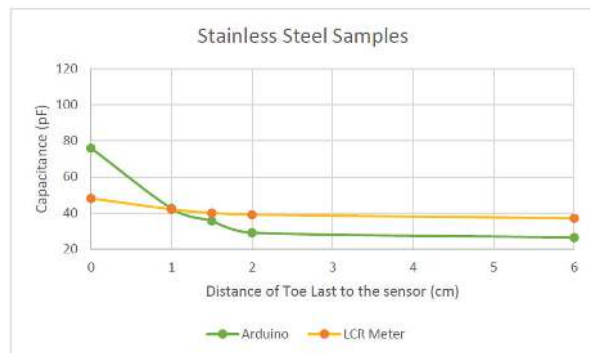


Figure 23. 80/20 polyester/stainless steel sample capacitance reading comparison between Arduino and LCR meter.

Although the readings from the Arduino are not the same as those from the LCR meter, both measurements followed the same trend/pattern. This is a good indication that the Arduino interfaced with MPR121 is good enough to measure the capacitance changes.

From the results in Table 1, Figures 22 and 23, as the foot phantom comes closer to the fabric sensor, the capacitance increases. This occurs regardless of what material it is made of or the equipment it is being measured with. All the knitted sensors can sense up to 2 cm away, as the graph starts to change its slope when the last reaches this part.

From the material evaluation point of view, as observed in Figure 24 below, the polyester/stainless steel (ss) sample (green line) showed a steeper increase in capacitance compared to the silver sample (blue line).

The SS sample also showed a more distinct increase when the toe last drew closer to the fabric sensor, whereas the silver sample exhibited an irregularity when it detected capacitance at the 1.5 cm mark. This may be because the sensor was not sitting flat in this area. As this sensor is shaped around a shoe last, there are areas in the sensor that may have more conductive material and so it bunched together, creating a thicker conductive plate. The more conductive material in this area means that it could have picked up more capacitance compared to an area before or after this. The bunching of sensor material is also possible as the knit material of the silver sample felt softer and more limp compared to

the SS sample. The texture of the knit could have affected the readings when the lasts were inserted. The last could have shifted the sensor, creating an irregularity in the reading.

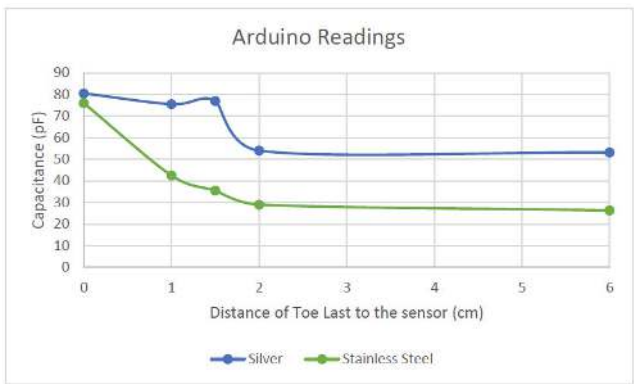


Figure 24. Capacitance readings using Arduino, comparing silver and 80/20 polyester/stainless steel knitted samples.

From Table 1, comparing the Arduino results, the two samples garnered two different capacitance ranges. The SS samples’ readings ranged from 26.3 to 76 pF, which corresponds to a capacitance range of 50 pF over 6 cm, whereas the silver samples’ varied between 53.19 to 80.5 pF—hence, a capacitance range of 27 pF over 6 cm. From these results, the sensitivity of the SS sample is greater (8.3 pF/cm) than that of the silver samples (4.5 pF/cm). This indicates that the SS sample has better sensitivity, which is more advantageous than the silver sample. A higher sensitivity means that the sensor can have more resolution in readings. This is suitable when measuring foot length because, especially from ages between 11 months old to 36 months old, the foot-length development is rapid. Therefore, a more sensitive sensor can pick up minor changes.

3.2. Sensor Prototype 2 Results and Evaluation

3.2.1. Foot-Length and Capacitance—Statistical Analysis

Referring to Table 2, the GH shoe-sensor results possessed a statistically significant result ($p < 0.05$) and a steeper relationship compared to those for the NGH ($p = 0.2$). The standard error (SE) of the GH was smaller, statistically indicating that it is unlikely for the true gradient to be negative.

Table 2. Linear Regression—Statistical Analysis for Capacitance and Foot-Length Relationship.

Linear regression model: $y = mx + c$ y = equation for both GH and NGH shoes; m = gradient of the equation; c = y intercept of the model			
	m (lower CI, upper CI, SE)	c (lower CI, upper CI, SE)	p
y_{ngh}	0.130 (−0.085, 0.346, 0.0991)	10.57 (5.669, 15.48, 2.252)	0.2135
y_{gh}	0.238 (0.057, 0.418, 0.0829)	8.158 (4.012, 12.2, 1.889)	0.0141

3.2.2. Foot-Length and Capacitance Relationship Results

For these results, the circumference of the foot was taken from the widest part of the foot. An assumption was made that the volume of the foot has approximately the section of an ellipse. The volume of the foot is assumed to have a cylindrical ellipse shape; the semi-minor axis is assumed to be one-fourth the semi-major axis.

Referring to Table 3, the GH shoe-sensor results possessed a statistically significant result ($p < 0.05$) and a steeper gradient relationship compared to NGH ($p = 0.06$). The

standard error (SE) of GH is smaller, statistically indicating that it is unlikely for the true gradient to be zero or negative.

Table 3. Linear Regression—Statistical Analysis for Foot Volume and Capacitance.

Linear regression model $yv = \text{foot volume regression equation} = mx + cyv = \text{equation for both GH and NGH shoes; } m = \text{gradient of the equation; } c = y \text{ intercept of the model}$			
	m (lower CI, upper CI, SE)	c (lower CI, upper CI, SE)	p
yv_{ngh}	7.54 (−0.60, 15.68, 3.73)	112.73 (−72.11, 297.6, 84.83)	0.066
yv_{gh}	10.06 (3.15, 16.97, 3.16)	52.84 (−104.6, 210.3, 72.25)	0.008

3.2.3. Prototype 2 Evaluation

The results shown in Figures 25–27 showed the sensors are functioning when real human feet with socks on are inside the shoes. Both GH and NGH showed the feasibility of the textile sensors measuring foot length. This was indicated by the positive gradients in Figure 25, which mean that, as longer feet are inserted inside the shoe, the capacitance readings increase.

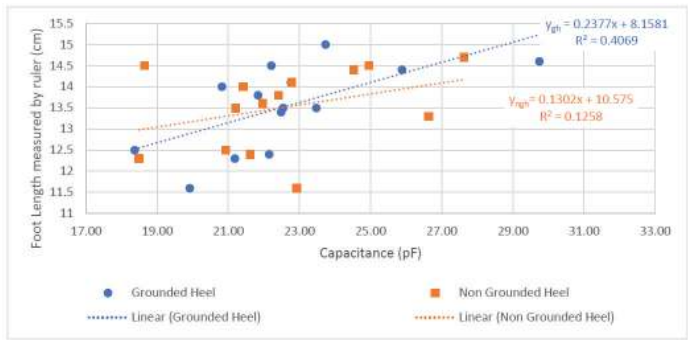


Figure 25. Graph of Foot Length and Capacitance.

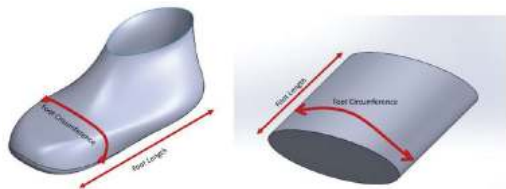


Figure 26. Assumption of the foot as an ellipse shape.

In Figure 25, y_{gh} showed a steeper gradient and a larger r^2 value compared to y_{ngh} . This may be due to the extra conductive pad in the GH sensor creating an effect that increases the sensitivity of the sensor. This extra conductive pad on the heel decreased the parasitic noise from the background. The result for y_{gh} in Table 3 also showed statistical significance, as the p -value is less than 0.05 ($p = 0.01$), whereas for y_{ngh} it is greater than 0.05 ($p = 0.2$)—hence, the GH shows a more credible result compared to the NGH sensor.

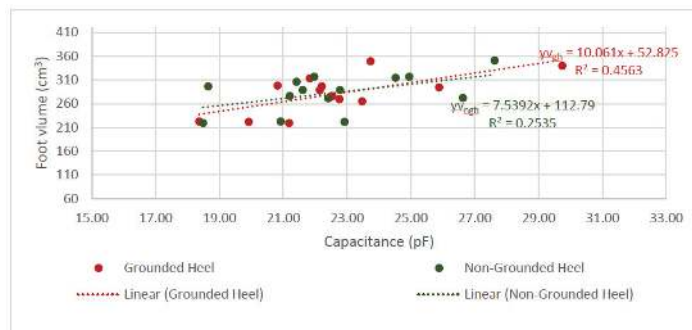


Figure 27. Graph of Foot Volume and Capacitance.

Additionally, both GH and NGH sensors showed feasibility in measuring foot volume. This can indicate how much of the human foot is expanding inside the shoe. This is indicated by the positive gradients in Figure 27, as a larger volume indicates an increase in capacitance. However, a similar effect was presented where y_{vgh} showed a steeper gradient and a larger r^2 value compared to y_{vng} in the p -values in this linear regression model. Table 3 depicted lower values compared to the foot length and capacitance in Table 2. Even though the values are lower for the GH shoe, y_{vgh} showed a more statistically significant result due to its p -value being 0.008, whereas for the NGH shoe, the y_{vng} p -value was 0.06. Again, this could have been the effect of the extra conductive pad on the heel of the GH shoe sensor.

The textile sensors appear feasible for sensing the foot length, especially the GH shoe sensor. However, the r^2 value of the GH shoe is less than 50% when capacitance is compared against foot length and foot value. This can mean that the linear regression model may not be a suitable model for characterizing the textile sensors. Although the p -values showed credibility and proved the concept, more information, such as more variability in the foot lengths, is needed to create a conclusive and more definite capacitance-to-foot-length model.

4. Conclusions

To conclude, there were two aspects to this work. Sensor Prototype 1 tested the feasibility of textile sensors and Sensor Prototype 2 further improved the design of Sensor Prototype 1. Both sensor prototypes went through an iterative design process to create final prototypes, and these were tested and validated throughout the design process for further verification.

Sensor Prototype 1 showed that the physical presence of a phantom foot increases the capacitance reading of the fabric sensor. This was proved by reading capacitance using the gold standard measurement, an LCR meter (Keysight Technologies, Santa Rosa, CA, USA, Agilent E4980A). A smaller capacitance reading system was then implemented using Arduino and MPR121, to enable portable capacitance logging from the fabric sensor. The capacitance readings from both measurement systems provided comparable results. An issue found in this design iteration was that it was not sensitive enough to measure incremental changes. This means that the reading from a phantom foot 2 cm away from the sensor was similar to that of one 0.5 cm away from it.

A different sensor design was implemented in Sensor Prototype 2, to increase the sensitivity of the fabric sensors. Two sensor configurations were tested, to see which design would work better. In vivo validation was performed to examine whether a child's foot would change the capacitance the same way a phantom foot would in the GH and NGH shoe. During this testing, the child's foot length and foot circumference were measured with a measuring tape. The foot length and the capacitance reading were plotted in a graph for the GH and NGH shoe prototypes. The GH shoe indicated better performance compared to the NGH shoe. This was indicated by the difference in slopes and difference

in standard errors, using the information analysed from linear regression. The p -value also indicated that the placement of an extra conductive pad in the heel part (GH shoe) increased the sensitivity of readings.

The positive increase in slope and small standard of error in the GH shoe indicated that textile sensors can measure a child's foot length.

Author Contributions: Methodology, S.D.G.; Formal analysis, A.L., C.W., A.K. and G.A.; Investigation, S.D.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Ministry of Business, Innovation and Employment, EXGware project, Endeavour Grant, New Zealand.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Foot Growth in Children Age One to Five Years—PubMed. Available online: <https://pubmed.ncbi.nlm.nih.gov/2307377/> (accessed on 21 September 2022).
2. Gould, N.; Moreland, M.; Alvarez, R.; Trevino, S.; Fenwick, J. Development of the child's arch. *Foot Ankle* **1989**, *9*, 241–245. [CrossRef]
3. Big Issues for Small Feet: Developmental, Biomechanical and Clinical Narratives on Children's Footwear | Journal of Foot and Ankle Research | Full Text. Available online: <https://footankleres.biomedcentral.com/articles/10.1186/s13047-018-0281-2> (accessed on 21 September 2022).
4. James, A.M.; Williams, C.M.; Haines, T.P. Heel raises versus prefabricated orthoses in the treatment of posterior heel pain associated with calcaneal apophysitis (Sever's Disease): Study protocol for a randomised controlled trial. *J. Foot Ankle Res.* **2010**, *3*, 3. [CrossRef] [PubMed]
5. Morrison, S.C.; Ferrari, J. Inter-rater reliability of the Foot Posture Index (FPI-6) in the assessment of the paediatric foot. *J. Foot Ankle Res.* **2009**, *2*, 26. [CrossRef]
6. Evans, A.M.; Nicholson, H.; Zakarias, N. The paediatric flat foot proforma (p-FFP): Improved and abridged following a reproducibility study. *J. Foot Ankle Res.* **2009**, *2*, 25. [CrossRef]
7. Evans, A.M. Growing pains: Contemporary knowledge and recommended practice. *J. Foot Ankle Res.* **2008**, *1*, 4. [CrossRef] [PubMed]
8. Price, C.; Morrison, S.C.; Hashmi, F.; Phethean, J.; Nester, C. Biomechanics of the infant foot during the transition to independent walking: A narrative review. *Gait Posture* **2018**, *59*, 140–146. [CrossRef] [PubMed]
9. Children Sport Shoes—Feelmax. Available online: <https://feelmax.com/research/children-sport-shoes/> (accessed on 21 September 2022).
10. Wegener, C.; Hunt, A.E.; Vanwanseele, B.; Burns, J.; Smith, R.M. Effect of children's shoes on gait: A systematic review and meta-analysis. *J. Foot Ankle Res.* **2011**, *4*, 3. [CrossRef] [PubMed]
11. Busscher, I.; Kingma, I.; Wapstra, F.H.; Bulstra, S.K.; Verkerke, G.J.; Veldhuizen, A.G. The value of shoe size for prediction of the timing of the pubertal growth spurt. *Scoliosis* **2011**, *6*, 1. [CrossRef]
12. Textiles—An Introduction—Textile School. Available online: <https://www.textileschool.com/119/textile-an-introduction/> (accessed on 21 September 2022).
13. Harms, H.; Amft, O.; Troester, G. Does Loose Fitting Matter? Predicting Sensor Performance in Smart Garments. In Proceedings of the 7th International Conference on Body Area Networks, Oslo, Norway, 24–26 September 2012; Available online: <https://eudl.eu/doi/10.4108/icst.bodynets.2012.249968> (accessed on 21 September 2022).



Article

The Development of a Built-In Shoe Plantar Pressure Measurement System for Children

Sarah De Guzman ¹, Andrew Lowe ¹, Cylie Williams ², Anubha Kalra ^{1,*} and Gautam Anand ¹

¹ School of Engineering, Computing and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand

² Faculty of Medicine, Nursing and Health Sciences, Monash University, Sydney 2109, Australia

* Correspondence: anubha.kalra@aut.ac.nz

Abstract: There is a rapid increase in plantar pressure from the infant to toddler stage, yet little is known about the reasons for this change. More information about plantar pressure distribution can help clinicians identify early-stage foot-related diseases that may occur during transitions from childhood to adulthood. This information also helps designers create shoes that adapt to different needs. This research describes the development of a low-cost, built-in shoe plantar pressure measurement system that determines foot pressure distribution in toddlers. The study aimed to improve and provide data on pressure distribution during foot growth. This was accomplished by implementing a plantar pressure capacitive measurement system within shoes. The capacitive sensors were laminated using a copper tape sheet on plastic backing with adhesive, elastomer layers, and a combination of conductive and non-conductive fabrics. Constructed sensors were characterized using compression tests with repeated loads. Results demonstrated that the sensors exhibited rate-independent hysteresis in the estimation of pressure. This enabled a calibration model to be developed. The system can mimic more expensive plantar pressure measurement systems at lower fidelity. This emerging technology could be utilized to aid clinicians, researchers, and footwear designers interested in how pressure distribution changes from infants to toddlers.

Keywords: sensor; foot; child; toddler

Citation: De Guzman, S.; Lowe, A.; Williams, C.; Kalra, A.; Anand, G. The Development of a Built-In Shoe Plantar Pressure Measurement System for Children. *Sensors* **2022**, *22*, 8327. <https://doi.org/10.3390/s22218327>

Academic Editor: José Machado Da Silva

Received: 16 September 2022

Accepted: 27 October 2022

Published: 30 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sensor technology in the medical and health industry has been increasing in demand with the evolution of mobile and home IoT health devices providing users remote analytics regarding the state of their health. Such examples of this technology are wristband heart rate monitors, glucose monitoring systems for diabetics, as well as sleep inspecting devices.

Children's foot pressure is highly dynamic due to growth, weight change, and skill demands [1]. At birth, some bones are present but the majority of the foot is cartilagenous, transitioning with rapid ossification in the first 12 months to 6 years [2]. A large fat pad is present underneath the foot up to the age of six. It is thought this structure distributes plantar pressures to protect delicate soft cartilage and tissues and help increase the surface contact with the ground during learning complex walking and running skills [3].

One way children's feet are assessed in clinical environments is through measuring and understanding dynamic plantar pressure patterns [4]. It is not typical that young children to have regular foot assessments [5] unless they have complex health conditions [6,7]. Due to this, there is little normative information about pressure distribution and changes for young children. This lack of knowledge leads to limited information on the variation in foot growth during early childhood, the impact of the different activities on change in pressure distribution, or for other industries such as footwear designers.

This means there is a need for an accessible database of pressures for all ages. Other known databases aid in understanding the variation during aging, however, these only start

from the age of three and do not include plantar pressures [8]. Bosch et al. [9] commenced curating normative values for plantar pressure data for 100 typically developing children over four years, as no information existed to determine the normality of an individual pressure pattern. Their findings provide valuable information about normal growth-related changes in foot shape and foot loading parameters (about plantar pressure). However, their information was over a limited timeframe and with limited appreciation and description for foot posture. Even with this information, there is limited transferable science from clinical research into children's footwear science to advance understanding of the interaction between the foot and the shoe in young children [10]. There are a small number of existing studies that discuss how shoes affect young children's gait [11–13]. Extremely soft-soled footwear has shown a limited effect on gait in a small number of toddlers [13]. Whereas footwear can increase velocity, step-time, and step length compared to barefoot walking in children [11,12,14]. Only one study has examined the differences in shoe flexibility on plantar pressure loading, founding a very flexible shoe showed the same plantar pressure as a barefoot pressure reading, whereas medium to low flexibility shoes gave lower plantar peak pressures [15].

Finding these differences between different shoe soles is important. However, how data has been collected is limited by age and expensive technology. Capacitive pressure sensors have been reliably used in various technologies such as touch screen devices, and digital audio players, as well as a substitute for mechanical switches. This means, there is potential for a suitable capacitive sensor to be placed within children's shoes, therefore identifying changes and patterns during foot growth. This simple method of data collection may assist in normative research dataset development or during monitoring of children with conditions known to impact foot health.

This study aimed to develop and validate an in-shoe monitoring tool using a low-cost sensor assembly that mimics pressure sensing insoles currently available on the market.

2. Materials and Methods

The overall system contains the following four components as shown in Figure 1:

- (a) Mechanical/Physical Sensor assembly—This includes the sensor pads that mainly detect the pressure change.
- (b) Connections and Routing system—This includes the connection from the sensor pads to the electronic system.
- (c) Electronics System—This includes the implementation of Arduino 101 and MPR121 which reads the capacitance of the sensor.
- (d) Data Visualisation and Data Logger—This includes LabVIEW Setup for visualization and logging of data for post-processing.

The proposed built-in-shoe foot pressure measurement system was developed through the following approach:

- (a) A qualitative analysis of existing manufacturing techniques was conducted on the fabrication of the plantar pressure sensor. A technique that would be inexpensive and still deliver a high-quality sensor was chosen.
- (b) Testing, Validation, and Calibration—After the sensor was manufactured, it was tested to see whether any connection faults were present. Data was logged and visualized using LabVIEW after the plantar pressure sensor was connected to Arduino and the MPR121 electronics unit. This test was referred to as Validation Testing 1 (VT1). Following calibration curves derived from the plantar pressure sensor system, calibration was performed. This was achieved with the repeated loading of the sensor using controlled instruments. Hounsfield H10KS Tensile Testing machine was configured as a compression testing machine to apply loads to the sensor. This stage was referred to as Validation Testing 2 (VT2).

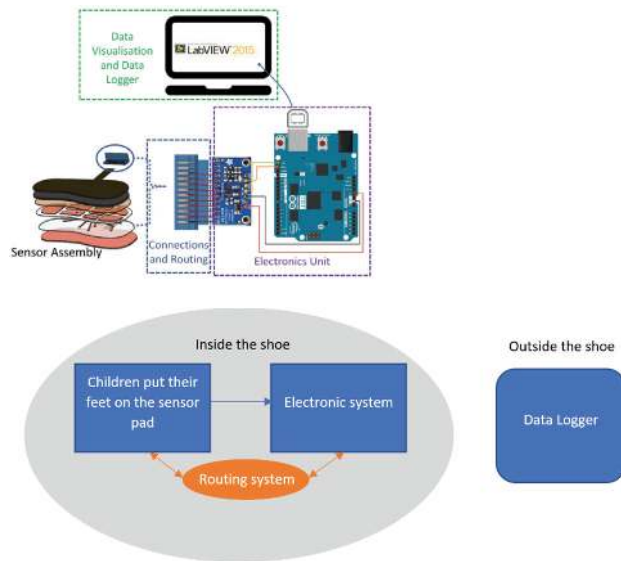


Figure 1. System layout of foot pressure measurement system.

2.1. Sensor Design Layout

This included the mechanical design and development of the sensor assembly comprising of the development of sensor topography/shapes and location, evaluation of materials and manufacturing techniques suitable for the creation of the sensors, and the assembly of the sensors. The plantar pressure sensor was aimed to measure the overall plantar pressure. To do this, the sensor needed to be fixed in a location where it is going to experience minimal movement during the child's dynamic activity. The plantar pressure sensor was placed in the Strobel area, located between the insole and the midsole of the shoes. Herein, the sensor was fixed (i.e., not loose) preventing it from moving around. Having the sensor fixed in place decreases the amount of environmental noise hence a cleaner signal. This area was also ideal as the sensors can be layered with other protective materials, such as EVA and other lining materials. In this work, the sensor was not tested while it is inside a shoe. Instead, the characteristics of the layers on top of the sensor are observed and characterized.

Pressure points of interest were informed by the patterns collected by researchers and clinicians using gold standard devices. These devices contain an array of sensors to capture all the pressure points of the plantar area and they have their own software. Masking is one of the functionalities in the accompanying software. Masking considers where plantar areas can be divided up to observe areas of interest. The masks can be setup differently depending on clinician's and researcher's interests.

2.2. Sensor Manufacturing and Routing

The sensor pads need to be connected to the main electronics unit so that they can be used as pressure sensors. Hence, different manufacturing techniques were explored and qualitatively evaluated. Some techniques are readily available, and these techniques were applied in making the sensors [16,17]. The physical sensors were visually assessed, and their effectiveness was tested by connecting the assembled sensors to MPR121 and testing whether the capacitance would change with a fingertip touch. Vinyl cutting is one of the techniques where there is a computerized blade that can cut the pattern.

The cut-out of the sensor areas shown in Figure 2 was done using a vinyl cutter, Roland CAMM-1 GS-24.



Figure 2. Cut-out of sensor areas using Vinyl Cutter.

The sheet-fed on the vinyl cutter was a copper tape sheet with plastic backing. The vinyl cutter is set up to cut with the right amount of force so that it only cuts through the copper tape and not the plastic backing. The unnecessary areas of copper tape are weeded away only keeping the wanted sensor areas. The sensor route layer was also cut using the vinyl cutting machine. These are on a separate sheet for the sensor pads. The sensor route layer and the sensor pads are connected by placing a conductive thread (Figure 3) on the sensor route and this is fed through the through hole (Figure 4) of the sensor pad sheet.

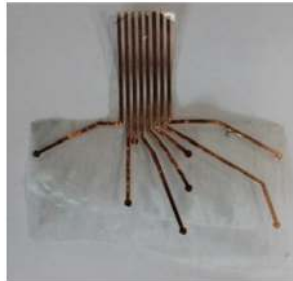


Figure 3. Sensor Route Layer.



Figure 4. Through holes in the sensor pad for the conductive thread to join with the route layer.

The thread fed through the hole is secured with conductive tape to ensure contact with the sensor pads (Figures 5 and 6).

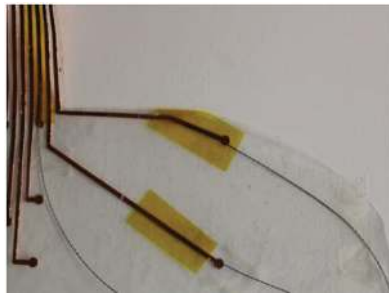


Figure 5. Route layer with conductive tape.



Figure 6. Assembled sensor installed on the outsole.

2.3. Sensor Layer Layout

The sensor layer layout consists of the manufactured sensor pads and route layer using the vinyl cutting machine as pictured in Figures 7 and 8.

The functions of these layers were as follows:

- a. Cotton Cloth–Sock liner—This served as the protection of the copper taffeta from the environment (e.g., sweaty skin). The copper taffeta tends to tarnish with moisture contact. This also functions as an insulator of the ground plane.
- b. Copper Taffeta–Ground Planes—These are present in both the top and bottom layers of the sensor. These reduce the amount of interference from the environment (such as Figure 7: Side View and Layers used for the sensor Figure 8: Isometric View of different sensor layers 127 as RF Interference). One Ground Plane was located on the bottom of the sensor to block any interference coming from the bottom. The top ground plane served as a part of the sensor: as the force applied changes, the top ground plane moves closer to the sensor pad (copper tape) changing the overall capacitance of the sensor.
- c. EVA Foam–Dielectric—This was the medium allowing space between the top ground plane and sensor plane to change. This material was chosen because a conventional shoe would have this as its shoe insole providing cushion support on the feet. In a capacitive sensor, the EVA Foam acts as the mechanical element allowing the distance

- between the two conductive plates to change when force is applied, and rebounds back to its original dimension when the force applied was removed.
- d. **Copper Tape–Sensor and Routes**—This material was chosen as the sensor and the routes due to the ease of cutting in the vinyl cutting machine. There are two layers of copper tape, and both are bonded to a plastic backing sheet. The top layer is for the sensor and the bottom layer is for the routes. This multilayer construction of the sensor is a common practice in flexible PCB Circuit printing to avoid any crosstalk and interference between the routes and the sensor pads. (Jia, 2012). The top sensor layer is contained through a hole so that it could be connected to the routes.
 - e. **Plastic backing–Insulator**—This mechanically supported the copper tape sensor and routes during the vinyl cutting process and after the copper has been cut. This plastic backing is chosen due to its flexibility and being tear resistant during the cutting process. This also serves as an insulator and prevents the top sensor pad and bottom route layers come in contact.
 - f. **Polyimide Tape–Insulator**—This mechanically supported both the copper taffeta and copper tape by keeping it in place and decreasing the stress on the copper tape's surface. This was especially useful on the joint when the copper tape was punched through the FCI Clincher, helping to resist the tear on the copper tape. This provided a protective layer from both the copper surfaces preventing them from tarnishing.

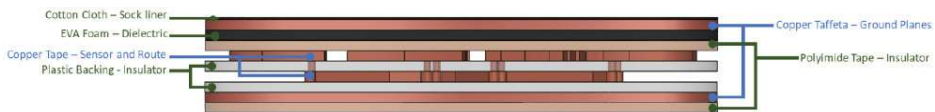


Figure 7. Side View and Layers used for the sensor.

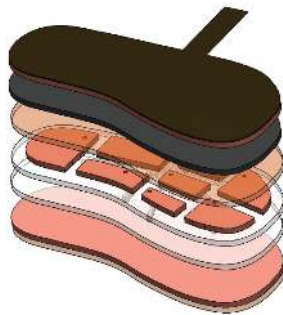


Figure 8. Isometric View of different sensor layers.

2.4. Sensor Assembly

The layers introduced above adhered to each other using pressure-sensitive adhesives. Some materials selected already contain a layer of pressure-sensitive adhesive, these being polyimide tape, plastic backing, and copper tape. The bottom copper taffeta is adhered to the plastic backing using polyimide tape. The top copper taffeta was adhered to the cotton cloth (sock liner) by an iron-on fabric glue interfacing. This is a heat-sensitive material that melts and is usually placed between two non-adhesive materials (mainly fabric) to bond them together. The top sensor layer and the bottom route layer were connected using a conductive thread which is secured on its corresponding track route by a piece of polyimide tape (Figure 5). This was then inserted in the through-hole and secured by copper tape on the top sensor layer. The completed sensor assembly was then connected using a 9-way connector to enable attachment to an electronics unit so that it could read the capacitance changes when force was applied. Moreover, this sensor is intended to build inside the shoe. This sensor assembly could adhere to any outsole or midsole of the shoe assembly.

2.5. Sensor Testing and Validation

2.5.1. VT1: Validation after Sensor Assembly

A bigger sensor was made using the same manufacturing technique as described in the section on sensor assembly for validation of the sensor assembly. To validate the sensor operation, a 50 kg person stood on the sensor and recorded the change in pressure. This was recorded by connecting the sensor to an Arduino with MPR121, this was then connected to visualization software programmed in LabVIEW. These results were compared to a static pressure recorded using an E-med pressure sensing platform. The E-med was considered a gold standard device for measuring plantar pressure. We referred to this validation testing as VT1.

2.5.2. VT2: Validation for Repeatability and Calibration

A testing jig was set up with the use of a machined foot phantom so that it can spread pressure evenly throughout the 8 sensors to validate the repeatability of the sensor. Before machining, a deflection analysis was carried out in SolidWorks to see whether three 18 mm panels of MDF would fail under 100 kg of load. 100 kg is chosen because a Factor of Safety of 2 is assumed together with a 50 kg body mass. A load simulation was done using SolidWorks Simulation studies, to verify that 54 mm thick of MDF is enough to withstand 100 kg loads. Figure 9 shows that the most displacement was 0.0157 mm. This small deflection ensured uniform load across the sensors during testing.

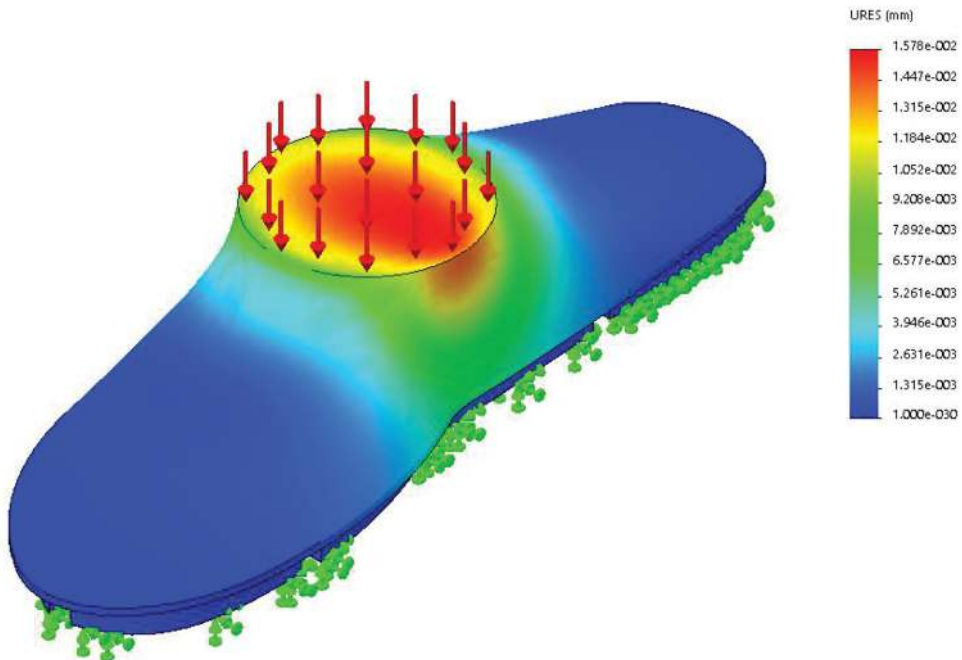


Figure 9. SolidWorks Loading with 100 kg Load.

A compression testing Machine (Hounsfield H10KS) was configured to apply a 100 kg load to apply an accurate load. This machine was programmed and paired to using LabVIEW for Data Acquisition. Figure 10 shows the testing setup for VT2.

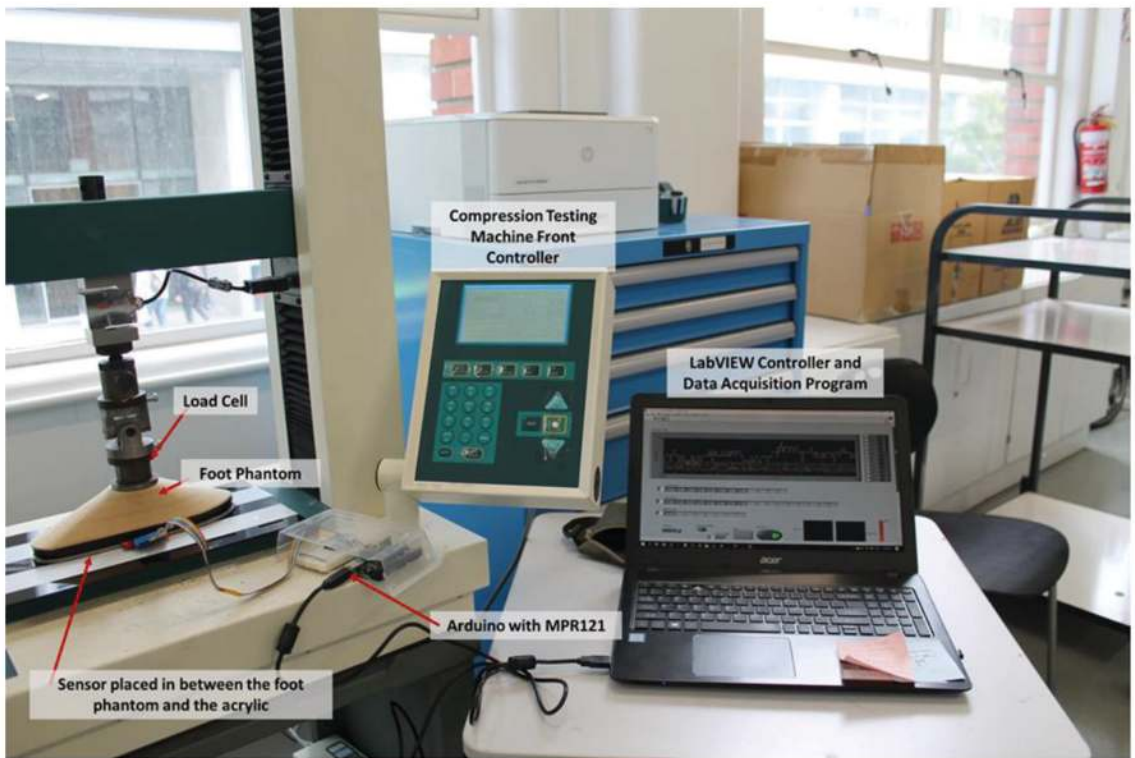


Figure 10. Testing Setup for VT2.

2.6. Test Methodology and Test Setup

Two tests were carried out for validating the repeatability of the sensors and for calibration:

- a. Loads were repeated 10 times, with the same loading and unloading rate. For this test, the load was repeated 10 times with a single dielectric. The maximum load applied for the compression testing machine was set at 1000 N, as this was an approximate equivalent of a 100 kg mass. The loading and unloading rates were set at 10 mm/min. The sensor was resting on an acrylic sheet to stop it from contacting the metal plate, at the base of the compression testing machine.
- b. Loads were repeated 11 times, with different loading and unloading rates. The load was repeated 11 times. This frequency was determined based because it was found on the previous testing, the first test would always be an offset to any remaining tests. The loading and unloading rates varied from 15–50 mm/min with increments of 5 mm/min for each test. A new dielectric was used whenever the loading rate changed and the sensor again rested on an acrylic sheet to stop direct contact with the metal plate base (Figure 11).

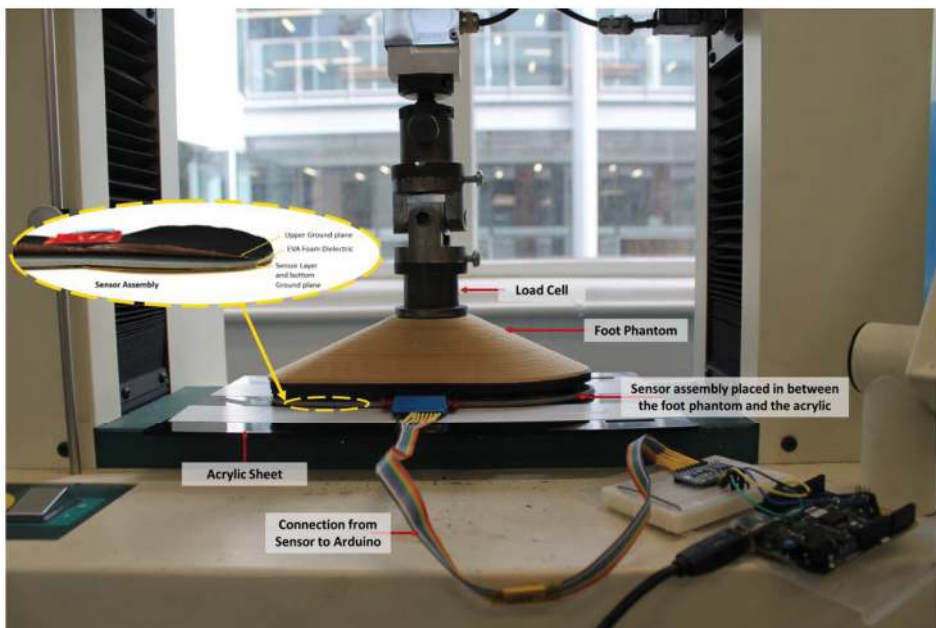


Figure 11. Test Setup for Tests 1 and 2.

3. Results

3.1. VT1: Validation after Sensor Assembly

A visual comparison was made between the results from the e-med pressure platform Figure 12, which is gold standard equipment for measuring pressure, and LabVIEW results, Figure 13. The E-med result contains more resolution due to it having more pixels and so the foot shape can be made out. The LabVIEW results are depicted as a line graph with 8 lines indicating each sensor pad. The mid-level pressure changes the most as the child grows.

The E-med and LabVIEW results were comparable to each other. The heel areas on both results show similar indications. The heel area on the E-med results shows a large red area, which means that there is a high concentration of force in that area. Similarly, the LabVIEW results show that the heel area has the highest peak, indicating a high relative capacitance reading. This means that the highest concentration of force is also in the heel area.

Although the E-med result shows a large red area on the heel, this does not mean that it is the highest force read. The E-med result also shows a magenta color located on the hallux and the forefoot area. Similarly, LabVIEW results exhibited the second-highest peaks after the heel readings for the plantar pressure sensor.

The LabVIEW results only show relative capacitance values and not the amount of force for the corresponding capacitance reading. Testing was performed for the second part of validation testing to further investigate the electromechanical behaviors of material combinations. This would then help to characterize the sensors, resulting in an equation to calculate force from capacitance.

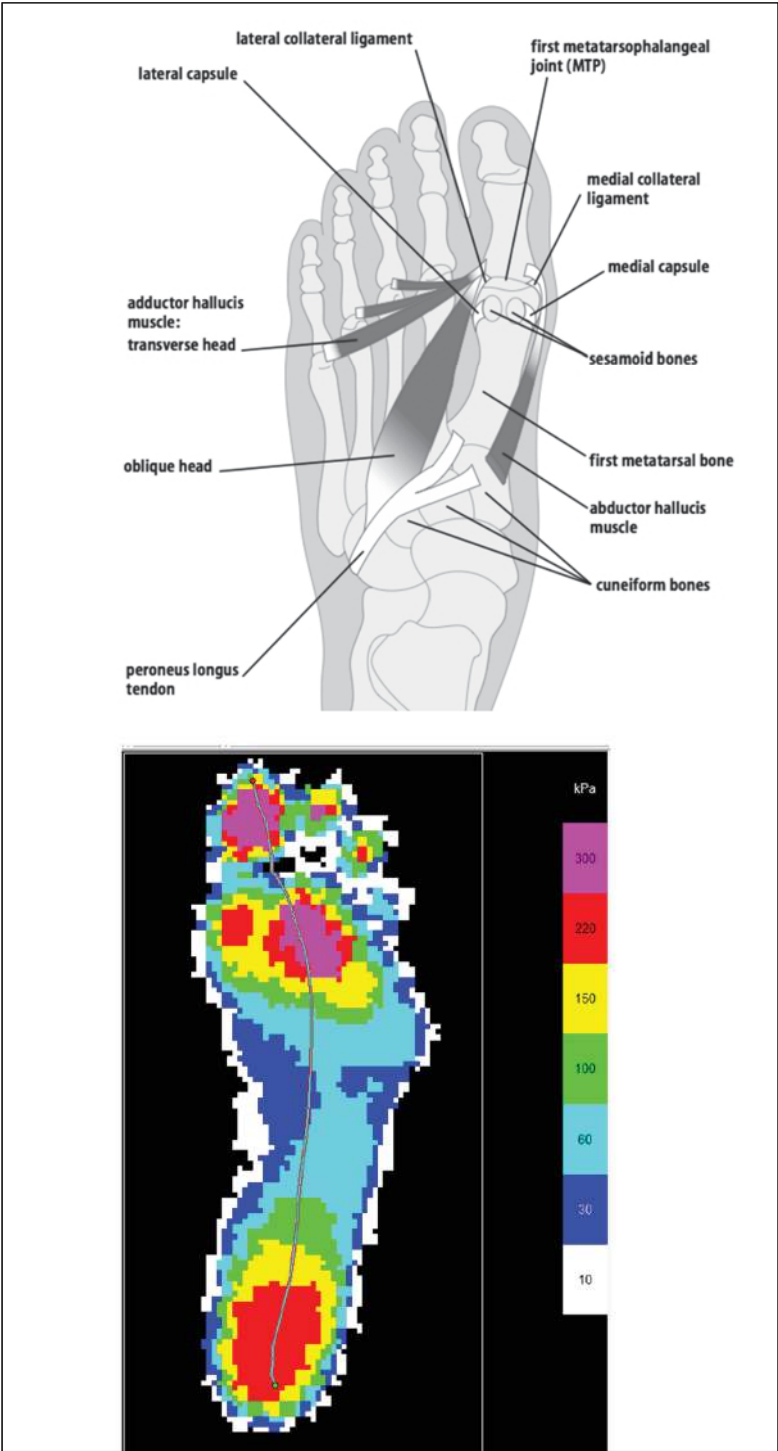


Figure 12. Result from E-med Pressure platform.

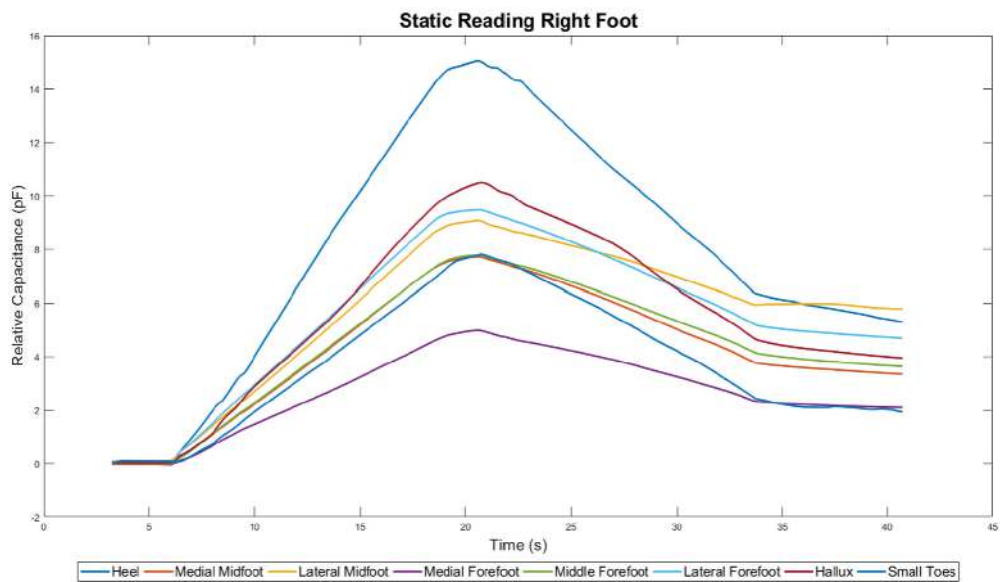


Figure 13. Line Graph of Sensor Prototype Readings from LabVIEW.

3.2. VT2: Validation for Repeatability and Calibration
3.2.1. VT2.1 Results

The graph in Figure 14 shows the time series plots of the change in absolute capacitance during loading and unloading (loading cycle). The peaks indicate one loading cycle of the sensor. The left part of the peak indicated the loading of the sensor, and the right part indicates unloading. In between these peaks, is an almost horizontal line. This indicates the time it took to reset and start the loading again.

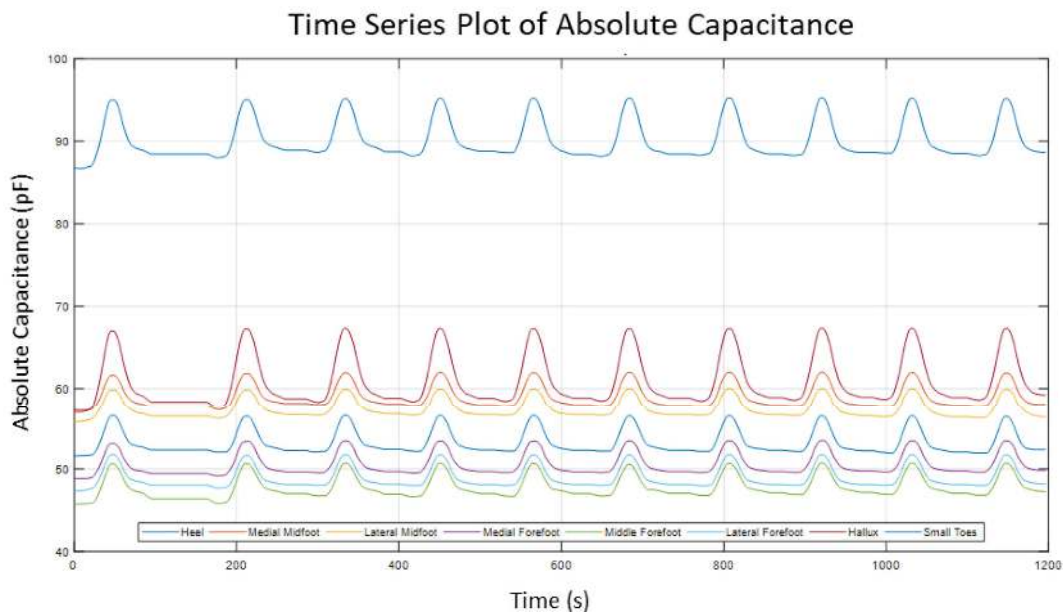


Figure 14. Time series graph of Absolute Capacitance of 8 sensor pads.

Figure 14 displays 10 peaks indicating 10 loading cycles and 8 sets of plots, representing the 8 areas in the sensor. Each sensor showed different absolute capacitance values, indicated by the plots not overlapping each other. This is due to the sensors having different areas from each other. The highest reading in the graph indicates the heel sensor which had the largest area, whereas the other areas like the midfoot, forefoot, and toes were twice as small as the heel area. Knowing that sensor areas are dissimilar to each other, results in the next set of tests scaled based on the area so that the sensors were characterized easily.

The graph in Figure 15 shows the capacitance and displacement plots of test 1. There are 8 lines representing each sensor. The hallux and the heel plot are above all other readings. The heel readings may be higher than the other sensors because it has the largest area among the other sensors. The hallux performed like the heel sensor, despite it being a smaller area and from the other end of the sensor assembly. This may be because of non-uniformity in the loading system on the foot shape.

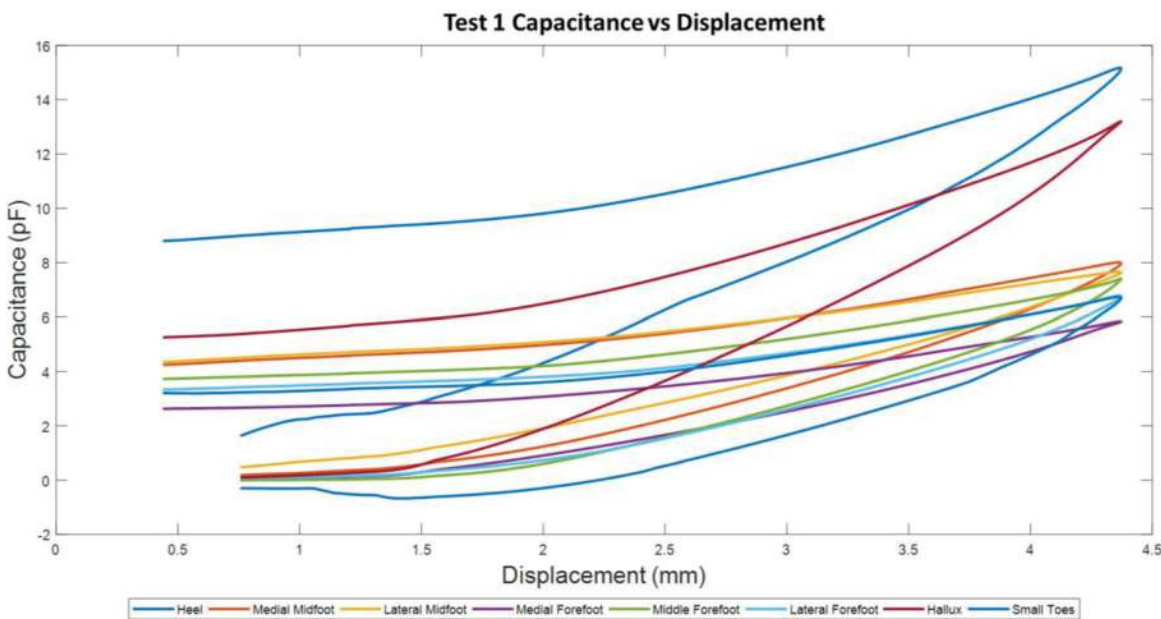


Figure 15. Test 1 Capacitance and Displacement plots for each sensor pad.

When tests 2–10 were performed, the curves now have relative capacitance readings between 0–6 pF (Figure 16). The test 1 result is disregarded because the data trend obtained by the hallux and heel does not follow the results obtained by tests 2–10 as highlighted in Figure 16.

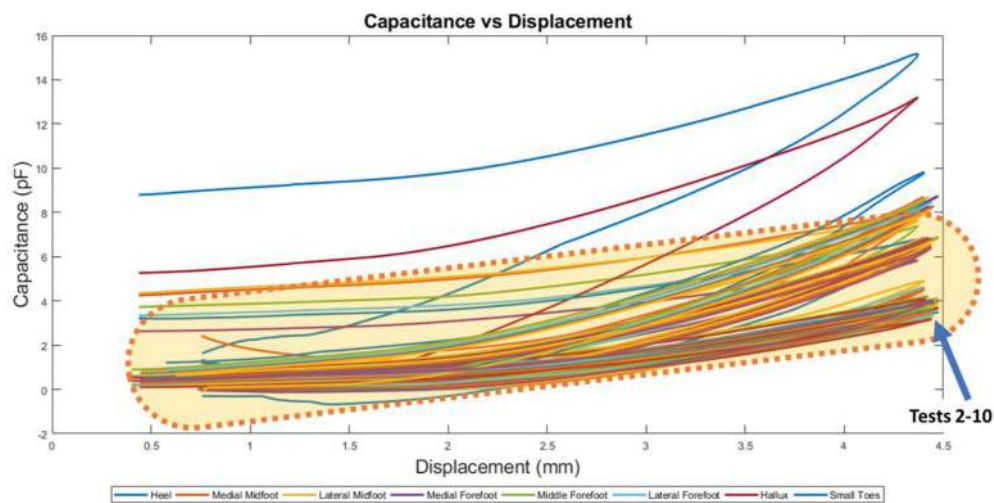


Figure 16. Capacitance and Displacement plots for 10 tests.

Comparable results were acquired and shown on the Capacitance and Force Plot (Figure 17), where the hallux and heel plot of test 1 is an outlier (Figure 18).

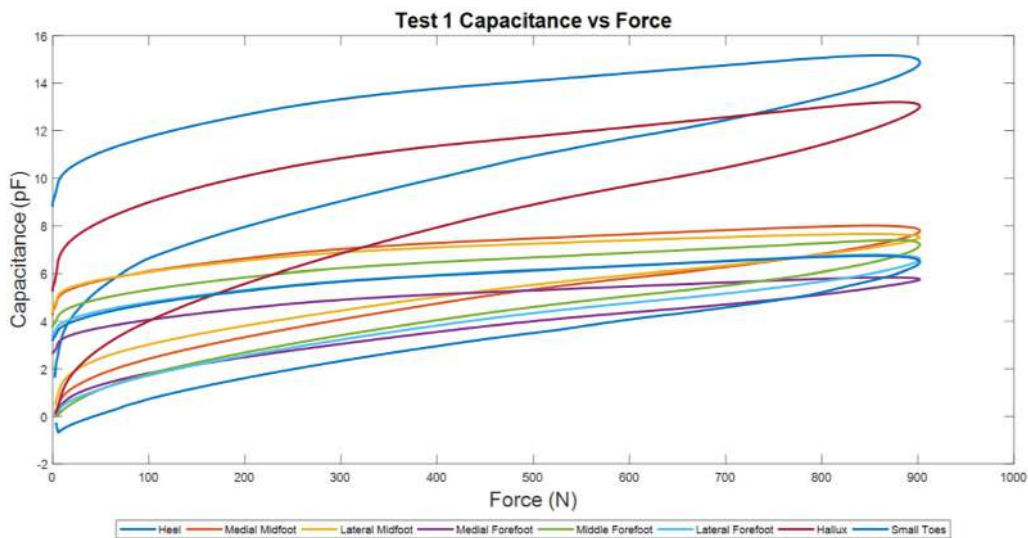


Figure 17. Test 1 Capacitance and Force Plot for each sensor.

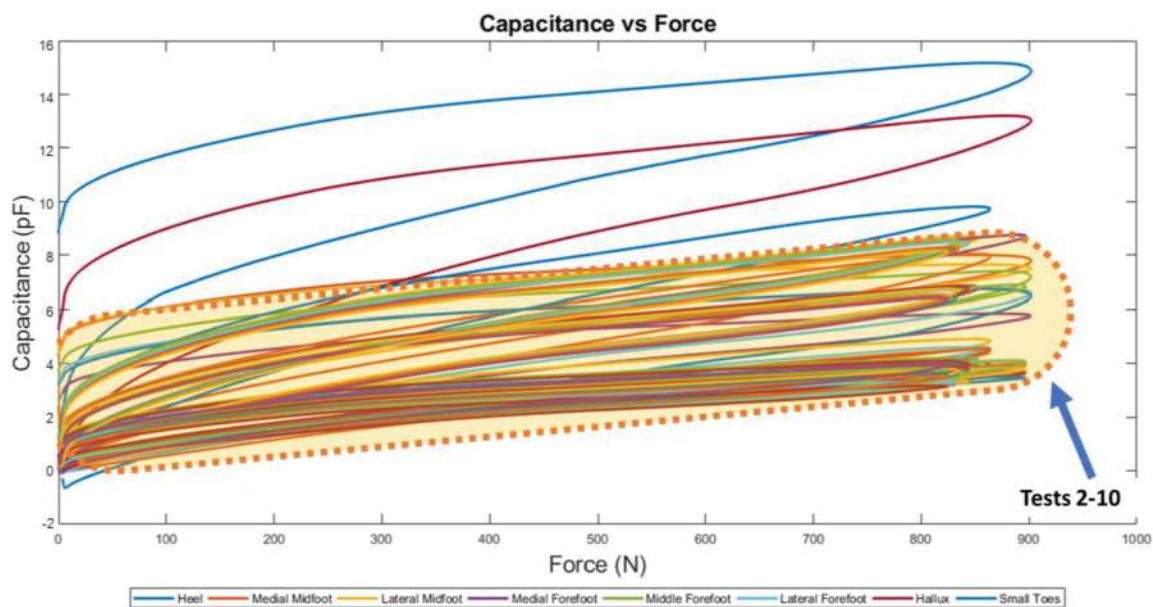


Figure 18. Capacitance and Force Plot for 10 Tests.

3.2.2. VT2.2 Results

For this part of the results, three data sets were analyzed. Each data set was analyzed in three ways—(a) Curves are fitted into each data set (Force and Displacement, Capacitance and Displacement, Capacitance and Force). Each data set has two curves, one for loading and unloading data; (b) coefficients for the equations of the curve fit in each data set are collated, and their variability was observed amongst the same tests and different load/unloading rates. These are represented as box plots; (c) mean values of each data spread in the box plot were acquired and these are statistically analyzed using linear regression techniques. The equation used to acquire Force and Displacement Coefficients is:

$$F = ad^b \tag{1}$$

where a and b are coefficients, F is force, and d is displacement.

The variability of coefficient values per loading rate showed fluctuating ranges as shown in the spread in Figure 19. In the same figure, coefficient ‘ a ’ showed broad variability at both loading and unloading rates of 30 mm/min, compared to the rest of the loading rates. The ‘ b ’ coefficients showed a less consistent range throughout the different loading and unloading rates.

The equation used to acquire Capacitance and Displacement Coefficients is:

$$C = ad^2 + bd + c \tag{2}$$

where a , b and c are coefficients, C is capacitance, and d is displacement.

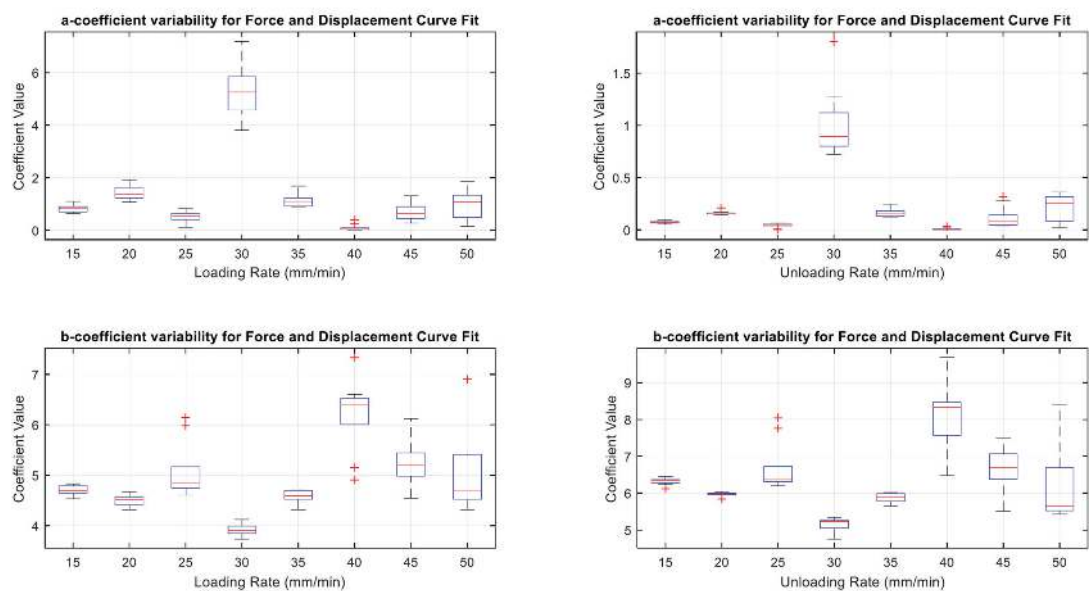


Figure 19. Box Plots showing coefficient variability for Equation (1).

3.3. Coefficient Variability for Capacitance and Displacement Curve Fit

The variability of all the coefficients at all loading rates in Figure 20 showed an inconsistent spread of coefficient values. However, the coefficient variability during unloading was more consistent. From 15–35 mm/min unloading rate, coefficients showed narrow variability, indicated by a narrow height of the box whereas, on higher unloading rates between 40–50 mm/min, coefficients showed a larger spread of values.

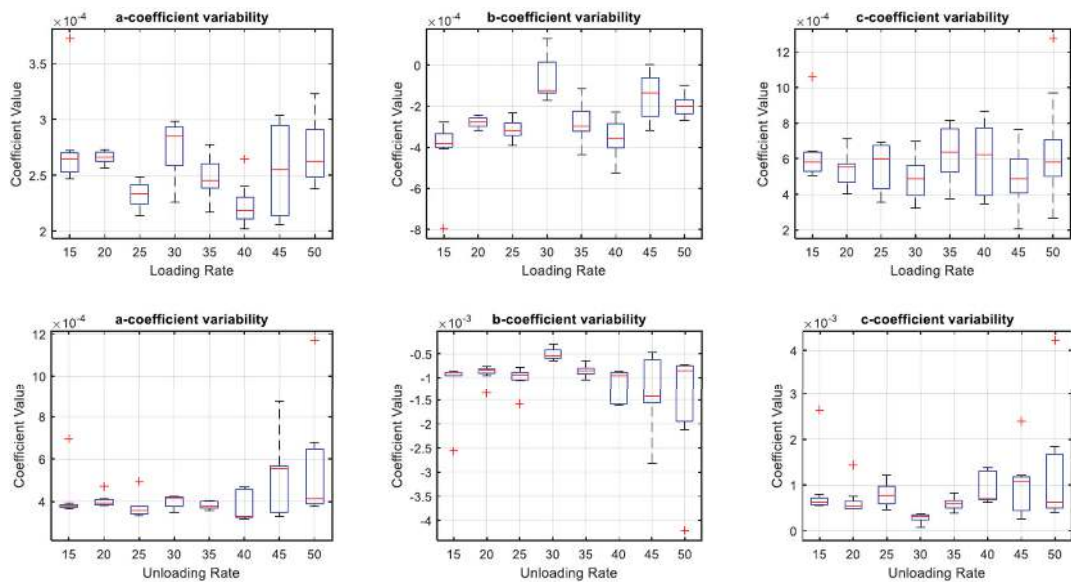


Figure 20. Box Plots of Coefficient Variability for Equation (2).

The equation used to acquire Capacitance and Force Coefficients is:

$$C = aF^b \quad (3)$$

where a and b are coefficients, C is capacitance and F is force.

3.4. Coefficient Variability for Capacitance and Force Curve Fit

The coefficient variability amongst both loading and unloading rates are inconsistent in Figure 21, this is indicated by the uneven heights of the boxes.

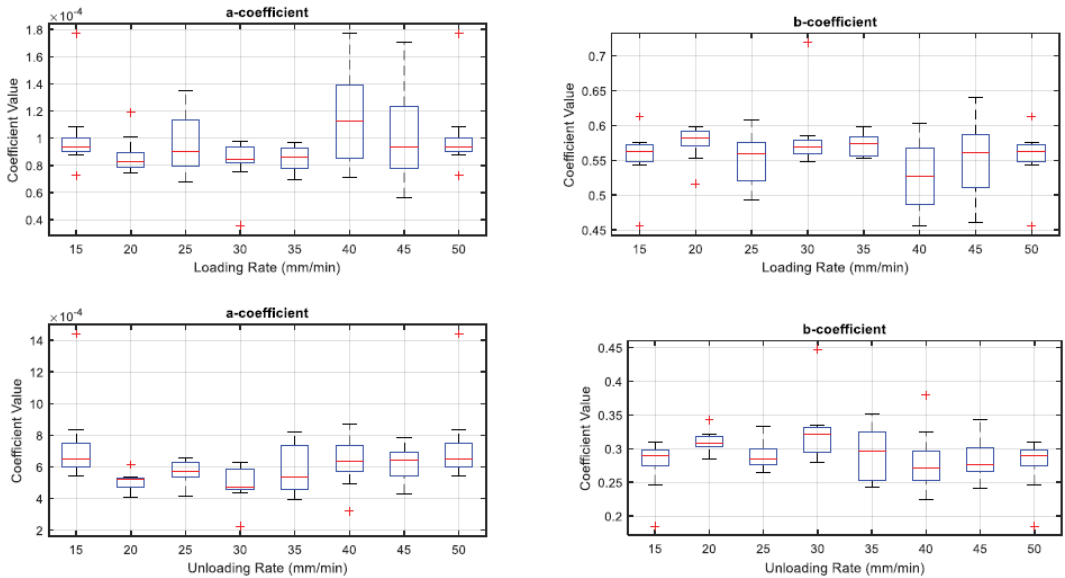
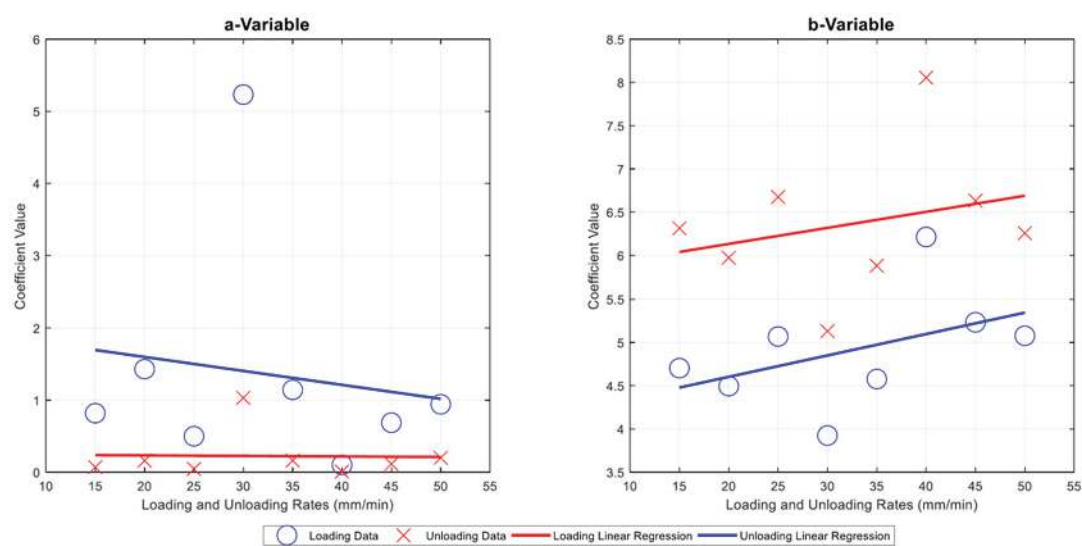


Figure 21. Box Plots for coefficient variability of Equation (3).

The results in Figures 22–24 show the relevance of the strain rates in the role of curve fits for Equations (1)–(3). Statistical analysis results were derived from MATLAB, and showed no p values less than 0.05, indicating that the loading and unloading rates do not affect the curve fits done on Force—Displacement, Capacitance—Displacement, and Capacitance—Force data sets with statistical significance.

Figures 22–24 show the plots of mean coefficient values for both loading and unloading data. There is a difference in the values of slopes between loading and unloading. This is indicated by the parameter ‘ m ’ (slope of the linear regression) shown in Figures 22–24. Additionally, there is an offset between the two linear regression lines indicated by different ‘ c ’ values (linear regression intercept) in Figures 22–24.

As the loading and unloading rates do not affect the coefficients derived from the curve fits, mean values of the coefficients shown in box plots of Figures 19–21 were acquired to obtain a suitable set of calibration curves. Figure 25 shows the curve fit graphs obtained using the mean coefficient values for each data set.



Linear Regression Model for evaluation of curve fit coefficients for $F = ad^b$ <i>m</i> being the slope of the linear regression line, <i>c</i> being the intercept of the linear regression line.			
	<i>m</i> (SE)	<i>c</i> (SE)	<i>p</i>
a_{load}	−0.019 (0.05)	1.99(1.84)	0.73
a_{unload}	−0.00074 (0.01)	0.25(0.38)	0.95
b_{load}	0.024(0.011)	4.11(0.69)	0.26
b_{unload}	0.019(0.027)	5.76(0.93)	0.52

Figure 22. Plot of mean values for Figure 19 with Linear Regression Model Fitted.

Linear Regression Model for evaluation of curve fit coefficients for $C = ad^2 + bd + c$ <i>m</i> being the slope of the linear regression line, <i>c</i> being the intercept of the linear regression line.			
	<i>m</i> (SE)	<i>c</i> (SE)	<i>p</i>
a_{load}	$-2.75 \times 10^{-7}(3.61 \times 10^{-7})$	$2.64 \times 10^{-4}(2.17 \times 10^{-5})$	0.68
a_{unload}	$3.66 \times 10^{-6}(1.77 \times 10^{-6})$	$3.07 \times 10^{-4}(6.10 \times 10^{-5})$	0.08
b_{load}	$4.29 \times 10^{-6}(3.41 \times 10^{-6})$	$-3.93 \times 10^{-4}(1.17 \times 10^{-4})$	0.25
b_{unload}	$-1.33 \times 10^{-5}(8.43 \times 10^{-6})$	$-6.02 \times 10^{-4}(2.9 \times 10^{-4})$	0.17
c_{load}	$4.15 \times 10^{-7}(2.05 \times 10^{-6})$	$5.58 \times 10^{-6}(7.06 \times 10^{-5})$	0.85
c_{unload}	$1.18 \times 10^{-5}(8.26 \times 10^{-6})$	$4.02 \times 10^{-4}(2.84 \times 10^{-4})$	0.20

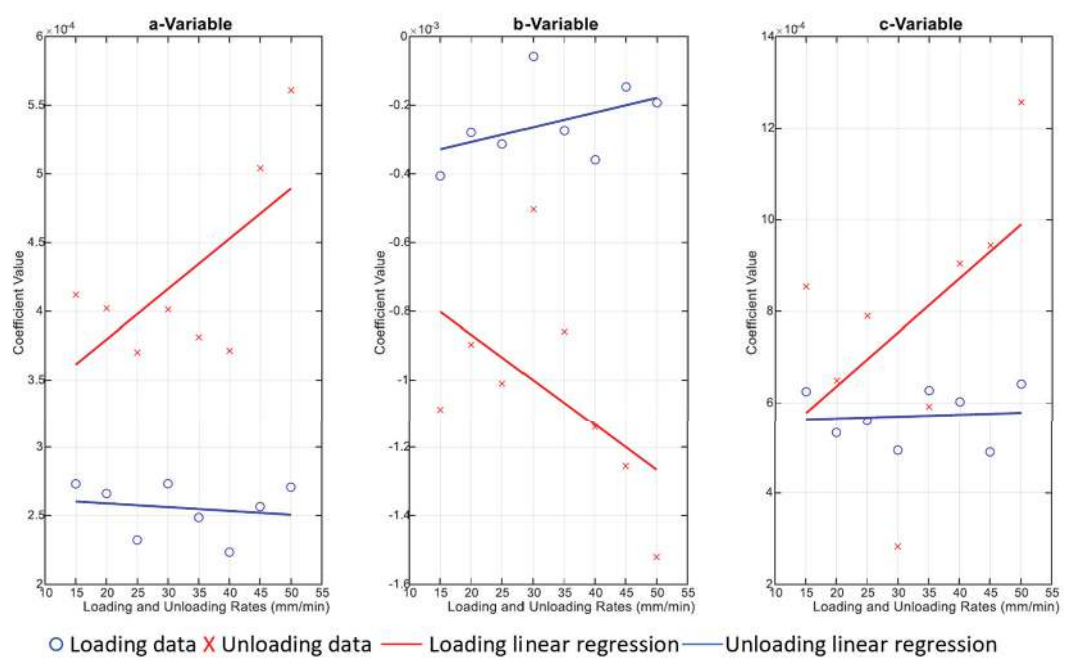


Figure 23. Plot of mean Values from Figure 20 with a fitted linear regression model.

Linear Regression coefficients for a mean of the coefficient values in Figure 24.

Linear Regression Model for evaluation of curve fit coefficients for $C = aF^b$ <i>m being the slope of the linear regression line, c being the intercept of the linear regression line.</i>			
	<i>m (SE)</i>	<i>c (SE)</i>	<i>p</i>
a_{load}	$3.66 \times 10^{-7} (3.61 \times 10^{-7})$	$8.47 \times 10^{-5} (1.24 \times 10^{-5})$	0.35
a_{unload}	$2.09 \times 10^{-6} (3.03 \times 10^{-6})$	$5.43 \times 10^{-4} (1.05 \times 10^{-4})$	0.51
b_{load}	$-5.10 \times 10^{-4} (5.67 \times 10^{-4})$	0.56 (0.02)	0.261
b_{unload}	$-4.48 \times 10^{-4} (5.60 \times 10^{-4})$	0.31(0.02)	0.45

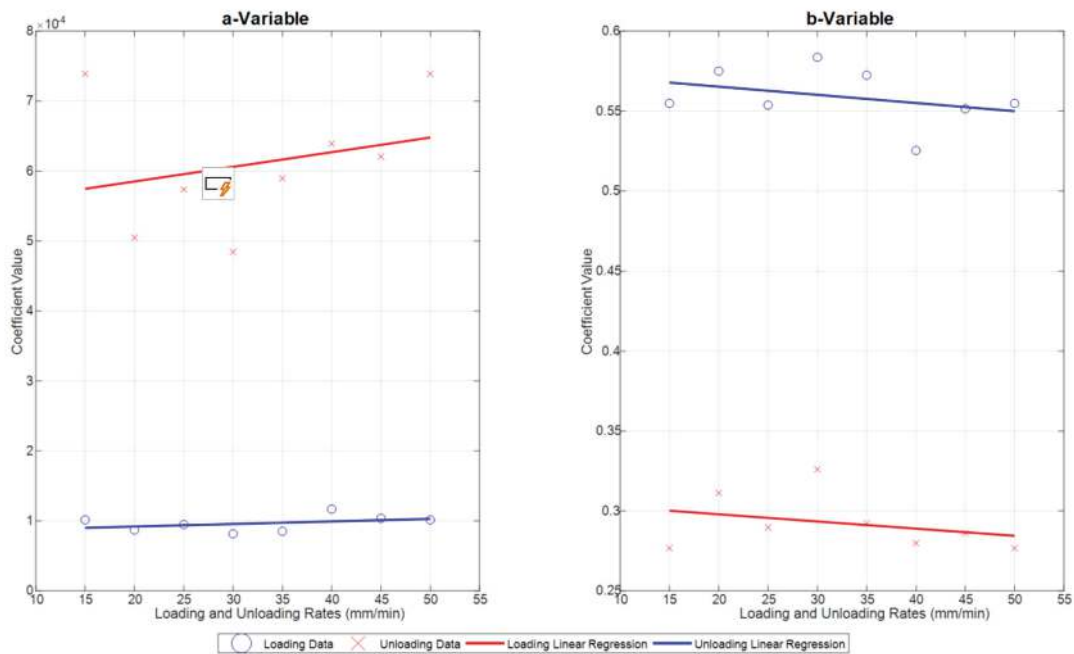


Figure 24. Plot of mean Values from Figure 21 with a fitted linear regression model.

Referring to Figure 25 there is a significant difference between the coefficients of loading and unloading parameters. It can be observed in the graph there is space between the loading and unloading curves in all the curve fits. This indicated a hysteretic behavior due to the EVA Foam being an elastomeric material. The force and displacement showed little space between the load and unload curves. These curves crossed over each other at the point when displacement is at 4 mm and the force is at 400 N. This is an artifact of the curve fitting of different data sets. The same cross-over feature is starting to show in higher displacement on the capacitance vs. displacement and force graphs.

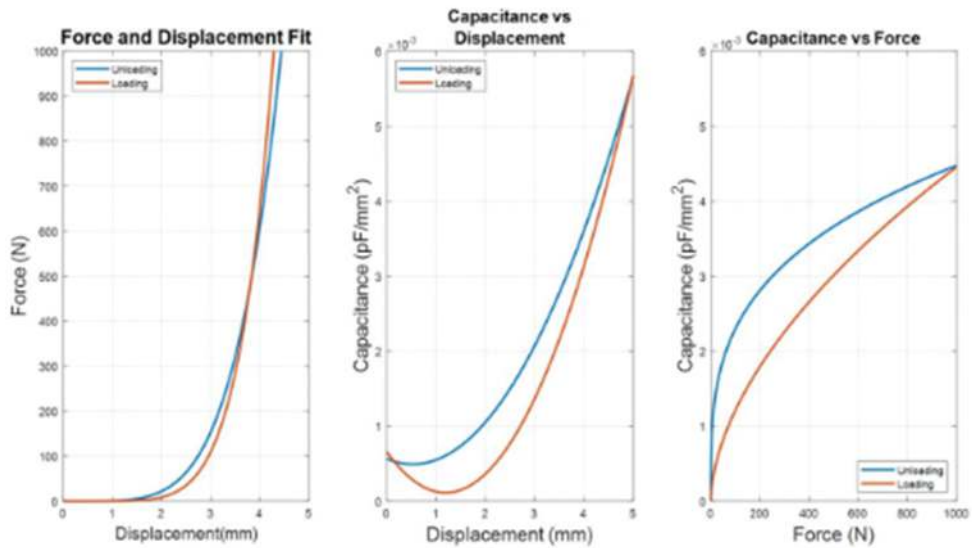


Figure 25. Calibration Curves for different data sets.

Although the best fit was chosen for each loading and unloading curve, the mean values of the coefficients do not necessarily represent the best curve fit for the overall data set. A constraint can be given when carrying out the curve fitting so that the loading curve would always be beneath the unloading curve within the operating range.

Both capacitances vs. displacement and capacitance vs. force showed a large space between the loading and unloading curves, even though unloading lines were expected to overlap and be in the same place as the loading lines for Figure 25. This behavior indicated that the readings from the sensor during loading are not the same during the unloading phase. The hysteretic condition is more evident when capacitance is plotted against force and displacement.

The capacitance vs. displacement curve fits took a quadratic form. A characteristic of this curve is that both loading and unloading curves show a dip in the graph between the displacement of 1 and 2 mm. They then show an almost constant slope with a positive linear relationship from the displacement of 2 to 5 mm (Figure 25). This could reflect the behavior of the EVA foam: when the EVA foam is in the early compression stage, the cell walls of the closed cells are bent but not completely collapsed. The uncompressed thickness of the EVA Foam is 3 mm, although this does not necessarily mean that the allowable travel of the head displacement of the load cell is 3 mm. At 3 mm, the cell walls could have collapsed or completely fractured. Beyond 3 mm, cell walls have crushed together, leaving a non-foam EVA material.

The capacitance and force relationship showed a power function. Both loading and unloading curves in this graph have a distinct shape, where the loading curve is less curved compared to the unloading curve. At low and high force readings the curves converge a point. Visually, this graph contained more area between the two curves depicting higher hysteretic behavior compared to force and displacement curves. This could also be due to the elastomeric behavior of the EVA Foam but may include electrical phenomena as well.

The hysteretic phenomenon exhibited in Figure 25 is usually present when an elastomeric foam is under compression loads. This is due to the densification of the materials. Densification is the collapse of cell walls throughout the materials, causing the load to distribute against opposite cell walls (i.e., cell walls are now touching each other bearing the load) (Gibson, 2005). Figure 26 shows a typical stress strain curve of an elastomer, where in the beginning of the curve, a linear stress strain relationship called linear elasticity is depicted. It is determined by bending of the cell walls of the closed cell. The plateau

stage corresponds to gradual collapse of the cell wall, through elastic buckling or plastic yielding. The cell wall collapse is dependent on the material.

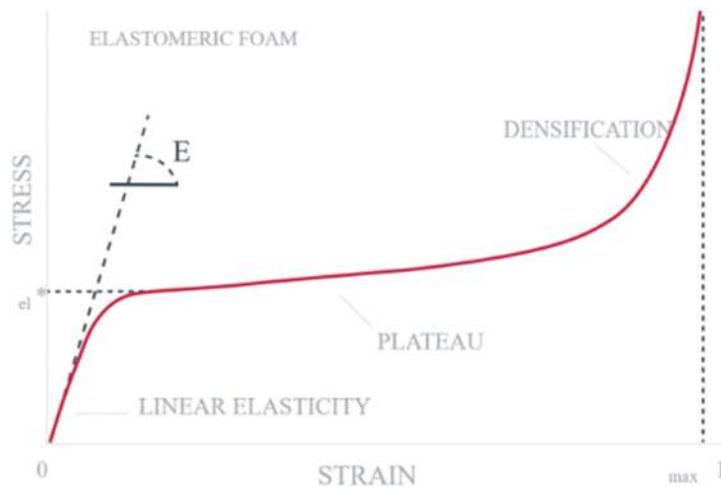


Figure 26. Stress vs. Strain Curve of an elastomeric foam (Carbon (R), 2017).

For this investigation using EVA closed cell foam, a similar non-linear phenomenon can be seen in Figure 25. The curve fits and graphs in Figure 25 may not be representative of the whole data set but are a representation of behavior during the loading and unloading stages.

4. Discussion

The final prototype for built-in-shoe pressure sensing was achieved through iterative design. Suitable manufacturing techniques were selected and the vinyl cutting technique on copper tape and plastic backing was the cheapest and most reliable option available. VT1 mimicked the pressure sensing characteristic of the E-med pressure platform. While VT2 is comprised of two types of testing, one for repeatability and the other with calibration. VT2.1 showed satisfactory results with 10 consistent peaks when the same load was applied 10 times. VT 2.2 determined the impact of different loading rates on calibration curves. These loading rates represented different landing rates for footfalls. The loading rates did not make a difference, with coefficient values of different parameters not showing a consistent range. Moreover, loading rates also were not found to not affect the sensor readings and were not correlated with each other.

Creating this new technology may be the first step forward to addressing issues clinicians, researchers, or shoe designers face in accessing shoe pressure data. The next step will be modeling its use in footwear and user testing with children technical results may not solve immediate problems that parents face, but this research increased foot health and footwear literacy.

This research is limited by design. It presents the feasibility of developing low-cost in-shoe sensor technology. This means it does not present clinicians and researchers with an immediately usable tool. In future research, several changes will support future design. In particular, the collapse of cell walls of the EVA could be modeled to better predict and anticipate the capacitance behavior as it is dependent on force and displacement. However, it presents new knowledge regarding new and innovative technologies that are available that can solve data collection issues in the future.

5. Conclusions

For future work, more accurate material modeling to better predict material and sensor properties would be helpful. This includes modeling the return of EVA foam during cyclic

loading and unloading. This can be combined with the study of electromagnetic fields of capacitive plates sandwiched between soft materials.

Uniform sensor shapes should be implemented, this would make it easier to calibrate the sensors, as they should all experience the same amount of change during the loading cycle.

Author Contributions: S.D.G.—Conceptualization, methodology, software, validation; A.L. and C.W.—Data Curation, supervision, visualization; A.K. and G.A.—Data curation, writing—review and editing, visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Ministry of Business, Innovation and Employment.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Montagnani, E.; Price, C.; Nester, C.; Morrison, S.C. Dynamic characteristics of foot development: A narrative synthesis of plantar pressure data during infancy and childhood. *Pediatr. Phys. Ther.* **2021**, *33*, 275–282. [CrossRef] [PubMed]
2. Hallemans, A.; De Clercq, D.; Van Dongen, S.; Aerts, P. Changes in foot-function parameters during the first 5 months after the onset of independent walking: A longitudinal follow-up study. *Gait Posture* **2006**, *23*, 142–148. [CrossRef] [PubMed]
3. Gould, N.; Moreland, M.; Alvarez, R.; Trevino, S.; Fenwick, J. Development of the child's arch. *Foot Ankle* **1989**, *9*, 241–245. [CrossRef] [PubMed]
4. Machado, D.B.; Hennig, E. The influence of daily activity movement patterns on the in-shoe plantar pressure distribution of children. In Proceedings of the 4th Symposium on Footwear Biomechanics, Canmore, AB, Canada, 5–7 August 1999.
5. Evans, A.M. Screening for foot problems in children: Is this practice justifiable? *J. Foot Ankle Res.* **2012**, *5*, 1–10. [CrossRef] [PubMed]
6. Merker, J.; Hartmann, M.; Haas, J.-P.; Schwirtz, A. Combined three-dimensional gait and plantar pressure analyses detecting significant functional deficits in children with juvenile idiopathic arthritis. *Gait Posture* **2018**, *66*, 247–254. [CrossRef] [PubMed]
7. Cardoso, J.; de Baptista, C.R.A.; Sartor, C.D.; Elias, A.H.N.; Júnior, W.M.; Martinez, E.Z.; Sacco, I.C.; Mattiello-Sverzut, A.C. Dynamic plantar pressure patterns in children and adolescents with Charcot-Marie-Tooth disease. *Gait Posture* **2021**, *86*, 112–119. [CrossRef] [PubMed]
8. McKay, M.J.; Baldwin, J.N.; Ferreira, P.; Simic, M.; Vanicek, N.; Hiller, C.E.; Nightingale, E.J.; Moloney, N.A.; Quinlan, K.G.; Pourkazemi, F.; et al. 1000 Norms Project: Protocol of a cross-sectional study cataloging human variation. *Physiotherapy* **2016**, *102*, 50–56. [CrossRef] [PubMed]
9. Bosch, K.; Gerss, J.; Rosenbaum, D. Preliminary normative values for foot loading parameters of the developing child. *Gait Posture* **2007**, *26*, 238–247. [CrossRef] [PubMed]
10. Morrison, S.C.; Price, C.; McClymont, J.; Nester, C. Big issues for small feet: Developmental, biomechanical and clinical narratives on children's footwear. *J. Foot Ankle Res.* **2018**, *11*, 1–5. [CrossRef] [PubMed]
11. Cranage, S.; Perraton, L.; Bowles, K.A.; Williams, C. The impact of shoe flexibility on gait, pressure and muscle activity of young children. A systematic review. *J. Foot Ankle Res.* **2019**, *12*, 55. [CrossRef] [PubMed]
12. Cranage, S.; Perraton, L.; Bowles, K.-A.; Williams, C. A comparison of young children's spatiotemporal gait measures in three common types of footwear with different sole hardness. *Gait Posture* **2021**, *90*, 276–282. [CrossRef] [PubMed]
13. Williams, C.; Kolic, J.; Wu, W.; Paterson, K. Soft soled footwear has limited impact on toddler gait. *PLoS ONE* **2021**, *16*, e0251175. [CrossRef] [PubMed]
14. Cranage, S.; Perraton, L.; Bowles, K.-A.; Williams, C. A comparison of young children's spatiotemporal measures of walking and running in three common types of footwear compared to bare feet. *Gait Posture* **2020**, *81*, 218–224. [CrossRef] [PubMed]
15. Hillstrom, H.J.; Buckland, M.A.; Slevin, C.M.; Hafer, J.F.; Root, L.M.; Backus, S.I.; Kraszewski, A.P.; Whitney, K.A.; Scher, D.M.; Song, J.; et al. Effect of shoe flexibility on plantar loading in children learning to walk. *J. Am. Podiatr. Med. Assoc.* **2013**, *103*, 297–305. [PubMed]
16. Wu, Z.; Zhang, B.; Zou, H.; Lin, Z.; Liu, G.; Wang, Z.L. Multifunctional Sensor Based on Translational-Rotary Triboelectric Nanogenerator. *Adv. Energy Mater.* **2019**, *9*, 1901124. [CrossRef]
17. Wu, Z.; Cheng, T.; Wang, Z.L. Self-Powered Sensors and Systems Based on Nanogenerators. *Sensors* **2020**, *20*, 2925. [CrossRef] [PubMed]



Article

Remotely Powered Two-Wire Cooperative Sensors for Biopotential Imaging Wearables

Olivier Chételat ^{1,*}, Michaël Rapin ¹, Benjamin Bonnal ¹, André Fivaz ¹, Josias Wacker ¹ and Benjamin Sporrer ²

¹ CSEM, Electronics/Systems/Digital Health, Jaquet-Droz 1, 2002 Neuchâtel, Switzerland

² CSEM, Integrated & Wireless Systems/System-on-Chip/ASIC for the Edge, Technopark, Technoparkstrasse 1, 8005 Zürich, Switzerland

* Correspondence: olivier.chetelat@csem.ch

Abstract: Biopotential imaging (e.g., ECGi, EEGi, EMGi) processes multiple potential signals, each requiring an electrode applied to the body's skin. Conventional approaches based on individual wiring of each electrode are not suitable for wearable systems. Cooperative sensors solve the wiring problem since they consist of active (dry) electrodes connected by a two-wire parallel bus that can be implemented, for example, as a textile spacer with both sides made conductive. As a result, the cumbersome wiring of the classical star arrangement is replaced by a seamless solution. Previous work has shown that potential reference, current return, synchronization, and data transfer functions can all be implemented on a two-wire parallel bus while keeping the noise of the measured biopotentials within the limits specified by medical standards. We present the addition of the power supply function to the two-wire bus. Two approaches are discussed. One of them has been implemented with commercially available components and the other with an ASIC. Initial experimental results show that both approaches are feasible, but the ASIC approach better addresses medical safety concerns and offers other advantages, such as lower power consumption, more sensors on the two-wire bus, and smaller size.

Keywords: biopotential imaging; body surface potential; active electrode; dry electrode; cooperative sensor; wearables; medical device

Citation: Chételat, O.; Rapin, M.; Bonnal, B.; Fivaz, A.; Wacker, J.; Sporrer, B. Remotely Powered Two-Wire Cooperative Sensors for Biopotential Imaging Wearables. *Sensors* **2022**, *22*, 8219. <https://doi.org/10.3390/s22218219>

Academic Editor: José Machado
Da Silva

Received: 5 September 2022

Accepted: 25 October 2022

Published: 27 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The technology of *biopotential* measurements has been known for decades [1] and is widely used to measure ECG (electrocardiogram), EEG (electroencephalogram), EMG (electromyogram), etc. Today, most products are still based on adhesive *gel electrodes* applied to specific areas of the body's skin and connected in a star arrangement to a *central unit* (e.g., a recorder or monitor) by shielded cables. Another solution, less commercialized but scientifically well known, features *active electrodes* [2]. Active electrodes have amplification electronics at the electrode to achieve a high input impedance and low output impedance. Although they do not require shielded cables, each electrode still requires at least two wires to be connected to the central unit in a star arrangement [3].

Implementing biopotential sensors in *medical wearables* is difficult for a few fundamental reasons. One of them results from the preferential use of *dry electrodes*. For dry electrodes to provide good quality signals, a higher input impedance of the amplification circuit (cables and amplifier) is required [4–6]. Active electrodes are a good solution to meet this requirement, but they must be powered. However, powering them from the central unit via their leads is not trivial in the context of medical devices. The basic safety standards [7–9] for MEDICAL ELECTRIC EQUIPMENT (words in small capitals have a defined meaning in the standards) require a maximum PATIENT LEAKAGE CURRENT of 10 μ A d.c. (or 10 mV across the 1 k Ω taken by the standards for skin) for ECG medical devices. In addition, the standards specify two MEANS OF PATIENT PROTECTION [7]. Reliable DOUBLE

INSULATION of conductive tracks in wearables is difficult to achieve, especially due to the possible presence of *body fluid* (e.g., sweat, urine, etc.). To be considered as one mean, for WORKING VOLTAGES lower than 60 V, solid insulation must pass the test of dielectric strength at a voltage of 500 V rms during 1 min under the worst expected condition (i.e., with body fluids and at the end of the EXPECTED SERVICE LIFE) [7]. Therefore, for wearable devices, other MEANS OF PATIENT PROTECTION than solid insulation, such as the electronic detection and active limitation of tiny PATIENT LEAKAGE CURRENTS, are desirable.

Another reason that wearables are difficult to design, especially for devices requiring many electrodes (e.g., >100) such as for *electrographic imaging* (e.g., ECGi [10], EEGi [11], EMGi [12]), is the ‘star arrangement’ of conventional wiring approaches. In addition to the complexity of routing many cables, connecting them to the central unit where they meet results in expensive and cumbersome *connectors*, which is a major integration challenge for waterproof and reliable wearables. To solve these problems, an architecture where each electrode includes electronics to be interfaced in parallel with a bus is attractive because such an architecture is virtually independent of the number of electrodes. Buses with a small number of lines, such as one or two at most, are the most desirable, but more difficult because the bus must perform several functions (potential reference, current return, synchronization, communication, and power supply). *Cooperative sensors* [13–19], defined as active electrodes connected to a parallel bus of up to two wires, have been proposed to address these challenges, except for the power supply. Adding the power supply function—without interfering with the other functions and addressing the medical safety issue in wearables—is the main contribution of this paper.

Section 2 presents the previously known cooperative sensor technique characterized by a dedicated power supply (battery) per sensor and a bootstrap circuit allowing high input impedance. This section is essential to understand the functions that must remain unaltered despite the introduction in Section 3 of the power supply on the same wires. Two solutions are presented, one called ‘*Legacy approach with 500 Hz powering and off-the-shelf components*’ (because directly built on the cooperative sensors of Section 2) and the other named ‘*Approach addressing the safety issue with powering at 1 MHz and ASIC for 250 sensors*’. Section 4 provides implementation details and experimental results for both approaches. The paper ends with a conclusion (Section 5).

2. State-of-the-Art Cooperative Sensors

Active electrodes connected by *up to two wires* in a parallel bus arrangement are called *cooperative sensors*—‘cooperative’ because co-operation of at least two sensors, each measuring one biopotential, is required to obtain a difference of potential (such as how two fingers, at least, must cooperate to pick up a golf ball). Compared to conventional approaches [1–3,20,21], cooperative sensors benefit from a *parallel bus arrangement*, which contributes to the scalability of the system. Moreover, unlike the multi-wire bus of direct multiplexing [22], the complexity of their connection is reduced to a minimum (only two wires for all functions).

2.1. Basic Circuit and Their Interconnections

Figure 1 shows a patented generic mechanism [13,14] for synchronization and control of the cooperative sensors and for transferring the acquired signal to the central unit. To measure the voltage e , a *current electrode* (left in Figure 1), a *potential electrode* for potential reference (middle in Figure 1), and one potential electrode per independent biopotential channel (right in Figure 1) are required. The current electrode is often referred to in the classical technique [1] as *right-leg electrode* and its purpose is to provide a path for currents capacitively coupled to the electronics (central unit and cables) from the environment (e.g., 50 Hz or low-frequency currents resulting from motion in the earth’s electric field). The central unit sends synchronization information and other commands simultaneously to all cooperative sensors with the voltage source U . The received signal is picked up by the cooperative sensors with the voltage across their current source, i.e., the voltage

between the two bus wires, and feeds the clock recovery and sync block S. The cooperative sensors use their current source to communicate their measured potential v_i (or other signals) to the central unit by means of a modulation M. The central unit receives the composite information from all the sensors by detecting the current in one of the bus wires. The demodulation D allows to recover the individual signals. Note that there are many possible modulations M. Amplitude modulation of a frequency carrier is a simple example. Digitization combined with phase shift keying could be another. The dark gray box symbolizes the power supply (e.g., a battery) of the cooperative sensors.

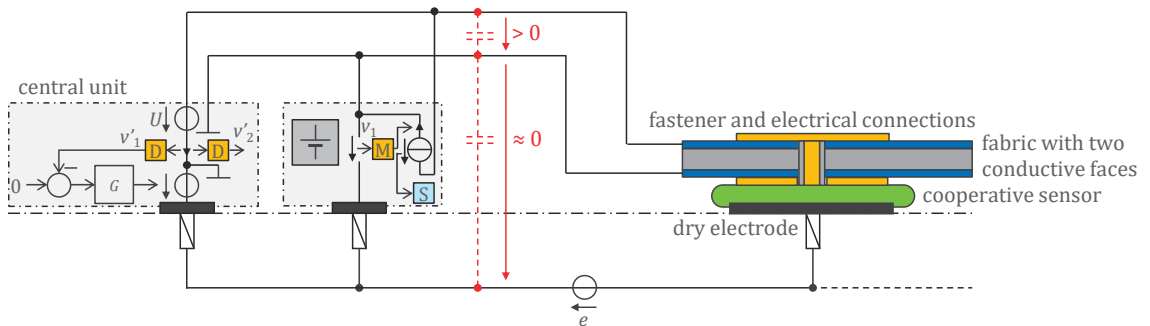


Figure 1. Cooperative sensors are active electrodes equipped with electronic circuits that allow them to be connected to a parallel bus of up to two wires. They are synchronized by the central unit to which they transmit their measured biopotential. For devices that do not need to be defibrillator-proof, the 2-wire bus can be made of a fabric with both sides made conductive. Insulation of the bottom side is not necessary because the voltage between the bottom side and the body is close to zero due to the G controller. The top side can easily be insulated with an additional layer of fabric (e.g., a regular garment), providing that excess of leakage currents is electronically detected. The small cooperative sensors are attached and connected to the fabric, making the assembly seamless while maintaining the usual properties of the fabric, i.e., flexibility, stretchability, breathability, and washability. Symbol legend in Appendix A.

Since the total gain of the acquisition chain (from the potential v_i to the demodulated signal v'_i) may be slightly different from sensor to sensor, the common-mode rejection when performing the difference $v'_i - v'_j$ might be insufficient. The problem can be neatly solved by the method described in [15], which relies on online identification of the v'_i/v_i transfer functions through an excitation common-mode voltage added to the voltage source driven by the G controller.

Figure 1 also shows that the two wires could be implemented as a textile spacer with both sides made conductive. Note that these conductive surfaces also implicitly act as ‘body shield’, in contrast to the classical approach where the shield must be explicitly added [23]. One side of the cooperative sensor is used as dry electrode and the other as bottom contact with the fabric. The upper contact is made on the other side of the fabric by the sensor attachment. Due to the integrated electronics (e.g., ASIC), the cooperative sensors can be very small (e.g., $4 \times 4 \text{ mm}^2$) while the dry electrode and the contacts to the fabric can be larger and flexible. As a result, the assembly of the sensors with the fabric is virtually cable-free and seamless, making it easier to meet the wearability constraints (flexibility, breathability, stretchability, washability).

The G controller ensures that the voltage between the lower wire and the body is approximately zero. As long as this voltage is always less than 10 mV for ECG and less than 100 mV for other biopotentials, e.g., EMG—which is the case—no insulation is needed, because the standards require ECG devices to be TYPE CF (limit to 10 mV) and other devices TYPE BF (limit to 100 mV) [7]. By default, the other side is not in contact with the skin and can only be touched intermittently, for example with the hand. This problem can be

solved by electronic detection (see below). However, the detection of leakage currents will stop the operation of the device. Therefore, one may consider adding a layer of fabric to insulate the top conductive side in most situations and relying on electronic detection only for exceptional situations. Optionally, the outer side of this additional layer can be made conductive and connected to the skin to provide one MEAN OF PATIENT PROTECTION in the same idea as CLASS I MEDICAL ELECTRICAL EQUIPMENT [7]. This conductive layer also shields the middle conductive layer with respect to EMC emission and immunity.

2.2. Floating Supply and Bootstrapping

Figure 2 shows in more detail an implementation of cooperative sensors [16–18] for biopotential measurement (the cooperative sensors described in [17,18] also measure bioimpedance for EIT, electrical impedance tomography, but in this paper, we focus only on biopotentials for simplicity). In this implementation, each cooperative sensor is powered by a battery. All batteries can be recharged simultaneously via the two-wire bus when the system is not worn. The safety issue related to the maximum leakage current of $10\ \mu\text{A}$ d.c. is not applicable when the system is not worn by the patient—the connector to the charger is made so that it is impossible for the patient to wear the system during charging. Once in the sensor, the stored charges are prevented from leaving the sensor by diodes (not shown). Thus, there is no possible harm resulting from insulation failure, for example, due to body fluid. Note that communication is not a problem because at higher frequencies, higher leakage currents are allowed, and communication requires lower voltages (which reduces possible leakage currents).

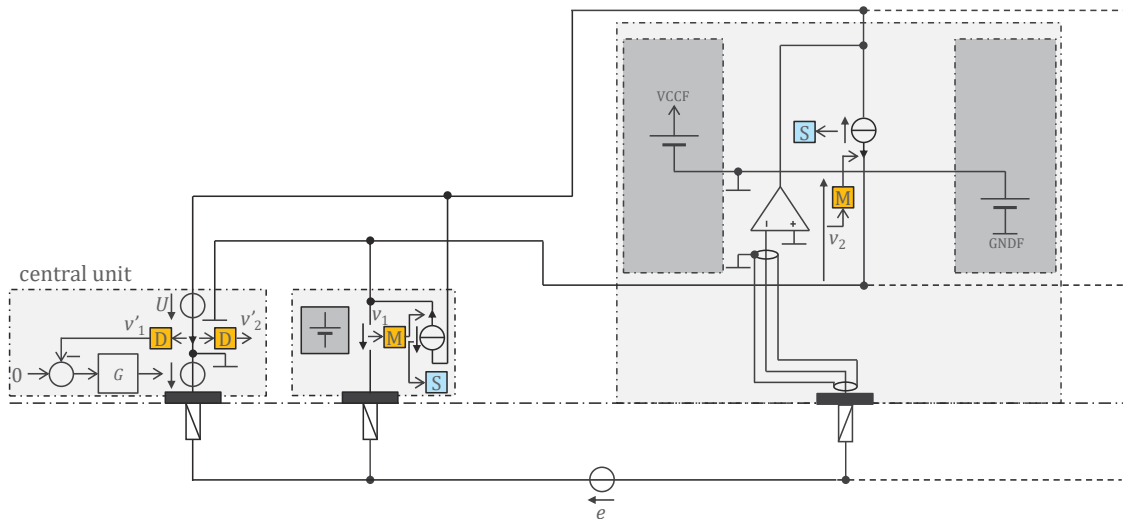


Figure 2. Implementation of cooperative sensors with extremely high input impedance achieved with a simple bootstrap circuit taking advantage of floating sensor batteries. The patient is protected against sensor leakage currents by diodes (not shown) that prevent stored charges from accidentally leaving the sensors but allow the 2-wire bus to be used to simultaneously recharge all batteries when the system is not worn. Symbol legend in Appendix A.

The input impedance of the cooperative sensors is significantly increased by a power supply bootstrapping approach [19] that takes advantage of the degree of freedom provided by the floating power supplies of cooperative sensors. The measured potential connected to the negative input and the middle potential of the supply connected to the positive input are made equal by the operational amplifier. This bootstrapping strategy is preferred to other well-known solutions, such as positive feedback [24], because the performance is higher and the risk of instability is lower. The electrode and its connection to the operational

amplifier are shielded with the middle potential of the power supply in a manner equivalent to the driven shielding of classical techniques.

3. Method

Cooperative sensors, with their very high input impedance suitable for dry electrodes (thanks to bootstrapping) and their connection via a 2-wire parallel bus, solve the integration difficulties related to wiring and the scalability issues of star-arrangement connections, thus paving the way for biopotential imaging wearables comprising many electrodes. Like any active electrode, cooperative sensors require power. Safety constraints resulting from poor and unreliable insulation in wearables due to wear and tear, as well as the possible presence of body fluids, make remote power from the central unit a real challenge for medical devices where the maximum leakage current is 10 μ A d.c. [7]. The cooperative sensor approach in Figure 2 avoided this problem by using one battery in each sensor and implementing a recharging strategy using the 2-wire bus when the vest is not worn (e.g., when placed on a hanger modified as a charger). However, one battery per sensor is expensive, heavy, and cumbersome.

This section presents two solutions for remote powering of cooperative sensors. The first solution (Section 3.1) solves the difficulty of adding the power supply function to the bus without increasing the noise of the measured biopotentials beyond the limit required by the standards, and without interfering with the synchronization and bidirectional communication between the sensors and the central unit. The second one (Section 3.2) also aims to solve the safety issue mentioned above, significantly reduce the volume of the sensors (e.g., from 7.5 cm³ to 0.3 cm³) and the power consumption (e.g., from 5.8 mA to 150 μ A), operate with more sensors (e.g., from 20 to 250), and implement a bootstrap that does not require floating batteries as in Figure 2.

3.1. Legacy Approach with 500 Hz Powering and Off-the-Shelf Components

Figure 3 shows the circuit in Figure 2 with modifications to allow for remote powering. The cooperative sensors ‘harvest’ their energy from the 2-wire bus powered by the voltage source U of the central unit. Capturing power from the bus is symbolized in the first cooperative sensor (middle of Figure 3) by a current source whose current is in phase with the voltage U (power is consumed when current and voltage have the same sign, as with a resistance).

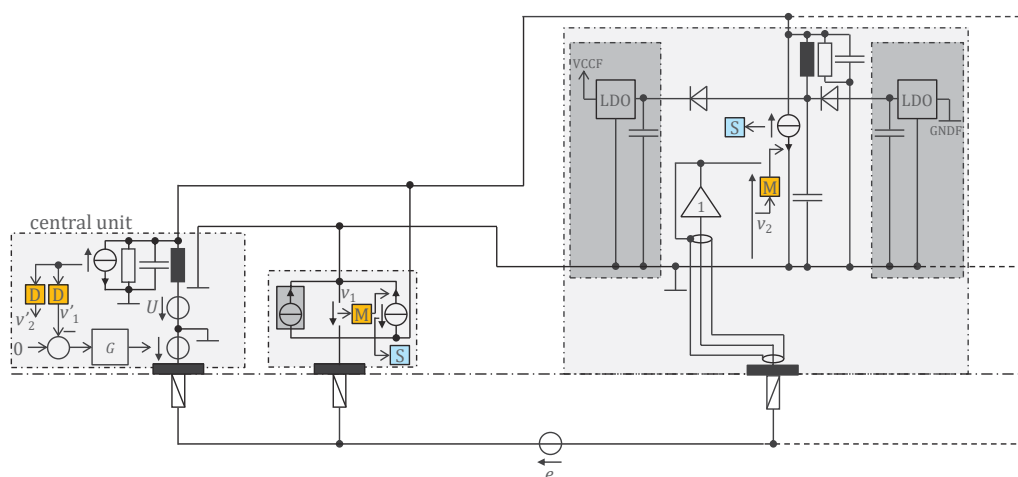


Figure 3. Remotely powered cooperative sensors for biopotential measurement with dry electrodes, with digital communication at 1.28 Mb/s in both directions (full duplex), and remote power supply at 500 Hz. Symbol legend in Appendix A.

The resistance of the bus wire is low but the supply current high, and the resulting voltage drop on the lower wire is in series with the potential e to be measured. Therefore, the supply and measurement must be in separate frequency bands to minimize interference with the measurement. The first-order bandwidth of biopotentials is, for instance, 0.05–150 Hz for ECG. Therefore, constant current sensors will give a disturbance spectrum theoretically at 0 Hz, but even if they are carefully designed with active control to keep the current constant, it is very difficult to avoid overlap with the ECG band. To avoid this interference and to easily obtain a bipolar supply in the sensors (see below), the voltage U was chosen as a square wave at 500 Hz corresponding to the Nyquist frequency of the 1 kHz sample rate. Interference with the fundamental (500 Hz) is first avoided by fine-tuning the phase of the supply square wave (at the Nyquist frequency, a sinewave has all its samples equal to zero; only the cosine has non-zero samples). Any remaining energy of the fundamental—in practice, it is impossible to have the same adjusted phase for all sensors—is removed with a first-order notch filter by signal processing (i.e., moving average of two consecutive samples). The harmonics (1 kHz, 1.5 kHz, 2 kHz, ...) must be removed before sampling, otherwise they create aliases at 0 (even harmonics) and 500 Hz (odd harmonics). An ideal square wave has energy at the odd harmonics and the 500 Hz notch filter is, therefore, necessary. To remove the even harmonics from a non-ideal square wave (an unavoidable situation in practice due to the asymmetry of power consumption for positive and negative currents), a third-order delta-sigma analog-to-digital converter was chosen, because the digital antialiasing filter of a delta-sigma converter is a comb filter with notches at multiples of 1 kHz (sample frequency).

The harmonics of a 500 Hz power supply are also low enough in the MHz range that they do not significantly disrupt the digital communication (modified to 1.28 Mb/s, in both directions, to conform to the frequency required by the delta-sigma converter). The bits just past the edges of the 500 Hz supply are disturbed and, thus, removed from the communication payload (in our prototype, 110 bits are removed, i.e., 17%).

In Figure 2, the upstream channel (from the central unit to the sensors) is realized as a voltage source and the downstream channel (from the sensors to the central unit) as current sources. An alternative is to interleave the two channels, each with its own time slots. In this way, both the up and down channels can use voltages (or currents). In Figure 3, the current sources have a resistance in parallel. Thevenin's equivalent, a voltage source with the resistance in series, is easier to implement since the voltage source is simply a digital output. The capacitance and inductance are chosen so that their resonance is at 2.56 MHz, and the resistance so that the RLC triplet implements the first-order bandpass filter used first to prevent the communication band from overlapping the biopotential band and second to avoid EMC problems that sharp edges in the digital signal can cause. Note that the RLC triplet is a parallel assembly of R , L , and C in both the central unit (voltage source U at 0) and in the sensors (the capacitance in series with the inductance is chosen so that its impedance is negligible at communication frequencies). The received signal is the voltage on the RLC triplet. After its reconstruction with a high-pass filter and Schmitt trigger, the digital signal is demodulated in the D-block to obtain v_i^j .

Since the LC blocks any current at the communication frequency, the voltage on the RLC triplet is the result of a voltage divider consisting of the emitter resistance and all other resistances (of the receivers) in parallel. The consequence is that the received voltage is the emitter voltage divided by the number of units (i.e., sensors and central unit). Therefore, this approach limits the number of units in practice to approximately 20.

The 500 Hz supply square wave is provided by the voltage source U , which is easily realized with switching transistors. The impedance of the central unit and sensor inductances is negligible for the supply current. Therefore, the supply square wave is rectified by diodes in the sensors to provide a positive and a negative voltage on the storage capacitors.

3.2. Approach Addressing the Safety Issue with Powering at 1 MHz and ASIC for 250 Sensors

The approach of Figure 3 solves the problem of remote power supply with respect to interference with biopotential measurement and communication on the 2-wire bus, but does not solve the safety problem for medical devices, because the allowed PATIENT LEAKAGE CURRENT at 500 Hz is also 10 μ A for ECG devices that must be TYPE CF [8]. For other biopotentials, there is already an appreciable advantage over d.c., since the devices can be TYPE BF and the allowed PATIENT LEAKAGE CURRENT is ten times higher (i.e., 100 μ A). Assuming a per sensor consumption of 8 mA (see next section) and 25 sensors, the supply current on the bus is 200 mA. Detecting a 100 μ A leak from monitoring the supply current at the central unit is a difficult task (1 part in 2000). In addition, all sensors must have a buffer current source that makes closed-loop adjustments to its current to ensure that the current of any sensor is exactly 8 mA (again with high accuracy).

To better address the safety issue, including the ten-fold increase in ECG requirement, and to increase the number of sensors by a factor of 10, i.e., to up to 250, we need to reduce the power consumption of the sensors (say, by a factor of 20, i.e., to 400 μ A) and move the power supply frequency to 1 MHz where the standards allow a patient leakage current of up to 10 mA (which is also the absolute maximum). This will make detecting a leakage current much easier (1 part in 10).

Note that the 8 mA and 400 μ A mentioned above are the current of a sensor as measured in the bus. The sensor itself consumes half of this current, i.e., 4 mA and 200 μ A, respectively. The factor of two comes from energy conservation, i.e., the bus supply voltage U is $\pm V_{CC}/2$ with a current of $\pm 2I$ (rms value $2I$ for square waves) which allows sensors with the dual half-wave rectifier (assuming perfect diodes) to have a VCC supply and I current for the electronics.

To reduce power consumption by a factor of 20, we developed an ASIC (application-specific integrated circuit) that optimized each electronic function. In addition, we eliminated the digitization (analog-to-digital converter). The transmission of analogue values instead of bits has also increased the throughput (required for 10 times more sensors). The 1 MHz power supply is now interleaved with the communication, i.e., every other period the power supply is replaced by the communication [25]. Figure 4 shows the principle of this implementation. The inductances are no longer needed (which is good because they cannot be integrated into silicon as passive components). A switch reroutes the 1 MHz square-wave signal either to the rectifier diodes and storage capacitors (harvesting period) or to the communication current source. Before transmission, the biopotential is amplified and filtered. A high-pass filter prevents the transmission of the electrode offset that can be as high as 300 mV according to the standards, and thus improves the signal-to-noise ratio of the analogue communication. The M modulator simply selects the right time slot for the sensor to transfer its value. All sensors sample their biopotential at the same time. The value to be transmitted is stored in a capacitor until transmission.

Bootstrapping is achieved with the regulated supply rails VCCF and GNDF (specific for each sensor) following the electrode potential (with offsets). This is obtained by the follower controlling the reference of the LDO voltage regulators. Assuming an LDO gain g (i.e., the LDO outputs a current $i = gu$ where u is the voltage error of the LDO output), the input impedance of the open loop circuit is magnified by gz at low frequencies. High gain at low frequencies can be achieved if z behaves like a capacitance at low frequency. At higher frequencies, for stability reasons, it is preferable for z to behave like a resistance. The open loop input impedance is essentially the input impedance of the follower (typically 10 pF). The bootstrap magnifies this impedance by gz , allowing the circuit to have a very high input impedance at low frequencies [25]. Compared to Figure 3 where bootstrapping is not implemented, this bootstrap also makes shielding of the sensor input more efficient and natural, as the ground and power rail planes provide implicit (driven) shielding.

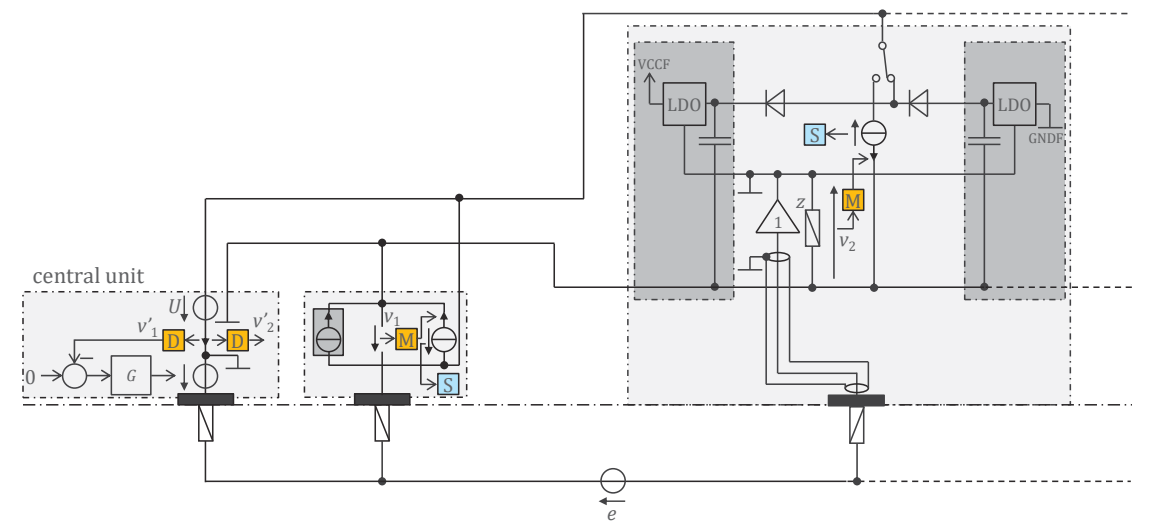


Figure 4. Remotely powered cooperative sensors for measuring biopotentials with dry electrodes, with digital communication at 500 000 samples per second and remote supply U voltage at 1 MHz. Left: schematic overview of the central unit circuit; middle: schematic overview of a sensor circuit; right: detailed circuit diagram of a sensor. Symbol legend in Appendix A.

This paper only describes the measurement of biopotentials. However, the developed ASIC is also capable of measuring bioimpedance (for EIT) and can be interfaced with an electret to pick up body sounds (stethoscope).

3.3. Comparison to Existing Work

Table 1 shows a comparison with existing work to further highlight the significance of the presented work. Compared to the closest state of the art, remotely powered cooperative sensors do not require a local power supply (e.g., a battery per sensor) which allows them to be miniaturized, among other things. Being able to monitor the leakage currents allows wearables to be safe (in the context of medical standards) without relying on the insulation of conductors in a garment and without the need for waterproof connectors.

Table 1. Comparison to existing work (the main contribution of the paper is highlighted in grey).

Technique/Features	[Ref], Section	Comment
Conventional star arrangement		Not suitable for wearables with many electrodes
Passive electrodes, shielded cables	[1]	Widespread
Active electrodes, two-wire cables	[2,3,20,21]	Well-known in the literature, but little used
Parallel bus arrangement		Scalable (connector size independent of nb. of electr.)
Bus with more than 2 wires	[22]	Not easily flexible, stretchable, breathable, washable
Two-wire bus (cooperative sensors)	Section 2	Simplest connection
Locally powered	Section 2.1	Easy to comply with safety (medical standards)
Bootstrapping	Section 2.2	Suitable for dry electrodes
Remotely powered	Section 3/Section 4	Sensors can be miniaturized
No monitoring of leakage currents	Section 3.1/Section 4.1	Requires reliable waterproof double insulation
No bootstrapping		Not ideal for dry electrodes
Monitorable leakage currents	Section 3.2/Section 4.2	Suitably flexible, stretchable, breathable, washable
Bootstrapping		Suitable for dry electrodes

4. Results

4.1. Legacy Approach with 500 Hz Powering and Off-the-Shelf Components

4.1.1. ECG Study Prototype

Figure 5 shows a study prototype with 224 sensors for ECGi based on the approach shown in Figure 3 with a 500 Hz power supply and commercially available components. The device is CF and is defibrillation-proof between two electrodes (verified according to [8]). The central unit has recording and wireless communication capabilities and is powered by IEC 62,133 batteries. A close look at Figure 5a shows that the sensors are not connected on a single two-wire bus. This is because the approach in Figure 3, as described in the previous section, cannot have more than approximately 20 sensors on the same bus. Instead, columns of eight sensors are each connected by a two-wire bus to repeaters, which in turn are connected to a two-wire bus with 14 repeaters, and the system is doubled with left and right banks from the central unit. So, it is not a fully parallel bus topology but a tree arrangement comprising $2 \times 14 \times 8 = 224$ sensors. The sensors are equipped with a CPLD (complex programmable logic device) to implement the PLL and internal clock reconstruction, a lossless compression scheme to reduce the 24 bits of delta sigma to 10 bits, and information exchange with the bus.

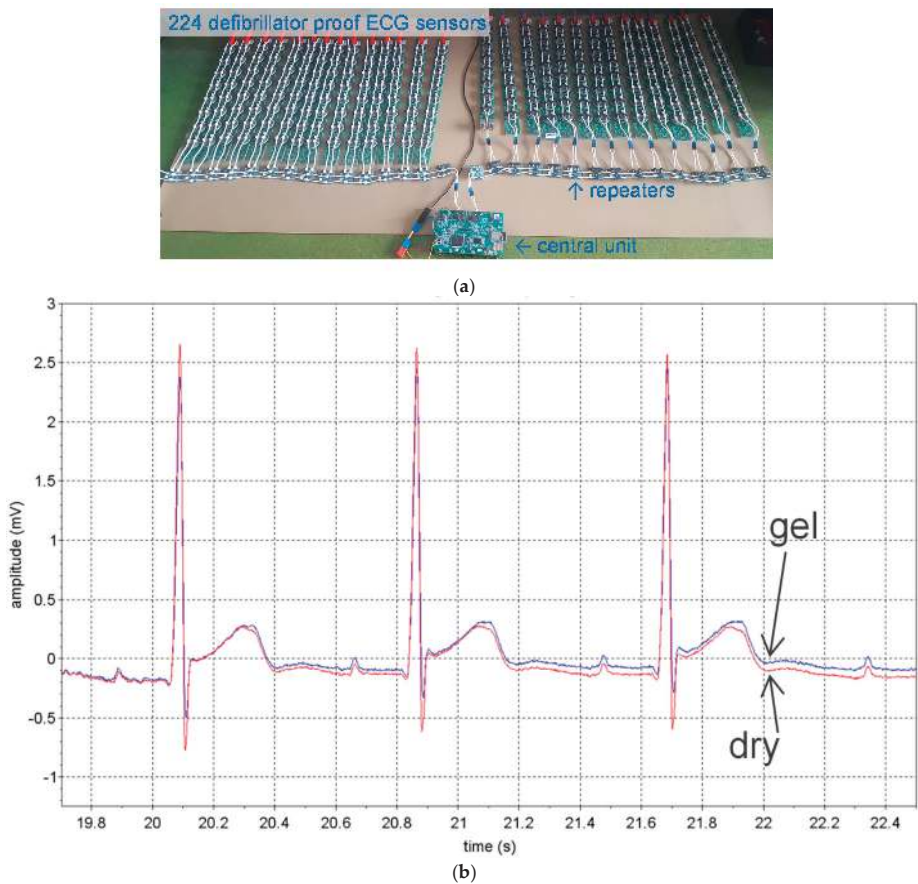


Figure 5. Remotely powered cooperative sensors based on the approach of Figure 3: (a) study prototype made of 224 sensors for ECGi with defibrillation protection between all electrodes; (b) onesubject trial to compare dry stainless steel electrodes with adhesive Ag/Ag⁺Cl⁻ gel electrodes showing identical noise level for both electrode types (measured from adjacent leads).

The study prototype passed all hardware performance tests [8], including tests related to defibrillator protection. The device was designed to have lower noise than required by ECG standards because its intended use is ECGi. The measured rms value of ECG noise is $2.5 \mu\text{V}$ and the peak-to-peak value $20 \mu\text{V}$ over 10 s. The recovery time after defibrillation is less than 300 ms (standards require a maximum of 5 s). The study prototype at this phase of development was not built to be worn. However, Figure 5b shows the signals measured on a subject with two pairs of electrodes on either side of the chest. One pair consisted of dry stainless-steel electrodes and the other of $\text{Ag}/\text{Ag}^+\text{Cl}^-$ gel electrodes. For both signals, noise is barely visible. Note that the two signals are not identical because they result from close but different electrode positions.

Figure 6 shows a simplified electronic schematic of the entire device (the safety and defibrillator protections are not shown) implementing the principle of Figure 3. The central unit (left) is connected to one of the 14 repeaters (middle) of the right bank, which in turn is connected to one of the eight cooperative sensors (right). The voltage source U of the central unit (see Figure 3) is implemented with switch transistors alternatively connecting the battery (4.6 V) or a bypass (0 V). The repeaters, cooperative sensors, and central unit harvest energy from the two-wire bus using dual half-wave rectifiers. A capacitor connected just before the half-wave rectifiers removes the offset of U (2.3 V) and provides a bipolar VCC/GND supply ($\pm 2.3 \text{ V}$) which is symmetrical with respect to REF, the potential of the lower line of the two-wire bus. The LCR triplet consists of an inductor, a parasitic capacitance and two resistors in parallel directly connected to digital outputs of a CPLD (Logic) operating in push-pull mode to send a signal. The edges of the received signal is regenerated by a Schmitt trigger.

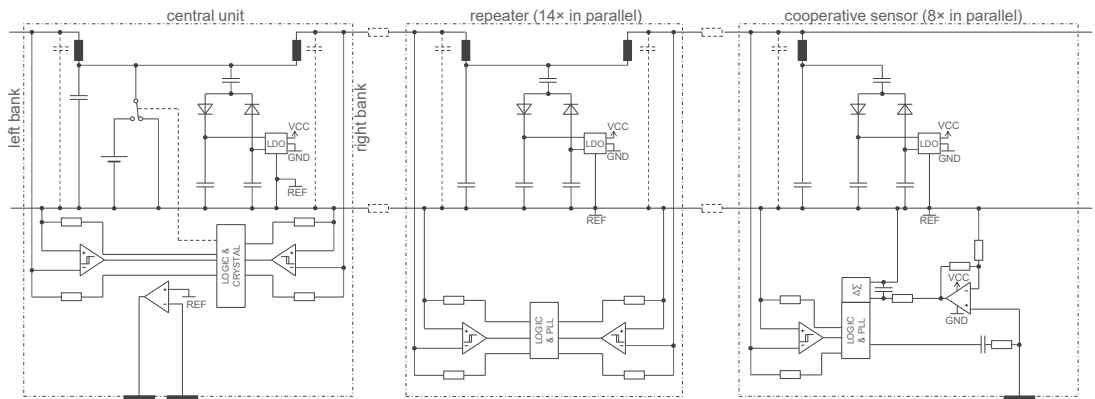


Figure 6. A possible implementation of the principle shown in Figure 3 (as prototyped in the device shown in Figure 5). Symbol legend in Appendix A.

In the implementation of Figure 6, to keep it simple, the central unit has two electrodes. Therefore, the G controller is analogue (the operational amplifier) since the v_1 signal comes from the extra electrode of the central unit. The repeater receives the digital signal on one side and retransmits it on the other side. This is achieved at the cost of a one-bit delay, which is not a problem because the shift is taken into account by the communication protocol. The same communication principle as described in [18] is used. The sensor electrode is connected to the positive input of an operational amplifier. Two resistances define the gain of the first amplification stage and an RC circuit at the output of the operational amplifier implements the coarse antialiasing filter required by the delta-sigma converter.

To measure skin impedance (e.g., for a lead-off detection function), the RC circuit connected to the electrode provides a simple way to inject a tiny current of $\pm 1 \text{ nA}$ at 500 Hz (square wave) with a digital output. The phase is chosen to be 90° offset from the sample times (cosine), because a sinewave at the Nyquist frequency has all its samples equal to zero.

The capacitance is chosen to have a corner frequency just below 500 Hz to maximize the input impedance at lower frequencies where the biopotential is measured. The voltage drop resulting from the injection of this current through the skin impedance is superimposed to the biopotential signal. However, it can be isolated by taking the difference of two consecutive samples and ‘demodulated’ by taking only the even samples (down sampling by two). This processing is conducted in the CPLD of each sensor before the data are transferred to the central unit. Taking the difference of two consecutive samples (filter with transfer function $1 - z^{-1}$) also allows to efficiently ‘compress’ the biopotential signal (for example, on the 24 bits of the delta-sigma converter, only 10 bits are enough to ‘encode’ a medical ECG). For the biopotential, the decompression is performed in the central unit by the inverse filter $1/(1 - z^{-1})$. This compression scheme is lossless (except at 0 Hz, but 0 Hz is outside the bandwidth of biopotential signals). The notch filter at 500 Hz (sum of two consecutive samples, i.e., the filter with the transfer function $1 + z^{-1}$) needed to remove the supply disturbance is also realized in the sensor CPLDs. Therefore, the central unit must apply this corresponding inverse filter, i.e., $1/(1 + z^{-1})$, to recover the skin-impedance signal. Again, this decompression is lossless, except for the frequency exactly at 500 Hz. Therefore, the 500 Hz current for measuring skin impedance is turned on and off every 1 s to make a differential measurement (corresponding in the frequency domain to a frequency line at 499 Hz that is not affected by the lossless compression/decompression process). The entire processing scheme is shown in Figure 7.

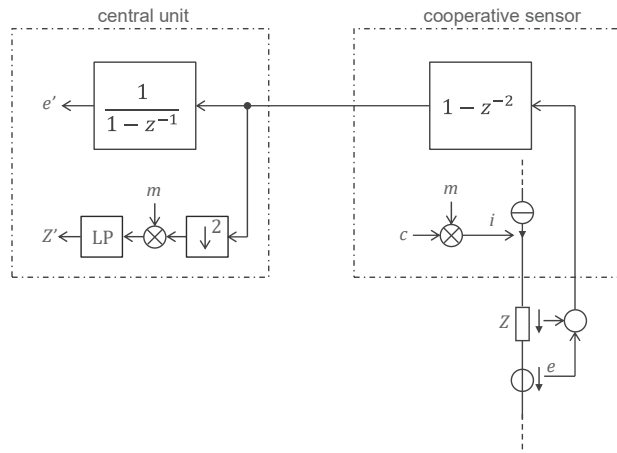


Figure 7. Processing scheme for the transmission of the signal from the cooperative sensor to the central unit. The sum of the biopotential e and the voltage drop across the skin impedance z is filtered by $1 - z^{-2} = (1 - z^{-1})(1 + z^{-1})$ to compress the biopotential signal ($1 - z^{-1}$) and eliminate the supply disturbance at the Nyquist frequency ($1 + z^{-1}$). The current i injected through the skin impedance Z is obtained from a carrier c (square wave at Nyquist frequency) modulated by the signal m (typically a square wave at 1 Hz). Decompression (lossless in the biopotential band, e.g., 0.05 to 150 Hz) with the filter $1/(1 - z^{-1})$ provides the measured biopotential e' for further processing by the central unit. In parallel, the received signal is subsampled by 2 and demodulated by multiplying it by the signal m and filtering it with the low-pass filter LP to obtain the skin impedance signal Z' . Symbol legend in Appendix A.

4.1.2. EEG Study Prototype

Figure 8 shows another study prototype based on the development presented in Figure 5 but reworked to handle dry electrode EEG. The modifications are mainly a higher gain for lower noise (0.7 μV rms, 5.4 μV pp over 10 s), a narrower bandwidth (0.5–50 Hz, 1st order), and a device limited to eight sensors. The noise is within the limit (6 μV pp over 10 s) accepted by the standard [26]. The obtained power spectra are displayed on the

right and show the expected changes in power at certain frequencies resulting from the closed eyes.

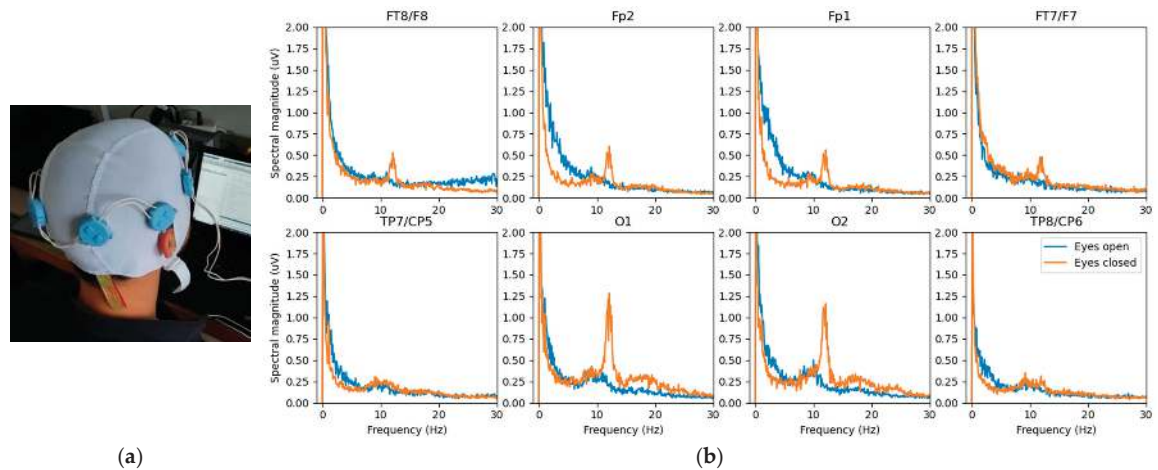


Figure 8. Remotely powered cooperative sensors based on the approach shown in Figure 3: (a) study prototype consisting of 8 EEG sensors; (b) power spectra for each electrode showing the increase in power at certain frequencies resulting from closed eyes (orange) versus open eyes (blue).

The implementation circuit is similar to that in Figure 6, but does not include the repeaters, and the central unit connects only a two-wire bus with eight cooperative sensors. Table 2 gives the measured noise on the prototype resulting from digital processing of the difference in potentials (sensor 1 as reference) in the EEG bandwidth of 0.5–50 Hz specified by the standard [26].

Table 2. Acquisition chain noise (bandwidth 0.5–50 Hz).

Sensor	Peak-to-Peak Noise over 10 s (μV)	Rms Noise (μV)
1	(potential reference)	(potential reference)
2	4.01	0.59
3	4.65	0.60
4	5.18	0.66
5	4.90	0.64
6	4.88	0.66
7	5.41	0.67
8	4.40	0.66

4.2. Approach Addressing the Safety Issue with Powering at 1 MHz and ASIC for 250 Sensors
4.2.1. ASIC Architecture

The ASIC can be mainly divided into two sections, as shown in Figure 9. The first section, in red, provides the interface to the two-wire bus. At startup, the power management unit uses the power square wave generated by the central unit to turn on all internal power supplies. After this is accomplished, a delay-locked loop in the clock and timing recovery block listens for the sync marker in the 1 MHz square wave. The sync marker is a periodicity break that marks the beginning of a sequence of 1,000,000 periods, i.e., there is sync marker every 1 s (see the illustration in Figure 10). The power management harvests current from the bus only every other period. The other periods are used for communication. Only the ASIC whose ID corresponds to a given communication slot transmits its acquired sample to the central unit.

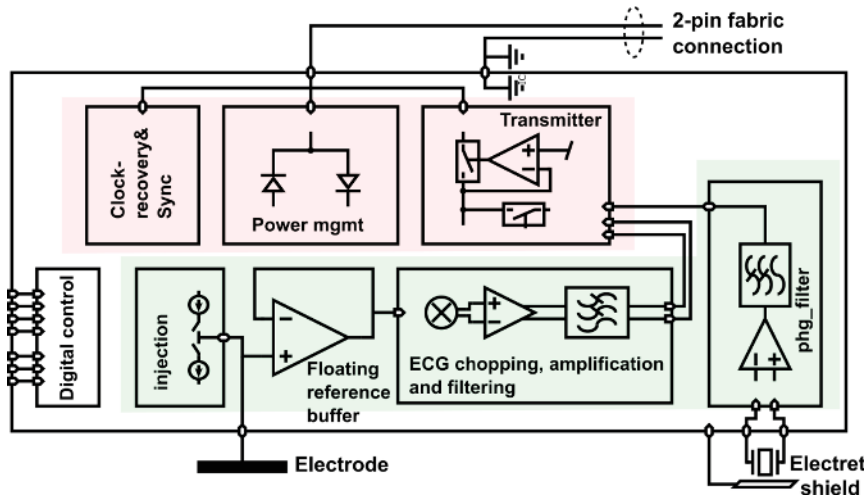


Figure 9. ASIC block diagram of biopotential cooperative sensor. The circuit blocks that interface with the 2-wire sensor bus are marked in red and the signal processing circuits in green. Symbol legend in Appendix A.

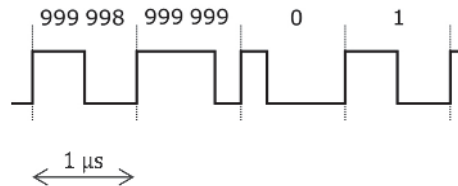


Figure 10. Supply voltage U consisting of a 1 MHz square wave with a sync marker (periodicity break) every 1 s (every 1,000,000 periods of the 1 MHz square wave).

The second section of the ASIC, shown in green in Figure 9, provides the signal chains for a biopotential electrode and additional sensing functionality that can be integrated into a sensor. In the example in Figure 9, the additional functionality is a stethoscope. In addition, a current injection block is used to check the contact impedance of the electrode.

The biopotential acquisition chain starts with a unity gain buffer, as shown in Figure 9. The output of the reference buffer is used as a ‘floating reference’, that is, it is used as a ground reference for a positive supply rail and a negative supply rail that feed the reference buffer itself (see Figure 11). With this bootstrapping approach, the buffer supply perfectly follows the (a.c.) biopotential to be measured. As a result, the voltage on the parasitic capacitances at the input is asymptotically close to zero and, thus, virtually no current flows. This is equivalent to a drastic increase in the input impedance.

The vcc and vee supply rails in Figure 11 follow the floating reference determined by the buffer itself fed by vcc and vee. Positive feedback is possible and to avoid instability, the loop gain must be kept below unity. In the ASIC, this is achieved by a single-stage design that has a large power supply rejection ratio (PSRR) for vcc and vee within the bandwidths of the low dropout regulators (LDOs) generating these power supplies.

After the high input impedance follower, the ECG signal is chopped at a frequency of 1–12.5 kHz to avoid flicker noise during amplification. Then it is filtered and sampled before the transmitter sends the analog signal to the central unit via the bus.

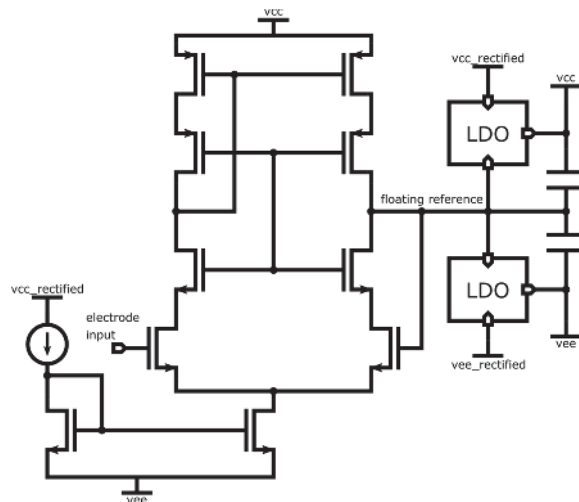


Figure 11. Transistor-level circuit of the floating reference buffer. The output of the unity gain buffer is used as ground for its own positive and negative supply. Symbol legend in Appendix A.

4.2.2. Central Unit Architecture

The central unit implementation based on the approach shown in Figure 4 is presented in Figure 12. The two wires A and B of the bus are powered by the voltage source U —a square wave at 1 MHz with a voltage of $\pm V_h$. Every other 1 MHz period, the sensors are not powered by the bus. Instead, one of them—as determined by its address—sends into the transimpedance capacitor, during the period when $U = V_h$, a quantity of electrical charge proportional to the measured potential. The voltage at the terminals of this capacitor is then sampled by the ADC to be demodulated by the microcontroller. After sampling, the transimpedance capacitor is reset by shorting its terminals with a switch. The controller G in Figure 4 is implemented with a pass-through as in Figure 6.

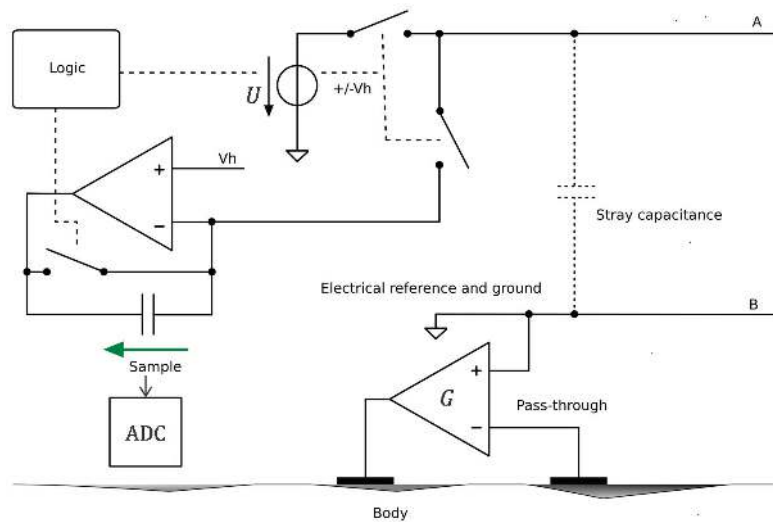


Figure 12. Implementation of the circuit for the central unit based on the approach shown in Figure 4. Symbol legend in Appendix A.

4.2.3. Implementation Results

Figure 13a shows the ASIC package mounted on a micro-PCB ($7 \times 7 \text{ mm}^2$). On the other side of the PCB (not shown in Figure 13), there are some capacitors and other components that were not practical to be part of the ASIC. However, their footprint is very small, and since the chip itself is only approximately $2 \times 2 \text{ mm}^2$ and the number of pins can be reduced to ≤ 20 , a final sensor implementation of $4 \times 4 \text{ mm}^2$ could be targeted. The first integration test setup is shown in Figure 13b. The central unit development board is on the left and is connected via the two-wire bus to a cooperative sensor (i.e., the ASIC). This setup acquired the signal of an ECG simulator applied between the sensor and GND. Figure 13c shows the acquired signal for an ECG of 2 mV (peak R).

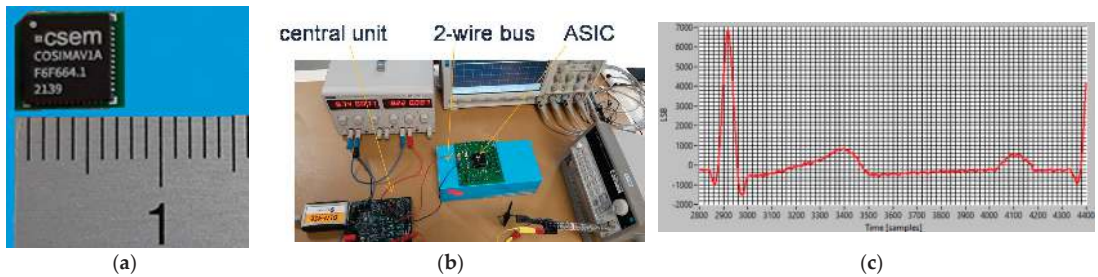


Figure 13. Remotely powered cooperative sensors based on the approach in Figure 4: (a) developed ASIC (mounted on a $7 \times 7 \text{ mm}^2$ micro PCB); (b) experimental setup during the integration process with (bottom left) the central unit and (middle) the ASIC of one sensor; (c) 2 mV ECG of simulator measured by a sensor.

Another important step demonstrating the feasibility of the proposed approach is shown in Figure 14 where two ASICs (see Figure 13a) were applied to stainless-steel dry electrodes on the body, as shown in Figure 14. Despite the use of an optocoupler (DLN-4SE), some 50 Hz problem remained, but at a low enough level to allow acquisition of a clear ECG (see Figure 14). The 50 Hz problem should be fully resolved when the central unit will have been made ‘wearable’ with a floating power supply and when a few other defects discovered during this exploration phase have been corrected.

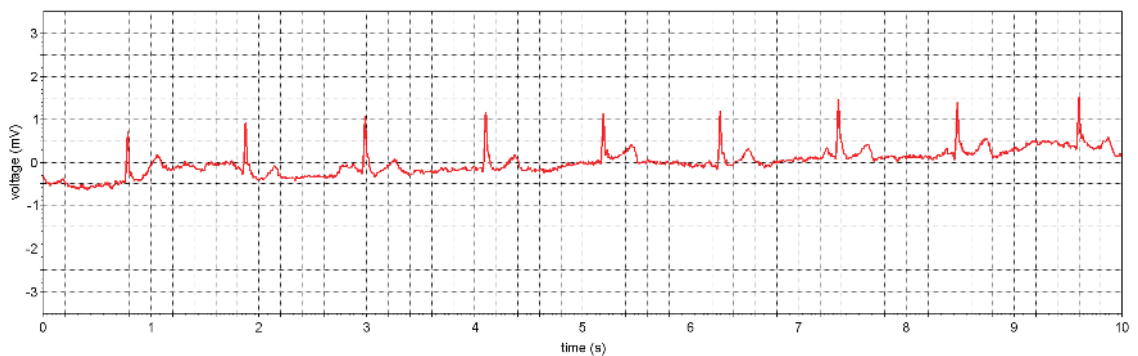


Figure 14. Same as Figure 13 but with two sensors applied to the body for actual ECG measurement (some 50 Hz noise remains due to limitations of the current setup).

5. Conclusions

The two developments presented in this paper have demonstrated the feasibility of cooperative sensors for medical imaging wearables for biopotentials. The main contribution is the disclosure of circuits that allow remote powering of cooperative sensors via their two-wire bus.

Cooperative sensors are ideal for biopotential imaging wearables because they can be deployed in large numbers (>100) without suffering from the wiring complexity of conventional star arrangements. In an initial development, we demonstrated that high-quality ECG signals can be acquired from up to 224 defibrillator-proof dry electrodes connected via a two-wire bus (or rather, a tree of two-wire buses).

A second development designed for up to 250 dry electrodes on a two-wire bus (not a tree of two-wire buses) in the form of high-frequency remotely powered ASICs to enable detection of hazardous leakage current has been demonstrated. So far, the demonstration is limited to 2 sensors on the body and 15 sensors on the test bench. A next iteration to correct some implementation errors and defects is needed to address more sensors. However, the tree arrangement strategy used in the first development could also be considered as a backup solution for the ASICs, if we have difficulties to reach the goal of 250 sensors on a single two-wire bus.

In addition to the challenge of combining remote power supply, microvolt biopotential measurements, synchronization, and communication on the same two-wire bus, a significant difficulty addressed in this paper is compliance with medical standards for leakage currents in the context of wearables that can hardly provide reliable waterproof double insulation of electrical connections. Although leakage current detection was not implemented during this development, the design was made to address this issue by proposing power at frequencies where a leakage current can be detected by the central unit by monitoring the current.

6. Patents

The work presented in this paper is based on patent [25].

Author Contributions: Conceptualization, O.C.; methodology, B.B., B.S., A.F., M.R. and O.C.; validation, M.R. and A.F.; writing—original draft preparation, O.C.; writing—review and editing, J.W. and B.S.; visualization, A.F.; supervision, O.C., A.F. and J.W.; project administration, O.C., A.F. and J.W.; funding acquisition, O.C., A.F. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partly funded by Innosuisse-Schweizerische Agentur für Innovationsförderung, grant number 42657.1 IP-ICT.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

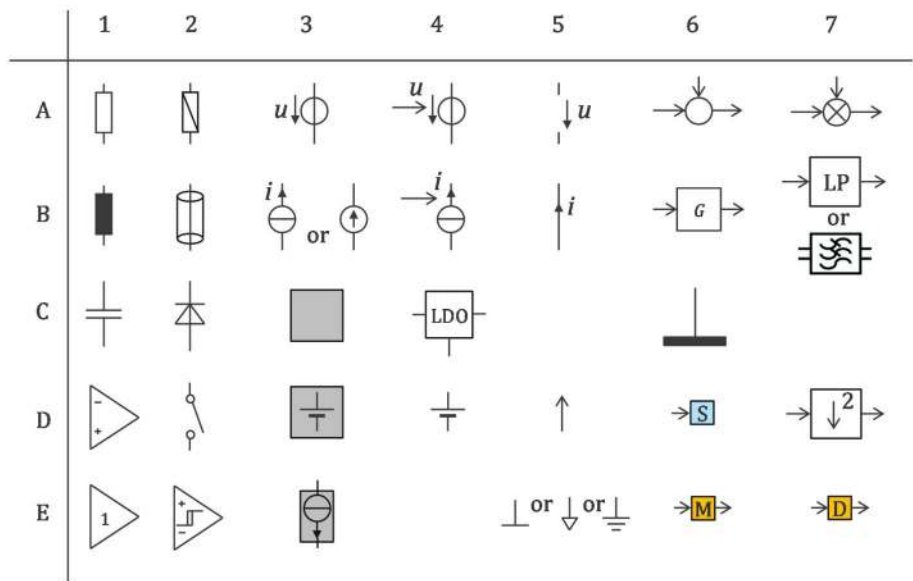


Figure A1. Symbol legends used in figures. A1: resistance, A2: impedance, A3: voltage source, A4: controlled voltage source, A5: voltage (between two conductors), A6: summator, A7: multiplier, B1: inductance, B2: shielded cable (e.g., coaxial cable), B3: current source, B4: controlled current source, B5: current (in a conductor), B6: transfer function, B7: low-pass filter, C1: capacitance, C2: diode, C3: power supply block, C4: LDO (low-dropout regulator), C6: electrode, D1: operational amplifier, D2: switch, D3: power supply including a battery, D4: battery, D5: connection to positive power-supply rail, D6: clock recovery and sync block, D7: down sampling by 2, E1: follower, E2: Schmitt trigger, E3: power supply block harvesting energy with controlled current, E5: common ground, E6: modulator, E7: demodulator.

References

1. Webster, J.G.; Clark, J.W. (Eds.) *Medical Instrumentation: Application and Design*, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA, 2010.

2. Boehm, A.; Yu, X.; Neu, W.; Leonhardt, S.; Teichmann, D. A Novel 12-Lead ECG T-Shirt with Active Electrodes. *Electronics* **2016**, *5*, 75. [CrossRef]

3. Guerrero, F.N.; Spinelli, E.M. A Two-Wired Ultra-High Input Impedance Active Electrode. *IEEE Trans. Biomed. Circuits Syst.* **2018**, *12*, 437–445. [CrossRef] [PubMed]

4. Zipp, P.; Ahrens, H. A model of bioelectrode motion artefact and reduction of artefact by amplifier input stage design. *J. Biomed. Eng.* **1979**, *1*, 273–276. [CrossRef]

5. Meziane, N.; Webster, J.G.; Attari, M.; Nimunkar, A.J. Dry electrodes for electrocardiography. *Physiol. Meas.* **2013**, *34*, R47–R69. [CrossRef] [PubMed]

6. Godin, D.T.; Parker, P.A.; Scott, R.N. Noise characteristics of stainless-steel surface electrodes. *Med. Biol. Eng. Comput.* **1991**, *29*, 585–590. [CrossRef] [PubMed]

7. IEC 60601-1; Medical Electrical Equipment—Part 1: General Requirements for Basic Safety and Essential Performance. IEC: Geneva, Switzerland, 2020.

8. IEC 60601-2-25; Medical Electrical Equipment—Part 2-25: Particular Requirements for the Basic SAFETY and essential Performance of Electrocardiographs. IEC: Geneva, Switzerland, 2011.

9. IEC 60601-2-47; Medical Electrical Equipment—Part 2-47: Particular Requirements for the Basic Safety and Essential Performance of Ambulatory Electrocardiographic Systems. IEC: Geneva, Switzerland, 2012.

10. Pereira, H.; Niederer, S.; Rinaldi, C.A. Electrocardiographic imaging for cardiac arrhythmias and resynchronization therapy. *EP Eur.* **2020**, *22*, 1447–1462. [CrossRef] [PubMed]

11. Gevins, A.; Smith, M.E.; McEvoy, L.K.; Leong, H.; Le, J. Electroencephalographic imaging of higher brain function. *Phil. Trans. R. Soc. Lond. B* **1999**, *354*, 1125–1134. [CrossRef] [PubMed]

12. Urbanek, H.; van der Smagt, P. iEMG: Imaging electromyography. *J. Electromyogr. Kinesiol.* **2016**, *27*, 1–9. [CrossRef] [PubMed]
13. Chételat, O. Synchronization and Communication Bus for Biopotential and Bioimpedance Measurement Systems. EP Patent 2567657 B1, 13 March 2013.
14. Chételat, O.; Correvon, M. Measurement Device for Measuring Bio-Impedance and/or a Bio-Potential of a Human or Animal Body. US Patent 2015/0173677 B2, 25 June 2015.
15. Rapin, M.; Regamey, Y.-J.; Chételat, O. Common-mode rejection in the measurement of wearable ECG with cooperative sensors. *at-Automatisierungstechnik* **2018**, *66*, 1002–1013. [CrossRef]
16. Rapin, M.; Proença, M.; Braun, F.; Meier, C.; Sola, J.; Ferrario, D.; Grossenbacher, O.; Porchet, J.-A.; Chételat, O. Cooperative dry-electrode sensors for multi-lead biopotential and bioimpedance monitoring. *Physiol. Meas.* **2015**, *36*, 767–783. [CrossRef] [PubMed]
17. Rapin, M.; Wacker, J.; Chételat, O. Two-Wire Bus Combining Full Duplex Body-Sensor Network and Multilead Biopotential Measurements. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 113–122. [CrossRef] [PubMed]
18. Rapin, M. A Wearable Sensor Architecture for High-Quality Measurement of Multilead ECG and Frequency-Multiplexed EIT. Ph.D. Thesis, ETH, Zürich, Switzerland, 2018. [CrossRef]
19. Chételat, O.; Carós, J.M.S.i. Floating front-end amplifier and one-wire measuring device. US Patent 8427181 B2, 16 September 2009.
20. Guermandi, M.; Bigucci, A.; Scarselli, E.F.; Guerrieri, R. EEG Acquisition System Based on Active Electrodes with Common-mode Interference Suppression by Driving Right Leg Circuit. In Proceedings of the Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, Milan, Italy, 25–29 August 2015; pp. 3169–3172. Available online: <http://ieeexplore.ieee.org/abstract/document/7319065/> (accessed on 10 October 2017).
21. Ko, W.H. Active Electrodes for EEG and Evoked Potential. In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Vol.20 Biomedical Engineering towards the Year 2000 and Beyond (Cat. No. 98CH36286), Hong Kong, China, 1 November 1998; Volume 4, pp. 2221–2224. [CrossRef]
22. Rowlandson, G.I.; Xue, J.Q.; Henderson, R.E.; Rao, C.S. System and Apparatus for Collecting Physiological Signals from a Plurality of Electrodes. US Patent 8326394 B2, 4 December 2012.
23. Prot, P.; Frouin, P.-Y. Textile Device for Measuring the Electro-Physiological Activity of a Subject. US Patent 2019/0380613 A1, 19 December 2019.
24. Chételat, O.; Bonnal, B.; Fivaz, A. *Low Power active Electrode ICs for Wearable EEG Acquisition*; Springer: Berlin/Heidelberg, Germany, 2018.
25. Xu, J.; Yazicioglu, R.F.; Hoof, C.V.; Makinwa, K. Remotely Powered Cooperative Sensor Device. US Patent 2021/0169543 A1, 10 June 2021.
26. IEC 60601-2-26; Medical Electrical Equipment—Part 2–26: Particular Requirements for the Basic Safety and Essential Performance of Electroencephalographs. IEC: Geneva, Switzerland, 2012.



Review

Biomechanical Assessment Methods Used in Chronic Stroke: A Scoping Review of Non-Linear Approaches

Marta Freitas ^{1,2,3,4,*}, Francisco Pinho ^{1,2}, Liliana Pinho ^{1,2,3,4}, Sandra Silva ^{1,2,5,6}, Vânia Figueira ^{1,2,4}, João Paulo Vilas-Boas ^{6,7} and Augusta Silva ^{3,8}

- ¹ Escola Superior de Saúde do Vale do Ave, Cooperativa de Ensino Superior Politécnico e Universitário, Rua José António Vidal, 81, 4760-409 Vila Nova de Famalicão, Portugal; francisco.pinho@ipsn.cespu.pt (F.P.); liliana.pinho@ipsn.cespu.pt (L.P.); sandra.silva@ipsn.cespu.pt (S.S.); vania.figueira@ipsn.cespu.pt (V.F.)
- ² H²M—Health and Human Movement Unit, Polytechnic University of Health, Cooperativa de Ensino Superior Politécnico e Universitário, CRL, 4760-409 Vila Nova de Famalicão, Portugal
- ³ Center for Rehabilitation Research (CIR), R. Dr. António Bernardino de Almeida 400, 4200-072 Porto, Portugal; augusta.silva@ess.ipp.pt
- ⁴ Porto Biomechanics Laboratory (LABIOMEPE), 4200-450 Porto, Portugal
- ⁵ Department of Medical Sciences, University of Aveiro, 3810-193 Aveiro, Portugal
- ⁶ School of Health Sciences, University of Aveiro, 3810-193 Aveiro, Portugal; jpvb@fade.up.pt
- ⁷ Centre for Research, Training, Innovation and Intervention in Sport (CIF2D), Faculty of Sport, University of Porto, 4200-450 Porto, Portugal
- ⁸ Department of Physiotherapy, School of Health, Polytechnic of Porto, 4200-072 Porto, Portugal
- * Correspondence: marta.goncalves@ipsn.cespu.pt

Abstract: Non-linear and dynamic systems analysis of human movement has recently become increasingly widespread with the intention of better reflecting how complexity affects the adaptability of motor systems, especially after a stroke. The main objective of this scoping review was to summarize the non-linear measures used in the analysis of kinetic, kinematic, and EMG data of human movement after stroke. PRISMA-ScR guidelines were followed, establishing the eligibility criteria, the population, the concept, and the contextual framework. The examined studies were published between 1 January 2013 and 12 April 2023, in English or Portuguese, and were indexed in the databases selected for this research: PubMed[®], Web of Science[®], Institute of Electrical and Electronics Engineers[®], Science Direct[®] and Google Scholar[®]. In total, 14 of the 763 articles met the inclusion criteria. The non-linear measures identified included entropy ($n = 11$), fractal analysis ($n = 1$), the short-term local divergence exponent ($n = 1$), the maximum Floquet multiplier ($n = 1$), and the Lyapunov exponent ($n = 1$). These studies focused on different motor tasks: reaching to grasp ($n = 2$), reaching to point ($n = 1$), arm tracking ($n = 2$), elbow flexion ($n = 5$), elbow extension ($n = 1$), wrist and finger extension upward (lifting) ($n = 1$), knee extension ($n = 1$), and walking ($n = 4$). When studying the complexity of human movement in chronic post-stroke adults, entropy measures, particularly sample entropy, were preferred. Kinematic assessment was mainly performed using motion capture systems, with a focus on joint angles of the upper limbs.

Keywords: chronic stroke; assessment; non-linear; kinetic; kinematic; EMG

Citation: Freitas, M.; Pinho, F.; Pinho, L.; Silva, S.; Figueira, V.; Vilas-Boas, J.P.; Silva, A. Biomechanical Assessment Methods Used in Chronic Stroke: A Scoping Review of Non-Linear Approaches. *Sensors* **2024**, *24*, 2338. <https://doi.org/10.3390/s24072338>

Academic Editor: José Machado da Silva

Received: 16 February 2024

Revised: 22 March 2024

Accepted: 4 April 2024

Published: 6 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The integrated concept of the human movement system (HMS) results from a broad physiological network comprising interactions between the neurological, musculoskeletal, cardiovascular, respiratory, tegumentary, and metabolic systems in order to move the body or parts of it [1]. This interaction fluctuates in a continuous, complex, and non-linear manner [2,3], acting on different time scales in response to different contextual conditions to adapt and maintain harmonious movement based on its characteristic variability [3].

Motor variability is an inherent feature of HMS, enabling a variety of solutions and strategies during one task [4], and providing flexibility to the response depending on context or unexpected environmental changes [2]. Thus, human movement is also endowed with complexity, consisting of the ability to perform a specific task with different strategies over a period and the regularity of the inherent variability pattern [5]. These properties should allow movements without rigid patterns [4] to respond to small perturbations and task demands [6]. Thus, all functional movements ideally have variability and complexity, translating into a state between high variability with multiple movement options and complete repeatability [7]. This needs to be described and studied objectively for better understanding.

Linear measures are limited in explaining the variability of human movement and cannot fully describe its characteristics [8], which are predominantly non-linear [9]. The use of non-linear measures allows the description of small and subtle changes that depend on time and the environment, which, according to a linear approach, could be diluted by the averaging of global data, making them imperceptible, despite their existence [10].

Both approaches, linear and non-linear, are complementary and should be considered in the analysis of human movement, as both are sensitive to small perturbations that occur and affect performance and motor capacity [6]. According to a linear perspective, any typical movement without deviations that could be considered as errors represents human movement with optimal variability [2]. This contrasts with the perspective of the non-linear approach, where movement with ideal variability is actually fundamental for individuals to perform a range of variations to adapt to small perturbations induced by the environment [2,11]. A non-linear approach appears to be more consistent with dynamic human systems, as suggested by Montesinos et al. [12], and with the way interventions should be advocated for in this domain [2,7,13].

Among neuromotor disorders, stroke is the second leading cause of death and the third leading cause of disability worldwide [14]. It is commonly characterized by impairments in movement complexity and variability, resulting in functional task difficulties and reduced patient autonomy due to limitations in motor and postural control [15,16]. Linear approaches may provide limited information about the motor control system's response to change and may not include inherent aspects of the complex movement system [4,17]. Recognizing that variability analysis can enhance our understanding of adaptive strategies and the overall behavior of dynamic systems in post-stroke recovery, the investigation of non-linear measurements has emerged as a novel approach for the better interpretation of the neuromotor expression of this condition [4].

To our knowledge, there has not been a comprehensive review of the non-linear methods used to study sensorimotor recovery after stroke, despite emerging studies utilizing non-linear analysis to study stroke populations [18–20]. Understanding sensorimotor recovery is crucial for neurorehabilitation, particularly post-stroke recovery. Identifying measures and methods can aid in compiling data, interpreting the mechanisms underlying motor recovery, and guiding more effective intervention strategies [21].

Thus, the main objective of this scoping review was to summarize the non-linear measures used in the analysis of kinetic, kinematic, and electromyographic (EMG) data of human movement after stroke. The secondary aim was to summarize the methodological considerations, namely participant characteristics, studied tasks, measurement instruments and kinetic, kinematic, and EMG variables.

Review questions

The main review question was “What non-linear measures are used in the processing of kinematic, kinetic, and EMG data in the assessment of human movement after chronic stroke?”

The review sub-questions are listed as follows:

- 1. What tasks were performed in the identified studies?
- 2. What laboratory instruments were used to collect kinematic, kinetic, and EMG data in the identified studies?
- 3. Which kinematic, kinetic, and EMG variables were included in the identified studies?

2. Materials and Methods

Our scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Reviewers and Meta-Analysis extension for Scoping Reviews (PRISMA-ScR) [22]. The Supplementary Material protocol was registered on the Open Science Framework, in which the review questions and the methodology were specified, <https://osf.io/r3pe9> (accessed 28 July 2023). The following supporting information can be downloaded from <https://doi.org/10.17605/OSF.IO/R3PE9> (accessed on 7 July 2022).

2.1. Eligibility Criteria

Eligibility criteria were defined a priori using the acronym PCC (population, concept, and context) according to the methodology of the Joanna Briggs Institute (JBI) [23] (Table 1).

Table 1. Eligibility criteria according to PCC.

Criteria	
Population	Chronic [24] poststroke adults (>19 years old)
Concept	Non-linear measures in kinetic, kinematic, or EMG data processing of human movement analysis
Context	Open

Experimental and epidemiological study designs published in English or Portuguese were eligible for inclusion. Systematic, narrative, or scoping reviews (to avoid duplication of data), letters, editorials, and qualitative methodological designs were excluded.

2.2. Search Strategy

Relevant studies were identified by searching through the PubMed®, Web of Science®, Institute of Electrical and Electronics Engineers® (IEEE) and Science Direct® databases. In addition to these databases, the scholarly literature web search engine, Google Scholar®, was also used. To identify other eligible studies, the reference lists of the original research articles and reviews on the topic were manually reviewed. Studies published from 1st January 2013 to 12 April 2023 were included.

The search strategy for PubMed® was as follows: stroke AND (measure OR measurement OR evaluation OR analysis OR assessment) AND (non-linear OR nonlinear OR entropy OR Lyapunov OR “nonlinear variables” OR tools OR dynamic OR variability) AND (“human movement” OR motion) NOT (EEG OR cardiac). Two reviewers independently carried out the search.

2.3. Selection of Evidence Sources

The selection of evidence sources considered the PCC acronym, purpose, and research questions. Both investigators (MF and LP) performed the search simultaneously in the same databases using the defined strategies. No discrepancies were found during the data extraction process, which was imported into Endnote® where duplicates were removed. To facilitate the screening process and confirm the presence of any duplicates not identified by the software (Endnote version 20.0), all extracted articles were imported into Microsoft Excel.

Two independent reviewers analyzed the same 25 abstracts that met the eligibility criteria in a pilot test (the first 25 titles/abstracts of the PubMed database) [25]. The screening process did not begin until a consensus of at least 75% was reached [25].

Two reviewers screened the titles and abstracts and categorized studies as either ‘included’ or ‘excluded’. This stage allowed for the identification of articles for full-text screening. Two independent reviewers (MF and LP) extracted the data, and any disagreements were resolved by a third reviewer (SS).

2.4. Data Extraction

Two authors (MF and LP) independently extracted the relevant data using a draft charting table adapted from the original JBI template. Data were extracted according to the following categories: authors, year of publication, study design, participant characteristics (n, sex, age, side lesion, stroke type, time post-stroke), tasks studied, assessment instruments, kinetic, kinematic and EMG variables used to obtain the non-linear measures, and non-linear measures. Disagreements were resolved by a third author.

2.5. Data Presentation

To summarize the extracted data, a narrative report was produced on the following outcomes: tasks, assessment instruments, kinetic, kinematic and EMG variables, non-linear measures, and non-linear parameters. The results were expressed in relation to the main question of the study, complemented by a tabular form summarizing the main findings.

3. Results

A total of 763 articles were identified—761 records via a database search and 2 additional articles via a manual search of the reference lists. After removing 2 duplicates, 761 records remained. The screening of titles and abstracts resulted in the removal of 714 articles, while 2 articles were not retrieved, leaving 45 for full-text analysis. Of these, 33 were excluded after full-text analysis because they did not meet the inclusion criteria, namely the population (n = 9) and concept (n = 24). The remaining 14 articles were included in this review. The results of the search were presented in a PRISMA-ScR flowchart, as shown in Figure 1.

Tables 2 and 3 summarize the details of the reviewed articles, including participant characteristics, tasks, assessment instruments, kinetic, kinematic and EMG variables, non-linear measures, and non-linear parameters.

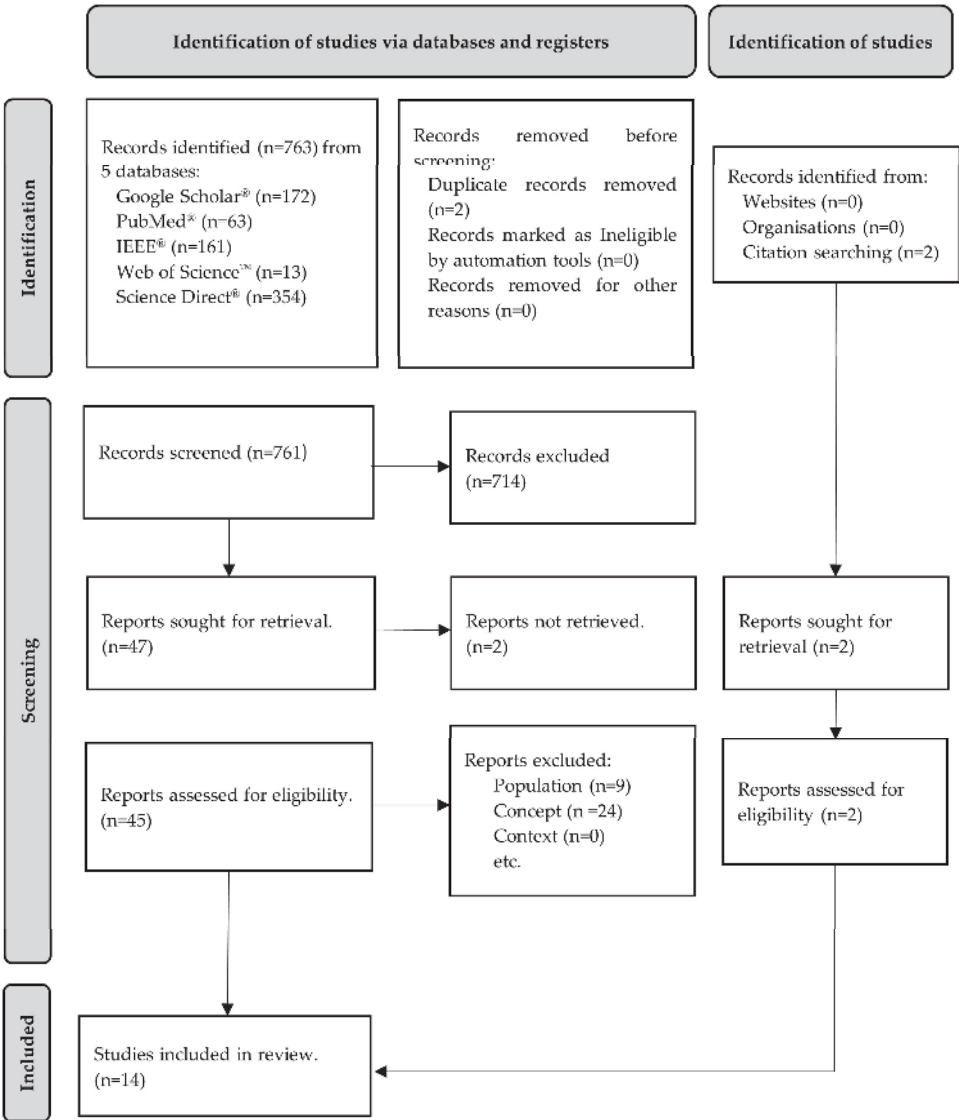


Figure 1. Flow diagram for the scoping review process adapted from the PRISMA-ScR statement [26].

Table 2. Characteristics of the participants, study design, and tasks.

Author, Year	Study Design	Participants	Task
Sethi et al., 2013 [27]	Observational, analytical study (with healthy control group)	n = 16 (11 M/5 F) 67.60 ± 8.1 years old 6 CLR/10 CLL 16 Isch 71.31 ± 48.84 months poststroke UE_FM 37.27 ± 9.27	Performed three valid reaching trials to grasp a soda can (56 mm diameter; 208 mm circumference) placed at 80° arm's length on a table in front of each shoulder with the paretic UL. All participants wore dark sleeveless shirts and sat on an adjustable backless bench with knees bent at 90° and feet flat on the floor. Their hands were placed palms down on a table in front of them and supported at 90° elbow flexion by armrests positioned flush with the table. They were instructed to reach for the can, lift it off the table, place it back on the table, and return to the starting position as quickly as possible.
		n = 6 (2 M/4 F) 67.00 ± 10.69 years old 2 CLR/4 CLL 6 Isch	Performed three valid reaching trials to grasp a soda can (56 mm diameter; 208 mm circumference) placed at 80° arm's length on a table in front of each shoulder with the paretic UL. All participants wore dark sleeveless shirts and sat on an adjustable backless bench with knees bent at 90° and feet flat on the floor. Their hands were placed palms down on a table in front of them and supported at 90° elbow flexion by armrests positioned flush with the table. They were instructed to reach for the can, lift it off the table, place it back on the table, and return to the starting position as quickly as possible.
Sethi et al., 2013 [28]	Quasi-experimental	41.83 ± 35.01 months poststroke UE_FM 32 ± 6.9 MAS < 2 MAL ≤ 3	Performed 2 trials of maximum isometric knee extension (10 s) of both legs, in random order, at different levels (10%, 20%, 30%, or 50% of MVC), as accurately and consistently as possible, in a seated position, with 60 s of rest in between. Subjects were asked to extend the knee and match the displayed torque signal (a horizontal line) to a specified target force, marked on the monitor and placed according to individual preference, usually 40–50 cm in front of the head.
Chow and Stokic, 2014 [29]	Randomized Controlled Trial (with healthy control group)	n = 23 (13 M/10 F) 64.90 ± 13.6 years old 9 CLR/14 CLL 16.6 ± 22.7 months poststroke LE_FM 26.0 ± 6.7 MAS < 2 RMI 13.4 ± 6.7	Two tasks were carried out: 1. Performance of the MIVE and the MIVF on the affected elbow at 90° of flexion, 3 times for 5 s each. 2. Performing 3 repetitive arm tracking trials, starting with the elbow at 90°. Two points were displayed in real-time on a computer screen in front of the subjects. Subjects attempted to control their elbow movement to track and match the target pointer. During the elbow extension, the robotic system continuously generates an assistive torque to support the elbow movement.
Sun et al., 2014 [30]	Quasi-experimental	n = 8 (7 M/1 F) 50.13 ± 9.13 years old 3 CLR/5 CLL 4 Hemo / 4 Isch 132.00 ± 101.76 months poststroke	Walking on a treadmill at 60%, 80%, and 100% of the preferred walking speed (determined from 5 trials of walking 10 m on the ground) and the fastest achievable speed (determined as the fastest speed the subjects could maintain for 1 min). Each speed was tested 3 times for 1 min.
Kao et al., 2014 [31]	Observational, analytical study (with healthy control group)	n = 9 (5 M/4 F) 60.8 ± 9.0 years old 40.8 ± 39.6 months poststroke LE_FM 27 ± 4	

Table 2. Cont.

Author, Year	Study Design	Participants	Task
Ao et al., 2015 [32]	Observational, analytical study (with healthy control group)	n = 11 (9 M/2 F) 47.00 ± 10.86 years old 4 CLR/7 CLL 4 Hemo/7 Isch 53.30 ± 50.9 months poststroke MAS ≤ 3	Performed 18 trials (36 s each) of elbow flexion, divided into 3 blocks (at 6 different speed levels), in a sinusoidal trajectory tracking task ranging from 30° to 90°, with 5 min of rest between each block and 30 s of rest between each trial. Subjects were seated at a table with the elbow at 30° of flexion and the shoulder at almost 90° of abduction, with the forearms attached to a lightweight aluminum manipulandum. Target and elbow angles were provided by real-time feedback on a computer screen placed in front of the subjects.
		n = 10 (9 M/1 F) 67.00 ± 8.90 years old 4 CLR/6 CLL 10 Isch (3 RMCA, 2 LMCA, 1 PVWM, 1 LLI, 1 LPI, 1 RCS, 1 RSI) 53.30 ± 50.9 months poststroke UE_FM 36.9 ± 7.9	Performed 3 trials of reach-to-point movements (drawn at 80% arm length) in three conditions, without randomization: - Voluntary reaching at preferred speed (PREFERRED). - Reaching as fast as possible (FAST). - Reaching with rhythmic auditory cues generated by a metronome (RHYTHM). Participants were seated on an adjustable backless bench with no trunk restraint, knees bent at 90°, and feet flat on the floor. Their hands were placed palms down on a table in front of them and supported at 90° elbow flexion by armrests level with the table. Post-stroke participants reached the target with a more impaired UL.
Zhang et al., 2017 [34]	Observational, analytical study	n = 10 (7 M/3 F) 62.40 ± 12.34 years old 5 CLR/5 CLL 49.20 ± 26.47 months poststroke MAS < 2	Performance of two MVC trials of submaximal isometric elbow flexion for 10 s each at different torque targets (10% to 80%), with the impaired and non-impaired arm separately. The target values were displayed as visual targets on the computer screen in a random order. During the period of isometric muscle contraction, which lasted from 2 s to 8 s, the subjects were verbally instructed to match the visual target and to maintain it as accurately as possible for all trials. Subjects were seated comfortably in a height-adjusted chair with the arm to be tested firmly attached to a custom-made device with the shoulder flexed approximately 30° and abducted 45°, the elbow flexed 90°, and the forearm in a neutral position. The other arm was placed next to the body.
		n = 7 (5 M/2 F) 62.57 ± 5.86 years old 4 CLR/3 CLL 42.30 ± 49.03 months poststroke LE_FM 25.71 ± 5.22	Walking in sports shoes for 2 min on an instrumented split-belt treadmill at an individual's self-selected speed (determined by a 10 m walk test) and long-distance walking speed (determined by a 6 min walk test).
Kempski et al., 2018 [35]	Observational, analytical study	n = 11 (7 M/4 F) 61.55 ± 12.05 years old 6 CLR / 5 CLL 46.00 ± 27.00 months poststroke UE_FM 43.14 ± 11.68 MAS < 2	Performed 3 trials of elbow flexion with isometric graded increasing force levels corresponding to 10%, 30%, 50%, 70%, submaximal (90%), and almost 100% of MVC for at least 3 s each. Sufficient rest was allowed between 2 consecutive trials to avoid mental or muscular fatigue. Subjects were seated comfortably in a mobile chair or wheelchair next to a height-adjustable desk, with the tested elbow flexed at 90 degrees on the paretic and contralateral side in a randomized order.
Tang et al., 2018 [36]	Observational, analytical study (with healthy control group)		

Table 2. Cont.

Author, Year	Study Design	Participants	Task
Dugan et al., 2020 [18]	Observational, analytical study (with healthy control group)	n = 11 (8 M/3 F) 57.91 ± 14.7 years old 177.34 ± 7.2 cm LE_FM 25.00 ± 2.93	Walking for 2 min at their preferred speed around a level, tiled 60 m indoor walking track, without any aids or LL orthoses.
Kim et al., 2020 [19]	Observational, analytical study	n = 14 (7 M/7 F) 63.80 ± 15.9 years old 4 CLR/10 CLL 2 Hemo / 12 Isch 38.7 ± 45.2 months poststroke Hand Function (%) in SIS (3.0) 52.9 ± 31.7 MMSE > 23	Performance of 2 isometric force control tasks in 3 different randomized conditions: unimanual paretic, unimanual non-paretic, and bimanual. - 2 Submaximal force control task of 3 MVC trials (5 s) with a rest interval of 60 s - 2 Maximal sustained force production task of 3 MVC trials (8 s) with a rest interval of 60 s. For both tasks, participants performed either isometric unimanual or bimanual wrist and finger extension upwards (lifting) towards the padded platforms. Participants sat 78 cm away from a 43.2 cm LCD monitor and placed their forearms on the table in a stable position, maintaining 15–20° of shoulder flexion and 20–40° of elbow flexion. They were instructed to place either unilateral or bilateral hands under the specially padded platforms and to adjust the height of the platform.
Tian et al., 2021 [37]	Observational, analytical study (with healthy control group)	n = 3 (2 M/1 F) 52.33 ± 24.19 years old 1 CLR/2 CLL 2 Hemo/1 Isch 12.67 ± 5.03 months poststroke UE_FM 33.42 ± 19.70 MAS < 2 MMSE > 23	Two tracking tasks of elbow flexion and extension were performed in which the activation of the agonist muscle increased from 0% to 15% MVC in 0–5 s, then gradually relaxed within 5–10 s and returned to 0. Each round movement was performed at a constant speed (15%MVC/5 s) from the origin to the target point, and these included pointing (0% to 15% amplitude) under elbow flexion and pointing (15% to 0% amplitude) under elbow extension. The above procedure was repeated 3 times as a trial. Subjects randomly performed 10 trials (5 flexion trials and 5 extension trials) and rested for 30 s after each trial. During the performance, they were instructed to sit in a height-adjusted chair and place their impaired arm on the horizontal armrest with the elbow flexed at 90° and the shoulders abducted at 90°. They were asked to grasp the handle attached to the armrest.
Xu et al., 2022 [38]	Observational, analytical study (with healthy control group)	n = 10 (8 M/2 F) 48.3 ± 12.8 years old	Performed 3 trials of walking on a treadmill at a comfortable speed for 3 min, with 5 min of rest in between, using special experimental shoes.

M: male; F: female; CLR: contralesional right; CLL: contralesional left; Isch: ischemic; Hemo: hemorrhagic; UE_FM: Fugl-Meyer Assessment Upper Extremity; LE_FM: Fugl-Meyer Assessment Lower Extremity; MAS: Modified Ashworth Scale; MAL: Motor Activity Log; RMI: Rivermead Mobility Index; SIS: Stroke Impact Scale; MMSE: Mini Mental State Examination; MVC: maximal voluntary contraction; MIVE: maximum isometric voluntary extension. MIVE: maximum isometric voluntary flexion; UL: upper limb; LL: lower limb.

Table 3. Assessment instruments, kinematic, kinetic, and EMG variables, non-linear measures, and their parameters.

Author, Year	Assessment Instrument	Kinematic, Kinetic, and EMG Variables	Non-Linear Measure	Non-Linear Measures Parameters
Sethi et al., 2013 [27]	Kinematics: ■ Twelve-camera VICON™ motion capture system (Vicon 612; Oxford Metrics Inc., Oxford, UK) ■ 67 reflective markers in UB were placed using a marker set described by the VICON Plug-In-UE.	Joint Kinematics: ■ Shoulder, elbow, wrist, and PIP index finger (°)	■ ApEn	■ Kinematic data not filtered and surrogation with Theiler's first algorithm (20 surrogate time series of each trial) ■ Each joint angle time series was analyzed from the start of the reach through the entire length of the respective time series including the pauses between the 3 trials. ■ 4 time series were obtained (one for each joint) ■ lag = 1, m = 2, and r = 0.2 times the SD of the time series ■ ApEn values of each participant were normalized to the length of their time series
		Joint Kinematics: ■ Shoulder, elbow, and wrist (°)	■ ApEn	■ Kinematic data not filtered and surrogation with Theiler's first algorithm (20 surrogate time series of each trial) ■ Each joint angle time series was analyzed from the start of the reach through the entire length of the respective time series including the pauses between the three trials. ■ 4 time series were obtained (one for each joint) ■ lag = 1, m = 2, and r = 0.2 times the SD of the time series ■ ApEn values of each participant were normalized to the length of their time series
Sethi et al., 2013 [28]	Kinematics: ■ Twelve-camera VICON™ motion capture system (Vicon 612; Oxford Metrics Inc., Oxford, UK) ■ 67 reflective markers in UB were placed using a marker set described by the VICON Plug-In-UE.	Joint Kinematics: ■ Shoulder, elbow, and wrist (°)	■ ApEn	■ Kinematic data not filtered and surrogation with Theiler's first algorithm (20 surrogate time series of each trial) ■ Each joint angle time series was analyzed from the start of the reach through the entire length of the respective time series including the pauses between the three trials. ■ 4 time series were obtained (one for each joint) ■ lag = 1, m = 2, and r = 0.2 times the SD of the time series ■ ApEn values of each participant were normalized to the length of their time series
		Joint Kinematics: ■ Shoulder, elbow, and wrist (°)	■ ApEn	■ Kinematic data not filtered and surrogation with Theiler's first algorithm (20 surrogate time series of each trial) ■ Each joint angle time series was analyzed from the start of the reach through the entire length of the respective time series including the pauses between the three trials. ■ 4 time series were obtained (one for each joint) ■ lag = 1, m = 2, and r = 0.2 times the SD of the time series ■ ApEn values of each participant were normalized to the length of their time series
Chow and Stokic, 2014 [29]	Kinetics: ■ Biodex System 3 isokinetic dynamometer (Biodex Medical Systems, Inc., New York, NY, USA) and a custom-built amplifier connected directly to the torque sensor of the dynamometer (overall sensitivity 57.5 mV/Nm)	Torque Measurement: ■ MVC torque of quadriceps (Nm) ■ MVC Power of quadriceps (Nm2)	■ SampEn	■ Torques were filtered using a second-order Butterworth low-pass filter with a 30-Hz cutoff ■ Only the middle 8 s of each 10 s trial were analyzed. Out of 2 trials collected at each force level, the one with a lower CV was used for statistical analysis. ■ N = 800, m = 3, and r = 0.2 SD of the time series
		Torque Measurement: ■ MVC torque of quadriceps (Nm) ■ MVC Power of quadriceps (Nm2)	■ SampEn	■ Torques were filtered using a second-order Butterworth low-pass filter with a 30-Hz cutoff ■ Only the middle 8 s of each 10 s trial were analyzed. Out of 2 trials collected at each force level, the one with a lower CV was used for statistical analysis. ■ N = 800, m = 3, and r = 0.2 SD of the time series

Table 3. Cont.

Author, Year	Assessment Instrument	Kinematic, Kinetic, and EMG Variables	Non-Linear Measure	Non-Linear Measures Parameters
Sun et al., 2014 [30]	Kinetics: ■ Robot-aided developed dynamometer ■ sEMG recording system (Noraxon, Scottsdale, AZ, USA)	Torque Measurement: ■ MIVE torque of TRI (Nm) ■ MIVF torque of BIC (Nm) Electromyographic activity: ■ Raw EMG data of TRI and BIC	■ fApEn	■ The EMG data were not filtered ■ $N = 1000$, $m = 2$, and $r = 0.2$ SD of the time series Torque Measurement: ■ A segment of 1000 samples (500 samples before and after the maximum value) was selected for calculating fApEn Electromyographic activity: - fApEn values were calculated with a 1000 ms sliding window and a 10 ms window increment in consideration of the computation power of the computer
Kao et al., 2014 [31]	Kinematics: ■ Eight-camera video system (Motion Analysis Corporation, Santa Rosa, CA, USA) ■ 46 reflective markers attached on the LB, trunk, and over the C7 vertebra	Spatiotemporal parameters: ■ Vertical, anteroposterior and mediolateral of C7 velocities (m/s)	■ Short-term LDE ■ maxFM	■ Delay-embedded state spaces were reconstructed independently from the 3 velocities of non-filtered C7 vertebral marker data ■ Data from 30 continuous strides were extracted for each trial ■ Short-term LDE: ■ Data were resampled to 3000 total data points, approximately 100 data points per stride ■ maxFM: ■ The state spaces were first divided for individual strides and then each stride was time normalized to 101 samples
Ao et al., 2015 [32]	Kinematics: ■ Motion capture system (OptiTrack®, NaturalPoint, USA) ■ Two reflective markers are attached to the elbow joint and at the end of the handle Electromyography: ■ Tele-EMG system (MyoSystem2400T, Noraxon, USA) ■ 2 Ag/AgCl bipolar surface electrodes (Noraxon, USA)	Joint Kinematics: ■ Elbow (°) Electromyographic activity: ■ EMG Amplitude (mv) TRI and BIC	■ fApEn ■ ApEn ■ SampEn	■ The EMG signals were band-pass filtered through a 4th-order, zero-phase Butterworth digital filter with a frequency band from 5 to 400 Hz and a 50 Hz digital notch filter ■ The performances of the three entropies were compared using both simulated signals MIX(p) and real EMG signals. Three simulated signals, MIX(0.1), MIX(0.5) and MIX(0.9) were generated and each one had 5000 samples ■ The 100-point segments were selected randomly from each of the 5000-sample simulated MIX(0.1), MIX(0.5), and MIX(0.9) signals, respectively ■ $m = 2$ ■ N and r were carried out to find a suitable complexity measurement with optimal N and r

Table 3. Cont.

Author, Year	Assessment Instrument	Kinematic, Kinetic, and EMG Variables	Non-Linear Measure	Non-Linear Measures Parameters
Sethi et al., 2017 [33]	Kinematics: ■ Twelve-camera VICON™ motion capture system cameras (Vicon 612/T40; Oxford Metrics, Oxford, England) ■ Sixty-seven reflective markers in UB were placed using a marker set described by the VICON Plug-In-UE	Joint Kinematics: ■ Shoulder, elbow, wrist, and PIP index finger (°) Kinematics Spatiotemporal Parameters: ■ Index finger velocity (m/s) ■ PV (m/s)	■ ApEn	■ Kinematic data not filtered and surrogation with Theiler's first algorithm (20 surrogate time series of each trial) ■ Each joint angle time series was analyzed from the start of the reach through the entire length of the respective time series including the pauses between the three trials. ■ 4 time series were obtained (one for each joint) ■ lag = 1, m = 2, and r = 0.2 times the SD of the time series ■ ApEn values of each participant were normalized to the length of their time series
		Kinetic: ■ One torque sensor (Model TRS 500, Transducers Techniques, CA) on BIC Electromyography: ■ Porti sEMG system (TMS International, The Netherlands)		■ The EMG signals were filtered using a zero-lag band-pass (20–450 Hz) fourth-order Butterworth filter ■ The torque signal was filtered using a zero-lag low-pass second-order Butterworth filter ■ SampEn ■ Both surface EMG and torque data correspond to the same 2 s time window selected for data analysis. ■ N = 4000, m = 2, and r = 0.25 times the SD of the time series
Kempski et al., 2018 [35]	Kinematics: ■ Eight-camera motion capture system (Motion Analysis Corporation) ■ Twenty-five single reflective markers and nineteen markers on shells placed on the bony landmarks and tracking segments	Joint Kinematics: ■ Hip, knee and ankle (°) Kinematics Spatiotemporal Parameters: ■ Gait velocity (m/s)	■ MLYE (Wolf's algorithm)	■ Not specified
		Electromyography: ■ Home-made multi-channel sEMG recording system ■ One large round reference electrode (Dermatode; American Imex, Irvine, CA) was placed on the arm fossa cubitalis	■ SampEn	■ The EMG signal was filtered with a fourth-order zero-lag non-causal Butterworth band-pass filter set at 20–500 Hz. Then, a set of second-order notch filters were used to remove the 50 Hz power line interference and its harmonics. ■ N: the EMG signal was divided into several non-overlapping epochs, each with a time length of 1 s (equivalent to 1000 sample points) ■ m = 2 and r = 0.25 times the SD of the time series

Table 3. Cont.

Author, Year	Assessment Instrument	Kinematic, Kinetic, and EMG Variables	Non-Linear Measure	Non-Linear Measures Parameters
Dugan et al., 2020 [18]	Kinematics: ■ Wireless Inertial Measurement Units (IMUs) (InterSense Inertia Cube BT) attached to each LL segment and pelvis	Joint Kinematics: ■ Hip, knee and ankle (°)	■ mMSE	■ The tri-axial accelerometer data from each IMU dataset were unfiltered ■ N: the last 6000 data points, corresponding to the last 100 s of the walking trials were used for analysis. ■ m = 2 and r = 0.2
		Kinetics: ■ Force transducers (MLP-75, Transducer Techniques, 4.16 × 1.27 × 1.90 cm, range = 75 lbs., 0.1% sensitivity) attached to the padded platforms	■ SampEn	■ The force data were filtered using a bidirectional fourth-order Butterworth filter with a cutoff frequency = 30 Hz ■ N: for the submaximal force control tasks, the middle 16 s of the force signals were used ■ N: For the maximal sustained force production tasks, the middle 5 s of the force data were analyzed ■ m = 2, and r = 0.2 SD of the force signals
Tian et al., 2021 [37]	Electromyography: ■ sEMG system (DAQ USB-6341, National Instruments, USA)	Electromyography activity: ■ Raw EMG data (V) BIC and TRI	■ C-FuzzyEn ■ i-FuzzyEn ■ FuzzyEn ■ iC-FuzzyEn	The raw EMG signals were bandpass filtered using a 4th-order Butterworth filter at 20–300 Hz and a 50 Hz notch filter C-FuzzyEn and iC-FuzzyEn: ■ N = 500 (0.5 s), m = 2, and r = 0.2 SD of the EMG signal i-FuzzyEn: ■ N = 500, m = 2, and r = 0.15 SD of the EMG signal
		Kinematics: ■ Six-camera motion capture system (OptiTrack®, NaturalPoint, USA) ■ Ten 12 mm reflective markers bilaterally in LL were placed on the second and third metatarsal space, the lateral malleolus, the midleg, the lateral knee, and the mid-thigh	■ MFDEA	■ Proper polynomial order = 2 ■ The short-scale exponent was computed within the segment size ranging from 20 < n < 100 sample. Meanwhile, the long-scale exponent was evaluated in the range of 170 < n < 1000 sample
Xu et al., 2022 [38]		Joint Kinematics: ■ Knee (°) ■ Ankle (°)		

UB: upper body; UE: upper extremity; LL: lower limb; sEMG: surface electromyography; PIP: proximal interphalangeal; PV: peak velocity; MVC: maximal voluntary contraction; MVE: maximum isometric voluntary extension. MVEF: maximum isometric voluntary flexion; BIC: biceps brachii; TRI: triceps; SD: standard deviation; V: volts.

3.1. Participant’s Characteristics

Of the 14 included studies, 6 only included post-stroke subjects [19,28,30,33–35], and 8 also included healthy subjects [18,27,29,31,32,36–38]. The sample size of chronic post-stroke participants ranged from 3 [37] to 23 [29].

A total of 150 chronic post-stroke subjects, 102 men and 51 women, with a mean age of 52.33 years, were included in the studies. All the subjects were at a chronic stage, ranging from 12.67 [37] to 177.34 [18] months after stroke, and presented with a single and unilateral [18,19,27,28,30,33,35,37,38] or multiple and unilateral strokes [29]. Two studies did not report lesion characteristics [32,39].

The side of the body most affected was reported in 11 studies [19,27–30,32–37], with a total of 48 right-sided and 71 left-sided strokes. According to the type of stroke, four trials included both ischemic and hemorrhagic stroke [19,30,32,37] and three trials analyzed only ischemic stroke [27,28,33]. Overall, seven trials did not report the type of stroke [18,29,31,34–36,38]. Only one study reported the location of the stroke [19,33].

The Mini Mental State Examination (MMSE) was used in two trials to ensure an understanding of the instructions in the required tasks [19,37]. To assess sensorimotor impairment, the Fugl–Meyer Upper Extremity (UE_FM) was used in five studies [27,28,33,36,37], while the Fugl–Meyer Lower Extremity (LE_FM) was used in four studies [18,29,31,35]. Other clinical measures were included in three trials, namely the Motor Activity Log [28], the Rivermead Mobility Index (RMI) [29], and the Stroke Impact Scale (SIS)—Hand Function [19]. The Modified Ashworth Scale (MAS) was used to quantify spasticity in six studies [27,29,32,34,36,37]. Only two studies did not identify a clinical measure to meet inclusion criteria or sample characterization [30,38].

3.2. Tasks

Considering the tasks (described in Table 2), nine studies focused on UL and four on LL assessment (Figure 2). Regarding UL tasks, two studies analyzed reaching to grasp [27,28] and one analyzed reaching to point [33]. Arm tracking movement was assessed by two studies [30,32]. Five studies focused on elbow flexion [30,32,34,36,37], one on extension [37], and one on wrist and finger extension upward (lifting) [19]. The remaining studies analyzed knee extension [29] and walking [18,31,35,38].

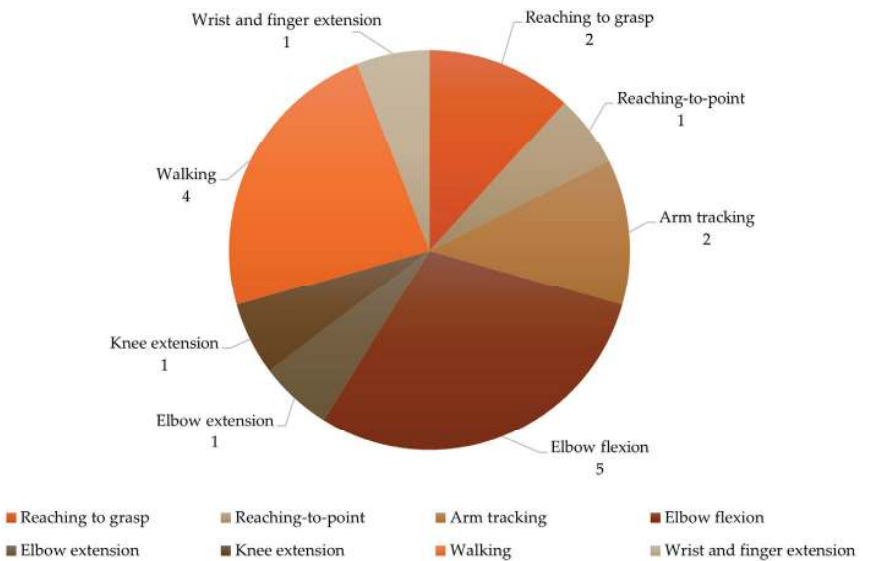


Figure 2. Characterization of the tasks.

3.3. Assessment Instruments

Table 3 shows the instruments used to explore kinematic, kinetic, and EMG variables in the 14 studies. Kinetic data were collected in four studies [19,29,30,34], kinematic data were collected in eight studies [18,27,28,31–33,35,38], and sEMG data were collected in five studies [30,32,34,36,37] (Figure 3).

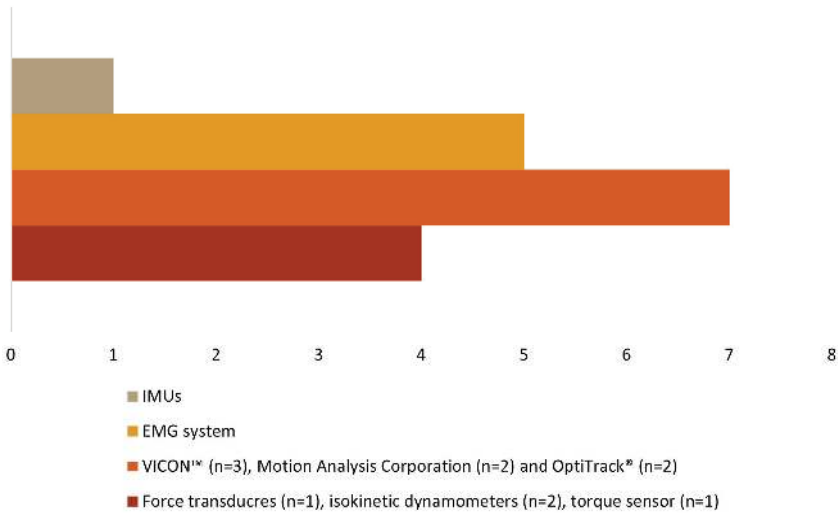


Figure 3. Types of assessment instruments.

Force transducers [19], isokinetic dynamometers [29,30], and a torque sensor [34] were used to evaluate the kinetic data. The muscular activity was assessed using a sEMG system [30,32,34,36,37]. Three different motion capture systems were used for kinematic data: VICON™ [27,28,33], Motion Analysis Corporation [31,35], and OptiTrack® [32,38], with several cameras, ranging from 6 [38] to 12 [27,28,33] in number. One study did not report the number of cameras [32]. The number of reflective markers ranged from 2 [32] to 67 [27,28,33], as different anatomical references were considered within each identified segment according to each author (see Table 3). One study [18] used wireless inertial units (IMUs) in lower limb segments and the pelvis to collect kinematic data.

3.4. Kinetic, Kinematic and EMG Variables

To analyze the human movement data, different instruments were used to collect several variables (Table 3). For kinetic data (Figure 4), three studies quantified the maximal voluntary contraction (MVC) torque and power of the quadriceps [29], elbow flexion [34], and wrist and finger extension [19]; one study used the maximum isometric voluntary extension (MIVE) torque of the elbow [30]; and another used the maximum isometric voluntary flexion (MIVF) torque of the elbow [30]. Regarding sEMG data (Figure 5), one study each measured muscle activity using the root mean square (RMS) of the biceps (BIC) [34], the EMG amplitude of the BIC and triceps (TRI) [32], the EMG integral of the BIC [36]; the remaining two studies used raw sEMG data of the BIC and TRI [30,37].

To analyze the kinematic data (Figure 6), four studies considered the range of motion (ROM) of the upper limb, namely the shoulder [27,28,33], elbow [27,28,32,33], wrist [27,28,33] and proximal interphalangeal (PIP) index finger [27,33]. The range of motion of the lower limb was considered in three studies, specifically the hip [18,35], knee [18,35,38], and ankle [18,35,38]. Only three studies included spatial–temporal kinematic parameters such as index finger velocity and peak velocity [33], C7 velocity [31], and gait velocity [35].

2

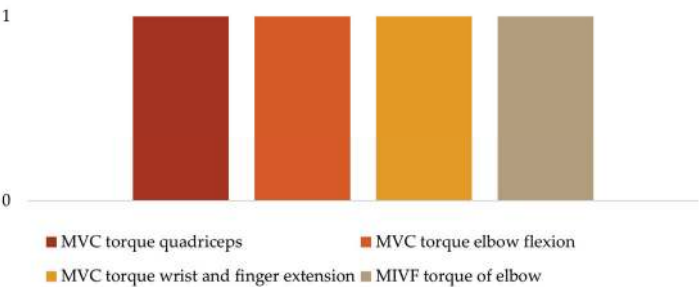


Figure 4. Kinetic variables.

3

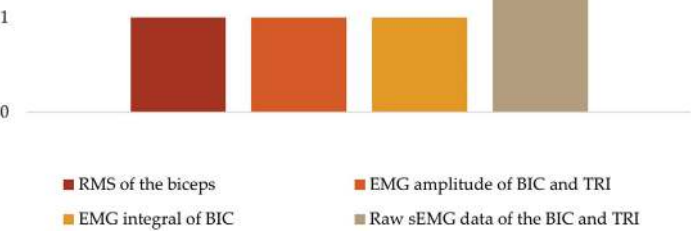


Figure 5. EMG variables.

4

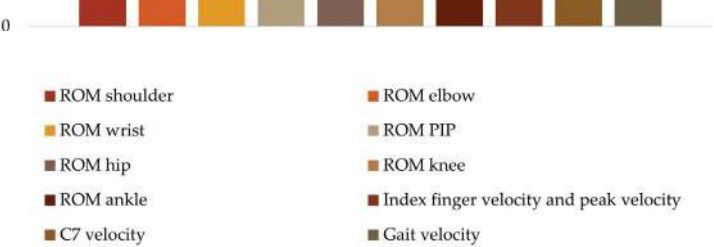


Figure 6. Kinematic variables.

3.5. Non-Linear Measures

Considering the units included in this review, 11 studies used entropy measures (Table 3). Approximate entropy (ApEn) was used in four studies [27,28,32,33]; sample entropy (SampEn) was in five studies [19,29,32,34,36]; fuzzy approximate entropy (fApEn) was used in two studies [30,32], and multivariate multiscale entropy (mMSE) was used in one study [18]. One study [37] considered the instantaneous fuzzy entropy (i-FuzzyEn), the cross-fuzzy entropy (C-FuzzyEn), and the instantaneous cross-fuzzy entropy (iC-FuzzyEn). The three remaining studies utilized the short-term local divergence exponent (LDE) and the maximum Floquet multiplier (maxFM) [31], the maximum Lyapunov exponent (MLyE) [35], and the multifractal detrended fluctuation analysis (MFDFA) [38]. Figure 7 summarizes the non-linear measures utilized in the studies included.

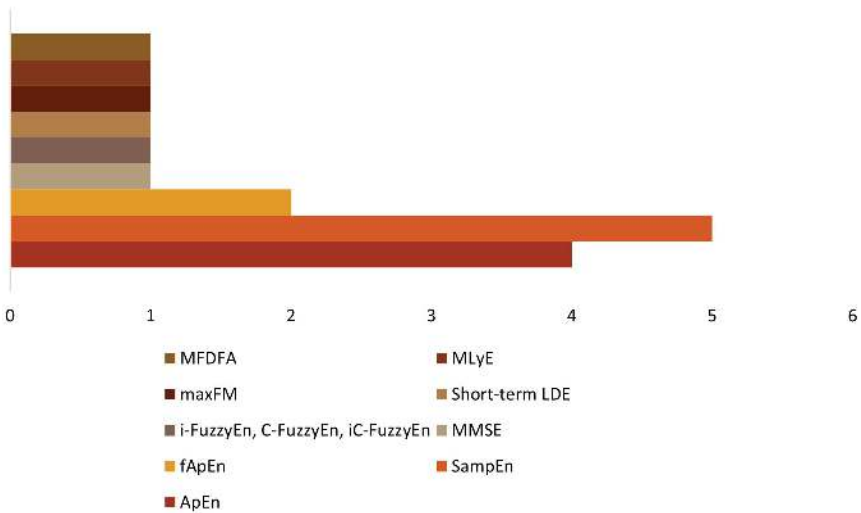


Figure 7. Non-linear measures.

3.6. Non-Linear Measures Parameters

For each of the non-linear measures, the included authors generally described their parameters in detail. Specifically, seven studies mentioned the dataset length (N) [18,19,29–31,34,36], eleven studies described the specific pattern length (m) [18,19,27–30,32–34,36,37], and ten studies specified the criterion of similarity (r) [18,19,27–30,33,34,36,37]. Time series were reduced to N points by analyzing between 20 [38] and 5000 points [32]. The time delay between related series of observations (lag) was described in three studies [27,28,33]. Furthermore, there is variation in the length of m , ranging from 2 ($n = 10$) [18,19,27,28,30,32–34,36,37] to 3 ($n = 1$) [29], and r ranging from 0.2 [18,19,27–30,33,37] to 0.25 [34,36]. In the studies that included lag ($n = 3$), the value remained constant at 1 [27,28,33]. Additionally, the studies had varying numbers of trials and lengths, ranging from 2 [29] to 18 trials/condition [32], and from 3 s [36] to 3 min [38], respectively. Only nine studies described the number of attempts/times analyzed in non-linear analyses [18,19,27–29,31,33,34,36].

4. Discussion

This scoping review aimed to summarize information on the current published research related to the non-linear measures used in the kinetic, kinematic, and EMG data analysis of human movement after stroke.

Considering that everyday functional tasks involve the continuous adaptation of movement patterns to meet their dynamic demands, the study of variability in everyday tasks is a key point in human movement analysis due to its non-linear behavior. This information can provide a greater opportunity to better understand the mechanisms under-

lying an impaired sensorimotor system and may suggest better interventions to improve movement adaptability while encouraging reflection on key findings to be considered in future research. Therefore, this discussion is divided into different aspects: (1) sample characterization; (2) tasks, non-linear measures, and non-linear parameters; (3) assessment instruments; and (4) kinematic, kinetic, and EMG variables.

4.1. Sample Characterization

In the included studies, the characteristics of sex, age, affected side, time after injury, and motor function were found to be the most common and frequently reported anthropometric variables. However, other important features such as stroke location, severity, ethnicity, pre-existing comorbid conditions (diabetes, cognitive impairment, or depression) [21], and pre-stroke lifestyle factors (social engagement and exercise) were not consistently considered [40]. Also, body function, structure, activity limitations, and participation restriction should be assessed using appropriate and adequate clinical measures. Indeed, although the Fugl–Meyer Assessment (FM) was commonly used in the present findings, it was not included in all studies, which may affect our understanding of participants' behavior [21].

It is important to note that limited or varied demographic and stroke-related data may hinder our understanding of stroke recovery trajectories [41]. Therefore, it is crucial to follow the core recommendations for stroke trials, which state that pre-stroke clinical, demographic, and stroke-related data are core measures that should be collected to improve the clinical prediction of recovery and characterization of patient cohorts. For instance, active hand movements and walking assessment are recommended, especially in studies with chronic stroke, when the NIHSS could not be collected at stroke onset [21].

Another important consideration is the presence or absence of non-disabled matched individuals, which facilitates comparisons. Comparing an individual's motor task performance with a reference population of non-disabled matched individuals has been shown to be advantageous for better understanding disabled performance [21]. Despite this assumption, this was only considered in about half of the present studies.

4.2. Tasks and Non-Linear Measures

The intrinsic variability of human movement can be assessed by considering two parameters: stability and complexity [9,42]. Both parameters can be measured using several non-linear measures, such as the MLyE, the maxFM, and fractal measures like MFDFA for stability assessment. In addition, entropy measures are used to quantify the irregularity that is associated with movement complexity [7,9,43].

Although entropy measures provide valuable insights into one aspect of variability, they do not capture the structural abundance and widespread component characterization of a complex system operating across multiple spatial and temporal scales [44]. Therefore, it is recommended to combine them with other non-linear measures for a more comprehensive assessment [7]. Entropy was the most used measure for processing kinetic, kinematic, and EMG data in the reviewed studies. It is worth noting that none of the studies combined entropy with other non-linear measures, and only two studies combined different entropy measures [32,37]. ApEn is a metric commonly used in the biomechanics literature to assess complexity [9,45]. However, SampEn was the most frequently reported measure in the studies reviewed, followed by ApEn. Despite its frequent use, ApEn may lead to inconsistent results due to its sensitivity to the length of the time series [9]. It should be noted that the studies included in this review used ApEn to quantify the temporal structure of variability during three trials of reaching-to-grasp [27,28] and reaching-to-point [33] tasks. Although the authors normalized the ApEn values to the length of their time series [27,28,33], the number of trials and time series length may not be sufficient for consistent results.

SampEn is generally considered a better option for analyzing time series due to its greater relative consistency and lesser dependence on the length of the dataset [20]. It has been widely used as a complexity measure in various fields of research to evaluate the

impact of age, illness [46], and performance [47]. In clinical contexts, this measure can help identify motor coordination disorders [48]. In this review, five of the studies used SampEn, with short-length tasks, to quantify the complexity of the sEMG signal during knee extension [29] and elbow flexion [32,34]; Tang et al. [36] used it as an indicator of the complexity and randomness of the sEMG signal during elbow flexion; and it was used as an indicator of the temporal structure of variability (force regularity) during wrist or finger extension [19]. Additionally, SampEn can differentiate specific characteristics of distinct populations [12].

Although SampEn demonstrates relative consistency and less dependence on data length, its similar definition of vectors is based on the Heaviside function, as in ApEn [49]. However, the use of the Heaviside function has inherent flaws, which can lead to problems in the validity of the entropy definition, particularly when small parameters are involved [50]. In this review, two studies used fApEn [30,32] to measure the time series regularity and complexity of sEMG signals of BIC and TRI during an arm tracking task. This measure was developed resulting from a combination of the concept of “fuzzy sets” and ApEn [51], showing better relative consistency and robustness to noisy physiological signals with short data length [52,53].

Because neuromuscular activities are inherently non-linear [54], cross-entropy-based measures are used to characterize pattern synchronization and quantify inter-muscular coupling [37]. Xie et al. [53] introduced the C-FuzzyEn, based on FuzzyEn. This method appears to be better suited to short time series and is more robust to noise, allowing the observation of the global intermuscular coupling between two separated but interconnected sEMG signals [53]. C-FuzzyEn was used in a study included in this review to characterize the global intermuscular coupling during a tracking task of elbow flexion and extension [37]. The iC-FuzzyEn and i-FuzzyEn are proposed by the authors to characterize the instantaneous intermuscular coupling and the dynamic complexity, respectively, in order to gain a better understanding of the neuromuscular mechanisms after a stroke [37].

Van Emmerik and colleagues [9] demonstrated that ApEn, SampEn, and their variants are limited in their ability to provide a detailed analysis of physiological signals because they only assess entropy on a temporal scale. As physiological systems exhibit structures across multiple time scales [55], mMSE was developed to quantify the intrinsic complexity of the system at small and large scales [55–57]. Nevertheless, only one of the studies included in this review used it to analyze the complexity of gait after stroke [18]. This may suggest an emerging interest in using measures that can better reflect the impact of complexity on the adaptability of the motor system after stroke.

According to the abovementioned studies, using different entropy measures shows a variety of approaches to the analysis of the data. The term complexity seems to have acquired synonyms such as ‘randomness’, ‘temporal structure of variability’, ‘predictability’, or ‘regularity’, depending on the research question of the studies. Therefore, the variety of entropy measures used in the studies analyzed in our review suggests an increasing interest in measures that can more accurately capture the complexity of movement.

Despite the growth of entropy measures, some authors question their validity in assessing the complexity of motion [42]. With our review, MFDFA [38], MLyE [35], short-term LDE [31] and MaxFM [31] were identified as non-linear measures to assess stability during a treadmill (TM) walking task. These are commonly used to assess the local (MLyE and short-term LDE) and orbital (MaxFM) stability of the neuromuscular system, as in the ability to attenuate small perturbations [43].

Xu et al. [38] considered MFDFA to assess gait variability. This measure allows one to capture the multifractal scaling elements in a nonstationary time series [58]. Fractal patterns are frequently observed in healthy and well-controlled movements, indicating a certain level of adaptability and flexibility in motor control [59], and providing information about the underlying control processes involved in movement [60]. As human gait is a cyclic and repetitive set of movements, fractal analysis can be a valuable tool in assessing gait adaptability and flexibility, identifying gait disorders, and predicting fall risk due to

central nervous system dysfunction [39,59,61]. Concerning the quantification of the local dynamic stability of complex non-linear systems, MLyE using Wolf's algorithm [35] and short-term LDE [31] were the measures applied. MLyE quantifies the rate of trajectory convergence or divergence in an n -dimensional state phase and how sensitive the system is to initial conditions [11,62,63]. Short-term LDE, on the other hand, is a measure of the local divergence or convergence of trajectories over a short period of time [11]. It focuses on how trajectories behave in the immediate neighborhood of a given point in phase space [11]. Hence, both concepts are related to the sensitivity of a dynamic system to initial conditions [64]. However, MLyE is typically concerned with long-term behavior, while short-term LyE focuses on local behavior over shorter time intervals, which is consistent with the included studies [31,35]. In the present review, it was identified that Wolf's method was used to calculate MLyE [35]. Nevertheless, the algorithm developed by Rosenstein is the most widely used in biomechanical studies [65]. Considering the length of the data in Kempski et al.'s [35] study, Wolf's algorithm appears to be more appropriate for small gait datasets [66]. The authors of this study considered, based on Moraiti et al. [67], that assessing two minutes of walking is sufficient for accurately calculating the LyE.

The MaxFM method, which quantifies the convergence/divergence of continuous gait variables towards a limit cycle [43], seems to be particularly relevant for analyzing long-distance walking or walking with many steps [68]. However, this condition was not observed in the study that used this measure [35]. In this case, the authors needed to adjust the protocol because the population included (post-stroke) was unable to walk long distances.

These measures address different aspects of stability and can complement each other, revealing underlying patterns, attractor states, and dynamical properties that are not evident through traditional linear methods.

Although non-linear measures can be helpful in interpreting variability dynamics, it is important not to underestimate the importance of appropriate parameter selection [69]. Inadequate selection of non-linear analysis parameters can impact the measurement of system complexity, leading to inconsistent outcomes [70]. Time series data from biological systems are typically non-stationary and noisy, containing extreme values [71]. Noise can have an impact on both the magnitude and structure of variability, potentially obscuring patterns and leading to inaccurate results [72]. To reduce noise and facilitate the accurate identification and analysis of non-linear features [60], filters were employed in eleven of the studies included, with frequencies that preserve relevant physiological information [18,19,27–29,31–34,36,37]. An important methodological consideration is the length of the data time series, as almost all non-linear measures are affected by this feature [69]. Regarding the specific parameters m , r , N , and lag, most of the included studies did not explain why they were chosen, and no article described all of them at the same time. Lag was described in only three studies written by the same study group [27,28,33]. It is important to note that all authors included the data collection time/number of trials in their evaluation protocol. However, two authors did not specify which data were included in the non-linear analysis [37,38]. To improve clarity, it is recommended to define the specific data used in the analysis.

Among the tasks analyzed, it is important to highlight that three studies were related to reaching [27,28,33] and four to walking [18,31,35,38], with both being considered functional daily tasks. The remaining protocols studied focused on isolated components of task execution, such as muscular strength [19,29,30,34,36], muscle activity [30,32,34,36,37], or range of motion [32]. However, research in movement sciences suggests that understanding integrated measures of task execution, such as coordination and adaptability, can provide the necessary information for the mechanisms underlying human movement [73]. Therefore, an incomplete assessment limited to the execution of isolated components diminishes the understanding of movement dysfunction [73]. The study of SMH should involve analyzing various motor tasks [74]. Although the reviewed studies analyzed several tasks (reaching and walking), it is important to also consider other core tasks such as sitting, standing,

transitioning from sitting to standing, transitioning from standing to sitting and climbing or descending a step [75]. These tasks can be used as a starting point for different clinical situations and can help to identify critical aspects of movement [75,76]. This may lead to the development of a standardized approach for analyzing human movement [75,76].

Considering the walking tasks, variability has been assessed in both overground (OG) [18] and TM walking [31,35,38]. The use of TM walking emphasizes the need for a significant amount of continuous data to calculate all non-linear dynamic measures, particularly LyE [77]. However, when using a TM, the constant speed imposes constraints on the gait, reducing the amount of movement variability. It appears to make the temporal structure of gait variability unrealistically regular compared to OG walking [78,79]. Sloot et al. [80] suggested that multiple episodes of shorter data can also be used to increase statistical precision. Therefore, OG walking appears to be a more accurate reflection of gait variability outcomes.

Another aspect to emphasize is that around nine of the included studies focused on the assessment of UL, rather than LL. It has been reported in the literature that the LL recovers more rapidly than the UL [81–85], and that about 40–50% of stroke survivors have UL dysfunction [86]. Considering this review in a chronic stroke population, the results obtained may reflect these assumptions. Also, the lower number of results with protocols involving LL tasks seems to reduce the fact that LL and gait may not express the higher CNS control centers. Thus, walking in humans is mainly produced by the combined roles of the reflex circuit, which produces motor patterns triggered by sensory feedback, and the central pattern generators (CPGs) [87,88]. The CPGs mainly innervate the muscles of the lower limb [89], and they may not be affected following a stroke [90]. Consequently, rhythmic pattern movement such as stepping during walking can be generated even in the absence of control of the higher CNS areas [87]. However, further investigation and analysis are needed to validate this hypothesis and understand the complex mechanisms underlying motor control in the lower limbs.

Incorporating the principles of variability and chaotic variation, together with non-linear tools, can provide additional insights to guide practice and assess human movement dysfunction. It is important to note that non-linear analysis complements traditional linear methods in the understanding of neuromotor control following central nervous system dysfunction.

4.3. Assessment Instruments and Kinematic, Kinetic and EMG Variables

In terms of kinematic data analysis instruments, three motion capture systems were used, namely VICON™ [27,28,33], Motion Analysis Corporation [31,35], and OptiTrack® [32,38], mostly for upper limb tasks. These systems feature high-precision optoelectronic cameras that are considered highly accurate measurement tools for kinematic analysis in the study of human movement [91]. In one of the included studies [18], inertial measurement units (IMUs) were used to evaluate kinematic data during walking. Despite the high accuracy and precision of motion capture systems, they are relatively costly and time-consuming and only allow data to be evaluated in a laboratory context [92]. IMUs can collect data outside the laboratory environment [93] to assess human movement in an easier, low-cost, valid, and reliable way [94,95]. Thus, in the context of research or clinical practice, they could be a viable alternative to motion capture systems. They allow evaluation in a simulated context, as described in the articles included in this review, but also over extended periods in real-world environments [73]. IMUs could contribute to the improvement of motor variability assessment by providing essential data required for the interpretation of real-world readiness for participation [73].

The accuracy of motion capture systems depends on the relative positions of the cameras, the position, the type and number of markers, and their movement within the capture volume [96]. The number of cameras in the included studies ranged from 6 [38] to 12 [27,28,33]. The use of a larger number of cameras can be seen as an advantage, knowing

that more cameras allow for better visual coverage [97], help to avoid marker occlusion by body parts, and allow for better reconstruction of motion in the observed time frame [98].

The optoelectronic systems mentioned in the included studies [27,28,31–33,35,38], used passive marker systems that reflect light back to the sensor. Although the active marker systems may provide more robust measurements, they require additional cables and batteries, limiting freedom of movement [98]. The passive marker systems appear to be less invasive and can provide highly accurate information and simultaneous 3D kinematic measurements during different tasks.

Different motion capture protocols have been found, with varying marker sets, positions, and numbers. The well-documented and consistent use of a set of 67 reflective markers for upper limb joint kinematics is found in only three studies [27,28,33]. Kinematic results are highly sensitive to anatomical markers [99,100]. The use of different marker sets can be a source of variability that can have a significant impact on the kinematic data generated [101]. The only study that used IMUs placed on the LL and pelvis did not specify the number and exact placement of the sensors. Sensor specifications, such as differences in sensor range, sampling rate, and placement, may be the reasons for variability in the reporting of non-linear measurements [102].

Furthermore, the lack of a standardized procedure for kinematic analysis may lead to discrepancies between studies. Data collection, processing, and analysis should be standardized and properly reported to facilitate comparisons between studies and the establishment of reference values for the field. The actual accuracy and precision achieved in practice may vary depending on the setup, user expertise, and environmental conditions [103]. Therefore, appropriate models for the specific research purpose, a single investigator to accurately place the markers [99] or sensors, prior testing of the experimental protocol, and improvement of the most accurate setups could be important factors in increasing the reliability of the data.

A wide variety of kinematic results were observed, with the most pronounced being those related to the joint kinematics of the upper [27,28,32,33] and lower limbs [35,38]. Kinematic assessment is a reliable and objective method and provides quantitative measures of movement control, including motor performance and movement quality [103,104]. Only two authors suggested combining joint kinematics with spatial-temporal parameters [33,35]. This approach integrates the complexity of joint motion and overall movement patterns in space and time, promoting objectivity and practicality by ensuring that assessments are directly related to real-world activities and functional tasks [105].

Regarding the kinetic data, the analyzed articles assessed the torque of several muscles. The UL torque muscle was the most assessed, using an isokinetic dynamometer [30], a torque sensor [34], or a force transducer [19]. An isokinetic dynamometer is frequently used in scientific research to assess muscle strength due to its simplicity and accuracy [106]. It is a reference in the comparison of measurements obtained with other instruments, with excellent levels of validity [107,108] and reliability of muscle strength assessment in UL and LL in chronic post-stroke patients [106,109–111].

The use of sEMG systems has been prevalent in assessing muscle activity during UL-isolated components of task execution [30,32,34,36,37]. Clinically, sEMG provides a non-invasive method to observe neuromotor activity and to assess the individual's neurological status and nervous system reorganization [112].

However, neuromuscular signals have limitations in providing insights into movement quality, which is often assessed via movement function tests or biomechanical performance [113]. Additionally, a broader range of assessment methods and metrics that integrate the benefits of biomechanical (kinematic and kinetic) and neuromuscular (sEMG) measures have been highlighted [113–115]. Although one study combined kinematic and sEMG data [32] and two studies combined kinetic and sEMG data [30,34], none of the others attempted this. Combining data from different domains can provide comprehensive insights by enabling the understanding of the origin of neuromuscular activity through the expression of biomechanical variables. The complexity of human movement's neuro-

motor behavior justifies a combined approach. Particularly in cases of neurological injury, kinematic and kinetic changes occur due to an altered ability to influence neural circuits, whose function is intrinsically linked to muscle activity [116]. Such changes can be more accurately documented using sEMG [117].

Some limitations need to be acknowledged. Firstly, the search was restricted to five databases, which means that it is possible that we may have overlooked some of the literature. Secondly, we intentionally developed a broad search strategy to prevent the possibility of missing important studies. Nevertheless, this scoping review and its findings seem to be indicative of the advances made in this field. The study of non-linear dynamics will allow us to understand the complexity of human movement and evaluate their biological adaptivity.

5. Conclusions

This scoping review provides a comprehensive overview of the current research landscape on non-linear measures in the analysis of human movement post-stroke. Entropy measures, particularly sample entropy, have been the preferred measures when investigating the complexity of human movement. The focus has been on UL tasks such as reaching and components of task execution, such as elbow flexion and extension. Regarding the analyzed variables, the joint kinematics and muscle torque and activity were the most prominent during the UL tasks. Assessment instruments employing motion capture systems and sEMG were the most used tools for kinematic and muscle activity analysis, respectively.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s24072338/s1>.

Author Contributions: The research authors provided the following contributions: Conceptualization, M.F., L.P., A.S., S.S., V.F. and F.P.; methodology, M.F., L.P., A.S., J.P.V.-B., S.S., V.F. and F.P.; validation, M.F., J.P.V.-B. and F.P.; formal analysis, M.F., J.P.V.-B. and F.P.; investigation, M.F., L.P., S.S. and F.P.; writing—original draft preparation, M.F., L.P. and S.S.; writing—review and editing, M.F., A.S., J.P.V.-B. and F.P.; visualization, M.F., A.S. and F.P.; supervision, M.F., J.P.V.-B. and F.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Saladin, L.; Voight, M. Introduction to the movement system as the foundation for physical therapist practice education and research. *Int. J. Sports Phys. Ther.* **2017**, *12*, 858–861. [CrossRef] [PubMed]
2. Cavanaugh, J.T.; Kelty-Stephen, D.G.; Stergiou, N. Multifractality, Interactivity, and the Adaptive Capacity of the Human Movement System: A Perspective for Advancing the Conceptual Basis of Neurologic Physical Therapy. *J. Neurol. Phys. Ther.* **2017**, *41*, 245–251. [CrossRef] [PubMed]
3. Ihlen, E.A.F.; van Schooten, K.S.; Bruijn, S.M.; van Dieën, J.H.; Vereijken, B.; Helbostad, J.L.; Pijnappels, M. Improved Prediction of Falls in Community-Dwelling Older Adults through Phase-Dependent Entropy of Daily-Life Walking. *Front. Aging Neurosci.* **2018**, *10*, 44. [CrossRef] [PubMed]
4. Stergiou, N.; Decker, L.M. Human movement variability, nonlinear dynamics, and pathology: Is there a connection? *Hum. Mov. Sci.* **2011**, *30*, 869–888. [CrossRef] [PubMed]
5. Bisi, M.C.; Stagni, R. Changes of human movement complexity during maturation: Quantitative assessment using multiscale entropy. *Comput. Methods Biomech. Biomed. Eng.* **2018**, *21*, 325–331. [CrossRef] [PubMed]
6. da Costa, C.S.; Batista, M.V.; Rocha, N.A. Quality and structure of variability in children during motor development: A systematic review. *Res. Dev. Disabil.* **2013**, *34*, 2810–2830. [CrossRef] [PubMed]

7. Harbourne, R.T.; Stergiou, N. Movement variability and the use of nonlinear tools: Principles to guide physical therapist practice. *Phys. Ther.* **2009**, *89*, 267–282. [CrossRef] [PubMed]
8. Barela, A.M.F.; Duarte, M. Use of force plate for acquisition of kinetic data during human gait. *Braz. J. Mot. Behav.* **2011**, *6*, 56–61. [CrossRef]
9. van Emmerik, R.E.A.; Ducharme, S.W.; Amado, A.C.; Hamill, J. Comparing dynamical systems concepts and techniques for biomechanical analysis. *J. Sport Health Sci.* **2016**, *5*, 3–13. [CrossRef] [PubMed]
10. Gorniak, S.L. The relationship between task difficulty and motor performance complexity. *Atten. Percept. Psychophys.* **2019**, *81*, 12–19. [CrossRef] [PubMed]
11. Stergiou, N. *Nonlinear Analysis for Human Movement Variability*; CRC Press: Boca Raton, FL, USA, 2016; pp. 1–388.
12. Montesinos, L.; Castaldo, R.; Pecchia, L. On the use of approximate entropy and sample entropy with centre of pressure time-series. *J. Neuroeng. Rehabil.* **2018**, *15*, 116. [CrossRef] [PubMed]
13. Kędziorek, J.; Błażkiewicz, M. Nonlinear Measures to Evaluate Upright Postural Stability: A Systematic Review. *Entropy* **2020**, *22*, 1357. [CrossRef] [PubMed]
14. Feigin, V.L.; Stark, B.A.; Johnson, C.O.; Roth, G.A.; Bisignano, C.; Abady, G.G.; Abbasifard, M.; Abbasi-Kangevari, M.; Abd-Allah, F.; Abedi, V.; et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet Neurol.* **2021**, *20*, 795–820. [CrossRef] [PubMed]
15. Hartman-Maeir, A.; Soroker, N.; Ring, H.; Avni, N.; Katz, N. Activities, participation and satisfaction one-year post stroke. *Disabil. Rehabil.* **2007**, *29*, 559–566. [CrossRef] [PubMed]
16. Adegoke, B.O.; Olaniyi, O.; Akosile, C.O. Weight bearing asymmetry and functional ambulation performance in stroke survivors. *Glob. J. Health Sci.* **2012**, *4*, 87–94. [CrossRef] [PubMed]
17. Strongman, C.; Morrison, A. A scoping review of non-linear analysis approaches measuring variability in gait due to lower body injury or dysfunction. *Hum. Mov. Sci.* **2020**, *69*, 102562. [CrossRef] [PubMed]
18. Dugan, E.L.; Combs-Miller, S.A. Physiological complexity of gait is decreased in individuals with chronic stroke. *Comput. Methods Biomech. Biomed. Eng.* **2019**, *22*, 658–663. [CrossRef]
19. Kim, H.J.; Kang, N.; Cauraugh, J.H. Transient changes in paretic and non-paretic isometric force control during bimanual submaximal and maximal contractions. *J. Neuroeng. Rehabil.* **2020**, *17*, 64. [CrossRef] [PubMed]
20. Liu, X.; Jiang, A.; Xu, N.; Xue, J. Increment Entropy as a Measure of Complexity for Time Series. *Entropy* **2016**, *18*, 22. [CrossRef]
21. Kwakkel, G.; Lannin, N.A.; Borschmann, K.; English, C.; Ali, M.; Churilov, L.; Saposnik, G.; Winstein, C.; van Wegen, E.E.; Wolf, S.L.; et al. Standardized measurement of sensorimotor recovery in stroke trials: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *Int. J. Stroke* **2017**, *12*, 451–461. [CrossRef] [PubMed]
22. Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef]
23. Peters, M.D.J.; Marnie, C.; Tricco, A.C.; Pollock, D.; Munn, Z.; Alexander, L.; McInerney, P.; Godfrey, C.M.; Khalil, H. Updated methodological guidance for the conduct of scoping reviews. *JBI Evid. Synth.* **2020**, *18*, 2119–2126. [CrossRef] [PubMed]
24. Bernhardt, J.; Hayward, K.S.; Kwakkel, G.; Ward, N.S.; Wolf, S.L.; Borschmann, K.; Krakauer, J.W.; Boyd, L.A.; Carmichael, S.T.; Corbett, D.; et al. Agreed Definitions and a Shared Vision for New Standards in Stroke Recovery Research: The Stroke Recovery and Rehabilitation Roundtable Taskforce. *Neurorehabil. Neural Repair.* **2017**, *31*, 793–799. [CrossRef] [PubMed]
25. Aromataris, E.; Munn, Z. Furthering the science of evidence synthesis with a mix of methods. *JBI Evid. Synth.* **2020**, *18*, 2106–2107. [CrossRef] [PubMed]
26. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ* **2021**, *372*, n160. [CrossRef] [PubMed]
27. Sethi, A.; Patterson, T.; McGuirk, T.; Patten, C.; Richards, L.G.; Stergiou, N. Temporal structure of variability decreases in upper extremity movements post stroke. *Clin. Biomech.* **2013**, *28*, 134–139. [CrossRef] [PubMed]
28. Sethi, A.; Davis, S.; McGuirk, T.; Patterson, T.S.; Richards, L.G. Effect of intense functional task training upon temporal structure of variability of upper extremity post stroke. *J. Hand Ther.* **2013**, *26*, 132–137, quiz 138. [CrossRef] [PubMed]
29. Chow, J.W.; Stokic, D.S. Variability, frequency composition, and complexity of submaximal isometric knee extension force from subacute to chronic stroke. *Neuroscience* **2014**, *273*, 189–198. [CrossRef] [PubMed]
30. Sun, R.; Song, R.; Tong, K.Y. Complexity analysis of EMG signals for patients after stroke during robot-aided rehabilitation training using fuzzy approximate entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2014**, *22*, 1013–1019. [CrossRef] [PubMed]
31. Kao, P.C.; Dingwell, J.B.; Higginson, J.S.; Binder-Macleod, S. Dynamic instability during post-stroke hemiparetic walking. *Gait Posture* **2014**, *40*, 457–463. [CrossRef] [PubMed]
32. Ao, D.; Sun, R.; Tong, K.Y.; Song, R. Characterization of stroke- and aging-related changes in the complexity of EMG signals during tracking tasks. *Ann. Biomed. Eng.* **2015**, *43*, 990–1002. [CrossRef] [PubMed]
33. Sethi, A.; Stergiou, N.; Patterson, T.S.; Patten, C.; Richards, L.G. Speed and Rhythm Affect Temporal Structure of Variability in Reaching Poststroke: A Pilot Study. *J. Mot. Behav.* **2017**, *49*, 35–45. [CrossRef] [PubMed]

34. Zhang, X.; Wang, D.; Yu, Z.; Chen, X.; Li, S.; Zhou, P. EMG-Torque Relation in Chronic Stroke: A Novel EMG Complexity Representation with a Linear Electrode Array. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1562–1572. [CrossRef] [PubMed]
35. Kempinski, K.; Awad, L.N.; Buchanan, T.S.; Higginson, J.S.; Knarr, B.A. Dynamic structure of lower limb joint angles during walking post-stroke. *J. Biomech.* **2018**, *68*, 1–5. [CrossRef] [PubMed]
36. Tang, X.; Zhang, X.; Gao, X.; Chen, X.; Zhou, P. A Novel Interpretation of Sample Entropy in Surface Electromyographic Examination of Complex Neuromuscular Alterations in Subacute and Chronic Stroke. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 1878–1888. [CrossRef] [PubMed]
37. Tian, N.; Chen, Y.; Sun, W.; Liu, H.; Wang, X.; Yan, T.; Song, R. Investigating the Stroke- and Aging-Related Changes in Global and Instantaneous Intermuscular Coupling Using Cross-Fuzzy Entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 1573–1582. [CrossRef] [PubMed]
38. Xu, P.; Yu, H.; Wang, X.; Song, R. Characterizing stroke-induced changes in the variability of lower limb kinematics using multifractal detrended fluctuation analysis. *Front. Neurol.* **2022**, *13*, 893999. [CrossRef] [PubMed]
39. Amirpourabasi, A.; Lamb, S.E.; Chow, J.Y.; Williams, G.K.R. Nonlinear Dynamic Measures of Walking in Healthy Older Adults: A Systematic Scoping Review. *Sensors* **2022**, *22*, 4408. [CrossRef] [PubMed]
40. Pongmoragot, J.; Lee, D.S.; Park, T.H.; Fang, J.; Austin, P.C.; Saposnik, G. Stroke and Heart Failure: Clinical Features, Access to Care, and Outcomes. *J. Stroke Cerebrovasc. Dis.* **2016**, *25*, 1048–1056. [CrossRef] [PubMed]
41. Bernhardt, J.; Raffelt, A.; Churilov, L.; Lindley, R.I.; Speare, S.; Ancliffe, J.; Katijahbe, M.A.; Hameed, S.; Lennon, S.; McRae, A.; et al. Exploring threats to generalisability in a large international rehabilitation trial (AVERT). *BMJ Open* **2015**, *5*, e008378. [CrossRef] [PubMed]
42. Stergiou, N.; Harbourne, R.; Cavanaugh, J. Optimal movement variability: A new theoretical perspective for neurologic physical therapy. *J. Neurol. Phys. Ther.* **2006**, *30*, 120–129. [CrossRef] [PubMed]
43. Bruijn, S.M.; Meijer, O.G.; Beek, P.J.; van Dieën, J.H. Assessing the stability of human locomotion: A review of current measures. *J. R. Soc. Interface* **2013**, *10*, 20120999. [CrossRef] [PubMed]
44. Yentes, J.M.; Raffalt, P.C. Entropy Analysis in Gait Research: Methodological Considerations and Recommendations. *Ann. Biomed. Eng.* **2021**, *49*, 979–990. [CrossRef] [PubMed]
45. Yentes, J.M.; Hunt, N.; Schmid, K.K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The appropriate use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.* **2013**, *41*, 349–365. [CrossRef] [PubMed]
46. Delgado-Bonal, A.; Marshak, A. Approximate Entropy and Sample Entropy: A Comprehensive Tutorial. *Entropy* **2019**, *21*, 541. [CrossRef] [PubMed]
47. Lubetzky, A.V.; Harel, D.; Lubetzky, E. On the effects of signal processing on sample entropy for postural control. *PLoS ONE* **2018**, *13*, e0193460. [CrossRef] [PubMed]
48. Harrison, S.J.; Stergiou, N. Complex Adaptive Behavior and Dexterous Action. *Nonlinear Dyn. Psychol. Life Sci.* **2015**, *19*, 345–394.
49. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef] [PubMed]
50. Chen, W.; Zhuang, J.; Yu, W.; Wang, Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. *Med. Eng. Phys.* **2009**, *31*, 61–68. [CrossRef] [PubMed]
51. Chen, W.; Wang, Z.; Xie, H.; Yu, W. Characterization of surface EMG signal based on fuzzy entropy. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2007**, *15*, 266–272. [CrossRef] [PubMed]
52. Hornero, R.; Abásolo, D.; Escudero, J.; Gómez, C. Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. *Philos. Trans. A Math. Phys. Eng. Sci.* **2009**, *367*, 317–336. [CrossRef] [PubMed]
53. Xie, H.-B.; Zheng, Y.-P.; Guo, J.-Y.; Chen, X. Cross-fuzzy entropy: A new method to test pattern synchrony of bivariate time series. *Inf. Sci.* **2010**, *180*, 1715–1724. [CrossRef]
54. Jin, S.H.; Lin, P.; Hallett, M. Linear and nonlinear information flow based on time-delayed mutual information method and its application to corticomuscular interaction. *Clin. Neurophysiol.* **2010**, *121*, 392–401. [CrossRef] [PubMed]
55. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy: A tool for complexity analysis of multichannel data. *Phys. Rev. E* **2011**, *84*, 061918. [CrossRef] [PubMed]
56. Busa, M.A.; van Emmerik, R.E.A. Multiscale entropy: A tool for understanding the complexity of postural control. *J. Sport. Health Sci.* **2016**, *5*, 44–51. [CrossRef] [PubMed]
57. Costa, M.; Goldberger, A.L.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906. [CrossRef] [PubMed]
58. Kantelhardt, J.W.; Zschiegner, S.A.; Koscielny-Bunde, E.; Havlin, S.; Bunde, A.; Stanley, H.E. Multifractal detrended fluctuation analysis of nonstationary time series. *Phys. A Stat. Mech. Appl.* **2002**, *316*, 87–114. [CrossRef]
59. Marmelat, V.; Torre, K.; Beek, P.J.; Daffertshofer, A. Persistent fluctuations in stride intervals under fractal auditory stimulation. *PLoS ONE* **2014**, *9*, e91949. [CrossRef] [PubMed]
60. Rhea, C.K.; Kiefer, A.W.; Wittstein, M.W.; Leonard, K.B.; MacPherson, R.P.; Wright, W.G.; Haran, F.J. Fractal gait patterns are retained after entrainment to a fractal stimulus. *PLoS ONE* **2014**, *9*, e106755. [CrossRef] [PubMed]
61. Phinyomark, A.; Larracy, R.; Scheme, E. Fractal Analysis of Human Gait Variability via Stride Interval Time Series. *Front. Physiol.* **2020**, *11*, 333. [CrossRef] [PubMed]

62. Liu, K.; Wang, H.; Xiao, J.; Taha, Z. Analysis of human standing balance by largest lyapunov exponent. *Comput. Intell. Neurosci.* **2015**, *2015*, 158478. [CrossRef] [PubMed]
63. Mehdizadeh, S. The largest Lyapunov exponent of gait in young and elderly individuals: A systematic review. *Gait Posture* **2018**, *60*, 241–250. [CrossRef] [PubMed]
64. Toebes, M.J.; Hoozemans, M.J.; Furrer, R.; Dekker, J.; van Dieën, J.H. Local dynamic stability and variability of gait are associated with fall history in elderly subjects. *Gait Posture* **2012**, *36*, 527–531. [CrossRef] [PubMed]
65. Mehdizadeh, S.; Sanjari, M.A. Effect of noise and filtering on largest Lyapunov exponent of time series associated with human walking. *J. Biomech.* **2017**, *64*, 236–239. [CrossRef] [PubMed]
66. Cignetti, F.; Decker, L.M.; Stergiou, N. Sensitivity of the Wolf's and Rosenstein's algorithms to evaluate local dynamic stability from small gait data sets. *Ann. Biomed. Eng.* **2012**, *40*, 1122–1130. [CrossRef] [PubMed]
67. Moraiti, C.O.; Stergiou, N.; Vasiliadis, H.S.; Motsis, E.; Georgoulis, A. Anterior cruciate ligament reconstruction results in alterations in gait variability. *Gait Posture* **2010**, *32*, 169–175. [CrossRef] [PubMed]
68. Ahn, J.; Hogan, N. Is estimation of Floquet multipliers of human walking valid? In Proceedings of the 2014 40th Annual Northeast Bioengineering Conference (NEBEC), Boston, MA, USA, 25–27 April 2014; pp. 1–2.
69. Caballero, C.; Barbado, D.; Moreno, F. Non-linear tools and methodological concerns measuring human movement variability: An overview. *Eur. J. Hum. Mov.* **2014**, *32*, 61–81.
70. Müller, W.; Jung, A.; Ahammer, H. Advantages and problems of nonlinear methods applied to analyze physiological time signals: Human balance control as an example. *Sci. Rep.* **2017**, *7*, 2464. [CrossRef] [PubMed]
71. Wijnants, M.L.; Bosman, A.M.; Hasselman, F.; Cox, R.F.; Van Orden, G.C. 1/f scaling in movement time changes with practice in precision aiming. *Nonlinear Dyn. Psychol. Life Sci.* **2009**, *13*, 79–98.
72. Costa, M.; Priplata, A.A.; Lipsitz, L.A.; Wu, Z.; Huang, N.E.; Goldberger, A.L.; Peng, C.K. Noise and poise: Enhancement of postural complexity in the elderly with a stochastic-resonance-based therapy. *Europhys. Lett.* **2007**, *77*, 68008. [CrossRef] [PubMed]
73. Kwakkel, G.; Stinear, C.; Essers, B.; Munoz-Novoa, M.; Branscheidt, M.; Cabanas-Valdés, R.; Lakičević, S.; Lampropoulou, S.; Luft, A.R.; Marque, P.; et al. Motor rehabilitation after stroke: European Stroke Organisation (ESO) consensus-based definition and guiding framework. *Eur. Stroke J.* **2023**, *8*, 880–894. [CrossRef] [PubMed]
74. Ozturk, A.; Tartar, A.; Ersoz Huseyinsinoglu, B.; Ertas, A.H. A clinically feasible kinematic assessment method of upper extremity motor function impairment after stroke. *Measurement* **2016**, *80*, 207–216. [CrossRef]
75. Hedman, L.D.; Quinn, L.; Gill-Body, K.; Brown, D.A.; Quiben, M.; Riley, N.; Scheets, P.L. White Paper: Movement System Diagnoses in Neurologic Physical Therapy. *J. Neurol. Phys. Ther.* **2018**, *42*, 110–117. [CrossRef] [PubMed]
76. Quinn, L.; Riley, N.; Tyrell, C.M.; Judd, D.L.; Gill-Body, K.M.; Hedman, L.D.; Packel, A.; Brown, D.A.; Nabar, N.; Scheets, P. A Framework for Movement Analysis of Tasks: Recommendations from the Academy of Neurologic Physical Therapy's Movement System Task Force. *Phys. Ther.* **2021**, *101*, pzab154. [CrossRef] [PubMed]
77. Kyvelidou, A.; Kurz, M.J.; Ehlers, J.L.; Stergiou, N. Aging and partial body weight support affects gait variability. *J. Neuroeng. Rehabil.* **2008**, *5*, 22. [CrossRef] [PubMed]
78. Hollman, J.H.; Watkins, M.K.; Imhoff, A.C.; Braun, C.E.; Akervik, K.A.; Ness, D.K. Complexity, fractal dynamics and determinism in treadmill ambulation: Implications for clinical biomechanists. *Clin. Biomech.* **2016**, *37*, 91–97. [CrossRef] [PubMed]
79. Terrier, P.; Dériaz, O. Kinematic variability, fractal dynamics and local dynamic stability of treadmill walking. *J. Neuroeng. Rehabil.* **2011**, *8*, 12. [CrossRef] [PubMed]
80. Sloom, L.H.; van Schooten, K.S.; Bruijn, S.M.; Kingma, H.; Pijnappels, M.; van Dieën, J.H. Sensitivity of local dynamic stability of over-ground walking to balance impairment due to galvanic vestibular stimulation. *Ann. Biomed. Eng.* **2011**, *39*, 1563–1569. [CrossRef] [PubMed]
81. Paci, M.; Nannetti, L.; Casavola, D.; Lombardi, B. Differences in motor recovery between upper and lower limbs: Does stroke subtype make the difference? *Int. J. Rehabil. Res.* **2016**, *39*, 185–187. [CrossRef]
82. Hamzat, T.K.; Owolabi, M.O.; Vincent-Onabajo, G.O. Trajectory of Motor Performance over Twelve Months in Nigerian Stroke Survivors. *Brain Impair.* **2014**, *15*, 43–50. [CrossRef]
83. Skurvydas, A.; Juodzbaliene, V.; Darbutas, T.; Brazaitis, M. One year after ischemic stroke: Changes in leg movement path stability in a speed-accuracy task but no major effects on the hands. *Hum. Mov. Sci.* **2018**, *57*, 50–58. [CrossRef]
84. Kong, K.H.; Lee, J. Temporal recovery and predictors of upper limb dexterity in the first year of stroke: A prospective study of patients admitted to a rehabilitation centre. *NeuroRehabilitation* **2013**, *32*, 345–350. [CrossRef] [PubMed]
85. Rand, D.; Eng, J.J. Disparity between functional recovery and daily use of the upper and lower extremities during subacute stroke rehabilitation. *Neurorehabil. Neural Repair.* **2012**, *26*, 76–84. [CrossRef] [PubMed]
86. Broeks, J.; Lankhorst, G.; Rumping, K.; Prevo, A. The long-term outcome of arm function after stroke: Results of a follow-up study. *Disabil. Rehabil.* **1999**, *21*, 357–364. [CrossRef] [PubMed]
87. Ryu, H.X.; Kuo, A.D. An optimality principle for locomotor central pattern generators. *Sci. Rep.* **2021**, *11*, 13140. [CrossRef] [PubMed]
88. Minassian, K.; Hofstoetter, U.S.; Dzeladini, F.; Guertin, P.A.; Ijspeert, A. The Human Central Pattern Generator for Locomotion: Does It Exist and Contribute to Walking? *Neuroscientist* **2017**, *23*, 649–663. [CrossRef] [PubMed]
89. Klarner, T.; Zehr, E.P. Sherlock Holmes and the curious case of the human locomotor central pattern generator. *J. Neurophysiol.* **2018**, *120*, 53–77. [CrossRef] [PubMed]

90. Kathe, C.; Skinnider, M.A.; Hutson, T.H.; Regazzi, N.; Gautier, M.; Demesmaeker, R.; Komi, S.; Ceto, S.; James, N.D.; Cho, N.; et al. The neurons that restore walking after paralysis. *Nature* **2022**, *611*, 540–547. [CrossRef] [PubMed]
91. Corazza, S.; Mündermann, L.; Gambaretto, E.; Ferrigno, G.; Andriacchi, T. Markerless Motion Capture through Visual Hull, Articulated ICP and Subject Specific Model Generation. *Int. J. Comput. Vis.* **2010**, *87*, 156–169. [CrossRef]
92. Hamacher, D.; Singh, N.B.; Van Dieën, J.H.; Heller, M.O.; Taylor, W.R. Kinematic measures for assessing gait stability in elderly individuals: A systematic review. *J. R. Soc. Interface* **2011**, *8*, 1682–1698. [CrossRef] [PubMed]
93. Usmani, S.; Saboor, A.; Haris, M.; Khan, M.A.; Park, H. Latest Research Trends in Fall Detection and Prevention Using Machine Learning: A Systematic Review. *Sensors* **2021**, *21*, 5134. [CrossRef] [PubMed]
94. Drapeaux, A.; Carlson, K. A comparison of inertial motion capture systems: DorsaVi and Xsens. *Int. J. Kinesiol. Sports Sci.* **2020**, *8*, 24–27. [CrossRef]
95. Chen, S.; Lach, J.; Lo, B.; Yang, G.Z. Toward Pervasive Gait Analysis with Wearable Sensors: A Systematic Review. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 1521–1537. [CrossRef] [PubMed]
96. Maletsky, L.P.; Sun, J.; Morton, N.A. Accuracy of an optical active-marker system to track the relative motion of rigid bodies. *J. Biomech.* **2007**, *40*, 682–685. [CrossRef] [PubMed]
97. van der Kruk, E.; Reijne, M.M. Accuracy of human motion capture systems for sport applications; state-of-the-art review. *Eur. J. Sport. Sci.* **2018**, *18*, 806–819. [CrossRef] [PubMed]
98. Stancic, I.; Supuk, T.; Panjkota, A. Design, development and evaluation of optical motion-tracking system based on active white light markers. *IET Sci. Meas.* **2013**, *7*, 206–214. [CrossRef]
99. Della Croce, U.; Leardini, A.; Chiari, L.; Cappozzo, A. Human movement analysis using stereophotogrammetry. Part 4: Assessment of anatomical landmark misplacement and its effects on joint kinematics. *Gait Posture* **2005**, *21*, 226–237. [CrossRef] [PubMed]
100. Kainz, H.; Modenese, L.; Lloyd, D.G.; Maine, S.; Walsh, H.P.J.; Carty, C.P. Joint kinematic calculation based on clinical direct kinematic versus inverse kinematic gait models. *J. Biomech.* **2016**, *49*, 1658–1669. [CrossRef] [PubMed]
101. Mantovani, G.; Lamontagne, M. How Different Marker Sets Affect Joint Angles in Inverse Kinematics Framework. *J. Biomech. Eng.* **2017**, *139*, 044503. [CrossRef] [PubMed]
102. Preatoni, E.; Bergamini, E.; Fantozzi, S.; Giraud, L.L.; Orejel Bustos, A.S.; Vannozzi, G.; Camomilla, V. The Use of Wearable Sensors for Preventing, Assessing, and Informing Recovery from Sport-Related Musculoskeletal Injuries: A Systematic Scoping Review. *Sensors* **2022**, *22*, 3225. [CrossRef] [PubMed]
103. Colyer, S.L.; Evans, M.; Cosker, D.P.; Salo, A.I.T. A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods towards Developing a Markerless System. *Sports Med. Open* **2018**, *4*, 24. [CrossRef] [PubMed]
104. Collins, K.C.; Kennedy, N.C.; Clark, A.; Pomeroy, V.M. Kinematic Components of the Reach-to-Target Movement after Stroke for Focused Rehabilitation Interventions: Systematic Review and Meta-Analysis. *Front. Neurol.* **2018**, *9*, 472. [CrossRef] [PubMed]
105. Roggio, F.; Ravalli, S.; Maugeri, G.; Bianco, A.; Palma, A.; Di Rosa, M.; Musumeci, G. Technological advancements in the analysis of human motion and posture management through digital devices. *World J. Orthop.* **2021**, *12*, 467–484. [CrossRef] [PubMed]
106. Rabelo, M.; Nunes, G.S.; da Costa Amante, N.M.; de Noronha, M.; Fachin-Martins, E. Reliability of muscle strength assessment in chronic post-stroke hemiparesis: A systematic review and meta-analysis. *Top. Stroke Rehabil.* **2016**, *23*, 26–36. [CrossRef] [PubMed]
107. Ekstrand, E.; Lexell, J.; Brogårdh, C. Isometric and isokinetic muscle strength in the upper extremity can be reliably measured in persons with chronic stroke. *J. Rehabil. Med.* **2015**, *47*, 706–713. [CrossRef] [PubMed]
108. Kwong, P.W.H.; Ng, S.S.M.; Chung, R.C.K.; Ng, G.Y.F. A structural equation model of the relationship between muscle strength, balance performance, walking endurance and community integration in stroke survivors. *PLoS ONE* **2017**, *12*, e0185807. [CrossRef] [PubMed]
109. Rabelo, M.; Fachin-Martins, E. Inter-rater and test/retest reliabilities of the isokinetic measurements: Assessing strength and endurance of the trunk muscles in two different protocols for able-bodied and post-stroke hemiparesis. *Top. Stroke Rehabil.* **2018**, *25*, 424–431. [CrossRef] [PubMed]
110. Abdollahi, I.; Taghizadeh, A.; Shakeri, H.; Eivazi, M.; Jaberzadeh, S. The relationship between isokinetic muscle strength and spasticity in the lower limbs of stroke patients. *J. Bodyw. Mov. Ther.* **2015**, *19*, 284–290. [CrossRef] [PubMed]
111. Karthikbabu, S.; Chakrapani, M. Hand-Held Dynamometer is a Reliable Tool to Measure Trunk Muscle Strength in Chronic Stroke. *J. Clin. Diagn. Res.* **2017**, *11*, Yc09–Yc12. [CrossRef] [PubMed]
112. Merletti, R.; Temporiti, F.; Gatti, R.; Gupta, S.; Sandrini, G.; Serrao, M. Translation of surface electromyography to clinical and motor rehabilitation applications: The need for new clinical figures. *Transl. Neurosci.* **2023**, *14*, 20220279. [CrossRef] [PubMed]
113. Belfatto, A.; Scano, A.; Chiavenna, A.; Mastropietro, A.; Mrakic-Sposta, S.; Pittaccio, S.; Molinari Tosatti, L.; Molteni, F.; Rizzo, G. A Multiparameter Approach to Evaluate Post-Stroke Patients: An Application on Robotic Rehabilitation. *Appl. Sci.* **2018**, *8*, 2248. [CrossRef]
114. Fang, C.; He, B.; Wang, Y.; Cao, J.; Gao, S. EMG-Centered Multisensory Based Technologies for Pattern Recognition in Rehabilitation: State of the Art and Challenges. *Biosensors* **2020**, *10*, 85. [CrossRef] [PubMed]
115. Pierella, C.; Pirondini, E.; Kinany, N.; Coscia, M.; Giang, C.; Miehlbradt, J.; Magnin, C.; Nicolo, P.; Dalise, S.; Sgherri, G.; et al. A multimodal approach to capture post-stroke temporal dynamics of recovery. *J. Neural Eng.* **2020**, *17*, 045002. [CrossRef] [PubMed]

116. Hallett, M.; DelRosso, L.M.; Elble, R.; Ferri, R.; Horak, F.B.; Lehericy, S.; Mancini, M.; Matsushashi, M.; Matsumoto, R.; Muthuraman, M.; et al. Evaluation of movement and brain activity. *Clin. Neurophysiol.* **2021**, *132*, 2608–2638. [CrossRef] [PubMed]
117. Maura, R.M.; Rueda Parra, S.; Stevens, R.E.; Weeks, D.L.; Wolbrecht, E.T.; Perry, J.C. Literature review of stroke assessment for upper-extremity physical function via EEG, EMG, kinematic, and kinetic measurements and their reliability. *J. Neuroeng. Rehabil.* **2023**, *20*, 21. [CrossRef] [PubMed]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Wearable Sensors as a Preoperative Assessment Tool: A Review

Aron Syversen ^{1,*}, Alexios Dosis ², David Jayne ² and Zhiqiang Zhang ³¹ School of Computing, University of Leeds, Leeds LS2 9JT, UK² School of Medicine, University of Leeds, Leeds LS2 9JT, UK; a.dosis@leeds.ac.uk (A.D.); d.g.jayne@leeds.ac.uk (D.J.)³ School of Electrical Engineering, University of Leeds, Leeds LS2 9JT, UK; z.zhang3@leeds.ac.uk

* Correspondence: scabs@leeds.ac.uk

Abstract: Surgery is a common first-line treatment for many types of disease, including cancer. Mortality rates after general elective surgery have seen significant decreases whilst postoperative complications remain a frequent occurrence. Preoperative assessment tools are used to support patient risk stratification but do not always provide a precise and accessible assessment. Wearable sensors (WS) provide an accessible alternative that offers continuous monitoring in a non-clinical setting. They have shown consistent uptake across the perioperative period but there has been no review of WS as a preoperative assessment tool. This paper reviews the developments in WS research that have application to the preoperative period. Accelerometers were consistently employed as sensors in research and were frequently combined with photoplethysmography or electrocardiography sensors. Pre-processing methods were discussed and missing data was a common theme; this was dealt with in several ways, commonly by employing an extraction threshold or using imputation techniques. Research rarely processed raw data; commercial devices that employ internal proprietary algorithms with pre-calculated heart rate and step count were most commonly employed limiting further feature extraction. A range of machine learning models were used to predict outcomes including support vector machines, random forests and regression models. No individual model clearly outperformed others. Deep learning proved successful for predicting exercise testing outcomes but only within large sample-size studies. This review outlines the challenges of WS and provides recommendations for future research to develop WS as a viable preoperative assessment tool.

Keywords: wearable sensors; exercise testing; preoperative assessment; perioperative pathway

Citation: Syversen, A.; Dosis, A.; Jayne, D.; Zhang, Z. Wearable Sensors as a Preoperative Assessment Tool: A Review. *Sensors* **2024**, *24*, 482. <https://doi.org/10.3390/s24020482>

Academic Editor: José Machado Da Silva

Received: 23 November 2023

Revised: 6 January 2024

Accepted: 9 January 2024

Published: 12 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Demand for general surgery is expected to increase in line with population ageing [1]. For various cancer types, abdominal surgery is considered first-line treatment [2,3]; for advanced-stage cancers, surgical treatment when combined with neoadjuvant treatment may be the only possible cure [2–4]. A common example of this is bowel/colorectal cancer, globally the third most common cancer [5]. Survival from bowel cancer has seen significant improvements in recent decades, with developments in surgical and perioperative care being suggested as reasons [6,7]. Further, mortality in the postoperative period has seen international decline and is frequently reported at below 5% [8–10]. However, postoperative complication rates in these populations have not seen the same decline.

Postoperative complications following abdominal surgery are frequently reported; up to a third of patients report some form of complication [11,12]. Common complications include surgical site infection, cardio-respiratory complications, gastrointestinal (GI) motility problems and anastomotic leak [13]. Postoperative complications have a profound impact not only on patients' quality of life but also on the hospital as a service provider [13]. They are strongly associated with prolonged length of hospital stay, readmission to ICU and hospital re-admissions [13]. One large cohort review concluded that 23.3% of patients undergoing colorectal surgery were readmitted to hospital [14]. This results in a huge

economic burden and has been evidenced widely across many studies [15]. With the population ageing, these costs are expected to increase up to 10% each year [16]. The identification of high-risk patients poses a crucial challenge for healthcare providers to optimise their allocation of resources and for the perioperative management of the patient. Preoperative assessment tools are used to support risk stratification of patients; however, none currently provide a precise assessment that is accessible to all patients.

1.1. Preoperative Assessment

Preoperative assessment occurs during the first stage of the perioperative period, as seen in Figure 1. Here, preoperative measurements are recorded from patients and used to stratify them into risk groups. The major goal of this is to identify the patients at the highest risk of perioperative morbidity and mortality [17]. As explained previously, it can also support the healthcare provider in resource allocation by estimating the support a patient may require across the perioperative pathway. However, as outlined by NICE guidelines, excessive preoperative testing is related to patient anxiety and significant delays to treatment [18]. Therefore, the benefits of testing should be carefully considered before implementation. The most common preoperative assessment tools are outlined here (see Figure 2).

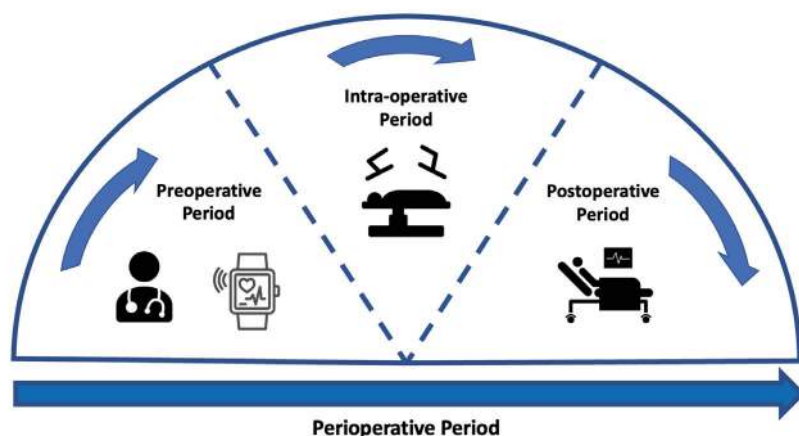


Figure 1. Figure to show the stages across the perioperative period. The perioperative pathway refers to the period that spans from the first point at which surgery is considered as a treatment option up until the full recovery [19]. This pathway has several sub-stages [20]. The preoperative period represents the period prior to surgery where any preoperative assessment takes place. The intra-operative period is representative of the period whilst the patient is undergoing treatment. The postoperative period relates to any period immediately following the operation and can continue after patient discharge.

Physical examinations build on an assessment of the patient's medical history. These pre-anaesthesia examinations include a physical assessment of the lungs, heart function and possible evaluation of the main vital signs using a variety of tests [17,18]. NICE guidelines provide a breakdown of recommendations for testing that vary depending on the severity of the surgical treatment and health status of the patient [18]. For example, an ECG is a common tool that has been shown to optimise risk stratification of cardiovascular complications for non-cardiac surgery [21]. However, for minor/intermediate treatments in young or patients considered healthy, a resting ECG is not recommended as part of routine preoperative assessment [18]. These physical assessments are routinely performed at a preoperative clinic appointment in a resting state.

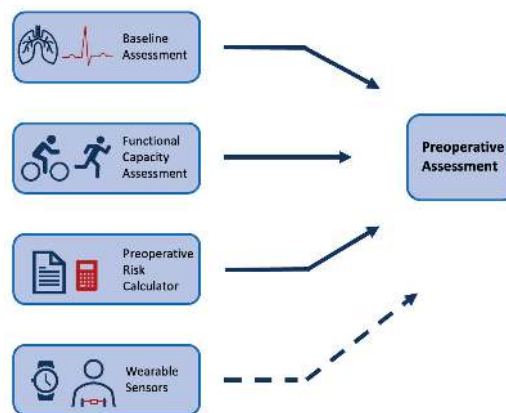


Figure 2. Figure to show common preoperative assessment tools used in practice. The top three boxes present common forms of preoperative assessment that are regularly used in practice (see Section 1.1), whilst the last box with a dashed arrow is included to show the potential for wearable sensors to be used alongside common methods in this context.

Multiple preoperative assessment tools exist that calculate patients' risk of adverse outcomes from routinely collected data. These tools are widely recommended by medical societies to be employed as a preoperative assessment tool [22]. The ASA system (American Society of Anaesthesiology), the APACHE II (Acute Physiology and Chronic Health Evaluation) and the POSSUM (Physiological and Operative Severity Score for Mortality and morbidity) scores have been shown to have good predictive value [23–25]. A review of common preoperative tools has shown that they have comparable predictive performance to machine learning (ML) techniques [26]. However, these tools are not consistently employed in practice. Lack of time and trust in the accuracy of measurements has frequently been reported by clinicians as a barrier to the implementation of risk calculators [27]. They present issues in that they can be open to subjectivity and sometimes require the input of variables that are not available in the preoperative period [25,28]. Further, the majority were originally developed with evidence that predates the last three decades of research [23–25].

Functional capacity assessment is a measure that aims to quantify the ability of a patient to undertake activities from a free-living environment that need 'sustained aerobic metabolism' [29]. Much research has identified the association between a higher functional capacity and a reduction in postoperative complications [30–32]. The 6-min walk test is a common exercise tolerance test but lacks accuracy [33]. In comparison, cardiopulmonary exercise testing (CPET) is considered the current gold standard for preoperative assessment. CPET is a non-invasive clinical tool that evaluates cardio-respiratory function to measure exercise capacity [34]. In the clinic, patients undergo an exercise test on a cycle-ergometer or a treadmill whilst ventilation and respiratory gas parameters are measured [35]. Multiple studies have been shown to support its use as a tool to identify patients at increased risk of developing postoperative complications following general surgery [36,37]. Although CPET is a proven tool for risk stratification and is routinely implemented, there are several barriers to CPET being accessible and precise. CPET requires trained specialists to complete testing with ready access to dedicated facilities; in 2018, only 53% of Trusts in the UK offered the service [38]. Further, CPET is an expensive test with costs estimated at £289 per unit of testing in 2018/2019 (NHS Improvement, 2019). Although the test measures direct oxygen consumption, there can be considerable subjectivity with one study reporting possible miss-classification of outcomes in up to 60% of tests completed [39]. Finally, CPET might be contraindicated meaning that patients who are at high risk of complications are not always able to achieve a representative score, or even complete the test [40]. Wearable

technology has been proposed as a tool that can overcome some of the barriers that are common in these preoperative assessment tools.

1.2. Wearable Sensors

The development of technologies in wearable sensors (WS) in the last decade has led to a significant increase in consumer uptake [41]. As a result, there is a vast quantity of data relating to individuals' health whilst in free-living environments. There are also many examples of WS being implemented in a clinical setting as a cost-effective tool to measure physiological signals [42]. There is evidence to suggest that these devices could hold a vast volume of data that can give clinicians a quantitative representation of patients' health in their day-to-day environment [43].

WS have multiple attributes that make them a suitable tool for preoperative assessment. Physiological signals have inherent biological variation and therefore, recording these signals over a longer time period may allow detection of abnormalities that present at irregular time periods [44,45]. A further advantage of collecting data over a longer time period is that the data may be more representative of normal routines. When collecting physical activity data, an increase in the number of days recorded is associated with a more reliable weekly estimate [46,47]. WS are usually autonomous devices that can record signals away from the clinic. This can provide a simpler alternative for clinicians who, under time constraints, cannot always complete preoperative physical assessments [27,48]. In some cases, measurements recorded away from the clinical environment may be more accurate; the 'White-coat' effect describes the increases in physiological measurements that are only seen when taking measurements in the clinical environment [49]. These attributes have led to multiple successful implementations of WS for diagnostics.

WS have high efficacy for continuous monitoring of numerous physiological variables [50]. Subsequently, this has been shown to be applicable to support the diagnosis of several diseases including Parkinson's, kidney failure and viral infections [51–53]. Particularly in the case of cardiovascular disease monitoring, WS can provide live monitoring capabilities of patients' medical status that can be used to alert clinicians [54,55]. In the postoperative period, wearable technology has shown consistent uptake and to be a particularly useful tool for monitoring recovery [56]. WS have been used to identify post-surgical cancer patients who are recovering slower than their predicted profile, this facilitates the determination of appropriate discharge dates and preventing re-admissions [57]. A wide range of wearable sensing technology has been shown to be useful for clinicians in the postoperative setting including chest patches and wrist-based fitness sensors [58–61]. In the preoperative period, initial research has reported similar successful applications of WS.

Some research reports utilising WS as a method to measure adherence to prehabilitation programmes rather than as a preoperative risk assessment tool [62,63]. In other cases, WS have been utilised specifically for preoperative assessment with an exclusive focus on accelerometer data [64,65]. More recently, there have been instances where research has combined Heart Rate (HR) data with accelerometer data to approximate outcomes of cardiovascular fitness testing [40,66]. This research suggests the potential for advanced computing methods to analyse a combination of HR and movement data. These papers highlight the variation in sensor modalities and analysis methods that exist across the field. This suggests that the field would benefit from a review of these factors.

1.3. Aims of Review

Wearable sensors are a widely researched tool across medicine and exercise science. They have been shown to provide accurate and representative measurements that can give an objective insight into an individual's health. In the postoperative period, multiple reviews have been conducted to provide an overview of WS both in the hospital and in the outpatient setting [56,67]. However, there is no review of research utilising WS in the preoperative period. This paper aims to investigate how WS have been used

in free-living environments in research to predict either preoperative measurements or postoperative outcomes.

Research investigating emergency surgery will not be included as preoperative evaluation for emergency surgery does not allow analysis of free-living data. Additionally, the review will focus on major abdominal surgery; research completed in a cardiac or orthopedic surgical setting will be excluded as these procedures are associated with unique complications that are not applicable to a wider range of medical contexts [68]. Given that preoperative risk assessments are a vital component in the perioperative pathway, this paper will include research that has attempted to predict clinical variables that are routinely collected at this stage. Although this is not a systematic review, search terms were employed to identify research from selected databases; a narrative synthesis was used to summarise findings. The findings of this review are presented across four sections: Sensor Modalities, Pre-processing, Feature Extraction and Predictive Models.

1.4. Literature Search

To identify key literature and ensure valuable research was not missed, search terms were identified and used to search major databases. MEDLINE (Ovid) and Web of Science were searched as well as the first 200 returns in Google Scholar [69]. Relevant papers from ARXiv were also included. Search terms can be seen in Table 1. Further articles were identified through backward chaining.

Table 1. Combination of search terms used for the review of the literature. These searches were combined with Boolean operators and entered into the databases MEDLINE and Web of Science in equal format. Initial investigation of search terms was completed to find the combination of terms that returned optimal results. A narrative review of results is completed in this paper.

Surgery	Preoperative Assessment	Wearable Sensor
major surgery	preoperative	wearable technology
general surgery	pre-surg *	wearable activity monitor
abdominal surgery	preoperative evaluation	heart rate monitor
elective surgery		accelerometer
		fitness tracker
		wearable fitness *

* The asterisk is used for truncation in the search, the asterisk added to end of a term allows the databases to search for all forms of the word to broaden the search.

Searches brought back a broad range of papers from which the most relevant were selected. Several inclusion criteria were outlined for the search. Research should analyse free-living data collected from WS in research. Data should be used to investigate the association of these signals with outcomes related to either clinical variables routinely collected preoperatively or postoperative outcomes. Clinical variables collected preoperatively varied from CPET outcomes to cardiovascular responses. A subsection of these papers was selected for in-depth analysis whilst multiple papers outside of this subset are referenced throughout this review. A compilation of the key features extracted from these papers selected for in-depth analysis can be found in Supplementary Table S2. A summary of the sample sizes and participant demographics can be seen in Supplementary Table S1. Sample sizes varied greatly across studies and had a range of 16 to 80,137. For research that had a sample size of under 1000, the mean sample size was 48.9 indicating that the majority of research in this field has a relatively small number of participants. This is likely due to having to provide hardware to each participant included in the research study rather than having access to pre-compiled data sets. There was also a broad range in the sex split across research; there was a slightly higher prevalence of male participants with an average of 57% males but this was not significantly imbalanced. Ethnicity was rarely reported with under 30% reporting the ethnic breakdown of participants. When ethnicity was reported there were significant imbalances with largely white participants.

For research collecting data from patients waiting to undergo surgery, participants were commonly approached immediately after being enlisted for surgery or at the preoperative evaluation clinic [70–75]. It was also common for patients to wear the device right up until the date of operation [70,71,76–79]. For all research apart from two papers, the data collection from the WS device took place within 33 days prior to surgical treatment. In two cases, it was unclear how far in advance of treatment patients wore their WS [64,80].

2. Hardware/Sensing Technologies

This review identified a range of different devices that have been utilised in the research in the preoperative setting. This paper does not give an analysis of each commercial device but presents an overview of the different sensor modalities that they employ to collect data. The accelerometer, ECG and PPG sensors were widely implemented; Figure 3a shows the breakdown of these sensor modalities across research. A full breakdown of the WS devices and models along with their sensor modalities can be found in Supplementary Table S2. The following sections outline the functionality of these sensors and their outputs.

2.1. Accelerometry

Raw accelerometer data is recorded as a signal of acceleration measured across axes; this can be analysed to detect regular patterns that represent movement [81]. This is most commonly measured across three axes in a tri-axis accelerometer: X, Y and Z (see Figure 4a). An accelerometer will measure linear acceleration across these three axes and by combining this with time, the signal can be used to quantify human movement and physical activity [81]. Specifically, movements including steps can be identified to track a step count whilst the general movement of the sensor can also be categorised into intensities of movement. The association between features extracted from an accelerometer with several health metrics has consistently been evidenced in research. Step count is one of the most commonly extracted features from accelerometer data (see Section 4.1.1) and has been linked to various health outcomes. A systematic review of over 13,000 adults identified that an increase in step count of 1000 daily steps above baseline was associated with a lower risk of all-cause mortality and cardiovascular disease (CVD) [82]. Similarly, there is a strong relationship between any recorded physical activity and a reduced risk of all-cause mortality [83]. Further, accumulated time spent in moderate to vigorous physical activity (MVPA) is linked to significant reductions in mortality risk [84]. This suggests that the collection of these variables prior to surgery may have a strong association with a patient's health and could be useful for predicting complications.

The majority of research in the preoperative period utilises WS from commercial providers that have their own internal proprietary systems. As a result, researchers rarely analyse raw accelerometer data in the preoperative period. Multiple projects utilised a Fitbit wearable from which they could extract step count and energy expenditure [65,70,71,74,85]. Similarly, multiple studies utilised Garmin-derived metrics including step count, distance travelled and energy expenditure [40,76,86]. Data generated from a combination of mobile devices and WS that had a common platform (Apple Health; Achievement) extracted similar variables [87,88]. Particularly amongst large-scale research with many participants, these pre-calculated features were utilised for analysis rather than raw data [88,89]. Some research reported utilising pedometers as their WS to collect patient data; these pedometers all included accelerometers rather than a traditional step counter allowing them to represent the intensity of movement [90].

Although uncommon, some researchers did utilise the raw accelerometer data from the devices. Three studies utilised raw accelerometer data to stratify the signals into different intensities of activity [64,72,91]. This allows researchers to quantify the intensity of movement and compare time spent in sedentary behaviour against higher intensities including MVPA. Figure 3b presents a breakdown of the location at which sensors were worn. The majority of all sensor devices, and therefore accelerometer sensors, were worn on the wrist in commercial devices (Fitbit/Garmin) [71,76,80,86]. However, many of the

research papers that only included an accelerometer sensor were worn on the participant’s hip, see Figure 3b [64,77–79].

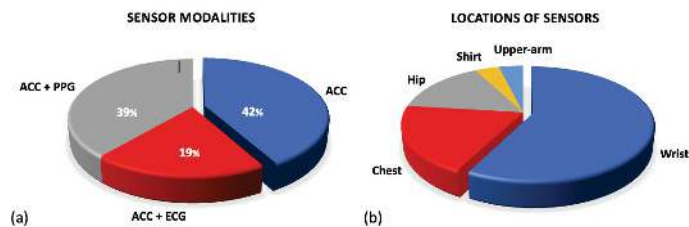


Figure 3. Figures for sensor modalities. (a) Shows the percentages of sensor modalities used across research. All research employed accelerometer sensors but a further subsection combine this with either ECG or PPG sensors. (b) Figure to show the variations in locations of sensor types. The common locations for sensors used in research applicable to the preoperative period are outlined.

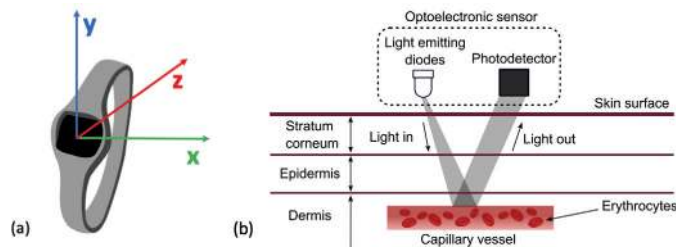


Figure 4. (a)Figure to show reference axes in a Tri–axes accelerometer. Presents the axes along which acceleration of movement can be measured across x, y and z. (b) The mechanism for HR detection in a PPG sensor by reflection. The LED can be seen emitting light which is reflected and then detected by the photo-detector and converted into a HR signal. This figure was taken from Moraes et al. (2018) with no changes made, Creative Commons Attribution International 4.0 License [92,93].

2.2. Photoplethysmography

Photoplethysmography (PPG), sometimes referenced as optical heart rate monitoring, is a sensor that can estimate an individual’s heart rate (HR). It utilises an optical emitter to give out light emitting diodes (LED) onto the skin that is attenuated from the pulse in the artery [94]. The reflection of the LED is captured by a photo-diode. The digital signal processor located in the device then translates this into heart rate data, as seen in Figure 4b. HR is one of the most commonly measured vital signs across medicine. It holds significant prognostic value for predicting general health and mortality. A lower HR is generally associated with a lower risk of cardiovascular mortality as well as a lower risk of all-cause mortality [95,96]. Research has also identified that a lower preoperative HR is associated with a lower risk of postoperative myocardial injury in patients undergoing non-cardiac surgery [97]. Therefore, cardiac assessments are common practice prior to non-cardiac surgery [98].

Previous research has investigated whether PPG produces comparable outputs to ECG. PPG sensors require both the LED and photo-detector to have contact with the surface of the skin and as a result, can be heavily impacted by movement or distance between the LED and skin resulting in optical noise [99]. A comparison of PPG to ECG signals can be seen in Figure 5a [100]. A large-scale study evidenced that HR estimates collected from PPG sensors correlate strongly with those from ECG signals [101]. Other research has concluded that at rest and at low HR levels, PPG has shown to have high accuracy but that this decreased with intensity of activity [102]. One publication reported the threshold for a reduction in accuracy of HR estimation to occur between 155–160 beats

per minute [94]. This high threshold is notably above estimated HR maximum values for elderly populations indicating potential suitability for this population and for monitoring low-intensity exercise [103]. Most WS that include accelerometry at the wrist also include a PPG sensor [40,65,70,71,74,75,85,88,104]. One popular device that included a PPG sensor reported a data storage limit of 7 days and a charge limit of 10 days within the device [73]. PPG is a widely utilised technology in the preoperative period due to its convenience as a tool to measure HR. Wrist-based wearables have few requirements and are a practical tool for researchers as they require little input from users and have long periods of storage.

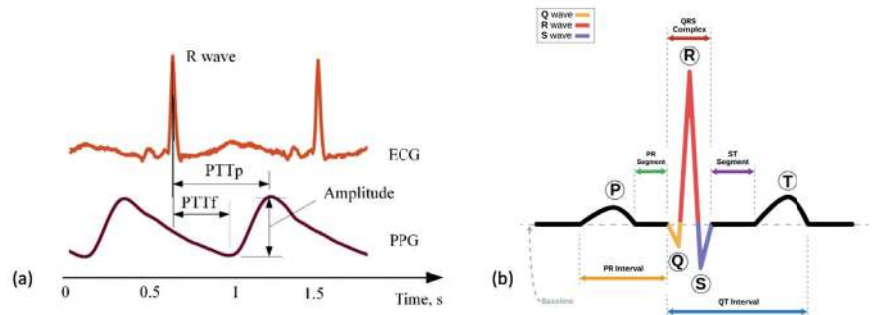


Figure 5. Figures to show the recordings produced from ECG and PPG recordings. (a) A comparison of the cardiac signals produced from a PPG versus ECG sensor over a period of 2 s. This figure was produced by Elgendi et al. (2019) and was taken from a larger figure with no changes made as part of the Creative Commons Attribution International 4.0 License [92,100]. (b) A segment of an ECG graph that has been portioned to show the stages in a normal cardiac cycle including the P wave, the QRS complex and the T-wave.

2.3. Electrocardiography

An ECG directly measures the electrical activity of the heart through electrodes placed upon the skin [105]. The electrodes measure electrical impulses from the heart that are then converted into an ECG graph. An ECG graph segment with annotations can be seen in Figure 5b. These annotations represent the detection of different stages in the cardiac cycle: the p-wave representing depolarization of the atria, the QRS complex representing the electrical impulse spreading to ventricular depolarization and the t-wave representing the re-polarization of the ventricles following contraction. The location and morphology of these annotations are used to extract several features including heart rate (HR) and heart rate variability (HRV) [106]. Using these features amongst others, an abnormal cardiac cycle can be identified from a signal and categorised. Heart disease is a leading cause of death worldwide making detection of cardiac abnormalities from the PQRST complex a key tool for preliminary diagnoses [107]. In clinical practice, a 12-lead ECG is common practice but this would be impractical for WS in a free-living environment. In this review, reduced lead ECG devices are considered wearable devices [108]. By analysing the ECG graph, HR can be calculated at any given time point. Additionally, using the RR interval (time portion between each R peak) other variables including heart rate variability (HRV) can be calculated.

An ECG sensor has several advantages as a prognostic tool compared to PPG sensors. The ECG is considered the ‘gold standard’ tool for measuring HR; the accurate identification of a heartbeat allows the calculation of HRV from which further inferences about health can be made [109]. HRV has been highlighted as a tool that has promising value for predicting complications during and after surgery [110]. Additionally, an ECG allows for further detection of potential abnormalities including atrial fibrillation [111]; preoperative atrial fibrillation has shown to be predictive of complications in patients undergoing non-cardiac surgery [112,113].

One large cohort study used the Actiheart wearable ECG, which places two leads on the sternum from which three papers analysed the HR data [66,89,114]. Other research utilised an ECG ‘necklace’, which involved placing electrodes in the II lead configuration on the chest (see Figure 6c) whilst a further project included an ECG sensor that was integrated into a smart shirt (Hexoskin), see Figure 6b [115,116]. Although not utilising all 12-leads, wearable ECG devices have been shown to have high accuracy for heartbeat detection [105].

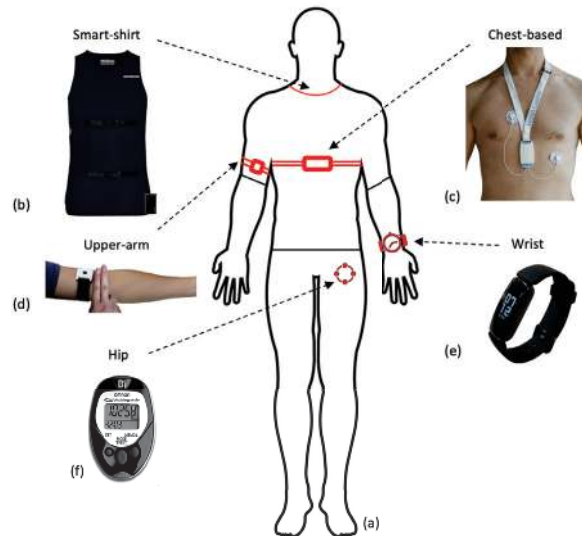


Figure 6. Figure to show wearable sensor devices used in research across the body (a) and where these are located (b–f). (b) The Hexoskin smart shirt that collects both ECG and activity data, used with permission from Hexoskin [117]. (c) An ECG wearable device that collects recordings from a single-lead ECG device and 3D-accelerometer data, used with permission from [118]. (d) An upper-arm PPG sensor utilising reflective PPG detection, similar to that used in preoperative monitoring research [119]. The figure is taken as part of a larger figure from Wang et al. (2023), Creative Commons Attribution International 4.0 License [92,120]. (e) The Fitbit Inspire collects a combination of accelerometer and PPG data from the user and is commonly used in preoperative research. The figure is taken from Li et al. (2023), Creative Commons Attribution International 4.0 License [92,121]. (f) The OMROM walking style pedometer that utilises a tri-axis accelerometer to collect step data, used in predicting VO₂ max [78]. This figure is taken from Bartlett et al. (2017) as part of a larger figure, Creative Commons Attribution International 4.0 License [92,122].

3. Pre-Processing of Signals

Pre-processing involves all changes to data that are made in order to prepare the data for analysis. Pre-processing can be the most vital stage in data processing and has a large impact on the inferences that can be made from a data set. Wearable data, even when collected in a controlled clinical environment, often requires heavy pre-processing due to the nature of the data. A wide range of pre-processing methods were implemented across the key papers and the most important techniques for each are outlined in Supplementary Table S2. There are two main challenges in WS data that pre-processing aims to overcome: missing data and noise.

3.1. Missing Data

Missing data is a frequently reported problem across research involving WS, particularly when using data from free-living environments [123]. Poor electrode placement, poor contact with skin or removal of device might lead to significant portions of poor quality or

missing data. Often, the underlying reasons for periods of missing data are unknown. The prevalence of missing data in WS used in the preoperative period is outlined.

There is a wide variety of missing data reported across studies using wearable sensors in the preoperative period. Missing data was frequently reported at ranges of up to 25% from WS in this context [70,71,119]. One study reported that across all accumulative days of collected data, only 0.25% of days had complete HR data [71]. For the majority of research, the reporting of missing data refers to HR rate, rather than movement data (see Section 3.1.1). The reporting of missing data differed between research; some researchers report the overall percentages of data that were missing whilst others report the number of participants excluded due to missing data [40]. In both of these cases, research rarely goes into detail as to the causes of missing periods and how to categorise these. Missing data can generally be classified into three separate categories: missing completely at random (MCAR) where no systematic relationship is present between values that are missing and existing values; missing at random (MAR) where missing data is systematically related to existing data that has been observed but not unobserved data and missing not at random (MNAR) where missing data is systematically related to unobserved data [124]. Depending on the category of missing data that is assumed for the data, different methods may be better suited for minimising the potential bias that may be introduced [125].

Across research applicable to the preoperative period, three different strategies for handling missing periods of data were employed. As seen in Figure 7, missing periods of data were either deleted entirely, tolerated or imputed. These techniques have been previously identified in research using WS and are common solutions for missing data across fields [124,126]. To build on this, several techniques were identified in the present review that involve overlap between categories. To differentiate between data that has missing portions but is still usable versus data that should be deleted, an extraction threshold can be identified. Further, some research employs imputation on only short-term segments of data (see Section 3.1.3).

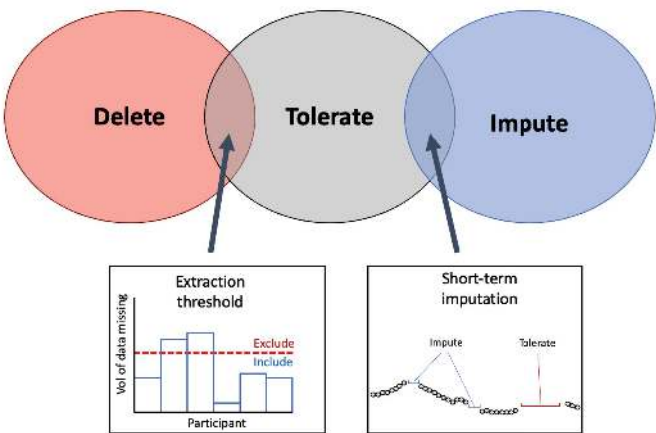


Figure 7. Venn-diagram to present the common methods for handling missing data from WS. The three techniques identified in the literature for handling the missing periods of data are presented in the Venn diagram. At the intersection between ‘delete’ and ‘tolerate’ the implementation of an extraction threshold was identified to delete data below the threshold and tolerate missing data above the threshold. At the intersection between ‘tolerate’ and ‘impute’, imputation on short-term segments of missing periods was identified as a solution that employs that imputation on select segments.

3.1.1. Extraction Threshold

The extraction threshold identifies a point at which a subject’s data will be included in final analyses or is abandoned/processed further. This extraction threshold is usually only applied to wearable devices that measure HR in some format. For research that only utilised

the pre-extracted step count, it is not possible to assess the exact volume of missing data. Step count data can appear as null values and still represent viable recordings indicating sedentary periods and so it is not always obvious to know whether this is as a result of non-wear, device malfunction or sedentary behaviour [127]. This is particularly true when step-count is only reported at the daily level [76,79,80].

These extraction thresholds differ widely between research. One study set a daily yield extraction threshold at a minimum of 8 h of collected data for that day to be included in analysis [71]. Other research set their extraction threshold at 10%, defining that any day with data of a daily yield above 10% would be included for analysis [70]. These studies indicate that the daily extraction threshold can be set at a relatively low value to allow for a high level of missingness in data and prevent this data from being abandoned. Research that used large data sets could set their extraction thresholds at a higher level; one study with over 80,000 participants only selected participants that had a minimum daily yield of 20 h [66]. However, this study did not report what percentage of participants had to be excluded as a result. Large data sets may have more flexibility in their extraction thresholds whilst a small research study may have to accept a higher level of missingness in order to prevent excluding a large portion of their sample.

A total yield extraction threshold can also be applied to the number of days in the monitoring period that have data [119]. This can be employed by only including participants that have above x number of days of data, with x indicating the threshold. The employment of a daily extraction threshold (i.e., 8 h of data) versus an extraction threshold for total data yield (i.e., 3 days of data needed) will depend on whether the data are subdivided into daily segments or kept as a total per participant.

3.1.2. Selecting an Extraction Threshold

Extraction thresholds should not be randomly selected. To investigate the influence of the extraction threshold on the predictive performance of analysis, one study varied their extraction threshold from 1 to 10 h and identified that between 8 and 10 h achieves the best performance [71]. This highlights the importance of identifying an optimal extraction threshold. Setting a high threshold for inclusion will result in less data available for analysis; a low threshold has the potential to allow data from days with large missing periods into analysis. If this is the case then the underlying reasons for missing periods should be assessed to prevent bias in the data (see Section 3.1).

Aside from reported extraction thresholds within their own data sets, very little research has focused on quantifying the volume of data that is needed to obtain reliable preoperative baseline measurements. It has been reported that when using a PPG sensor combined with an accelerometer, a minimum of three days of monitoring should be completed; however, an extraction threshold within each day was not specified and whether the location and type of sensors have an impact on this threshold was not discussed [128]. An appropriate extraction threshold is a useful tool for selecting data for analysis but does not provide a solution to missing periods. To overcome the missing periods of data that remain, several imputation methods can be employed.

3.1.3. Imputation

Data imputation in WS data is a complex process. Individuals will often have varying levels of missingness between them due to compliance with wearing the device. Further, there may be missing portions caused by technical issues in a device. Therefore, imputation techniques in WS data should generalise to both the participants' behaviour and the device's patterns [129]. Often in research collecting data in the preoperative period, data that was identified as missing was abandoned. Little effort is made to impute the missing values and why they are missing; this could reduce the size of the data sample and in some cases may introduce bias [130]. The imputation techniques that were employed in the research are outlined in the sections below.

The simplest method to replace missing data points in HR signals is with the mean HR values of activities at waking periods [88]. However, if the mechanism for missing portions is known to occur during sedentary periods or periods of vigorous activity then this may lead to under or over-estimations of daily HR. One particular study substituted missing HR values for HR recorded during a hospital visit [73]. This technique was not common across research, likely due to inconsistencies that may be present between free-living data and data collected in-clinic (see Section 1.2). For the studies that utilised the temporal aspects of data to employ imputation, they both did this using a two-layered pipeline [70,71].

The k-nearest neighbours (KNN) technique was shown to be a common method to impute missing HR values [70,71]. KNN has previously been implemented as a technique to address missing data in a range of applications [131]. The KNN algorithm is implemented as a 'sliding window' that allows missing HR data to be calculated from a combination of recent step count and HR data [70]. This method is rationalised by explaining that imputation is useful for short-term missing segments where previous values of step count and HR have a high correlation with future values. One technique employed a k-nearest neighbours (KNN) algorithm for an entire day of data if the daily yield for the relevant day was above 10% [70]. A different study utilised the same technique but for all portions of missing data that were shorter than 10 min long, regardless of daily yield. If the segments of missing data were less than 10 min, a KNN sliding window (length of 5) utilised recent HR and step counts to predict HR values [71], see Figure 8. Using this method, the feature vector is imputed to the KNN algorithm where hr_t and $step_t$ represent HR and step data at time t .

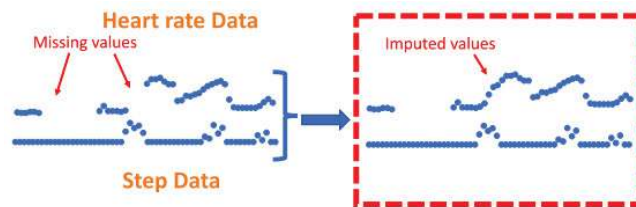


Figure 8. Figure to show imputation using K-nearest-neighbours. Zhang et al. (2023) utilise the KNN technique to impute on short-term segments of missing data under 10 min in length by utilising previous values from both the step count and heart rate signals to calculate missing values. This figure was produced by Zhang et al. (2023) and was taken from a larger figure but had no changes made, taken as part of the Creative Commons Attribution International 4.0 License [71,92].

To assess the performance of various imputation techniques, several metrics can be employed. Root Mean Square Error (RMSE) is a commonly used metric for assessing imputation in signals from WS [129,132]. RMSE measures the average magnitude of the errors of the imputed values. It is particularly useful for assessing imputation because it penalises larger errors more heavily making it especially sensitive to values that are big outliers from the predicted values. Mean Absolute Error (MAE) provides an alternative to the RMSE in that it is not sensitive to outliers. It presents the absolute differences between the imputed and original values and is focused on the overall accuracy of the algorithm. It may be useful to compare the outcomes of the MAE with the RMSE and the two have previously been used in combination [129]. For the MAE and RMSE, lower values indicate better performance. However, in order to assess imputation techniques using these metrics, ground truth values are needed. One method to do this is to add missing periods of data into a signal and perform imputation on these fabricated missing periods to compare outputs against original values as ground truths. It should be noted that this has been very rarely implemented in research across this field.

3.1.4. Feature Level Imputation

In instances where data has been abandoned due to significant periods of missing data, this can be imputed by employing feature-level imputation techniques. After abandoning days with a daily yield below the extraction threshold, a feature-level imputation technique can be employed to compute the features that represent the days with high portions of missing periods. In one case, researchers again utilise the KNN method to impute statistical and semantic features based on the neighbouring features that are available for that participant, rather than attempting to impute the missing values in the signal [70]. A further technique to deal with missing data that fell below the daily yield of 8 h was to employ imputation on high-level features that have been calculated from daily features [71]. Further application of Detrended Fluctuation Analysis was used to reduce the incomplete data.

It is important to note that research only employed imputation for HR data, this was not performed for other features extracted from accelerometer data. For research where HR signals are collected alongside step count data, the proportion of missing step count data can be extrapolated from the time periods with missing data points of the HR signal. Step count data is reported as being significantly less correlated and so less predictable than HR data [71]. Instead, these data were normalised by dividing the step count by the daily yield so as to prevent the step count from being drastically increased just for those patients with more data accumulated.

3.2. Noise

Aside from missing data in the signal, noise can also prevent meaningful features from being extracted. Accelerometry data can be plagued by white noise, altered by human motion or vibration whilst both ECG and PPG signals can be corrupted by motion artefact, baseline wander and electromyography (EMG) noise [133,134]. Few papers utilised raw signal data (see Section 2) meaning filtering of signals was not commonly reported. When techniques were employed to filter noise, this was performed on the accelerometer and HR data separately.

In the present review, many WS devices only report pre-calculated HR values from internal algorithms meaning raw cardiac signals were rarely processed. In certain cases, some removal of noise from signals was completed. Previously, Gaussian process robust regression has shown to be successful for noisy HR data and was implemented by one study to utilise prior knowledge of the HR data to reduce noise [66,135]. In comparison, a simpler method to limit the noise in HR data was to average the HR extracted from R-R intervals over a set time period, this varied from between 15 s to 15 min [114,115]. When employing this technique, all inaccurate HR values were identified and removed from the data where consecutive HR values varied by more than 20% [115]. Cardiac signals were very rarely passed through a low-pass filter; however, one study resampled the HR to 1 Hz before passing HR through a 0.01 Hz low-pass filter to remove high-frequencies affected by non-linearities introduced from circulatory distortions [116].

For accelerometry, to convert the raw signal data into magnitude of acceleration the Euclidean norm minus one and high-passed filtered vector magnitude were used [66]. Altini et al. (2016) employed a different filter technique where a low-pass filter (1Hz) was used to isolate the static component in the signal due to gravity and a band-pass filter (0.1 Hz, 10 Hz) was used to isolate dynamic components due to body noise [115]. As mentioned above, Beltrame et al. (2017) used a similar low-pass filter at 0.01 Hz for accelerometer as well as HR data [116]. The only implementation of a fast Fourier transformation (FFT) was to integrate the frequency in accelerometer data between 1 Hz and 10 Hz [91]. One particular paper reports a method for outlier detection within data by removing values that are greater than 3 standard deviations from the mean [104]. Building on the techniques commonly employed on HR data by averaging values over a short period, research employing a pedometer categorised each accelerometer period of 10 s into either lying, stationary or active periods [91].

For research using large-scale cohort data, after normalising their data through standard scaling with unit variance, researchers applied Principle Component Analysis (PCA) onto the original training data set that retained the components that explained 99.9% of variance [89]. To prevent any information leakage across the data sets, the fitted PCA scaler was applied individually to the test set. In a different project utilising the same large dataset, researchers attempted to reduce the noise that is seen in the labelling of HR data from the ECG wearable using deep learning [114]. The authors propose UDAMA (Unsupervised Domain Adaptation and Multi-discriminator Adversarial) training network [66]. However, these techniques utilising deep learning will require a large pool of data and may not be suitable for smaller single-centre studies. In order to reduce the noise that is present in daily features, one research study utilised singular spectrum analysis to further extract high-level features from daily features [71]. This allows trends to be extracted from the noisy data with missing portions by computing the mean, variance and slope from each time series of daily features.

3.3. Encoding Time

The temporal aspect of the HR and accelerometer data can provide useful insights but is infrequently used in analysis. To investigate the temporal aspect or reduce the bias that can be related to the time of recordings, it is possible to encode timestamps from sensors [66,89]. This is performed to help capture any periodic nature of certain behaviours, particularly those that may exhibit daily or monthly habits. By incorporating these cyclical components the model can better understand the cyclical nature of time. This was completed by encoding either the month of the year or hour of the day as (x, y) co-ordinates on a circle using Equations (1) and (2) where t represents the temporal aspect.

$$T_{f1} = \sin\left(\frac{2 * \pi * t}{\max(t)}\right) \quad (1)$$

$$T_{f2} = \cos\left(\frac{2 * \pi * t}{\max(t)}\right) \quad (2)$$

This is to ensure that when including time in analyses neighbouring timestamps are always consistent; December is only one month from January and 23:59 is only one minute from 00:00. This is particularly important for large data sets that are likely to be collecting data over an extended time period.

4. Feature Extraction

Feature extraction is an important step in signal processing to convert the signal data into numerical features that can be processed in a model [136]. It can also be useful to reduce the dimensions of the data when a large amount of data is collected [137]. Most research performed feature extraction from each signal separately but where possible features were extracted from multiple signals. The resulting features are outlined, and a full breakdown of the relevant features extracted from selected papers can be found in Supplementary Table S2.

4.1. Features Extracted from Accelerometer Signals

4.1.1. Step Count

Step count is a frequently reported feature that is used in WS research. Extracting step count from raw accelerometer signals involves implementing an algorithm to detect a pattern in the data, the choice of algorithm might depend on the computational complexity and resources that are available. Common algorithms include the peak detection algorithm, which identifies the local maxima and minima (see Figure 9) or simpler thresholding algorithms that identify optimal thresholds to detect steps [138,139]. Machine learning techniques also have shown good application but may require labelled data to train models, which are not often available [140,141].

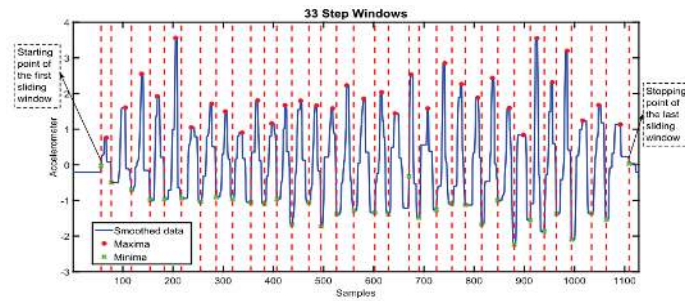


Figure 9. Graph to show the implementation of a maxima and minima step-counting algorithm that counts step number based on the number of steps windows detected. Each red line indicates the stopping point of each step window and the start of the next corresponding window, the length of time between each red vertical line indicates the step window size [142]. This figure was produced by Ho N et al. (2016) and was taken with no changes made, used as part of the Creative Commons Attribution International 4.0 License [92,142].

Much of the research that utilised commercial devices employed internal proprietary resulting in pre-calculated features. If the raw signal is automatically converted into step count then any further features relating to the intensity of movement are not able to be extracted. Frequently, steps are reported as an average measure across all days meaning any useful temporal aspect of this data will be lost [76,77,80,86].

Patients can be further classified into groups based on their average step counts to then produce features labelling individuals as active/inactive. Often research used pre-determined thresholds to classify participants into these groups. Threshold values range from 2500 to 5000 daily steps and there is no consensus on the threshold that should identify an individual as ‘active’ [73,77,80]. Another technique is to define these thresholds based on the variance of step count within the research cohort so that the split of participants in each group is even [79]. It is not conclusive which of these methods to stratify patients into activity groups is most conclusive but this should be relevant to the research cohort and context of the research [143].

4.1.2. Movement Intensity

Where raw signal data is available, the intensity of movement can be calculated from the acceleration and as a result, the time spent in sedentary, moderate to vigorous and vigorous activity can be reported [64,72]. This is calculated from time spent at activity counts above a specified threshold. Activity counts are calculated from ActiGraph’s proprietary algorithm; this algorithm has been widely used across research employing accelerometers and has been published as open access software [144]. However, the selection of a cutoff point for different intensities of physical activity results in significant differences in total MVPA between research and there is no standardised cutoff points [145]. Therefore, care should be taken when selecting a cutoff that is suitable for the research population.

4.1.3. Distance Covered

By utilising the features that can be extracted directly from the accelerometer signal, further features can be inferred. The distance that is covered by a participant can be calculated from a combination of the number of steps times that are taken in a day multiplied by the stride length of the participant [65,78]. It must be noted that stride length should be adjusted for using further participant information including sex and height information.

4.2. Features Extracted from Cardiac Signals

4.2.1. Heart Beat Detection

HR signals are complex and vary depending on sensor modality. As outlined previously, PPG and ECG signals are the most commonly collected cardiac signals that are used. To calculate HR from these signals, a process of data cleaning and beat detection is employed. The Pan-Tompkins algorithm (PT) is the most widely used beat detection algorithm which has been shown to be capable of detecting the location of a QRS complex in the signal across both clean and noisy data [146]. The PT algorithm employs a band-pass filter to isolate the relevant frequency before using a combination of thresholding and dynamic adjustment to identify the R-peaks. Other popular beat detection algorithms include wavelet-transform-based methods that analyse wavelet coefficients or simple algorithms that look for peaks in local maxima [147].

4.2.2. Heart Rate

By using the locations for the QRS complex that have been detected, see Figure 5b, HR at any given time can be calculated. As previously mentioned, the majority of research that is applicable to the preoperative period utilises commercial devices where HR values are often pre-calculated from the detected beats using an internal proprietary algorithm in the WS. As a result, it is rare that research in the preoperative period has to detect the location of a QRS complex. Instead, HR values are given at a varying frequency. The update period of the heart rate signal will dictate how regularly the heart rate is updated, commonly this is updated every beat. As a result, HR signals from WS are often extended signals that require further processing to extract meaningful features.

4.2.3. Resting HR

Resting HR (RHR) is a term that does not have a consistent definition but generally refers to the HR of an individual when they are inactive [148]. RHR is widely considered an important bio-marker of physical health and has been shown to be associated with both mortality and morbidity after non-cardiac surgery [97,149,150]. Although RHR is widely accepted as an important biomarker, there are also no set guidelines in medical literature for calculating RHR. Recent literature has suggested that when employing WS in research, a minimum four-minute rest time is required for a reliable RHR measurement [148]. From these values, a resting HR can be calculated. One research project reported calculating resting HR over a 24-h period but made a distinction between resting HR recorded during the night [73].

4.2.4. HR Changes

From the 24-h period, time spent in different HR zones can also be extracted as an indicator of activity throughout the daily period [73]. Other research utilised HR signals to create a new variable by assessing the difference between a current HR value and a previous value at a 1 s lag to represent 'dynamic changes' in HR and cardiac activity [116]. Other research also employed a two-level feature extraction pipeline where first-order statistical features like skewness and kurtosis of the HR are extracted and high-level features are then taken from the daily level features including the slope, mean and variance [71].

4.2.5. HR Variability

HR variability (HRV) is a metric that is calculated from the variations in intervals between detected heartbeats. It is an accepted tool for measuring the function of the autonomic nervous system, a vital factor in cardiovascular health [151]. Recent research has confirmed that preoperative HRV can be a useful predictor of postoperative outcomes [110]. However, in the preoperative setting, HRV has traditionally been calculated in a clinical setting using only ECG signals rather than employing WS [152].

Recent WS research has shown promising efficacy in calculating HRV from noisy signals in both PPG and ECG signals [153,154]. This highlights the potential for HRV to

be used as a preoperative tool calculated from WS. One study did calculate HRV from participants' ECG data by differencing the second-shortest and the second-longest inter-beat intervals [66]. There are variations in how HRV is calculated depending on whether they are time or frequency domain features. A common time domain frequency measure is the square root of the mean of the sum of the square of differences between NN intervals (RMSSD); both Garmin and Fitbit devices employ RMSSD to measure HRV in their commercial devices [155,156]. Although these devices may be able to report HRV, it is not always available to be extracted for use in research; one particular study reported that HRV, although measured by the sensor, was not able to be extracted from the device for use in research [73]. RMSSD is presented in Equation (3) where NN_i is the length of time of the i th interval and N is the total number of NN intervals in the dataset.

$$RMSSD = \sqrt{\frac{\sum_{i=1}^{N-1} (NN_i - NN_{i+1})^2}{N - 1}} \quad (3)$$

Frequency domain features are calculated using estimation of power spectral density and include a range of features such as low frequency (lf) and high frequency (hf) [157]. Further non-linear properties of HRV can be analysed and extracted using Poincare Plot and Sample Entropy; however, these are less frequently calculated.

4.3. Multi-Modal Sensor Feature Extraction

4.3.1. HR Recovery

Where possible, a combination of signals can be used to form features. HR recovery was calculated by measuring the decrease in HR one minute after exercise cessation which can be identified from movement data in accelerometer signals [65]. HR recovery has previously been shown to be predictive of cardiovascular health and it would be unknown whether the activity has completely stopped without the incorporation of accelerometer data [158].

4.3.2. Respiration Rate

Respiration rate (RR), also known as breathing rate, is a parameter that is commonly collected from WS but has been infrequently used in preoperative research. Across multiple clinical settings, RR has been identified as a valuable bio-marker of health status and has even been suggested as a predictor of in-hospital mortality [159]. On commercial devices, the most common implementation is to calculate breathing rate using the HRV data from PPG sensors [155,156]. Estimating respiration rates that are calculated from PPG sensor data is a well-established technique but has come under critique due to its sensitivity to noise [160]. There are a wide range of mathematical techniques that are proposed to estimate RR from PPG and ECG signals but the majority follow a general pathway. This involves extracting portions of the original signal that are dominated by respiratory modulation before fusing these signals to output one signal from which RR can be estimated [159]. A further method is to incorporate respiratory bands into a wearable device; Beltrame and colleagues (2017) [116] used RR calculated from respiratory bands to successfully predict oxygen uptake for participants.

When worn on the chest or torso, an accelerometer is another non-invasive method for estimating RR. RR is calculated by measuring movement and angular changes in the chest [161]. Some research has suggested that wrist-worn acceleration devices can also be used to accurately predict RR, but that this is only possible during periods of non-physical activity [162]. A further paper proposed a multi-sensor solution that employs sensors at several locations to predict respiratory rate [163]. There is no threshold for the acceptability of these devices when compared against gold-standard measurements (breaths counted by a specialist) meaning it is unclear which devices are best, although one study proposed a difference of under 2 breaths per minute to be acceptable [164]. Although this research suggests that there are multiple options for estimating RR from devices, research that has

application to the preoperative setting has very rarely employed this feature in its analyses, similar to HRV.

5. Data Analysis

Research extracting data from individuals’ behaviour and vital signs from WS use a large variety of computational techniques, this can be seen in Figure 10. The chosen models range from simple statistical analyses to complex models with high computational requirements. The variation in models may be explained by several factors including the outcome that is being predicted, the size of the data set and the nature of the features that are extracted from signals. The models employed are separated by their complexity in the following sections, a breakdown of the methods employed within key papers can be found under Supplementary Table S2:

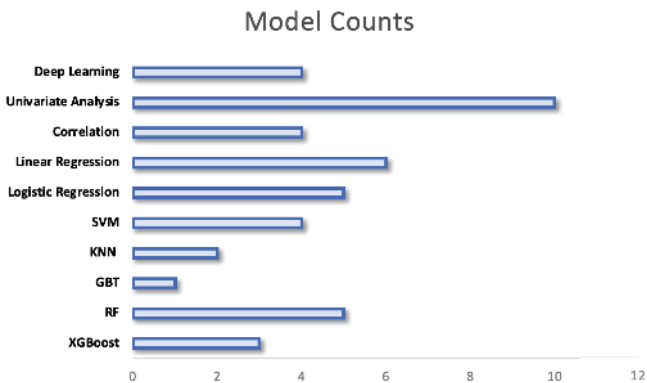


Figure 10. Figure to show the prevalence of each model of analysis across research. In research where multiple models are compared, all models are counted.

5.1. Feature Selection

Due to the vast range of features that can accumulate from free-living WS data, feature selection is the first stage in model development. This is important to assess the relevance and the effectiveness of the extracted features from the WS in relation to predicting outcomes. The most common feature selection method was to perform univariate analyses on features. Using this method, the relationship between each extracted feature and the target outcome is assessed. This evaluates each feature independently to determine its significance to the outcome variable. The features that are then found to have a strong relationship with the outcome are selected for further modelling. Pearson’s correlation was tested between all physical activity features and cardio-respiratory measurements, those variables with a *p*-value below 0.05 were excluded [87]. Within the remaining features, those with a covariance of over 0.7 were compared and the feature with the highest Pearson correlation would be included in the final multivariate model. A similar technique was employed but measured the correlation with wearable data and postoperative complications [72].

For studies that separated participants into groups based on thresholds in their WS data, chi-squared tests were used to test for differences in categorical variables and Fishers, Kruskal–Wallis or a Wilcoxon sum-ranked tests were used to test for differences in continuous variables [77,79]. Variables with a *p*-value that was below the set significance level for inclusion in final models were then included in multivariate analysis; the significance level for inclusion varied ($p < 0.05$ – $p < 0.1$). A similar technique was employed by instead grouping patients by their postoperative outcomes (re-admitted patients versus non-re-admitted patients) [85]. Where multivariate regression models were employed, all preoperative variables were included in the initial model. A backward-stepwise selection using the Akaike Information Criterion (AIC) was then used to assess each variable’s importance in the model [80,165]. Feature selection is an important stage in model development as it can

lead to enhanced model performance, reduced data dimensions or enhanced interpretation. This reduction in dimensions may also help prevent over-fitting of the model to the training data.

5.2. Univariate Analysis

The earliest study to investigate the relationships between preoperative physical activity (PA) levels and postoperative complications using a WS stratified patients based on their preoperative PA levels [64]. They performed independent *t*-tests within PA groups between those who did or did not develop a complication. More recent research employs similar univariate analysis between groups but utilises a Chi-squared test for categorical variables or a Wilcoxon sum-rank test for continuous variables [65]. This is the simplest method to assess if there are significant differences between groups and therefore was most commonly employed (see Figure 10). Although useful in a preoperative setting to stratify patients into high and low risk based on identified thresholds, the models themselves do not allow any predictions to be made regarding patients' risk of complication.

5.3. Correlation Analysis

Correlation analyses were a further common tool to investigate WS data. Greco et al. (2023) performed correlation analysis on preoperative daily step counts and the outputs of a 6MWT [65]. They showed that a strong correlation was present between the daily steps and clinical measurements taken preoperatively. Jones et al. (2021) also investigated the association between wearable variables and preoperative fitness measurements using correlation analysis but built on this using linear regression models [40]. The calculation of a correlation coefficient was used to investigate the relationship between preoperative step count and postoperative complications [76]. These models on their own do not allow predictions to be made in a clinical setting but are often useful as a preliminary tool to test for the presence of a relationship.

5.4. Machine Learning (ML)

Amongst papers investigating the link between preoperative WS data and outcomes, ML techniques are widely implemented. ML methods are well suited to non-linear, large-scale data. However, because the majority of research in the preoperative period has a relatively small number of participants and pre-processed data, not all ML models will be useful. Machine Learning methods are sub-categorised below:

5.4.1. Logistic Regression

Logistic regression was the most commonly implemented technique used in the clinical setting. Patient groups were either split by their preoperative activity levels or by their postoperative outcomes. Multiple research investigated the ability of preoperative PA to predict postoperative readmission as a binary outcome [79,85,86]. Internal validation techniques were reported by Rossi et al. (2021) using four-fold patient cross-validation and regularisation in the model to prevent the model learning noise and overfitting [86]. In comparison, another study employed leave-one-patient-out cross-validation (LOPO) to allow a change in the distribution of the data and prevent the model from overfitting to the training data [75]. This technique is a process where one patient is left out as the validation set and the validation is performed *k*-number of times, where *k* is the number of patients [166]. Other papers used a similar technique but where multiple outcomes were recorded, multiple models were used. To assess model performance in the case of multivariate logistic regression the C-index (Harrell's concordance index) was reported to be a useful measure [80].

5.4.2. Multivariate Regression

Rather than a logistic regression, multiple studies used multivariate regression techniques to predict continuous outcomes. Bille et al. (2020) used multivariate regression to

predict the number of absolute complications per patient [79]. Mylius et al. (2021) used this to predict the time to functional recovery as an overall measure of recovery alongside the odds ratio of the occurrence of complications [72]. Novoa et al (2011) used two separate linear regression models with bootstrap robust estimation (1000 iterations) of the standard error of regression coefficients. The first model was tested with mean daily distance walked as the independent variable whilst the second model incorporated distance travelled [78].

5.4.3. Ensemble Models

Ensemble models represent a development in Machine Learning where a combination of models are used to obtain better predictive performance [167]. Random Forest models were shown to significantly outperform the LASSO (Least Absolute Shrinkage and Selection Operator) models when predicting laboratory-based measurements [104,168]. Random Forest models use a combination of decision trees for solving a problem by aggregating them together to perform as a group rather than alone [169]. This may be related to the structure of the data and indicates that these models are superior when utilising non-linear data. When comparing models, there is no consensus on a particular ML model that is most suited to free-living wearable data and is likely dependent on a large variety of factors.

5.4.4. Support Vector Machines (SVM)

Ensemble models were not always shown to be superior to other individual ML methods. One study implemented a range of models: random forest (RF), k-nearest neighbours, XGBoost and SVM [71]. In this case, the SVM outperformed the remaining models by a larger margin for predicting postoperative complications. However, in a similar research case where SVM and RF models were compared against Gradient Boosted Trees (GBT), GBT models were found to have the highest performance [70]. In this case, the LOPO internal validation technique was implemented which may have influenced the performance of prediction models.

Other applications of the SVM technique were as a precursor to the final model. This was used to identify the portions in the extended HR data that would be most suitable for predictions. Data collected from simulated activities in the lab were used to recognise activities that are then completed in free-living environments [170]. SVM was used for pattern recognition in the free-living data. Acceleration was particularly useful for classifying high-intensity activities. SVM was used to find the optimal discriminatory boundary between activity clusters.

The range of ML models implemented across research displays how there are several options that may be suitable when predicting outcomes from WS data. However, there is no consensus on which ML techniques are best suited to this data type; particularly in research comparing the application of several ML models there have been confounding conclusions [70,71]. This is likely due to the range of factors that will influence the performance of these models, including the sections discussed previously.

5.5. Deep Learning

No deep learning (DL) techniques employing neural networks were utilised in the research investigating the associations between preoperative wearable sensor data and postoperative outcomes. However, there were multiple studies that utilised DL when predicting clinical fitness measurements. These were employed when analysing large cohort data sets [66,88,89].

5.5.1. Predicting HR Response

One paper employed an attentional convolutional neural network to identify and learn the signatures of different cardiovascular responses to data collected [88]. Specifically, a HR auto-encoder was trained to produce the given HR from physical activity and sleep stages. To learn the personalised cardiovascular response functions from the wearable data, the HR encoder was trained on physical activity and sleep stages. The encoder stage learns

the signature of an individual from their HR responses to exercising or movement and the decoder employs this learned signature to predict HR based on movement.

Further research uses DL techniques in a similar format to the aforementioned research but instead proposes a general-purpose model that does not require a historical input of one month [66]. This paper proposes the 'Step2Heart' receiving high-dimensional activity inputs to predict HR response which similarly uses the accelerometer data to predict HR. Stacked CNN and RNN layers are combined where the CNN learns spacial features and the RNN learns temporal features of the data. Aside from predicting a future HR response, this model has also shown to have further clinical value in the preoperative period in that it can be utilised to predict VO₂ max values.

5.5.2. Predicting VO₂ Max

Using a large WS data set, a deep neural network was able to successfully predict VO₂ Max measurements from cardio-respiratory testing. The network employed two feed-forward layers with 128 units that are densely connected [89]. Batch normalisation and dropout were included to help prevent over-fitting. The final layer in the model is a single unit layer; the model was trained using a pre-trained optimiser (Adam optimiser) to minimise the mean squared error [89]. The model was able to outperform other models and was also able to predict future changes to VO₂ max recordings.

Similar findings were found when deep learning was used to both clean noisy data and to predict cardio-respiratory fitness as a real value [114]. Due to the large nature of the data when WS are worn for an extended period, features were extracted from the raw signals and represented as feature vectors. The deep neural network trained on these inputs was able to achieve high performance when compared to the ground truth values from cardio-respiratory testing [114].

6. Future Challenges and Opportunities

This literature review has identified research utilising WS that is applicable to the preoperative period. Several findings were made that relate to the four subsections of the review. These are outlined below and described in the context of future challenges for research.

6.1. Comparison of Sensor Modalities

As is apparent from Figure 4a, the use of accelerometers is widespread across WS in research as a health assessment tool; a smaller subset of research employed an additional sensor to record cardiac signals (ECG or PPG sensors). Although less data is recorded from participants, there are several benefits associated with employing only an accelerometer in the wearable sensor. An accelerometer device will rarely require input from the user and advantages include simplicity and decreased power consumption. The RT3 accelerometer used by Feeney et al., (2011) collected data for up to 21 days without input whereas the ECG necklace device used by Altini et al., (2016) required daily charging [64,115,171]. Reducing periods of non-wear for charging throughout the study may reduce the likelihood of missing data. Further, accelerometer devices are not reliant on skin contact like PPG and ECG sensors which might further decrease periods of missing data (see Figure 4b). In applications where cardiac data may not be necessary, accelerometer sensors are simple, unobtrusive devices that collect valuable recordings regarding participants' activity levels.

However, when accelerometer signals are combined with signals from an ECG or PPG it allows for more comprehensive health monitoring. Cardiac data from PPG/ECG signals has consistently shown to have strong associations with CVD and poor clinical outcomes; this allows for interpretation of cardiac health that is not possible from accelerometer devices alone [172]. The incorporation of cardiac signals also allows for the detection of specific health conditions including arrhythmia's [172]. Further, multi-modal signals can improve the evaluation of free-living data. In research directly comparing the application of sensor modalities, it was found that activity recognition had superior accuracy when

combining an ECG recording with accelerometer data [173]. Other benefits that might arise when combining the two signals include an improved assessment of signal quality. Combining a PPG with accelerometer signals has previously been shown to allow for signal quality analysis that is not possible with each signal independently [174]. By including cardiac signals alongside movement data, a combination of sensor modalities can allow for a more holistic assessment of patient health. This includes the extraction of further features that require multi-model signals (see Section 4.3) and an in-depth analysis of an individual's cardiac health whilst also optimising the classification of activities. Although this adds value, it does come with the expense of added complexity both when collecting the data itself through the limitations associated with multi-modal sensors and the analysis of added complex signals. The attributes of these devices should be compared and sensor modality should be selected based on the context of the research.

In the research included in this review, patients were commonly recruited once surgical treatment had been scheduled or at the preoperative assessment clinic. As a result, any WS data from patients was usually collected in the immediate weeks preceding surgical treatment (see Section 1.4). There was no variation in the data collection period between sensor modalities; however, the preoperative period is by definition a broad label for any preoperative assessment (see Figure 1). The varying characteristics of sensor modalities outlined above indicate that simple devices with large data storage and battery capabilities may lend themselves to data collection periods extending over a longer period whilst more complex devices that incorporate multiple signals may be useful at shorter intervals to assess vital signs. Therefore, different sensor modalities may be suitable at different stages prior to surgical treatment but there is no existing literature discussing this.

A further area of interest within the field of preoperative wearable sensors is the development of new sensor modalities that can be relevant to the field. Many of the issues relating to the causes of missing data may have the potential to be overcome through the use of new materials. For example, poor electrode positioning of an ECG lead may lead to missing periods due to excessive movement. A sub-class of flexible sensors known as Piezo-Resistive sensors have shown to have a linear response to bending and elongation and as a result, are highly sensitive [175]. This means they have shown to be capable of measuring precise human movements like breathing and exercising whilst also being able to measure a pulse [176]. The development of these materials as sensors that can collect vital signs from patients whilst allowing significant movement may provide a better and more comfortable alternative than current solutions.

6.2. Missing Data Periods

When collecting WS data, there are often significant periods of missing data. This was seen in the frequency of studies reporting missing periods in signals (Section 3.1) and several solutions to this are outlined. Pre-processing of data is a key component of the analysis timeline and missing data is not always well addressed. Across the studies, there was no set protocol for handling missing data. The underlying reasons for missing data periods are rarely thoroughly investigated and this highlights a significant problem in the literature. This highlights a challenge in the field that should be addressed; a protocol for assessing the reasons for missing data might help research better address the missing periods with more appropriate solutions. Within the studies that do address missing data, there is some mention of the underlying reasons. This includes non-wear time or poor electrode positioning. One research paper refers to the frequent need to replace electrodes and this highlights an issue that may lead to an increase in missing periods of signals [115]. In order to overcome this, future research should compare the missing periods that are associated with different sensor modalities and whether specific techniques can be employed to address this.

Extraction thresholds are often used to identify when there is a suitable volume of data collected either over the course of each day in the monitoring period or over the course of the entire recording. Although some research justifies the selection of an extraction

threshold through testing, this is not always the case and an arbitrary unit is selected. Some research reports abandoning data whilst others report using imputation techniques. Future research should focus on identifying an appropriate protocol for overcoming missing data in WS research as there are significant variations across the field.

6.3. Raw Signal Data

The majority of research that is applicable to the preoperative setting utilises commercially available sensors that employ their own internal proprietary algorithms, a common example being the Fitbit Inspire [65,71,73,85]. Researchers are provided with pre-extracted features; for example, heart rate and step count, where no pre-processing is required. This can improve access to research as data from these sensors is computationally simpler to work with; however, this limits researchers in the agency that they have for pre-processing and feature extraction. In research predicting VO₂ Max, step-count and floors-climbed were extracted by an internal algorithm from a wrist-worn sensor [40]. Other research has shown that acceleration-derived Metabolic equivalent of Task (METs) and raw acceleration alongside step count data has been shown to be predictive of cardio-respiratory fitness [89]. Further, a variety of pre-processing techniques can be applied to accelerometer signals (see Section 3.2) and when using proprietary algorithms, research may be limited in how it can filter signals. Proprietary algorithms may limit researchers in the features they can use from the raw accelerometer signals.

Similar conclusions are particularly pertinent for cardiac signals from ECG or PPG sensors (see Section 4.2.2). Often HR is the only extracted feature from these signals and although HR is a useful measure with strong associations for health (see Section 2.2), there is a vast amount of information that can be further extracted from cardiac signals. HRV, calculated from the location of the QRS complex (see Section 4.2.5), has been a popular metric of health but is not able to be extracted from a HR signal. Research has suggested that not only ECG but PPG signals are adequate for estimating specific HRV features that are relevant for assessing patient deterioration [177]. Further, recent research has identified the potential for these cardiac signals from wearable sensors to identify patients at high risk of suffering cardiac abnormalities [178]. This highlights a large gap in useful data from both ECG and PPG signals that is not being utilised within research applicable to the preoperative setting. Although proprietary algorithms can simplify data access to wearable sensor data by removing the barriers of pre-processing data, it may in turn limit research capabilities by reducing the dimensions of the data and preventing extraction of features. Future research should investigate the added prognostic value that these underutilised features from raw signal data can bring to the predictive performance.

6.4. Predictive Models

There is a wide range of models being employed to investigate differences in outcomes. This ranges from models that compare characteristics across groups (active/inactive) to predictive models that predict the risk of complication/readmission/CPET. Machine learning models that utilise pre-processing and feature extraction techniques have been shown to be successful at predicting both postoperative outcomes and clinical fitness measurements. Due to the nature of data when combining HR with accelerometry, there are many different potential features for extraction. However, most feature extraction techniques in this review do not utilise the temporal aspect of the signals. Investigating how temporal data can be incorporated into feature extraction (e.g., HR recovery, see Section 4.2.4) and how this impacts the predictive performance of models should be researched further. Further, ML models have unique advantages that allow non-linear relationships to be analysed. All of these factors should be considered when selecting a model and where appropriate, multiple models should be compared. When using data from the preoperative period, there is no single ML model that clearly outperforms others. This indicates that future research should consider multiple models for predicting outcomes.

Deep learning methods were shown to be successful at predicting clinical fitness measurement research using large sample sizes. Although this suggests that DL methodologies may have application to wearable sensor data in this setting, there are several points to consider. Research utilising preoperative data requires approaching prospective patients to wear a sensor prior to undergoing treatment rather than utilising pre-existing data sets. As a result, preoperative research often has a low average sample size; DL models in this setting with low sample sizes will be disposed to over-fit the data [179]. Additionally, proprietary algorithms that limit the number of features that are extracted from signals may exacerbate this (see Section 6.4). A combination of these factors suggests that DL models may struggle to generalise these patterns effectively outside of the training data. Other considerations when implementing DL are the high computational requirements and access to this; the hardware required to integrate DL into predictive models is not accessible to all research groups. Whilst this is the case, shallow machine learning models have been shown to be suitable for predicting postoperative outcomes. Future work should investigate how to incorporate DL to large volumes of raw data from WS. Huge data sources are available when utilising raw data and therefore, deep learning techniques including Generative Adversarial networks may be suited to cleaning and pre-processing of data.

6.5. Considerations for Implementation

For research involving the collection of free-living data, much consideration should be given to the ethical implications that come with using this. Firstly, wearable sensors can collect a vast amount of highly personal data over a long period of time, referred to as Patient-Generated Health Data (PGHD) [180]. This results in a large stored set of sensitive data; ensuring that this is stored in a secure environment and is used in a method that will not impact negatively on participants is vital. Previous research has shown that understanding privacy concerns and legislation in WS data can have a significant impact on privacy concerns amongst patients and that this in turn may impact the self-disclosure of participant data. Rarely did the research in this review discuss the privacy and storage considerations for research but this was likely not a main research aim across the studies. Several regulatory frameworks exist that provide guidance on how to safely utilise individual data, including the General Data Protection Regulation [181]. One particular recent publication outlines recommendations for addressing privacy concerns regarding WS data and suggests that amplifying user agency over health data is a key measure to address this [181]. Given that this may be a determining factor in participants contributing data to research, future studies should consider the understanding that patients have over their data and any concerns that they might have.

A further point that is not particularly discussed across research is clinicians' views regarding WS data. Although patient uptake is key to collecting data for research purposes, without clinician input and buy-in for the technology the benefits will be limited. Preliminary research suggests that clinicians value the benefits that WS can bring to the healthcare setting, particularly as a tool to provide feedback [182]. However, other qualitative research has identified an overarching concern regarding trust and lack of clarity over data control. Increased regulatory efforts have been discussed as a solution and this was highlighted as a priority for future work [182]. It is possible that this may become a component of future regulatory and clinical validation criteria. Across the research discussed in the review, there is limited reference to clinical validation of devices. Under current legislation, the FDA does not classify wearable sensors that 'maintain or encourage a healthy lifestyle' as a device that requires approval and as a result, many of the devices referenced throughout the review do not require FDA approval [183]. Given the increasing evidence of the benefits that these devices can provide to the perioperative pathway, the regulatory frameworks and clinical validation process for patient data from WS may change. This may also impact the research requirements that are required to ensure that the findings from research in this field are valid. For example, much research in the field did not report on participant ethnicity, and when reported it was largely white participants (see Section 1.4; Supplementary Table S1).

Unrepresentative study samples highlight a pitfall in the research field that should be addressed before findings can be generalised to the wider population.

Finally, if wearable sensors were to be implemented as a tool across the perioperative pathway, a cost-benefit analysis would be required. At the current time point, limited literature has evaluated the costs associated with WS in perioperative medicine, particularly at the preoperative stage. Given that much as the research in the review is regarded as feasibility research into the application of WS, it is likely difficult to estimate the savings in healthcare costs that are associated. Having said this, recent research has suggested that for patients with limited access to healthcare facilities following surgical treatment; for example, rural communities or those reluctant to return to facilities for repeat care, WS can provide an alternative that is cost-effective [184]. Other research has highlighted that WS can be a cost-effective tool to monitor patients across the perioperative pathway but mirrors the conclusions from this review that more research is required to conduct a cost-benefit analysis.

6.6. Conclusions

This review presents the recent literature relating to WS in the preoperative period. Several sensor modalities are discussed in detail alongside common pre-processing methods. The wide range of features that can be extracted are outlined alongside the models that most effectively utilise these features to analyse WS data. Several important points are highlighted in the final chapter alongside the gaps in the research field. Particularly, WS data is often not analysed in its raw form and this may limit the capabilities for pre-processing and feature extraction. From this review, several directions for future research are suggested including a strong focus on utilising raw signal data for analysis.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/s24020482/s1>, Table S1: This table provides a breakdown of the samples used in each of the key research papers included in this review for analysis. Within each column, the sample size is reported alongside participant demographics including: Average age, gender split, ethnicity and health outcomes.; Table S2 : This table provides a comprehensive breakdown of the most relevant papers included in this review. The columns present the data extracted from each paper including the Outcome variable, sensor modality, pre-processing methods, features extracted and the model of analysis.

Author Contributions: Conceptualization, A.S., Z.Z., D.J. and A.D.; investigation, A.S.; methodology, A.S. and Z.Z.; project administration, A.S.; supervision, Z.Z. and D.J.; visualization, A.S.; writing—original draft, A.S.; writing—review and editing, A.S., A.D., D.J. and Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by EPSRC grant number EP/S024336/1.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CPET	Cardiopulmonary Exercise Testing
HR	Heart Rate
WS	Wearable Sensor
RHR	Resting Heart Rate
HRV	Heart Rate Variability
PA	Physical Activity

MVPA	Moderate-Vigorous Physical Activity
LOPO	Leave-one-patient-out
ML	Machine Learning
DL	Deep Learning
MET	Metabolic Equivalent of Task
LASSO	Least Absolute Shrinkage and Selection Operator

References

- Liu, J.H.; Etzioni, D.A.; O'Connell, J.B.; Maggard, M.A.; Ko, C.Y. The Increasing Workload of General Surgery. *Arch. Surg.* **2004**, *139*, 423–428. [CrossRef] [PubMed]
- Shinji, S.; Yamada, T.; Matsuda, A.; Sonoda, H.; Ohta, R.; Iwai, T.; Takeda, K.; Yonaga, K.; Masuda, Y.; Yoshida, H. Recent Advances in the Treatment of Colorectal Cancer: A Review. *J. Nippon Med. Sch.* **2022**, *89*, 246–254. [CrossRef]
- Brunner, M.; Wu, Z.; Krautz, C.; Pilarsky, C.; Grützmann, R.; Weber, G.F. Current Clinical Strategies of Pancreatic Cancer Treatment and Open Molecular Questions. *Int. J. Mol. Sci.* **2019**, *20*, 4543. [CrossRef] [PubMed]
- Orditura, M.; Galizia, G.; Sforza, V.; Gambardella, V.; Fabozzi, A.; Laterza, M.M.; Andreozzi, F.; Ventriglia, J.; Savastano, B.; Mabilia, A.; et al. Treatment of gastric cancer. *World J. Gastroenterol. Wjg* **2014**, *20*, 1635. [CrossRef] [PubMed]
- World Cancer Research Fund. *Colorectal Cancer Statistics*; World Cancer Research Fund: London, UK, 2023.
- Cutsem, E.V.; Borràs, J.M.; Castells, A.; Ciardiello, F.; Ducreux, M.; Haq, A.; Schmoll, H.J.; Tabernero, J. Improving outcomes in colorectal cancer: Where do we go from here? *Eur. J. Cancer* **2013**, *49*, 2476–2485. [CrossRef]
- Cancer Research UK. *Bowel Cancer*; Cancer Research UK: London, UK, 2023.
- Morris, E.J.; Taylor, E.F.; Thomas, J.D.; Quirke, P.; Finan, P.J.; Coleman, M.P.; Rachet, B.; Forman, D. Thirty-day postoperative mortality after colorectal cancer surgery in England. *Gut* **2011**, *60*, 806–813. [CrossRef]
- Wells, C.I.; Varghese, C.; Boyle, L.J.; McGuinness, M.J.; Keane, C.; O'Grady, G.; Gurney, J.; Koea, J.; Harmston, C.; Bissett, I.P. "Failure to Rescue" following Colorectal Cancer Resection: Variation and Improvements in a National Study of Postoperative Mortality. *Ann. Surg.* **2023**, *278*, 87–95. [CrossRef]
- Ketelaers, S.H.; Orsini, R.G.; Burger, J.W.; Nieuwenhuijzen, G.A.; Rutten, H.J. Significant improvement in postoperative and 1-year mortality after colorectal cancer surgery in recent years. *Eur. J. Surg. Oncol. J. Eur. Soc. Surg. Oncol. Br. Assoc. Surg. Oncol.* **2019**, *45*, 2052–2058. [CrossRef]
- Silva, S.D.; Ma, C.; Proulx, M.C.; Crespín, M.; Kaplan, B.S.; Hubbard, J.; Prusinkiewicz, M.; Fong, A.; Panaccione, R.; Ghosh, S.; et al. Postoperative Complications and Mortality Following Colectomy for Ulcerative Colitis. *Clin. Gastroenterol. Hepatol.* **2011**, *9*, 972–980. [CrossRef]
- Alves, A.; Panis, Y.; Mathieu, P.; Manton, G.; Kwiatkowski, F.; Slim, K. Postoperative Mortality and Morbidity in French Patients Undergoing Colorectal Surgery: Results of a Prospective Multicenter Study. *Arch. Surg.* **2005**, *140*, 278–283. [CrossRef]
- Tevis, S.E.; Kennedy, G.D. Postoperative complications and implications on patient-centered outcomes. *J. Surg. Res.* **2013**, *181*, 106. [CrossRef] [PubMed]
- Wick, E.C.; Shore, A.D.; Hirose, K.; Ibrahim, A.M.; Gearhart, S.L.; Efron, J.; Weiner, J.P.; Makary, M.A. Readmission rates and cost following colorectal surgery. *Dis. Colon Rectum* **2011**, *54*, 1475–1479. [CrossRef] [PubMed]
- Louis, M.; Johnston, S.A.; Churilov, L.; Ma, R.; Christophi, C.; Weinberg, L. Financial burden of postoperative complications following colonic resection: A systematic review. *Medicine* **2021**, *100*, e26546. [CrossRef]
- Ludbrook, G.L. The Hidden Pandemic: The Cost of Postoperative Complications. *Curr. Anesthesiol. Rep.* **2022**, *12*, 1–9. [CrossRef]
- Zambouri, A. Preoperative evaluation and preparation for anesthesia and surgery. *Hippokratia* **2007**, *11*, 13.
- Royal College of Physicians Developed by the National Guideline Centre. *Preoperative Tests (Update) Routine Preoperative Tests for Elective Surgery*; National Guideline Centre UK: London, UK, 2016.
- Grocott, M.P.; Pearce, R.M. Perioperative medicine: The future of anaesthesia? *Bja Br. J. Anaesth.* **2012**, *108*, 723–726. [CrossRef]
- Grocott, M.P.; Edwards, M.; Mythen, M.G.; Aronson, S. Peri-operative care pathways: Re-engineering care to achieve the 'triple aim'. *Anaesthesia* **2019**, *74*, 90–99. [CrossRef] [PubMed]
- Noordzij, P.G.; Boersma, E.; Bax, J.J.; Feringa, H.H.; Schreiner, F.; Schouten, O.; Kertai, M.D.; Klein, J.; Urk, H.V.; Elhendy, A.; et al. Prognostic Value of Routine Preoperative Electrocardiography in Patients Undergoing Noncardiac Surgery. *Am. J. Cardiol.* **2006**, *97*, 1103–1106. [CrossRef]
- Vernooij, J.E.; Koning, N.J.; Geurts, J.W.; Holewijn, S.; Preckel, B.; Kalkman, C.J.; Vernooij, L.M. Performance and usability of pre-operative prediction models for 30-day peri-operative mortality risk: A systematic review. *Anaesthesia* **2023**, *78*, 607–619. [CrossRef]
- Wagner, D.P.; Draper, E.A. Acute physiology and chronic health evaluation (APACHE II) and Medicare reimbursement. *Health Care Financ. Rev.* **1984**, *1984*, 91.
- Jones, H.J.S.; Cossart, L.D. Risk scoring in surgical patients. *Br. J. Surg.* **1999**, *86*, 149–157. [CrossRef] [PubMed]
- Copeland, G.P. The POSSUM System of Surgical Audit. *Arch. Surg.* **2002**, *137*, 15–19. [CrossRef]
- Dosis, A.; Helliwell, J.; Syversen, A.; Tiernan, J.; Zhang, Z.; Jayne, D. Estimating postoperative mortality in colorectal surgery—A systematic review of risk prediction models. *Int. J. Color. Dis.* **2023**, *38*, 155. [CrossRef]

27. Pradhan, N.; Dyas, A.R.; Bronsert, M.R.; Lambert-Kerzner, A.; Henderson, W.G.; Qiu, H.; Colborn, K.L.; Mason, N.J.; Meguid, R.A. Attitudes about use of preoperative risk assessment tools: A survey of surgeons and surgical residents in an academic health system. *Patient Saf. Surg.* **2022**, *16*, 13. [CrossRef] [PubMed]
28. Goffi, L.; Saba, V.; Ghiselli, R.; Necozone, S.; Mattei, A.; Carle, F. Preoperative APACHE II and ASA scores in patients having major general surgical operations: Prognostic value and potential clinical applications. *Eur. J. Surg. Acta Chir.* **1999**, *165*, 730–735. [CrossRef]
29. Arena, R.; Myers, J.; Williams, M.A.; Gulati, M.; Kligfield, P.; Balady, G.J.; Collins, E.; Fletcher, G. Assessment of Functional Capacity in Clinical and Research Settings. *Circulation* **2007**, *116*, 329–343. [CrossRef]
30. Ferreira, V.; Lawson, C.; Ekmekjian, T.; Carli, F.; Scheede-Bergdahl, C.; Chevalier, S. Effects of preoperative nutrition and multimodal prehabilitation on functional capacity and postoperative complications in surgical lung cancer patients: A systematic review. *Support. Care Cancer* **2021**, *29*, 5597–5610. [CrossRef]
31. Mayo, N.E.; Feldman, L.; Scott, S.; Zavorsky, G.; Kim, D.J.; Charlebois, P.; Stein, B.; Carli, F. Impact of preoperative change in physical function on postoperative recovery: Argument supporting prehabilitation for colorectal surgery. *Surgery* **2011**, *150*, 505–514. [CrossRef]
32. Makker, P.G.S.; Koh, C.E.; Solomon, M.J.; Steffens, D.; BMedSc, P.G.S.M.; MBBS, C.E.K.; BCH, M.J.S.M.; BPhyt, D.S. Preoperative functional capacity and postoperative outcomes following abdominal and pelvic cancer surgery: A systematic review and meta-analysis. *Anz J. Surg.* **2022**, *92*, 1658–1667. [CrossRef]
33. Silvapulle, E.; Darvall, J. Objective methods for preoperative assessment of functional capacity. *Bja Educ.* **2022**, *22*, 312–320. [CrossRef]
34. Levett, D.Z.; Jack, S.; Swart, M.; Carlisle, J.; Wilson, J.; Snowden, C.; Riley, M.; Danjoux, G.; Ward, S.A.; Older, P.; et al. Perioperative cardiopulmonary exercise testing (CPET): Consensus clinical guidelines on indications, organization, conduct, and physiological interpretation. *Br. J. Anaesth.* **2018**, *120*, 484–500. [CrossRef] [PubMed]
35. Albouaini, K.; Egred, M.; Alahmar, A.; Wright, D.J. Cardiopulmonary exercise testing and its application. *Heart* **2007**, *93*, 1285. [CrossRef] [PubMed]
36. Hennis, P.J.; Meale, P.M.; Grocott, M.P. Cardiopulmonary exercise testing for the evaluation of perioperative risk in non-cardiopulmonary surgery. *Postgrad. Med. J.* **2011**, *87*, 550–557. [CrossRef]
37. Levett, D.Z.; Grocott, M.P. Cardiopulmonary exercise testing for risk prediction in major abdominal surgery. *Anesthesiol. Clin.* **2015**, *33*, 1–16. [CrossRef] [PubMed]
38. Reeves, T.; Bates, S.; Sharp, T.; Richardson, K.; Bali, S.; Plumb, J.; Anderson, H.; Prentis, J.; Swart, M.; Levett, D.Z.H. Cardiopulmonary exercise testing (CPET) in the United Kingdom—a national survey of the structure, conduct, interpretation and funding. *Perioper. Med.* **2018**, *7*, 2. [CrossRef]
39. Rose, G.A.; Davies, R.G.; Appadurai, I.R.; Williams, I.M.; Bashir, M.; Berg, R.M.; Poole, D.C.; Bailey, D.M. ‘Fit for surgery’: The relationship between cardiorespiratory fitness and postoperative outcomes. *Exp. Physiol.* **2022**, *107*, 787–799. [CrossRef] [PubMed]
40. Jones, L.; Tan, L.; Carey-Jones, S.; Riddell, N.; Davies, R.; Brownsdon, A.; Kelson, M.; Williams-Thomas, R.; Busse, M.; Davies, M.M.; et al. Can wearable technology be used to approximate cardiopulmonary exercise testing metrics? *Perioper. Med.* **2021**, *10*, 9. [CrossRef]
41. Ferreira, J.J.; Fernandes, C.I.; Rammal, H.G.; Veiga, P.M. Wearable technology and consumer interaction: A systematic review and research agenda. *Comput. Hum. Behav.* **2021**, *118*, 106710. [CrossRef]
42. Vijayan, V.; Connolly, J.; Condell, J.; McKelvey, N.; Gardiner, P. Review of Wearable Devices and Data Collection Considerations for Connected Health. *Sensors* **2021**, *21*, 5589. [CrossRef]
43. Smuck, M.; Odonkor, C.A.; Wilt, J.K.; Schmidt, N.; Swiernik, M.A. The emerging clinical role of wearables: Factors for successful implementation in healthcare. *npj Digit. Med.* **2021**, *4*, 45. [CrossRef]
44. Ontario, H.Q. Long-Term Continuous Ambulatory ECG Monitors and External Cardiac Loop Recorders for Cardiac Arrhythmia: A Health Technology Assessment. *Ont. Health Technol. Assess. Ser.* **2017**, *17*, 1–56.
45. Turakhia, M.P.; Hoang, D.D.; Zimetbaum, P.; Miller, J.D.; Froelicher, V.F.; Kumar, U.N.; Xu, X.; Yang, F.; Heidenreich, P.A. Diagnostic Utility of a Novel Leadless Arrhythmia Monitoring Device. *Am. J. Cardiol.* **2013**, *112*, 520–524. [CrossRef]
46. Yao, J.; Tan, C.S.; Lim, N.; Tan, J.; Chen, C.; Müller-Riemenschneider, F. Number of daily measurements needed to estimate habitual step count levels using wrist-worn trackers and smartphones in 212,048 adults. *Sci. Rep.* **2021**, *11*, 9633. [CrossRef]
47. Ricardo, L.I.C.; Wendt, A.; Galliano, L.M.; Muller, W.D.A.; Cruz, G.I.N.; Wehrmeister, F.; Brage, S.; Ekelund, U.; Silva, I.C.M. Number of days required to estimate physical activity constructs objectively measured in different age groups: Findings from three Brazilian (Pelotas) population-based birth cohorts. *PLoS ONE* **2020**, *15*, e0216017. [CrossRef]
48. Böhmer, A.B.; Wappler, F.; Zwißer, B. Preoperative Risk Assessment—From Routine Tests to Individualized Investigation. *Dtsch. Ärzteblatt Int.* **2014**, *111*, 437. [CrossRef]
49. Lequeux, B.; Uzan, C.; Rehman, M.B. Does resting heart rate measured by the physician reflect the patient’s true resting heart rate? White-coat heart rate. *Indian Heart J.* **2018**, *70*, 93–98. [CrossRef]
50. Deshmukh, S.D.; Shilaskar, S.N. Wearable sensors and patient monitoring system: A Review. In Proceedings of the 2015 International Conference on Pervasive Computing (ICPC), Pune, India, 8–10 January 2015; pp. 1–3. [CrossRef]

51. Maetzler, W.; Domingos, J.; Srulijes, K.; Ferreira, J.J.; Bloem, B.R. Quantitative wearable sensors for objective assessment of Parkinson's disease. *Mov. Disord.* **2013**, *28*, 1628–1637. [CrossRef] [PubMed]
52. Promphet, N.; Hinestroza, J.P.; Rattanawaleedirojn, P.; Soatthiyanon, N.; Siralermukul, K.; Potiyaraj, P.; Rodthongkum, N. Cotton thread-based wearable sensor for non-invasive simultaneous diagnosis of diabetes and kidney failure. *Sens. Actuators Chem.* **2020**, *321*, 128549. [CrossRef]
53. Quer, G.; Radin, J.M.; Gadaleta, M.; Baca-Motes, K.; Ariniello, L.; Ramos, E.; Kheterpal, V.; Topol, E.J.; Steinhubl, S.R. Wearable sensor data and self-reported symptoms for COVID-19 detection. *Nat. Med.* **2020**, *27*, 73–77. [CrossRef]
54. Kakria, P.; Tripathi, N.K.; Kitipawang, P. A real-time health monitoring system for remote cardiac patients using smartphone and wearable sensors. *Int. J. Telemed. Appl.* **2015**, *2015*, 373474. [CrossRef]
55. Huang, J.D.; Wang, J.; Ramsey, E.; Leavey, G.; Chico, T.J.A.; Huang, J.D.; Wang, J.; Ramsey, E.; Leavey, G.; Chico, T.J.A.; et al. Applying Artificial Intelligence to Wearable Sensor Data to Diagnose and Predict Cardiovascular Disease: A Review. *Sensors* **2022**, *22*, 8002. [CrossRef]
56. Wells, C.I.; Xu, W.; Penfold, J.A.; Keane, C.; Gharibans, A.A.; Bissett, I.P.; O'Grady, G. Wearable devices to monitor recovery after abdominal surgery: Scoping review. *BMJ Open* **2022**, *6*, e003031. [CrossRef] [PubMed]
57. van der Stam, A.; van Riel, N.A.W.; Machine, L.; van den Eijnden, M.A.C.; van der Stam, J.A.; Bouwman, R.A.; Mestrom, E.H.J.; Verhaegh, W.F.J.; van Riel, N.A.W.; Cox, L.G.E. Machine Learning for Postoperative Continuous Recovery Scores of Oncology Patients in Perioperative Care with Data from Wearables. *Sensors* **2023**, *23*, 4455. [CrossRef]
58. Daskivich, T.J.; Houtman, J.; Lopez, M.; Luu, M.; Fleschner, P.; Zaghiyan, K.; Cunneen, S.; Burch, M.; Walsh, C.; Paiement, G.; et al. Association of Wearable Activity Monitors With Assessment of Daily Ambulation and Length of Stay Among Patients Undergoing Major Surgery. *JAMA Netw. Open* **2019**, *2*, e187673. [CrossRef]
59. Esteban, P.A.; Hernández, N.; Novoa, N.; Varela, G. Evaluating patients' walking capacity during hospitalization for lung cancer resection. *Interact. Cardiovasc. Thorac. Surg.* **2017**, *25*, 268–271. [CrossRef]
60. Rinne, J.K.; Miri, S.; Oksala, N.; Vehkaoja, A.; Kössi, J. Evaluation of a wrist-worn photoplethysmography monitor for heart rate variability estimation in patients recovering from laparoscopic colon resection. *J. Clin. Monit. Comput.* **2023**, *37*, 45–53. [CrossRef] [PubMed]
61. Breteler, M.J.; Huizinga, E.; Loon, K.V.; Leenen, L.P.; Dohmen, D.A.; Kalkman, C.J.; Blokhuis, T.J. Reliability of wireless monitoring using a wearable patch sensor in high-risk surgical patients at a step-down unit in the Netherlands: A clinical validation study. *BMJ Open* **2018**, *8*, e020162. [CrossRef] [PubMed]
62. Waller, E.; Sutton, P.; Rahman, S.; Allen, J.; Saxton, J.; Aziz, O. Prehabilitation with wearables versus standard of care before major abdominal cancer surgery: A randomised controlled pilot study (trial registration: NCT04047524). *Surg. Endosc.* **2022**, *36*, 1008. [CrossRef]
63. Beilstein, C.M.; Krutkyte, G.; Vetsch, T.; Eser, P.; Wilhelm, M.; Stanga, Z.; Bally, L.; Verra, M.; Huber, M.; Wuethrich, P.Y.; et al. Multimodal prehabilitation for major surgery in elderly patients to lower complications: Protocol of a randomised, prospective, multicentre, multidisciplinary trial (PREHABIL Trial). *BMJ Open* **2023**, *13*, e070253. [CrossRef]
64. Feeney, C.; Reynolds, J.V.; Hussey, J. Preoperative physical activity levels and postoperative pulmonary complications post-esophagectomy. *Dis. Esophagus* **2011**, *24*, 489–494. [CrossRef]
65. Greco, M.; Angelucci, A.; Avidano, G.; Marelli, G.; Canali, S.; Aceto, R.; Lubian, M.; Oliva, P.; Piccioni, F.; Aliverti, A.; et al. Wearable Health Technology for Preoperative Risk Assessment in Elderly Patients: The WELCOME Study. *Diagnostics* **2023**, *13*, 630. [CrossRef] [PubMed]
66. Spathis, D.; Perez-Pozuelo, I.; Brage, S.; Wareham, N.J.; Mascolo, C. Self-supervised transfer learning of physiological representations from free-living wearable data. In Proceedings of the ACM CHIL 2021–2021 ACM Conference on Health, Inference, and Learning, Virtual, 8–9 April 2021; Volume 21, pp. 69–78. [CrossRef]
67. Soon, S.; Svavarsdottir, H.; Downey, C.; Jayne, D.G. Wearable devices for remote vital signs monitoring in the outpatient setting: An overview of the field. *BMJ Innov.* **2020**, *6*, 55. [CrossRef]
68. Taylor, J.M.; Gropper, M.A. Critical care challenges in orthopedic surgery patients. *Crit. Care Med.* **2006**, *34*, S191–S199. [CrossRef] [PubMed]
69. Haddaway, N.R.; Collins, A.M.; Coughlin, D.; Kirk, S. The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLoS ONE* **2015**, *10*, e0138237. [CrossRef] [PubMed]
70. Cos, H.; Li, D.; Williams, G.; Chininis, J.; Dai, R.; Zhang, J.; Srivastava, R.; Raper, L.; Sanford, D.; Hawkins, W.; et al. Predicting Outcomes in Patients Undergoing Pancreatectomy Using Wearable Technology and Machine Learning: Prospective Cohort Study. *J. Med. Internet Res.* **2021**, *23*, e23595. [CrossRef] [PubMed]
71. Zhang, J.; Li, D.; Dai, R.; Cos, H.; Williams, G.A.; Raper, L.; Hammill, C.W.; Lu, C. Predicting Post-Operative Complications with Wearables. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2022**, *6*, 1–27. [CrossRef]
72. Mylius, C.F.; Krijnen, W.P.; Takken, T.; Lips, D.J.; Eker, H.; van der Schans, C.P.; Klaase, J.M. Objectively measured preoperative physical activity is associated with time to functional recovery after hepato-pancreato-biliary cancer surgery: A pilot study. *Perioper. Med.* **2021**, *10*, 33. [CrossRef]
73. Angelucci, A.; Greco, M.; Canali, S.; Marelli, G.; Avidano, G.; Goretti, G.; Cecconi, M.; Aliverti, A. Fitbit Data to Assess Functional Capacity in Patients Before Elective Surgery: Pilot Prospective Observational Study. *J. Med. Internet Res.* **2023**, *25*, e42815. [CrossRef]

74. Hedrick, T.L.; Hassinger, T.E.; Myers, E.; Krebs, E.D.; Chu, D.; Charles, A.N.; Hoang, S.C.; Friel, C.M.; Thiele, R.H. Wearable Technology in the Perioperative Period: Predicting Risk of Postoperative Complications in Patients Undergoing Elective Colorectal Surgery. *Dis. Colon Rectum* **2020**, *63*, 538–544. [CrossRef]
75. Doryab, A.; Dey, A.K.; Kao, G.; Low, C. Modeling Biobehavioral Rhythms with Passive Sensing in the Wild. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2019**, *3*, 1–21. [CrossRef]
76. Sun, V.; Dumitra, S.; Ruel, N.; Lee, B.; Melstrom, L.; Melstrom, K.; Woo, Y.; Sentovich, S.; Singh, G.; Fong, Y. Wireless Monitoring Program of Patient-Centered Outcomes and Recovery Before and after Major Abdominal Cancer Surgery. *JAMA Surg.* **2017**, *152*, 852–859. [CrossRef]
77. Nakajima, H.; Yokoyama, Y.; Inoue, T.; Nagaya, M.; Mizuno, Y.; Kayamoto, A.; Nishida, Y.; Nagino, M. How Many Steps Per Day are Necessary to Prevent Postoperative Complications Following Hepato-Pancreato-Biliary Surgeries for Malignancy? *Ann. Surg. Oncol.* **2020**, *27*, 1387–1397. [CrossRef]
78. Novoa, N.M.; Varela, G.; Jiménez, M.F.; Ramos, J. Value of the average basal daily walked distance measured using a pedometer to predict maximum oxygen consumption per minute in patients undergoing lung resection. *Eur. J. Cardio-Thorac. Surg.* **2011**, *39*, 756–762. [CrossRef] [PubMed]
79. Billé, A.; Buxton, J.; Viviano, A.; Gammon, D.; Veres, L.; Routledge, T.; Harrison-Phipps, K.; Dixon, A.; Minetto, M.A. Preoperative Physical Activity Predicts Surgical Outcomes Following Lung Cancer Resection. *Integr. Cancer Ther.* **2021**, *20*, 1534735420975853. [CrossRef] [PubMed]
80. Richards, S.J.G.; Jerram, P.M.; Brett, C.; Falloon, M.; Frizelle, F.A. The association between low pre-operative step count and adverse post-operative outcomes in older patients undergoing colorectal cancer surgery. *Perioper. Med.* **2020**, *9*, 20. [CrossRef] [PubMed]
81. Yang, C.C.; Hsu, Y.L. A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring. *Sensors* **2010**, *10*, 7772. [CrossRef] [PubMed]
82. Hall, K.S.; Hyde, E.T.; Bassett, D.R.; Carlson, S.A.; Carnethon, M.R.; Ekelund, U.; Evenson, K.R.; Galuska, D.A.; Kraus, W.E.; Lee, I.M.; et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. *Int. J. Behav. Nutr. Phys. Act.* **2020**, *17*, 78. [CrossRef]
83. Zhao, M.; Veeranki, S.P.; Magnussen, C.G.; Xi, B. Recommended physical activity and all cause and cause specific mortality in US adults: Prospective cohort study. *BMJ* **2020**, *370*, 2031. [CrossRef]
84. Saint-Maurice, P.F.; Troiano, R.P.; Matthews, C.E.; Kraus, W.E. Moderate-to-Vigorous Physical Activity and All-Cause Mortality: Do Bouts Matter? *J. Am. Heart Assoc.* **2018**, *7*, e007678. [CrossRef]
85. Kane, W.J.; Hassinger, T.E.; Myers, E.L.; Chu, D.L.; Charles, A.N.; Hoang, S.C.; Friel, C.M.; Thiele, R.H.; Hedrick, T.L. Wearable technology and the association of perioperative activity level with 30-day readmission among patients undergoing major colorectal surgery. *Surg. Endosc.* **2022**, *36*, 1584–1592. [CrossRef]
86. Rossi, L.A.; Melstrom, L.G.; Fong, Y.; Sun, V. Predicting post-discharge cancer surgery complications via telemonitoring of patient-reported outcomes and patient-generated health data. *J. Surg. Oncol.* **2021**, *123*, 1345–1352. [CrossRef] [PubMed]
87. Eades, M.T.; Tsanas, A.; Juraschek, S.P.; Kramer, D.B.; Gervino, E.; Mukamal, K.J. Smartphone-recorded physical activity for estimating cardiorespiratory fitness. *Sci. Rep.* **2021**, *11*, 14851. [CrossRef]
88. Hallgrímsson, H.T.; Jankovic, F.; Althoff, T.; Allen, P.G.; Foschini, L. Learning Individualized Cardiovascular Responses from Large-scale Wearable Sensors Data. *arXiv* **2018**, arXiv:1812.01696. <https://doi.org/10.48550/arxiv.1812.01696>.
89. Spathis, D.; Perez-Pozuelo, I.; Gonzales, T.I.; Wu, Y.; Brage, S.; Wareham, N.; Mascolo, C. Longitudinal cardio-respiratory fitness prediction through wearables in free-living environments. *npj Digit. Med.* **2022**, *5*, 176. [CrossRef] [PubMed]
90. Saris, W.H.; Binkhorst, R.A. The use of pedometer and actometer in studying daily physical activity in man. Part I: Reliability of pedometer and actometer. *Eur. J. Appl. Physiol. Occup. Physiol.* **1977**, *37*, 219–228. [CrossRef]
91. Cui, H.W.; Kirby, G.S.; Surmacz, K.; Hargrove, C.; Griffiths, J.; Turney, B.W. The association of pre-operative home accelerometry with cardiopulmonary exercise variables. *Anaesthesia* **2018**, *73*, 738–745. [CrossRef] [PubMed]
92. Commons, C. CC BY 4.0 Deed. Available online: <https://creativecommons.org/licenses/by/4.0/> (accessed on 24 November 2023).
93. Moraes, J.L.; Rocha, M.X.; Vasconcelos, G.G.; Filho, J.E.V.; de Albuquerque, V.H.C.; Alexandria, A.R. Advances in Photoplethysmography Signal Analysis for Biomedical Applications. *Sensors* **2018**, *18*, 1894. [CrossRef]
94. Weiler, D.T.; Villajuan, S.O.; Edkins, L.; Cleary, S.; Saleem, J.J. Wearable heart rate monitor technology accuracy in research: A comparative study between PPG and ECG technology. In Proceedings of the Human Factors and Ergonomics Society, Rome, Italy, 28–30 September 2017; pp. 1292–1296. [CrossRef]
95. Kannel, W.B.; Kannel, C.; Paffenbarger, R.S.; Cupples, L.A. Heart rate and cardiovascular mortality: The Framingham study. *Am. Heart J.* **1987**, *113*, 1489–1494. [CrossRef]
96. Zhang, D.; Shen, X.; Qi, X. Resting heart rate and all-cause and cardiovascular mortality in the general population: A meta-analysis. *CMAJ* **2016**, *188*, E53–E63. [CrossRef]
97. Abbott, T.E.; Ackland, G.L.; Archbold, R.A.; Wragg, A.; Kam, E.; Ahmad, T.; Khan, A.W.; Niebrzegowska, E.; Rodseth, R.N.; Devereaux, P.J.; et al. Preoperative heart rate and myocardial injury after non-cardiac surgery: Results of a predefined secondary analysis of the VISION study. *BJA Br. J. Anaesth.* **2016**, *117*, 172. [CrossRef]
98. Freeman, W.K.; Gibbons, R.J. Perioperative Cardiovascular Assessment of Patients Undergoing Noncardiac Surgery. *Mayo Clin. Proc.* **2009**, *84*, 79. [CrossRef] [PubMed]

99. Castaneda, D.; Esparza, A.; Ghamari, M.; Soltanpur, C.; Nazeran, H. A review on wearable photoplethysmography sensors and their potential future applications in health care. *Int. J. Biosens. Bioelectron.* **2018**, *4*, 195. [CrossRef]
100. Elgendi, M.; Fletcher, R.; Liang, Y.; Howard, N.; Lovell, N.H.; Abbott, D.; Lim, K.; Ward, R. The use of photoplethysmography for assessing hypertension. *npj Digit. Med.* **2019**, *2*, 60. [CrossRef] [PubMed]
101. Avram, R.; Tison, G.H.; Aschbacher, K.; Kuhar, P.; Vittinghoff, E.; Butzner, M.; Runge, R.; Wu, N.; Pletcher, M.J.; Marcus, G.M.; et al. Real-world heart rate norms in the Health eHeart study. *npj Digit. Med.* **2019**, *2*, 58. [CrossRef] [PubMed]
102. Wang, R.; Blackburn, G.; Desai, M.; Phelan, D.; Gillinov, L.; Houghtaling, P.; Gillinov, M. Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiol.* **2017**, *2*, 104–106. [CrossRef]
103. Tanaka, H.; Monahan, K.D.; Seals, D.R. Age-predicted maximal heart rate revisited. *J. Am. Coll. Cardiol.* **2001**, *37*, 153–156. [CrossRef] [PubMed]
104. Dunn, J.; Kidzinski, L.; Runge, R.; Witt, D.; Hicks, J.L.; Rose, S.M.S.F.; Li, X.; Bahmani, A.; Delp, S.L.; Hastie, T.; et al. Wearable sensors enable personalized predictions of clinical laboratory measurements. *Nat. Med.* **2021**, *27*, 1105–1112. [CrossRef]
105. Kumar, S.; Victoria-Castro, A.M.; Melchinger, H.; O'Connor, K.D.; Psotka, M.; Desai, N.R.; Ahmad, T.; Wilson, F.P. Wearables in Cardiovascular Disease. *J. Cardiovasc. Transl. Res.* **2022**, *16*, 557–568. [CrossRef]
106. What Is an electrocardiogram (ECG)?—InformedHealth.org—NCBI Bookshelf. National Library of Medicine. Available online: <https://www.ncbi.nlm.nih.gov/books/NBK536878/#:~:text=An> (accessed on 24 November 2023).
107. World Health Organization. *Cardiovascular Diseases*; World Health Organization: Geneva, Switzerland, 2023.
108. Meek, S.; Morris, F. ABC of clinical electrocardiography: Introduction. I—Leads, rate, rhythm, and cardiac axis. *BMJ Br. Med. J.* **2002**, *324*, 415. [CrossRef]
109. Sammito, S.; Böckelmann, I. [Options and limitations of heart rate measurement and analysis of heart rate variability by mobile devices: A systematic review]. *Herzschrittmachertherapie Elektrophysiologie* **2016**, *27*, 38–45. [CrossRef]
110. Frandsen, M.N.; Mehlsen, J.; Foss, N.B.; Kehlet, H. Preoperative heart rate variability as a predictor of perioperative outcomes: A systematic review without meta-analysis. *J. Clin. Monit. Comput.* **2022**, *36*, 947–960. [CrossRef] [PubMed]
111. Steinhubl, S.R.; Waalen, J.; Edwards, A.M.; Ariniello, L.M.; Mehta, R.R.; Ebner, G.S.; Carter, C.; Baca-Motes, K.; Felicione, E.; Sarich, T.; et al. Effect of a Home-Based Wearable Continuous ECG Monitoring Patch on Detection of Undiagnosed Atrial Fibrillation: The mStoPS Randomized Clinical Trial. *JAMA* **2018**, *320*, 146–155. [CrossRef]
112. Prasada, S.; Desai, M.Y.; Saad, M.; Smilowitz, N.R.; Faulx, M.; Menon, V.; Moudgil, R.; Chaudhury, P.; Hussein, A.A.; Taigen, T.; et al. Preoperative Atrial Fibrillation and Cardiovascular Outcomes After Noncardiac Surgery. *J. Am. Coll. Cardiol.* **2022**, *79*, 2471–2485. [CrossRef] [PubMed]
113. McAlister, F.A.; Youngson, E.; Jacka, M.; Graham, M.; Conen, D.; Chan, M.; Szczeklik, W.; Alonso-Coello, P.; Devereaux, P.J. A comparison of four risk models for the prediction of cardiovascular complications in patients with a history of atrial fibrillation undergoing non-cardiac surgery. *Anaesthesia* **2020**, *75*, 27–36. [CrossRef] [PubMed]
114. Wu, Y.; Spathis, D.; Jia, H.; Perez-Pozuelo, I.; Gonzales, T.I.; Brage, S.; Wareham, N.; Mascolo, C. Turning Silver into Gold: Domain Adaptation with Noisy Labels for Wearable Cardio-Respiratory Fitness Prediction. *arXiv* **2022**, arXiv:2211.10475.
115. Altini, M.; Casale, P.; Penders, J.; Amft, O. Cardiorespiratory fitness estimation in free-living using wearable sensors. *Artif. Intell. Med.* **2016**, *68*, 37–46. [CrossRef]
116. Beltrame, T.; Amelard, R.; Wong, A.; Hughson, R.L. Prediction of oxygen uptake dynamics by machine learning analysis of wearable sensors during activities of daily living. *Sci. Rep.* **2017**, *7*, 45738. [CrossRef]
117. HEXOSKIN. *HEXOSKIN PROSHIRT—MEN'S*; HEXOSKIN: Montreal, Canada, 2023.
118. Altini, M. Personalization of energy expenditure and cardiorespiratory fitness estimation using wearable sensors in supervised and unsupervised free-living conditions. Ph.D. Thesis, 1(Research TU/e/Graduation TU/e); Eindhoven University of Technology, Eindhoven, The Netherlands, 2015.
119. Haveman, M.E.; van Melzen, R.; Schuurmann, R.C.; Hermens, H.J.; Tabak, M.; de Vries, J.P.P. Feasibility and patient's experiences of perioperative telemonitoring in major abdominal surgery: An observational pilot study. *Expert Rev. Med. Devices* **2022**, *19*, 515–523. [CrossRef]
120. Wang, C.F.; Wang, T.Y.; Kuo, P.H.; Wang, H.L.; Li, S.Z.; Lin, C.M.; Chan, S.C.; Liu, T.Y.; Lo, Y.C.; Lin, S.H.; et al. Upper-Arm Photoplethysmographic Sensor with One-Time Calibration for Long-Term Blood Pressure Monitoring. *Biosensors* **2023**, *13*, 321. [CrossRef]
121. Li, Y.; Wong, A.; Chung, W.M.; Li, M.I.; Molasiotis, A.I.; Bressington, D.; Ma, C.Z.H.; Kor, P.P.K.; Yeung, W.F. Evaluation of a Physical-Psychological Integrative (PPI) intervention for community-dwelling spinal cord injury survivors: Study protocol of a preliminary randomized controlled trial. *PLoS ONE* **2023**, *18*, e0282846. [CrossRef]
122. Bartlett, Y.K.; Webb, T.L.; Hawley, M.S. Using Persuasive Technology to Increase Physical Activity in People With Chronic Obstructive Pulmonary Disease by Encouraging Regular Walking: A Mixed-Methods Study Exploring Opinions and Preferences. *J. Med. Internet Res.* **2017**, *19*, e124. [CrossRef]
123. Cho, S.; Ensari, I.; Weng, C.; Kahn, M.G.; Natarajan, K. Factors Affecting the Quality of Person-Generated Wearable Device Data and Associated Challenges: Rapid Systematic Review. *Jmir Mhealth Uhealth* **2021**, *9*, e20738. [CrossRef] [PubMed]
124. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; John Wiley & Sons.: Hoboken, NJ, USA, 2019; pp. 1–449. [CrossRef]
125. Jakobsen, J.C.; Gluud, C.; Wetterslev, J.; Winkel, P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—A practical guide with flowcharts. *BMC Med. Res. Methodol.* **2017**, *17*, 162. [CrossRef] [PubMed]

126. Darji, J.; Biswas, N.; Jones, L.D.; Ashili, S. Handling missing data in the time-series data from wearables. In *Time Series Analysis—Recent Advances, New Perspectives and Applications [Working Title]*; IntechOpen: London, UK, 2023. [CrossRef]
127. Tackney, M.S.; Cook, D.G.; Stahl, D.; Ismail, K.; Williamson, E.; Carpenter, J. A framework for handling missing accelerometer outcome data in trials. *Trials* **2021**, *22*, 379. [CrossRef] [PubMed]
128. Haveman, M.E.; van Melzen, R.; Moumni, M.E.; Schuurmann, R.C.L.; Hermens, H.J.; Tabak, M.; de Vries, J.P.P.M. Determining the Reliable Measurement Period for Preoperative Baseline Values With Telemonitoring Before Major Abdominal Surgery: Pilot Cohort Study. *Jmir Perioper. Med.* **2022**, *5*, e40815. [CrossRef] [PubMed]
129. Lin, S.; Wu, X.; Martinez, G.; Chawla, N.V. Filling missing values on wearable-sensory time series data. In Proceedings of the 2020 SIAM International Conference on Data Mining, Cincinnati, OH, USA, 7–9 May 2020; pp. 46–54. [CrossRef]
130. Lin, J.Y.; Lu, Y.; Tu, X. How to avoid missing data and the problems they pose: Design considerations. *Shanghai Arch. Psychiatry* **2012**, *24*, 181. [CrossRef]
131. Meesad, P.; Hengpraprom, K. Combination of KNN-based feature selection and KNN-based missing-value imputation of microarray data. In Proceedings of the 3rd International Conference on Innovative Computing Information and Control, ICICIC'08, Dalian, China, 18–20 June 2008. [CrossRef]
132. Chakrabarti, S.; Biswas, N.; Karnani, K.; Padul, V.; Jones, L.D.; Kesari, S.; Ashili, S. Binned Data Provide Better Imputation of Missing Time Series Data from Wearables. *Sensors* **2023**, *23*, 1454. [CrossRef]
133. Bashar, S.K.; Ding, E.; Walkey, A.J.; McManus, D.D.; Chon, K.H. Noise Detection in Electrocardiogram Signals for Intensive Care Unit Patients. *IEEE Access Pract. Innov. Open Solut.* **2019**, *7*, 88357. [CrossRef]
134. Kang, S.; Paul, A.; Jeon, G. Reduction of mixed noise from wearable sensors in human-motion estimation. *Comput. Electr. Eng.* **2017**, *61*, 287–296. [CrossRef]
135. Stegle, O.; Fallert, S.V.; MacKay, D.J.; Brage, S. Gaussian process robust regression for noisy heart rate data. *IEEE Trans. Bio-Med. Eng.* **2008**, *55*, 2143–2151. [CrossRef]
136. Krishnan, S.; Athavale, Y. Trends in biomedical signal feature extraction. *BioMed. Signal Process. Control.* **2018**, *43*, 41–63. [CrossRef]
137. Mary, Q.; Khalid, S.; Khalil, T.; Nasreen, S. A Survey of Feature Selection And Feature Extraction Techniques in Machine Learning. In Proceedings of the SAI, 2014 Enhanced Framework for Recognizing Indoor Daily Life Activities View Project Shamila Nasreen a Survey of Feature Selection and Feature Extraction Techniques in Machine Learning, London, UK, 27–29 August 2014.
138. Yang, X.; Huang, B. An accurate step detection algorithm using unconstrained smartphones. In Proceedings of the 2015 27th Chinese Control and Decision Conference, CCDC 2015, Qingdao, China, 23–25 May 2015; pp. 5682–5687. [CrossRef]
139. Ducharme, S.W.; Lim, J.; Busa, M.A.; Aguiar, E.J.; Moore, C.C.; Schuna, J.M.; Barreira, T.V.; Staudenmayer, J.; Chipkin, S.R.; Tudor-Locke, C. A Transparent Method for Step Detection using an Acceleration Threshold. *J. Meas. Phys. Behav.* **2021**, *4*, 311. [CrossRef] [PubMed]
140. Mannini, A.; Sabatini, A.M. Machine learning methods for classifying human physical activity from on-body accelerometers. *Sensors* **2010**, *10*, 1154–1175. [CrossRef] [PubMed]
141. Luu, L.; Pillai, A.; Lea, H.; Buendia, R.; Khan, F.M.; Dennis, G. Accurate Step Count with Generalized and Personalized Deep Learning on Accelerometer Data. *Sensors* **2022**, *22*, 3989. [CrossRef] [PubMed]
142. Ho, N.H.; Truong, H.; Jeong, G.M.; Sabatini, A.M. Step-Detection and Adaptive Step-Length Estimation for Pedestrian Dead-Reckoning at Various Walking Speeds Using a Smartphone. *Sensors* **2016**, *16*, 1423. [CrossRef] [PubMed]
143. Tudor-Locke, C.; Craig, C.L.; Aoyagi, Y.; Bell, R.C.; Croteau, K.A.; Bourdeaudhuij, I.D.; Ewald, B.; Gardner, A.W.; Hatano, Y.; Lutes, L.D.; et al. How many steps/day are enough? For older adults and special populations. *Int. J. Behav. Nutr. Phys. Act.* **2011**, *8*, 80. [CrossRef]
144. Neishabouri, A.; Nguyen, J.; Samuelsson, J.; Guthrie, T.; Biggs, M.; Wyatt, J.; Cross, D.; Karas, M.; Migueles, J.H.; Khan, S.; et al. Quantification of acceleration as activity counts in ActiGraph wearable. *Sci. Rep.* **2022**, *12*, 11958. [CrossRef]
145. dos Santos, C.E.S.; Rech, C.R. Association between different cutoff points for objectively measured moderate-to-vigorous physical activity and cardiometabolic markers in older adults. *Arch. Gerontol. Geriatr.* **2020**, *91*, 104238. [CrossRef]
146. Pan, J.; Tompkins, W.J. A Real-Time QRS Detection Algorithm. *IEEE Trans. BioMed. Eng.* **1985**, *BME-32*, 230–236. [CrossRef]
147. Merino-Monge, M.; Castro-García, J.A.; Lebrato-Vázquez, C.; Gómez-González, I.M.; Molina-Cantero, A.J. Heartbeat detector from ECG and PPG signals based on wavelet transform and upper envelopes. *Phys. Eng. Sci. Med.* **2023**, *46*, 597–608. [CrossRef]
148. Speed, C.; Arneil, T.; Harle, R.; Wilson, A.; Karthikesalingam, A.; McConnell, M.; Phillips, J. Measure by measure: Resting heart rate across the 24-h cycle. *PLoS Digit. Health* **2023**, *2*, e0000236. [CrossRef] [PubMed]
149. Liu, Z.; Zhen, Y.; Lin, F.; Zheng, X.; Liu, X.; Sun, G.; Ye, Z.; Wen, J.; Liu, P. Resting heart rate as a preoperative predictor of postoperative atrial fibrillation after pulmonary thromboendarterectomy. *J. Card. Surg.* **2022**, *37*, 1644–1650. [CrossRef] [PubMed]
150. Ladha, K.S.; Beattie, W.S.; Tait, G.; Wijesundera, D.N. Association between preoperative ambulatory heart rate and postoperative myocardial injury: A retrospective cohort study. *Br. J. Anaesth.* **2018**, *121*, 722–729. [CrossRef] [PubMed]
151. Singh, N.; Moneghetti, K.J.; Christle, J.W.; Hadley, D.; Plews, D.; Froelicher, V. Heart Rate Variability: An Old Metric with New Meaning in the Era of using mHealth Technologies for Health and Exercise Training Guidance. Part One: Physiology and Methods. *Arrhythmia Electrophysiol. Rev.* **2018**, *7*, 193. [CrossRef] [PubMed]

152. Reimer, P.; Máca, J.; Szturcz, P.; Jor, O.; Kula, R.; Ševčík, P.; Burda, M.; Adamus, M. Role of heart-rate variability in preoperative assessment of physiological reserves in patients undergoing major abdominal surgery. *Ther. Clin. Risk Manag.* **2017**, *13*, 1223–1231. [CrossRef] [PubMed]
153. Foroozan, F.; Mohan, M.; Wu, J.S. Robust Beat-To-Beat Detection Algorithm for Pulse Rate Variability Analysis from Wrist Photoplethysmography Signals. In Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing—Proceedings, Calgary, AB, Canada, 15–20 April 2018; pp. 2136–2140. [CrossRef]
154. Aygun, A.; Ghasemzadeh, H.; Jafari, R. Robust interbeat interval and heart rate variability estimation method from various morphological features using wearable sensors. *IEEE J. BioMed. Health Inform.* **2020**, *24*, 2238–2250. [CrossRef] [PubMed]
155. Garmin. *Understanding HRV Status on Your Garmin Device*; Garmin: Olathe, Kansas, USA, 2023.
156. Fitbit. *How Do I Track Heart Rate with My Fitbit Device?* Fitbit: San Francisco, CA, USA, 2023.
157. Clifford, G.D.; Tarassenko, L. Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE Trans. Bio-Med. Eng.* **2005**, *52*, 630–638. [CrossRef]
158. Qiu, S.; Cai, X.; Sun, Z.; Li, L.; Zuegel, M.; Steinacker, J.M.; Schumann, U. Heart Rate Recovery and Risk of Cardiovascular Events and All-Cause Mortality: A Meta-Analysis of Prospective Cohort Studies. *J. Am. Heart Assoc.* **2017**, *6*, e005505. [CrossRef]
159. Charlton, P.H.; Birrenkott, D.A.; Bonnici, T.; Pimentel, M.A.; Johnson, A.E.; Alastruey, J.; Tarassenko, L.; Watkinson, P.J.; Beale, R.; Clifton, D.A. Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Rev. BioMed. Eng.* **2018**, *11*, 2. [CrossRef]
160. Lei, R.; Ling, B.W.K.; Feng, P.; Chen, J. Estimation of Heart Rate and Respiratory Rate from PPG Signal Using Complementary Ensemble Empirical Mode Decomposition with both Independent Component Analysis and Non-Negative Matrix Factorization. *Sensors* **2020**, *20*, 3238. [CrossRef]
161. Liu, G.Z.; Guo, Y.W.; Zhu, Q.S.; Huang, B.Y.; Wang, L. Estimation of Respiration Rate from Three-Dimensional Acceleration Data Based on Body Sensor Network. *Telemed. J.-Health* **2011**, *17*, 705. [CrossRef] [PubMed]
162. Leube, J.; Zschocke, J.; Kluge, M.; Pelikan, L.; Graf, A.; Glos, M.; Müller, A.; Bartsch, R.P.; Penzel, T.; Kantelhardt, J.W. Reconstruction of the respiratory signal through ECG and wrist accelerometer data. *Sci. Rep.* **2020**, *10*, 14530. [CrossRef] [PubMed]
163. Mohammadzadeh, F.; Nam, C.S.; Lobaton, E. Prediction of Physiological Response over Varying Forecast Lengths with a Wearable Health Monitoring Platform. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Honolulu, HI, USA, 18–21 July 2018; pp. 437–440. [CrossRef]
164. Smith, I.; Mackay, J.; Fahrid, N.; Krucke, D. Respiratory rate measurement: A comparison of methods. *Br. J. Healthc. Assist.* **2013**, *5*, 18–23. [CrossRef]
165. Bozdogan, H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* **1987**, *52*, 345–370. [CrossRef]
166. Rushing, C.; Bulusu, A.; Hurwitz, H.I.; Nixon, A.B.; Pang, H. A leave-one-out cross-validation SAS macro for the identification of markers associated with survival. *Comput. Biol. Med.* **2015**, *57*, 123–129. [CrossRef]
167. Pintelas, P.; Livieris, I.E. Special Issue on Ensemble Learning and Applications. *Algorithms* **2020**, *13*, 140. [CrossRef]
168. Belgiu, M.; Drăgu, L. Random forest in remote sensing: A review of applications and future directions. *Isprs J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [CrossRef]
169. Wan, S.; Yang, H. Comparison among methods of ensemble learning. In Proceedings of the 2013 International Symposium on Biometrics and Security Technologies, ISBAST 2013, Chengdu, China, 2–5 July 2013; pp. 286–290. [CrossRef]
170. Altini, M.; Casale, P.; Penders, J.; Velde, G.T.; Plasqui, G.; Amft, O. Cardiorespiratory fitness estimation using wearable sensors: Laboratory and free-living analysis of context-specific submaximal heart rates. *J. Appl. Physiol.* **2016**, *120*, 1082–1096. [CrossRef]
171. Powell, S.M.; Rowlands, A.V. Intermonitor Variability of the RT3 Accelerometer during Typical Physical Activities. *Med. Sci. Sport. Exerc.* **2004**, *36*, 324–330. [CrossRef]
172. Williams, G.J.; Al-Baraikhan, A.; Rademakers, F.E.; Ciravegna, F.; van de Vosse, F.N.; Lawrie, A.; Rothman, A.; Ashley, E.A.; Wilkins, M.R.; Lawford, P.V.; et al. Wearable technology and the cardiovascular system: The future of patient assessment. *Lancet Digit. Health* **2023**, *5*, e467–e476. [CrossRef]
173. Arani, M.S.A.; Costa, D.E.; Shihab, E. Human Activity Recognition: A Comparative Study to Assess the Contribution Level of Accelerometer, ECG, and PPG Signals. *Sensors* **2021**, *21*, 6997. [CrossRef] [PubMed]
174. Morelli, D.; Bartoloni, L.; Colombo, M.; Plans, D.; Clifton, D.A. Profiling the propagation of error from PPG to HRV features in a wearable physiological monitoring device. *Healthc. Technol. Lett.* **2018**, *5*, 59–64. [CrossRef]
175. Zazoum, B.; Batoo, K.M.; Khan, M.A.A. Recent Advances in Flexible Sensors and Their Applications. *Sensors* **2022**, *22*, 4653. [CrossRef] [PubMed]
176. Huang, Y.; Zhang, J.; Pu, J.; Guo, X.; Qiu, J.; Ma, Y.; Zhang, Y.; Yang, X. Resistive pressure sensor for high-sensitivity e-skin based on porous sponge dip-coated CB/MWCNTs/SR conductive composites. *Mater. Res. Express* **2018**, *5*, 065701. [CrossRef]
177. Antink, C.H.; Mai, Y.; Peltokangas, M.; Leonhardt, S.; Oksala, N.; Vehkaoja, A. Accuracy of heart rate variability estimated with reflective wrist-PPG in elderly vascular patients. *Sci. Rep.* **2021**, *11*, 8123. [CrossRef]
178. Lubitz, S.A.; Faranesh, A.Z.; Selvaggi, C.; Atlas, S.J.; McManus, D.D.; Singer, D.E.; Pagoto, S.; McConnell, M.V.; Pantelopoulos, A.; Foulkes, A.S. Detection of Atrial Fibrillation in a Large Population Using Wearable Devices: The Fitbit Heart Study. *Circulation* **2022**, *146*, 1415–1424. [CrossRef]

179. Martínez, J.P.; Zhang, S.; Li, Y.; Zhang, S.; Shahabi, F.; Xia, S.; Deng, Y.; Alshurafa, N. Deep Learning in Human Activity Recognition with Wearable Sensors: A Review on Advances. *Sensors* **2022**, *22*, 1476. [CrossRef]
180. Talebi, N.; Hallam, C.; Zanella, G. The new wave of privacy concerns in the wearable devices era. In Proceedings of the PICMET 2016—Portland International Conference on Management of Engineering and Technology: Technology Management For Social Innovation, Honolulu, HI, USA, 4–8 September 2017; pp. 3208–3214. [CrossRef]
181. Mone, V.; Shakhlo, F. Health Data on the Go: Navigating Privacy Concerns with Wearable Technologies. *Leg. Inf. Manag.* **2023**, *23*, 179–188. [CrossRef]
182. Simpson, L.A.; Menon, C.; Hodgson, A.J.; Mortenson, W.B.; Eng, J.J. Clinicians’ perceptions of a potential wearable device for capturing upper limb activity post-stroke: A qualitative focus group study. *J. Neuroeng. Rehabil.* **2021**, *18*, 135. [CrossRef]
183. FDA. *General Wellness: Policy for Low Risk Devices*; FDA: Rockville, MD, USA, 2019.
184. Yang, C.; Shang, L.; Yao, S.; Ma, J.; Xu, C. Cost, time savings and effectiveness of wearable devices for remote monitoring of patient rehabilitation after total knee arthroplasty: Study protocol for a randomized controlled trial. *J. Orthop. Surg. Res.* **2023**, *18*, 461. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Sensors Editorial Office
E-mail: sensors@mdpi.com
www.mdpi.com/journal/sensors



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editor. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editor and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.



Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-3538-6