UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Advancing the Clinical Relevance and Discovery of Smartphone-Derived Biomarkers

Permalink

https://escholarship.org/uc/item/8w4476g6

ISBN

9798193298755

Author

Barry, Colin

Publication Date

2026-10-15

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

ADVANCING THE CLINICAL RELEVANCE AND DISCOVERY OF SMARTPHONE-DERIVED BIOMARKERS

A Dissertation submitted in partial satisfaction of the requirements for the degree

Doctor of Philosophy

in

Electrical Engineering (Signal and Image Processing)

by

Colin Barry

Committee in charge:

Professor Edward Jay Wang, Chair
Professor Tauhidur Rahman, Co-Chair
Professor Sujit Dey
Professor Rajesh Gupta
Professor Xinyu Zhang

2026

Copyright

Colin Barry, 2026

All rights reserved.

The dissertation of Colin Barry is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego
2026

Table of Contents

Dissertation Approval Page iii

Table of Contents iv

List of Figures x

List of Tables xii

Acknowledgements xiv

Vita xvi

Abstract of the Dissertation xviii

1 Introduction 1

1.1 The Vision: An App Store of Medical Sensors 1

1.2 Challenges for Software-Enabled Medical Sensing on Smartphones 4

1.2.1 The Ever-Changing Sensor Package 4

1.2.2 Heterogeneity Within "Standard" Sensors 5

1.2.3 APIs Not Designed for Health Sensing 6

1.2.4 Physical Impedance Mismatch 7

1.3 Design Philosophy and Scope 9

1.3.1 Anchor to Universal Sensors 9

1.3.2 Validate Against Gold Standards, Not Just Against Prior Work . . . 9

1.3.3 Prioritize Equity Across Demographics 10

1.3.4 Design for Deployment, Including to Older Adults 11

1.4 Enabling Infrastructure: Camera Photoplethysmography Calibration 12

1.5 Research Contributions Overview 14
1.5.1 Neurological Assessment via Smartphone Pupillometry 14
1.5.2 Cardiovascular Monitoring via Oscillometric Blood Pressure 14
1.5.3 Musculoskeletal Assessment via Functional Strength Testing 15
1.6 Dissertation Organization 16

2 Neurological Assessment via Smartphone Pupillometry 17
2.1 Clinical Significance of Pupillometry 17
2.2 Why RGB Cameras Fail: The Physics of the Problem 18
2.2.1 Limitations of Prior Smartphone Pupillometry 19
2.3 NIR Facial Recognition Camera Repurposing 20
2.3.1 Concept and Hardware 20
2.3.2 Distance Estimation and Absolute Pupil Measurement 20
2.3.3 Pupil Detection Neural Network 21
2.3.4 Smartphone Application 22
2.3.5 Validation Against a Clinical Pupillometer 22
2.3.6 Limitations of the NIR Approach 24
2.4 Far-Red Spectrum Pupillometry for Universal RGB Cameras 26
2.4.1 Concept 26
2.4.2 Prototype Design 27
2.4.3 Cross-Device Contrast Study 28
2.4.4 Accuracy Validation 28
2.4.5 Limitations of the Far-Red Approach 29
2.5 Usability Study: Deployment to Older Adults 30
2.5.1 Motivation and Design 30
2.5.2 Results 31
2.5.3 Design Implications 31
2.6 Synthesis: Toward a Universal Smartphone Pupillometer 32

2.7 Limitations and Future Work 32

3 Cardiovascular Monitoring via Oscillometric Blood Pressure 34

3.1 The Hypertension Access Problem 34

3.2 Background: Oscillometry and the Two-Measurement Problem 35

3.2.1 The Oscillometric Method 35

3.2.2 Prior Work on Smartphone Blood Pressure 36

3.2.3 Camera Photoplethysmography for Blood Volume Sensing 37

3.3 Vibrometric Force Estimation 38

3.3.1 Physical Principle 38

3.3.2 Signal Processing 38

3.3.3 Per-Device-Model Calibration 40

3.3.4 Cross-Device Reliability 40

3.3.5 Intra-Device Reliability 41

3.4 Oscillometric BP Measurement Pipeline 41

3.4.1 Measurement Protocol 41

3.4.2 Signal Processing and Oscillogram Construction 43

3.4.3 Blood Pressure Feature Extraction and Model 43

3.5 Blood Pressure Proof-of-Concept Study 44

3.5.1 Participants and Protocol 44

3.5.2 Measurement Exclusion 45

3.5.3 Blood Pressure Results 45

3.5.4 Oscillogram Group Analysis 46

3.5.5 Equity Analysis 46

3.6 BPClip: Contextualizing an Alternative Implementation 48

3.6.1 Overview and Mechanism 48

3.6.2 Performance and Convergent Validation 50

3.7 The VibroBP Application 50

3.8 Translational Pathway 52
3.9 Limitations and Future Work 53
3.9.1 Measurement Performance 53
3.9.2 Usability 53
3.9.3 Population and Conditions 54
3.9.4 Device Dependence 54

4 Smartphone-Augmented Functional Assessment for Musculoskeletal Health 55

4.1 The Information Discarded by Clinical Scoring 55
4.2 Clinical Context: Sarcopenia and Functional Assessment 56
4.3 Handgrip Strength: Replicating an Established Measurement 57
4.3.1 Clinical Significance and the Access Gap 57
4.3.2 Smartphone-Based HGS Measurement 58
4.3.3 HGS as the Replication Case Study 60
4.4 The 30-Second Chair Stand Test: Augmenting Beyond the Count 60
4.4.1 The Information Loss Argument 60
4.4.2 Study Design and Data Collection 62
4.4.3 Signal Processing and Cycle Detection 63
4.4.4 Feature Engineering: Six Movement Categories 63
4.4.5 Inverse Kinematics Model 65
4.4.6 Statistical Analysis 66
4.4.7 Results 67
4.5 The Augmentation Principle 72
4.6 Toward At-Home Deployment 74
4.7 Limitations and Future Work 75
4.7.1 Cohort Representativeness 75

4.7.2 Kinematic Model Validity 75
4.7.3 Single-Visit Cross-Sectional Design 76
4.7.4 DXA System Variability 76

5 Discussion and Conclusion **77**

5.1 Revisiting the Thesis 77
5.1.1 Expanding Access to Established Clinical Measurements 77
5.1.2 Facilitating Novel Digital Biomarker Discovery 78
5.1.3 Prioritizing Usability for Target Populations 79
5.2 Cross-Cutting Technical Contributions 80
5.2.1 Vibrometric Force Estimation as a General Sensing Primitive 80
5.2.2 Far-Red Spectral Filtering as a General Solution to Melanin-Confounded Optical Measurement 81
5.2.3 The Postable Paradigm for Distributed Clinical Accessories 81
5.2.4 Camera PPG Calibration as Foundational Infrastructure 82
5.3 The Equity Dimension 82
5.4 Translational Pathway 83
5.5 Honest Limitations of the Body of Work 85
5.6 Future Directions 86
5.7 Closing Statement 88

A Camera Photoplethysmography Calibration **89**

A.1 Motivation and Background 89
A.2 The Two ISP Nonlinearities 90
A.2.1 Tone Mapping 90
A.2.2 Zero Light Offset 90
A.3 Calibration Procedure 91
A.3.1 Tone Mapping Linearization 91

A.3.2 Zero Light Offset Measurement 92
A.3.3 Pre-Measurement Exposure Calibration 92
A.4 Validation and Impact on PPG Quality 93
A.5 Implementation Notes 94

B IRB Protocols and Study Demographics 95

B.1 Chapter 2: Pupillometry Studies 95
B.1.1 NIR Validation and Far-Red Contrast Studies 95
B.1.2 Older Adult Usability Deployment 96
B.1.3 MCI Biomarker Analysis: VETSA Dataset 97
B.2 Chapter 3: Blood Pressure Study 97
B.3 Chapter 4: Musculoskeletal Studies 97
B.3.1 30-Second Chair Stand Study 97
B.3.2 Handgrip Strength Study 98

C Signal Processing Code Documentation 100

C.1 Pupillometry Processing Pipeline 100
C.1.1 NIR Pupil Detection and Distance Estimation 100
C.1.2 Far-Red PPG-Based Pupil Tracking 101
C.2 Blood Pressure Processing Pipeline 102
C.2.1 PPG Signal Extraction 102
C.2.2 Vibrometric Force Estimation 102
C.2.3 Oscillogram Construction and Feature Extraction 103
C.3 Chair Stand Processing Pipeline 103
C.3.1 IMU Preprocessing 103
C.3.2 Cycle Detection 104
C.3.3 Feature Computation 104
C.3.4 Model Training and Evaluation 104

References **106**

List of Figures

1.1 Representative medical applications enabled by smartphone sensors 3

1.2 The physical impedance mismatch between smartphone form factor and medical measurement postures. 8

1.3 Illustration of the two nonlinearities that corrupt smartphone camera photoplethysmography 12

2.1 NIR smartphone pupillometry system overview. 22

2.2 Bench-top distance calibration setup and results. 23

2.3 Representative PLR and digit span traces: smartphone (NIR system) versus clinical pupillometer (Neuroptics PLR-2000) for one light-iris and one dark-iris participant. 24

2.4 NIR smartphone pupillometry validation. 25

2.5 Additional regression and Bland-Altman analysis for the NIR pupillometry validation. 25

2.6 Overview of the far-red smartphone pupillometry system. 26

2.7 Melanin absorption spectrum across 400–900 nm. 27

2.8 Far-red pupillometry prototype. 28

2.9 Side-by-side camera images of a dark iris under full-spectrum RGB illumination (left column) versus far-red illumination (right column), across four smartphone models (Google Pixel 4, Apple iPhone 12, Samsung Galaxy A53, Motorola Moto G Power 2022). 29

2.10 Far-red pupillometry accuracy validation against the Neuroptics PLR-3000. 30

2.11 Representative frames from the older adult usability study. 31

3.1 VibroBP system overview. 39
3.2 VFE cross-device force estimation validation. 42
3.3 Blood pressure proof-of-concept validation. 47
3.4 Mean oscillogram by systolic BP category (<100, 100–120, 120–140, >140 mmHg). 48
3.5 BPClip mechanism and design. 49
3.6 BPClip systolic BP validation. 51

4.1 End-to-end 30-second chair stand study pipeline. 61
4.2 Leave-one-out R^2 grouped bar chart for four predictor conditions (IMU-only, all features, count-only, and height-and-weight) across all three regression targets (lean mass, percent body fat, and dominant handgrip strength). . . . 68
4.3 Leave-one-out predicted versus actual percent body fat under the IMU-only condition ($N = 78$ participants). 70
4.4 Leave-one-out predicted versus actual dominant handgrip strength (right hand) under the IMU-only condition. 71
4.5 Raw Pearson r versus demographic-adjusted partial correlation coefficient for each IMU feature category, for total lean mass (left) and percent body fat (right). 71
4.6 Smartphone-based pinch-force estimation validated against a reference force sensor (Google Pixel 4, $N = 209{,}603$ six-channel IMU samples, 15-fold hold-one-out cross-validation). 72

List of Tables

1.1 ZLO characterized for different phone models 12

B.1 Pupillometry validation study participant demographics. 96

B.2 Older adult usability study participant demographics. 96

B.3 Blood pressure study participant demographics ($N = 24$). 98

B.4 Chair stand study participant demographics ($N = 78$). 99

Acknowledgements

I would like to thank my advisor, Professor Edward Jay Wang, for his mentorship, guidance, and support throughout the course of this research. I would also like to thank the members of my doctoral committee for their time, feedback, and guidance in shaping this dissertation.

I am grateful to my collaborators and coauthors whose contributions made the work in this dissertation possible, including Yinan Xuan, Ava J. Fascetti, Nick Antipa, Alison A. Moore, David Wing, Jessica de Souza, Jason Holden, and Eric Granholm. Their partnership across the studies described in this dissertation was invaluable.

This effort would not have been possible without the support and love of my family and close friends. I am forever appreciative of my family's support and understanding of me pursuing my interests. A special thanks to Megan Barry, Cameron Barry, and Marius Ruh for getting me through to the end.

Chapter 1, in part, is a reprint of material as it appears in Xuan, Y., Barry, C., Antipa, N., and Wang, E. J. (2023), "A calibration method for smartphone camera photoplethysmography," Frontiers in Digital Health. The dissertation author was a co-investigator and coauthor of this material.

Chapter 2, in part, is a reprint of material as it appears in Barry, C., de Souza, J., Xuan, Y., Holden, J., Granholm, E., and Wang, E. J. (2022), "At-Home Pupillometry using Smartphone Facial Identification Cameras," Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, and in Barry, C. and Wang, E. J. (2023), Scientific Reports. The dissertation author was the primary investigator and author of both papers.

Chapter 3, in part, is a reprint of material as it appears in Barry, C., Xuan, Y., Fascetti, A., Moore, A., and Wang, E. J. (2024), Scientific Reports. The dissertation author was the primary investigator and author of this material.

Chapter 4, in part, is currently being prepared for submission as Barry, C., Wang, E. J., and Wing, D., and as Barry, C. and Wang, E. J. The dissertation author was the primary investigator and author of both manuscripts.

Vita

Education

2020–2026 Ph.D., Electrical Engineering (Signal and Image Processing), University of California San Diego. Advisor: Professor Edward Jay Wang.

2020–2022 M.S., Electrical Engineering (Signal and Image Processing), University of California San Diego.

2016–2020 B.S., Bioengineering (Biosystems), University of California San Diego.

Employment

2022–Present
Cofounder and Chief Technology Officer, Billion Labs Inc., La Jolla, CA.

2020–Present
Graduate Student Researcher, Digital Health Laboratories, University of California San Diego.

Publications

1. C. Barry, E. J. Wang, and D. Wing, "Smartphone instrumented 30-second chair stand recovers body composition and strength biomarkers," manuscript in preparation, 2026.

2. C. Barry and E. J. Wang, "Smartphone-based hand grip strength measurement," manuscript in preparation, 2026.

3. A. J. Fascetti, A. Pestonjamasp, C. Barry, and E. J. Wang, "Postable spirometry: user-assembled vortex whistle spirometer," in *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2024.

4. C. Barry, Y. Xuan, A. Fascetti, A. Moore, and E. J. Wang, "Oscillometric blood pressure measurements on smartphones using vibrometric force estimation," *Scientific Reports*, vol. 14, p. 26206, 2024.

5. Y. Xuan, C. Barry, N. Antipa, and E. J. Wang, "A calibration method for smartphone camera photoplethysmography," *Frontiers in Digital Health*, vol. 5, 2023.

6. C. Barry and E. J. Wang, "Racially fair pupillometry measurements for RGB smartphone cameras using the far red spectrum," *Scientific Reports*, vol. 13, p. 13841, 2023.

7. Y. Xuan, C. Barry, J. De Souza, J. H. Wen, N. Antipa, A. A. Moore, and E. J. Wang,

"Ultra-low-cost mechanical smartphone attachment for no-calibration blood pressure measurement," *Scientific Reports*, vol. 13, p. 8105, 2023.

8. C. Barry, J. de Souza, Y. Xuan, J. Holden, E. Granholm, and E. J. Wang, "At-home pupillometry using smartphone facial identification cameras," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2022.

9. C. Barry, T. Rahman, and E. J. Wang, "Tenets towards smartphone-based medical tricorders," in *Proceedings of the 2021 Workshop on Future of Digital Biomarkers*, 2021.

Honors and Awards

2025 Best Demo Award, Stanford E-Wear Competition / UbiComp 2025 / IEEE BSN 2025.

2024 Don Norman Design Award, Don Norman Design Award and Summit.

2023 A2Collective Pilot Award, Co-Investigator, Massachusetts AI and Technology Center.

2022–2023 La Verne Noyes Endowed Fellowship, La Verne Noyes Foundation.

2022 Best Paper Honorable Mention, ACM CHI Conference on Human Factors in Computing Systems.

2020–2021 Powell Fellowship, Charles Lee Powell Foundation.

Patents

2024 RGB Camera Pupillometry for All Eye Colors Utilizing Far Red Light. Non-exclusively licensed to Vitamin Greens.

2023 Smartphone-Based Hand Grip Strength Measurement. Licensed to Billion Labs Inc.

2023 Mobile Device Based Oscillometric Blood Pressure Measurement. Licensed to Billion Labs Inc.

ABSTRACT OF THE DISSERTATION

Advancing the Clinical Relevance and Discovery of Smartphone-Derived Biomarkers

by

Colin Barry

Doctor of Philosophy in Electrical Engineering (Signal and Image Processing)

University of California San Diego, 2026

Professor Edward Jay Wang, Chair

Professor Tauhidur Rahman, Co-Chair

Commodity smartphones carry an accelerometer, camera, and microphone that were designed for photography and communication, not medicine, yet these same sensors can be repurposed to capture physiological signals of clinical relevance without additional hardware. This dissertation develops and validates smartphone-based sensing methods across three domains of clinical assessment: neurological, cardiovascular, and musculoskeletal.

In the neurological domain, this work presents an at-home pupillometry system that uses a smartphone's front-facing camera to track pupillary light reflex, enabling remote assessment relevant to conditions such as schizophrenia, and extends this approach with a low-cost far-red attachment that improves measurement accuracy across a range of skin tones and ambient lighting conditions.

In the cardiovascular domain, rather than the pulse-transit-time or waveform-based estimation common to consumer wearables, this work replicates the oscillometric method used by clinical blood pressure cuffs, for its established clinical validation. The central contribution is Vibrometric Force Estimation (VFE), a method for measuring the force applied by a fingertip to a smartphone camera using only the smartphone's built-in vibration motor and camera, without any physical attachment or per-user calibration. Combined with camera photoplethysmography, VFE enables the first absolute oscillometric blood pressure measurement on an unmodified smartphone. This chapter also contextualizes VFE alongside BPClip, an alternative mechanical implementation of the same oscillometric principle, and includes a calibration method for smartphone camera photoplethysmography that improves blood pressure and other cardiovascular estimates across diverse skin tones.

In the musculoskeletal domain, this work augments the 30-second chair stand test, an assessment already administered in millions of clinical encounters, with smartphone inertial sensing to recover movement information that the standard repetition count discards, rather than introducing a new assessment unrelated to existing practice; it separately estimates a pinch-force proxy for grip strength from smartphone sensors, validated against a benchtop reference force sensor. The chair stand work is validated against dual-energy X-ray absorptiometry and clinical handgrip dynamometry, together with a user-assembled vortex whistle spirometer for pulmonary function screening.

Across all three domains, this dissertation demonstrates that careful signal processing and calibration can extract clinically meaningful physiological measurements from sensors already present in commodity smartphones, expanding the reach of at-home and point-of-

care health monitoring without requiring specialized or costly hardware.

Chapter 1

Introduction

1.1 The Vision: An App Store of Medical Sensors

The mobile health community has long held a vision for the medical tricorder: a single device capable of performing a broad range of clinical measurements, available to anyone, anywhere, at any time. For most of the twentieth century, that vision was constrained by the cost and complexity of medical hardware. The stethoscope, the sphygmomanometer, the spirometer, and the pupillometer each required their own dedicated device, their own trained operator, and their own institutional setting. The idea that these measurements could converge into a single pocket-sized instrument seemed, for most of that history, firmly in the realm of science fiction.

The smartphone changed the terms of that problem. Not by being a medical device, it was never designed as one, but by placing a remarkable collection of sensors into the hands of billions of people worldwide. A modern smartphone contains a high-resolution camera, a near-infrared illuminator, an inertial measurement unit with accelerometer and gyroscope, a vibration motor, a microphone, a speaker, a bright LED flashlight, and a powerful general-purpose processor, all in a device that costs less than a clinical pupillometer, a hospital-grade blood pressure cuff, or a spirometer. More importantly, it is a device people already own. The barrier to access is not the hardware. It is the software.

The vision articulated in this dissertation is grounded in that observation. If the sensors needed to measure the pupil, the blood pressure, and the functional strength of the human body are already present in the smartphones that people carry in their pockets, then the problem of expanding access to those measurements becomes, in principle, a software

problem. A developer can write an application, distribute it through an app store, and instantaneously transform hundreds of millions of existing devices into medical sensors. No new hardware needs to be manufactured. No distribution network needs to be built. No patient needs to purchase or be issued a dedicated device. The measurement arrives as a software update.

This vision carries significant implications for healthcare equity. Hypertension is the leading preventable cause of premature death and disability globally, and yet an estimated 46% of the more than 1.56 billion adults living with hypertension remain unaware of their condition[1]. Clinical pupillometers, which measure the pupillary light reflex as a biomarker for neurological diseases including Alzheimer's disease, Parkinson's disease, and traumatic brain injury, cost several thousand dollars per unit and are confined to high-resource institutions; the consequence is that longitudinal, at-home, and population-scale pupillometry research is essentially infeasible with current infrastructure[2]. Spirometry, the most common pulmonary function test for diagnosing asthma and chronic obstructive pulmonary disease, requires patients to travel to a clinic equipped with a spirometer that costs between one thousand and five thousand dollars[3]. Sarcopenia, the age-related loss of skeletal muscle mass and function, affects between five and fifty percent of adults over sixty years of age, yet the gold-standard body composition assessment, dual-energy X-ray absorptiometry, is available only in specialized clinical or research facilities[4]. In each of these cases, the measurement itself is well-established. The barrier is not scientific uncertainty about what to measure. It is access.

The smartphone offers a path around that barrier. But realizing the vision requires more than recognizing that the hardware is present. Every smartphone is different. An application developed for one phone model has no guarantee of functioning on the next. Sensors that exist in one generation of devices are removed in the next. The programming interfaces through which developers access sensor data were designed for photography and gaming, not for clinical measurement. The physical form factor of a phone, optimized

for calls, media consumption, and touch interaction, is poorly matched to the ergonomic requirements of medical sensing. And the regulatory and validation standards that govern medical devices are not designed for software that runs on consumer hardware that changes every year.

This dissertation documents a research program that engages directly with those challenges. It does not claim that smartphones can replace dedicated medical devices in all contexts. It claims something more specific and more tractable: that by repurposing universal smartphone sensors, deployable software provides opportunities for expanding access to established clinical measurements and facilitating the discovery of novel digital biomarkers, while prioritizing usability for older adults and other target populations. That claim is evaluated across three sensing domains, neurological assessment, cardiovascular monitoring, and musculoskeletal assessment, through a series of system developments, clinical validations, and deployment studies that span the full research arc from proof of concept to real-world use.

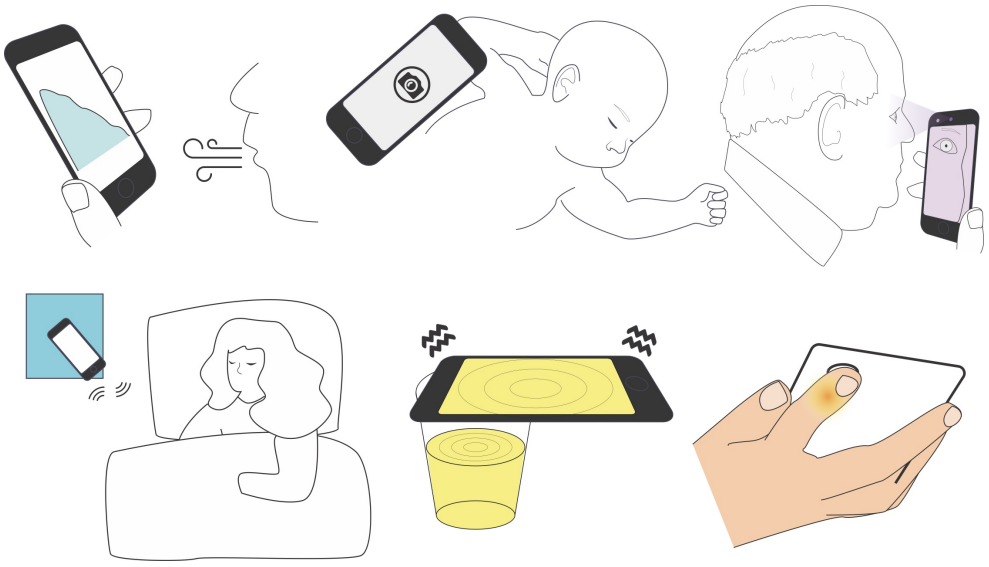


Figure 1.1: Representative medical applications enabled by smartphone sensors. The smartphone serves as the hardware substrate for a growing array of clinical measurement modalities, spanning neurological, cardiovascular, and musculoskeletal assessment.

1.2 Challenges for Software-Enabled Medical Sensing on Smartphones

The gap between the vision described above and the practical reality of smartphone-based medical sensing is substantial. Understanding that gap is essential context for the technical contributions of this dissertation. The challenges are not primarily algorithmic. They are structural, stemming from the fundamental mismatch between the economic logic that drives smartphone design and the requirements of clinical measurement.

1.2.1 The Ever-Changing Sensor Package

Since the introduction of the first iPhone in 2007, smartphones have accumulated an impressive array of sensors: GPS, front-facing cameras, flash LEDs, nine-degree-of-freedom inertial measurement units, fingerprint readers, near-infrared facial recognition cameras, LiDAR scanners, pressure-sensitive screens, and millimeter-wave gesture sensors, among others. Each new sensor generation opens new possibilities for health sensing. The near-infrared facial recognition camera introduced in the iPhone X and Google Pixel 4, for example, enables pupillometry measurements that are robust to ambient lighting conditions and work equitably across eye colors of varying melanin content. The pressure-sensitive screen introduced in the iPhone 7 enabled novel blood pressure measurement approaches. The LiDAR scanner in the iPhone 12 Pro opens possibilities for depth-based physiological assessment.

The problem is that not all sensors introduced in one generation survive to the next. The near-infrared facial recognition camera has been phased out of most major smartphone lines, including the Google Pixel and most Motorola devices, and Apple has never exposed its equivalent camera to third-party developers. Pressure-sensitive screens were removed from the iPhone lineup after the iPhone 6S. IR blasters appeared in Samsung's Galaxy S4 through S6 and then disappeared. The economic incentive driving these decisions is entirely rational from a consumer electronics perspective, components that do not directly enhance photography, battery life, or the core user experience are removed to reduce cost

and thickness. But for medical sensing applications that depend on a specific sensor, the removal of that sensor renders the application obsolete and strands the research investment that went into validating it.

This instability has a direct implication for the design philosophy of this dissertation. Where possible, the sensing approaches developed here are anchored to sensors that are present in essentially every smartphone, the RGB camera, the inertial measurement unit, the vibration motor, and the microphone, rather than relying on sensors that may be present in some devices but absent in others. When a sensor that is not universally present is used, as is the case with the near-infrared camera in the pupillometry work described in Chapter 2, the limitation is acknowledged explicitly, and a complementary approach that works on universal hardware is developed alongside it.

1.2.2 Heterogeneity Within "Standard" Sensors

Even for sensors that are present in essentially every smartphone, the hardware is far from uniform. The RGB camera varies in lens width, aperture, focal length, and distance from the LED flashlight. The LED flashlight varies in spectral output and luminance. The microphone opening may be a single pinhole, a slot, or a perforated grille, each with different acoustic frequency responses and different sensitivities to turbulent airflow. The inertial measurement unit may use an eccentric rotating mass motor or a linear resonant actuator, with fundamentally different vibration characteristics that affect any measurement that relies on the vibration signal. The front-facing camera may be embedded in a bezel, a notch, or a pinhole cutout in a bezel-less display, each affecting the geometry of any stereo-vision calculation.

These differences are not incidental. They matter for the physics of the measurement. If the spectral output of the LED flashlight differs across phone models, the illumination conditions during fingertip photoplethysmography change, affecting the ratio-of-ratios calculation that underlies pulse oximetry and hemoglobin estimation. If the distance between

the camera and the flashlight varies, the illumination geometry of the measurement changes. If the vibration motor characteristics differ, a force estimation model trained on one phone model may not transfer to another without recalibration.

The research in this dissertation addresses hardware heterogeneity in two complementary ways. The first is through calibration: the camera photoplethysmography calibration work described in Appendix A demonstrates that the nonlinear tone mapping applied by smartphone computational photography pipelines can be linearized through a per-model zero light offset correction, making measurements comparable across devices. The second is through cross-device validation: each sensing system developed in this dissertation is validated not on a single phone model but across multiple devices from different manufacturers, providing empirical evidence of generalizability rather than relying on theoretical arguments.

1.2.3 APIs Not Designed for Health Sensing

The programming interfaces through which developers access smartphone sensors were designed for photography, gaming, navigation, and haptic feedback, not for physiological measurement. Camera APIs expose controls for focus, white balance, and exposure that are useful for taking good photographs but introduce nonlinear processing pipelines that distort the amplitude information needed for clinical optical measurements. The inertial measurement unit API returns acceleration and angular velocity at a sampling rate that may be throttled by the operating system based on processor load or thermal management. The vibration motor API accepts a command to vibrate but does not expose information about the actual vibratory motion produced. Proximity sensors return binary presence values rather than the continuous distance measurements that would be useful for medical applications.

Critically, the timing of sensor data acquisition is not precisely controlled. The APIs are asynchronous: a developer submits a command to begin recording, and frames or samples are returned based on the parameters submitted, but the exact timing of when

recording begins is not guaranteed. For measurements that require synchronization between two sensor subsystems, for example, synchronizing inertial measurement unit data with camera photoplethysmography data during a blood pressure measurement, this timing uncertainty is a source of measurement error that must be addressed in the signal processing pipeline rather than eliminated at the hardware level.

1.2.4 Physical Impedance Mismatch

The physical form factor of a smartphone is designed around its primary use cases: holding it to the ear for phone calls, viewing its screen at arm's length for media consumption, and tapping and swiping its touchscreen for interaction. It was not designed to be pressed against the chest for cardiac sensing, held over the eye for pupillometry, pressed down on by a fingertip for blood pressure measurement, or held at shoulder height during a chair stand test. The ergonomics of these medical measurement postures are awkward relative to the natural handling of a phone, and the awkwardness introduces measurement noise, user error, and data quality problems that are independent of the underlying sensing algorithm.

Dedicated medical devices address this mismatch through purpose-built physical interfaces: the finger clip of the pulse oximeter, the nose clip and mouthpiece of the spirometer, the eye cup of the clinical pupillometer, the arm cuff of the blood pressure monitor. These interfaces are not incidental to the measurement, they are essential to achieving the positioning, contact, and ambient light exclusion that the measurement requires. Smartphone-based medical sensing systems must either tolerate the noisier and less controlled measurement conditions that result from a suboptimal physical interface, or they must introduce a physical attachment that compensates for the mismatch. Both strategies are employed in this dissertation. The attachment-free approaches, vibrometric force estimation for blood pressure, inertial measurement unit analysis for the chair stand test, are designed to be robust to the natural variation in how a user holds and positions the phone. The attachment-based approaches, the far-red pupillometry scope, the BPClip, are designed to be as simple and

inexpensive as possible, ideally distributable through the postal system, so that the introduction of an attachment does not recreate the access barriers that the smartphone was meant to eliminate.



Figure 1.2: The physical impedance mismatch between smartphone form factor and medical measurement postures. The smartphone is optimized for calls, media, and touch interaction; medical sensing requires holding postures, over the eye for pupillometry, pressed to the chest for cardiac sensing, fingertip-down for blood pressure, that are awkward relative to normal phone use and introduce measurement noise and user error.

1.3 Design Philosophy and Scope

The work in this dissertation is organized around a consistent set of design principles that reflect the challenges described in the preceding section. These principles are not constraints imposed from outside but conclusions drawn from experience across multiple years of developing, deploying, and validating smartphone-based medical sensing systems.

1.3.1 Anchor to Universal Sensors

Each system developed in this dissertation is designed to function on the broadest possible range of smartphone hardware. The primary sensing modalities employed, the RGB camera, the inertial measurement unit, the vibration motor, and the microphone, are present in essentially every smartphone currently in use worldwide, from budget devices costing under one hundred dollars to flagship models costing over one thousand. When a sensor that is not universally present is leveraged, as in the near-infrared pupillometry system, the scope of applicability is explicitly characterized and an alternative approach for universal hardware is developed.

This principle also guides the treatment of smartphone camera photoplethysmography, which underlies the blood pressure measurement in Chapter 3. The computational photography pipeline of a smartphone camera introduces nonlinear transformations, most importantly, a nonlinear tone mapping curve and a sensor-specific zero light offset, that vary across devices and corrupt the amplitude information that oscillometric blood pressure estimation requires. Appendix A presents a calibration methodology that linearizes the camera measurement, making it device-agnostic. This calibration is designed to be performed once per phone model by the application developer rather than once per device by the end user.

1.3.2 Validate Against Gold Standards, Not Just Against Prior Work

A persistent limitation of prior work in smartphone-based medical sensing is the tendency to validate new systems against other smartphone systems or against human

labeling rather than against dedicated clinical instruments. The pupillometry literature, for example, includes studies that report accuracy in millimeters but derive those millimeter values by applying a fixed pixel-to-millimeter conversion constant without a ground truth distance measurement, and that validate accuracy against human image labeling rather than against a clinical pupillometer[5]. The blood pressure literature includes systems that report performance in terms of correlation with a reference measurement without establishing that the reference measurement was itself obtained under controlled conditions[6].

Each system in this dissertation is validated against a recognized clinical gold standard. The near-infrared pupillometer is validated against the Neuroptics PLR-2000, a clinical-grade pupillometer used in neurological research and intensive care settings. The far-red RGB pupillometer is validated against the Neuroptics PLR-3000. The vibrometric blood pressure system is validated against an FDA-approved oscillometric arm cuff. The chair stand body composition model is validated against dual-energy X-ray absorptiometry. These are not merely methodological choices, they are the minimum standard of evidence required before a measurement can be described as clinical-grade.

1.3.3 Prioritize Equity Across Demographics

A recurring failure mode in prior smartphone-based optical sensing is the dependence of measurement accuracy on skin tone or iris color. In the RGB camera domain, the melanin in darker irises absorbs visible light such that the boundary between the iris and the pupil is nearly invisible in standard full-spectrum images; prior smartphone pupillometry systems have reported failure rates of up to 23% for dark-iris participants, with the majority of failures concentrated in participants with dark brown irises[7]. Similar issues affect photoplethysmography measurements, where darker skin tones reduce the amplitude of the optical signal, and colorimetric measurements of blood biomarkers, where melanin in the fingernail bed interferes with the ratio calculations.

The sensing approaches developed in this dissertation address these issues at the level

of physics rather than by training models on more diverse datasets. The far-red pupillometry system works because the far-red spectrum, approximately 630 to 700 nanometers, falls at a wavelength where melanin absorption is substantially lower than in the full visible spectrum, causing the iris to appear lighter and greatly increasing the contrast between the iris and the pupil regardless of baseline iris color. The vibrometric blood pressure system uses force estimation and oscillometric analysis that are independent of skin optical properties. Cross-device and cross-demographic validation is included in each study to provide empirical rather than theoretical evidence of equity.

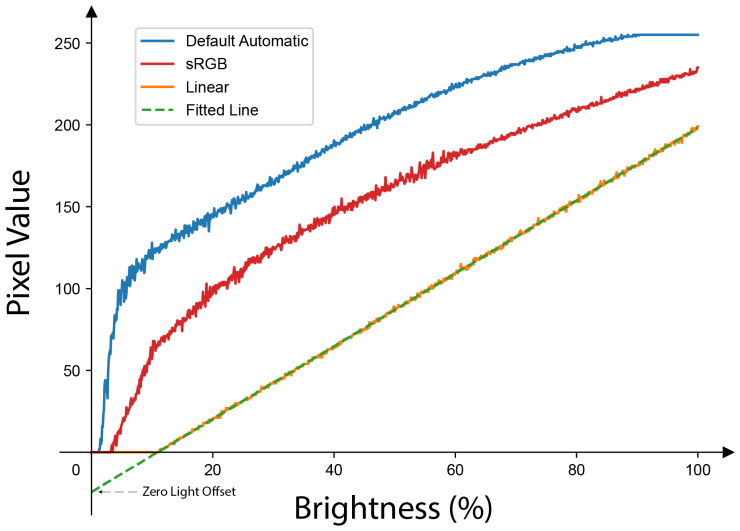
1.3.4 Design for Deployment, Including to Older Adults

A medical sensing system that achieves high accuracy in a laboratory setting with young, technically experienced participants but fails in the hands of an older adult using it at home has not solved the access problem. The populations that stand to benefit most from accessible medical sensing, older adults at risk for Alzheimer's disease, hypertensive patients in low-resource communities, individuals with sarcopenia who have difficulty traveling to clinical facilities, are often the populations with the least experience using smartphones and the greatest vulnerability to usability barriers.

This dissertation incorporates usability considerations throughout, not as an afterthought. The pupillometry usability study described in Chapter 2 deployed smartphone pupillometers to fifteen older adults with little to no prior smartphone experience, administering the system remotely via video call without researcher control over the device. The lessons from that deployment, eyelid and eyelash occlusion obscuring the pupil, difficulty operating the touch interface, confusion about scope attachment, directly informed subsequent system design. The blood pressure proof-of-concept study in Chapter 3 identifies the high exclusion rate for exercise measurements and the fingertip placement problem as the primary usability limitations, with specific design modifications proposed to address each. The chair stand and handgrip systems in Chapter 4 are designed with older adults as the

target population from the outset, given that sarcopenia is a geriatric condition. Usability is treated as a dimension of clinical validity, not a secondary concern.

1.4 Enabling Infrastructure: Camera Photoplethysmography Calibration



Phone Model	ZLO
Pixel 4	-22.5
Pixel 7	-14
Samsung S22	-19.6
Motorola G 2022	-14.9

Table 1.1: ZLO characterized for different phone models

Figure 1.3: Illustration of the two nonlinearities that corrupt smartphone camera photoplethysmography: (left) the nonlinear tone mapping applied by the computational photography pipeline, which distorts the AC/DC amplitude ratio needed for optical biomarker estimation; and (right) the per-model zero light offset, a nonzero pixel value output in the absence of incident light that introduces a systematic bias in the DC amplitude. The calibration methodology developed by Xuan and Barry et al. [8] corrects both effects.

Before describing the three application domains that constitute the core of this dissertation, it is necessary to introduce a cross-cutting methodological contribution that underlies two of them: a calibration method for smartphone camera photoplethysmography.

Smartphone camera photoplethysmography, the use of the smartphone camera to measure the pulsatile variation in blood volume at the fingertip or other vascularized tissue, is the signal foundation of the blood pressure measurement in Chapter 3 and is relevant to a broader class of optical biomarker measurements including pulse oximetry, hemoglobin concentration estimation, and vascular aging. The measurement relies on computing the ratio of alternating current (AC) to direct current (DC) signal components across two or more color channels of the camera image. This ratio of ratios encodes information about the

optical absorption properties of blood and tissue at different wavelengths, and it is the basis for clinical pulse oximetry.

The problem is that smartphone cameras do not behave as linear photosensors. The computational photography pipeline that processes raw sensor data before delivering a video frame to a developer's application applies nonlinear tone mapping, typically a logarithmic or gamma-curve transformation designed to mimic the nonlinear perception of human vision, that distorts the DC and AC amplitudes in ways that corrupt the ratio of ratios calculation. The default automatic tone mapping in Android's Camera2 API adapts dynamically based on the scene, producing transformations that are not predictable or reversible in post-processing. Additionally, each smartphone model exhibits a sensor-specific zero light offset, a nonzero pixel value that is output even in the absence of incident light, that introduces a systematic bias in the DC amplitude measurement.

The calibration methodology developed in collaboration with Xuan et al.[8] addresses both of these issues. By setting the camera to a linear tone mapping mode via the Camera2 API and characterizing the zero light offset for each phone model using a photonic test bench with pulse-width-modulated LED illumination, the camera's response can be linearized and made comparable across devices. Across four smartphone models, this calibration reduces the mean absolute error in ratio of ratios measurements by 74% relative to default settings and improves the coefficient of determination from 0.81 to 0.97. The calibration is transferable across phones of the same model, meaning that the application developer can perform the calibration once on a representative device and deploy the correction factor to all users of that model.

This work is presented in full in Appendix A. It is introduced here because it represents a solved enabling problem that subsequent chapters build upon. Chapter 3 describes in Section 3.2.3 how the calibration is applied in the context of oscillometric blood pressure measurement, and the implications for other optical measurement applications are discussed in Chapter 5.

1.5 Research Contributions Overview

This dissertation makes contributions across three clinical sensing domains. Each domain is anchored by one or more first-author publications, with supporting contributions from collaborative work that provides additional context, validation, or design principles. The three domains are described briefly here and in detail in Chapters 2 through 4.

1.5.1 Neurological Assessment via Smartphone Pupillometry

Chapter 2 presents two complementary approaches to smartphone pupillometry. The first repurposes the near-infrared facial recognition camera present in some recent smartphones to perform absolute pupil size measurements with sub-millimeter accuracy, validated against a clinical pupillometer across a participant population weighted toward dark iris colors[9]. The second develops a universal approach that works on any smartphone with an RGB camera, using a far-red longpass optical filter and dedicated 625-nanometer LED illumination in a low-cost attachment to achieve the high iris-to-pupil contrast that the near-infrared approach provides natively[10]. Together, these contributions define a roadmap for deployable smartphone pupillometry: absolute measurements on devices that expose near-infrared cameras, and clinically accurate relative measurements on all others.

1.5.2 Cardiovascular Monitoring via Oscillometric Blood Pressure

Chapter 3 presents the first absolute oscillometric blood pressure measurement on an unmodified smartphone, requiring no physical attachment, no hardware modification, and no individual calibration against a reference cuff measurement. The enabling contribution is Vibrometric Force Estimation (VFE): a technique that uses the smartphone's vibration motor and inertial measurement unit to measure the force applied by the user's fingertip to the smartphone camera, without any specialized hardware[11]. Combined with camera photoplethysmography as the blood volume sensing modality, VFE enables a complete oscillometric blood pressure measurement pipeline in a downloadable software application.

Chapter 3 also contextualizes the VFE-based approach by describing BPClip, an alternative implementation of the same oscillometric principle using a mechanical spring-loaded clip as the force-sensing mechanism[12]. The convergence of two completely different implementation strategies, software-only force estimation via VFE and passive mechanical force encoding via BPClip, on similar clinical accuracy strengthens the case for the oscillometric approach itself as a viable platform for smartphone blood pressure measurement.

1.5.3 Musculoskeletal Assessment via Functional Strength Testing

Chapter 4 presents a smartphone-based approach to functional strength and body composition assessment centered on the 30-second chair stand test. The key contribution is the extraction of inverse kinematics features from inertial measurement unit data recorded during the test, enabling estimation of knee and hip joint angles and the derivation of force, power, and coordination surrogates that go beyond the simple stand count that the test traditionally produces[13]. Validated against dual-energy X-ray absorptiometry in a cohort of 78 participants, the smartphone-derived features achieve strong performance for total lean mass estimation ($R = 0.92$) and more modest performance for percent body fat estimation ($R = 0.63$), establishing a foundation for a smartphone-based sarcopenia screening tool.

Chapter 4 also describes a smartphone-based handgrip strength measurement system[14] and Postable Spirometry[3], a spirometer self-assembled from a laser-cut postcard and used with a smartphone microphone. Together, these contributions constitute a smartphone-based functional assessment battery, chair stand, handgrip, and spirometry, that could be administered at home and distributed at scale, without requiring any clinic visit or specialized equipment.

1.6 Dissertation Organization

Chapter 2 presents the neurological assessment work, covering both the near-infrared and far-red RGB smartphone pupillometry systems and their application toward mild cognitive impairment screening. Chapter 3 presents the cardiovascular monitoring work, covering the camera photoplethysmography signal integrity problem, the vibrometric force estimation approach, the oscillometric blood pressure proof of concept, and the BPClip contextual comparison. Chapter 4 presents the musculoskeletal assessment work, covering the 30-second chair stand inverse kinematics analysis, the handgrip strength system, and postable spirometry. Chapter 5 synthesizes the findings across all three domains, evaluates the thesis claim against the evidence assembled, identifies the cross-cutting technical contributions of the body of work, and discusses the translational pathway toward clinical deployment and regulatory approval.

Appendix A presents the full methodology and validation results for the camera photoplethysmography calibration work. Appendix B contains study IRB protocols and participant demographic tables. Appendix C provides signal processing pipeline documentation for the chair stand and blood pressure systems.

Chapter 2

Neurological Assessment via Smartphone Pupillometry

2.1 Clinical Significance of Pupillometry

The pupil is the aperture of the eye: a circular opening in the colored iris that controls the amount of light reaching the retina. Like the aperture of a camera, the pupil dilates in dim conditions and constricts under bright illumination, optimizing the retinal image across a wide range of ambient light levels. But the pupil does not respond to light alone. It is innervated by both branches of the autonomic nervous system, the parasympathetic system drives constriction, and the sympathetic system drives dilation, and through this dual innervation it is sensitive to a broad range of neurological states including cognitive load, emotional arousal, pain, fatigue, pharmacological effects, and the early functional changes associated with neurodegenerative disease [15, 16].

Pupillometry, the measurement of pupil size and its dynamic response to stimuli, has been a tool of neurological assessment since the nineteenth century. Its modern clinical applications span a wide range of contexts. In the intensive care unit, automated pupillometry of the pupillary light reflex (PLR) is used to monitor intracranial pressure and detect neurological deterioration in patients with traumatic brain injury. In neurological research, the amplitude and velocity of the PLR have been identified as biomarkers for Parkinson's disease [17], schizophrenia [18], and mild cognitive impairment (MCI) as a precursor to Alzheimer's disease [19, 20]. In cognitive neuroscience, task-evoked pupil dilation recorded during working memory tasks reflects the activity of the locus coeruleus, a brainstem nucleus that is among the first regions to exhibit pathological changes in Alzheimer's disease [21].

Despite this rich clinical and research literature, the tools available to measure the

pupil remain largely confined to high-resource institutional settings. Clinical pupillometers cost several thousand dollars per unit and require trained operators. The consequence is a significant gap between what pupillometry research has established is measurable and what is routinely measured. Longitudinal studies of pupillary biomarkers over weeks or months are infeasible in clinic-bound settings. Population-scale screening programs for MCI are not practical with devices that require institutional ownership. At-home measurement in naturalistic conditions is essentially impossible.

This chapter presents a research program directed at closing that gap. The goal is not merely to replicate what clinical pupillometers do on cheaper hardware, but to deploy measurement at scales and in settings that clinical devices cannot reach. The chapter presents two complementary technical approaches that together constitute a roadmap for universal smartphone pupillometry, followed by an application of the resulting measurement platform toward a clinically meaningful biomarker discovery problem.

2.2 Why RGB Cameras Fail: The Physics of the Problem

The fundamental challenge of smartphone pupillometry is one of contrast. To measure the pupil accurately, the imaging system must reliably distinguish the dark pupil aperture from the surrounding iris tissue. In a clinical pupillometer, this is achieved by imaging in the near-infrared (NIR) spectrum (approximately 800–1200 nm). In the NIR, ocular melanin reflects light strongly regardless of its concentration, causing even dark brown irises to appear bright. The dark pupil aperture, which absorbs light across all wavelengths, stands in sharp contrast against this bright background [2].

In the full visible spectrum (400–700 nm) captured by a standard RGB smartphone camera, the situation is fundamentally different. Melanin absorbs visible light strongly, with absorption highest at short wavelengths and decreasing progressively toward longer wavelengths. For a light blue iris, low melanin content means the iris reflects visible light adequately and the pupil-iris boundary is visible in standard color photographs. For a dark

brown iris, high melanin concentration causes the iris to absorb most visible light and appear nearly as dark as the pupil itself. The pupil-iris boundary effectively disappears. This is not a limitation of imaging algorithms, it is a physical limitation of the optical signal.

The practical consequence is severe. Prior feasibility studies of smartphone-based pupillometry reported failure rates of up to 23% for dark-iris participants, with the majority of failures concentrated in individuals with dark brown irises [7]. Even in successful detections, the contrast-to-noise ratio is fundamentally lower for dark irises, introducing systematic measurement errors that cannot be fully corrected by more sophisticated segmentation. The failure is in the signal, not in the software.

2.2.1 Limitations of Prior Smartphone Pupillometry

Beyond the iris color problem, prior smartphone pupillometry systems exhibited several additional limitations. First, all prior systems used the RGB camera's flash LED as the sole illumination source, constraining the measurement to the pupil constriction phase of the PLR and preventing recording of the recovery phase. A complete PLR test requires both phases. Second, prior systems measured only relative pupil change, change expressed as a percentage of baseline, rather than absolute pupil diameter in millimeters. Converting pixel measurements to millimeters requires a known camera-to-pupil distance, which prior systems estimated either with an unvalidated fixed constant or by scaling to the ground truth measurement [5]. Third, accuracy validation was performed against human image labeling rather than against a clinical pupillometer, establishing only that the algorithm agreed with a human observer, not that the measurements were clinically accurate.

The two systems presented in this chapter address all of these limitations. Both are validated against clinical gold-standard pupillometers. Neither requires the flash to be directed at the eye, enabling full PLR tests including the recovery phase. Both address the iris color problem at the level of the optical signal rather than the algorithm.

2.3 NIR Facial Recognition Camera Repurposing

2.3.1 Concept and Hardware

The first approach exploits the NIR facial recognition camera system present in some smartphones released between approximately 2017 and 2022, including the Google Pixel 4, Apple iPhone X through 14, and Samsung Galaxy S9 through S21. Designed for biometric authentication, this system contains the same optical hardware that makes clinical pupillometry possible: a NIR illuminator and a NIR-sensitive camera, in which the pupil-iris contrast is high regardless of iris color.

What makes this hardware particularly useful for pupillometry is its geometric relationship to the RGB selfie camera on the same device. Both cameras are embedded in the same front bezel at a fixed, known horizontal separation, creating a stereo camera pair. By measuring the horizontal disparity between the pupil's position in the NIR image and its position in the RGB image, the distance from the phone to the pupil can be calculated. That distance measurement converts the pixel-based pupil diameter from the NIR camera into an absolute diameter in millimeters. This represents, to our knowledge, the first smartphone pupillometry system to perform and validate absolute pupil size measurements against a clinical pupillometer [9].

2.3.2 Distance Estimation and Absolute Pupil Measurement

The distance estimation pipeline relies on simultaneous capture from both cameras. Due to API limitations on the devices tested, simultaneous NIR and RGB video is not possible; instead, the application captures RGB images at the beginning and end of each test while NIR video is recorded throughout. The baseline distance is calculated from the first RGB frame and the first stable NIR frame.

The disparity between the normalized horizontal pupil center positions in the two

camera images is:

$$D = \left| \frac{x^*_{\text{RGB}}}{\text{RGB image width}} - \frac{x^*_{\text{NIR}}}{\text{NIR image width}} \right| \tag{2.1}$$

where x^* denotes the horizontal pixel coordinate of the pupil center in each image, normalized by image width to account for differing camera resolutions. The distance Z from the camera plane to the pupil is then:

$$Z = \frac{K}{D} \tag{2.2}$$

where K is an experimentally determined calibration constant, established via a bench-top experiment positioning the phone at known distances above a printed test target using a 3D-printed stair structure.

With Z estimated, the absolute pupil diameter is:

$$d_{\text{pupil}} = 2\, r_{\text{px}} \times \frac{2Z\tan(\theta/2)}{N_{\text{px}}} \tag{2.3}$$

where θ is the camera angle of view, N_{px} is the number of horizontal pixels in the NIR image, and r_{px} is the continuous-valued pupil radius output by the neural network. Because r_{px} is non-integer, this formulation supports sub-pixel pupil diameter estimation, enabling accuracy beyond the physical resolution limit of the NIR camera.

2.3.3 Pupil Detection Neural Network

Pupil detection in NIR and RGB frames uses the convolutional neural network of Eivazi *et al.* [22], selected for its validated performance across multiple eye tracking datasets and its suitability for real-time deployment. The network outputs pupil center coordinates and radius per frame. For NIR frames, the output radius enters the diameter calculation. For RGB frames, only the pupil center coordinates are needed for the disparity calculation, which avoids the iris color problem in the RGB domain entirely.

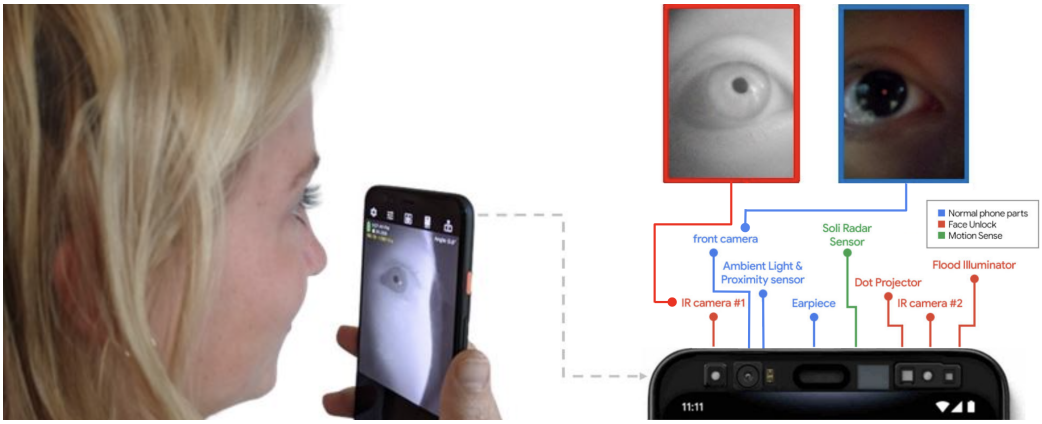


Figure 2.1: NIR smartphone pupillometry system overview. The stereo NIR+RGB camera geometry of the Google Pixel 4 is shown. Disparity between the pupil position in the NIR image and the RGB image yields the camera-to-eye distance Z, which converts the pixel-domain pupil radius into an absolute diameter in millimeters.

2.3.4 Smartphone Application

The system was implemented as an Android application targeting the Google Pixel 4 using the Camera2 API. Two versions were developed: a validation version for simultaneous recording alongside a clinical pupillometer under researcher supervision, and a deployment version for self-administration by older adults. Both guide the user through the test via verbal audio instructions. The application monitors the proximity sensor to verify correct camera-to-eye positioning and pauses the test if the phone is moved too far away. Data processing runs off-device after collection.

2.3.5 Validation Against a Clinical Pupillometer

Study Design

The validation study compared the smartphone pupillometer to the Neuroptics PLR-2000, a recognized clinical gold standard for pupil measurement in neurological research and intensive care. Eleven participants were recruited with a deliberate overrepresentation of dark iris colors (five blue, one hazel, three brown, two dark brown) to provide an explicit test of the system's equity properties.

Two tests were administered to each participant. The first was a PLR test in which

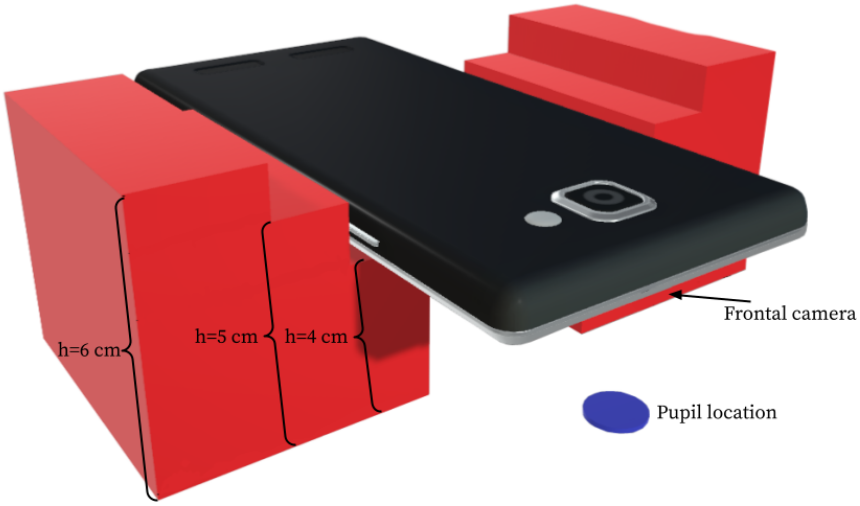


Figure 2.2: Bench-top distance calibration setup and results. (Left) A 3D-printed stair structure positions the smartphone at known, incrementally varying distances above a printed test target, enabling the calibration constant K to be determined from ground-truth distance measurements. (Right) Distance estimation accuracy across the calibration range, showing the relationship between the measured disparity D and the known distance Z.

the phone screen brightness shifted from minimum to maximum for five seconds and then returned to minimum, eliciting a large dynamic range of pupil sizes (mean range 2.86 mm). The second was a digit span recall test in which the phone read a nine-digit sequence aloud; the participant attempted to memorize and repeat the digits, eliciting task-evoked cognitive load dilation (mean range 1.41 mm). The participant's right eye was simultaneously recorded by the clinical pupillometer while the left eye was recorded by the smartphone, using the well-established consensual pupil response as justification for bilateral measurement.

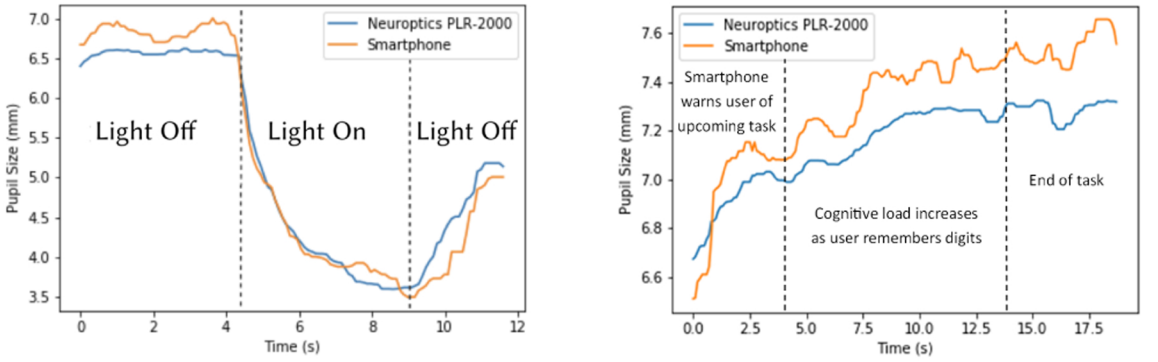
Results: Pupillary Light Reflex

Across eleven participants, the smartphone achieved a median absolute MAE of 0.27 mm and a median relative MAE of 3.52% on the PLR test. Mean absolute error was 0.39 mm ($\sigma = 0.29$ mm). Participant P4 was an outlier with absolute MAE more than three standard deviations above the mean, attributed to eye movement during the baseline distance calibration. Excluding P4, the mean absolute MAE was 0.27 mm. A Student's t-test comparing light-iris ($n = 5$) and dark-iris ($n = 6$) subgroups found no statistically significant difference in either absolute ($p = 0.55$) or relative error ($p = 0.36$), confirming

equity across iris colors.

Results: Digit Span Recall

The digit span test produced higher error, consistent with its lower signal-to-noise ratio. Median absolute MAE was 0.62 mm (9.37% relative). Mean absolute error was 0.96 mm. The elevated error is partially attributable to head movement during the cognitive task: participants tended to lean toward the phone as they concentrated on remembering digits, progressively decreasing the camera-to-eye distance and causing the single-baseline distance estimate to overestimate pupil dilation. This issue is discussed as a limitation in Section 2.7.



(a) Pupillary Light Reflex Example **(b) Digit Span Recall Trial Example**

Figure 2.3: Representative PLR and digit span traces: smartphone (NIR system) versus clinical pupillometer (Neuroptics PLR-2000) for one light-iris and one dark-iris participant. The smartphone and clinical device traces are closely aligned in both the PLR constriction and the task-evoked dilation amplitude, with the dark-iris participant (right column) showing no degradation in trace quality or accuracy relative to the light-iris participant (left column).

2.3.6 Limitations of the NIR Approach

The primary technical limitation is distance estimation sensitivity to eye movement during calibration. Because the distance is measured once at the start of each recording, any subsequent change in camera-to-eye distance propagates as a proportional error in all diameter measurements. Future work will explore continuous distance estimation using the iris-to-sclera ratio as a reference landmark.

The hardware availability constraint is also significant. The NIR facial recognition

Participant	Absolute MAE (mm)	Absolute Std. Dev.	Relative MAE (%)	Relative Std. Dev.	Range (Min - Max) (mm)
1	0.65	0.24	5.03	3.41	3.35 (3.27 - 6.62)
2	0.15	0.59	3.08	4.68	1.58 (4.55 - 6.13)
3	0.35	0.53	4.06	6.0	4.22 (2.90 - 7.13)
4	1.25	0.35	20.7	11.22	1.29 (6.07 - 7.36)
5	0.34	0.19	3.53	2.45	3.29 (4.45 - 7.74)
6	0.14	0.17	2.37	4.25	2.37 (4.29 - 6.66)
7	0.25	0.32	4.30	3.90	3.56 (3.1 - 6.66)
8	0.27	0.26	4.49	4.24	2.33 (3.88 - 6.21)
9	0.12	0.13	2.07	2.29	2.73 (3.94 - 6.67)
10	0.11	0.14	2.36	3.80	2.09 (3.75 - 5.84)
11	0.64	0.32	3.27	4.50	4.68 (2.34 - 7.01)
Mean	0.39	0.29	5.02	4.61	2.86 (3.87 - 6.73)
Median	0.27	0.26	3.52	4.24	2.73 (3.88 - 6.66)

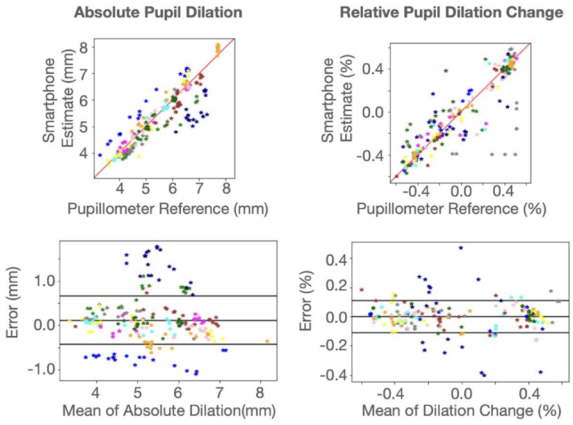


Figure 2.4: NIR smartphone pupillometry validation. (Top row) Linear regression and Bland-Altman plots for absolute pupil diameter accuracy on the PLR test across all 11 participants. (Bottom row) Regression and Bland-Altman plots for relative pupil change on the digit span recall test. Bland-Altman plots show limits of agreement (dashed lines) and mean bias (solid line). Error bars represent participant-level MAE.

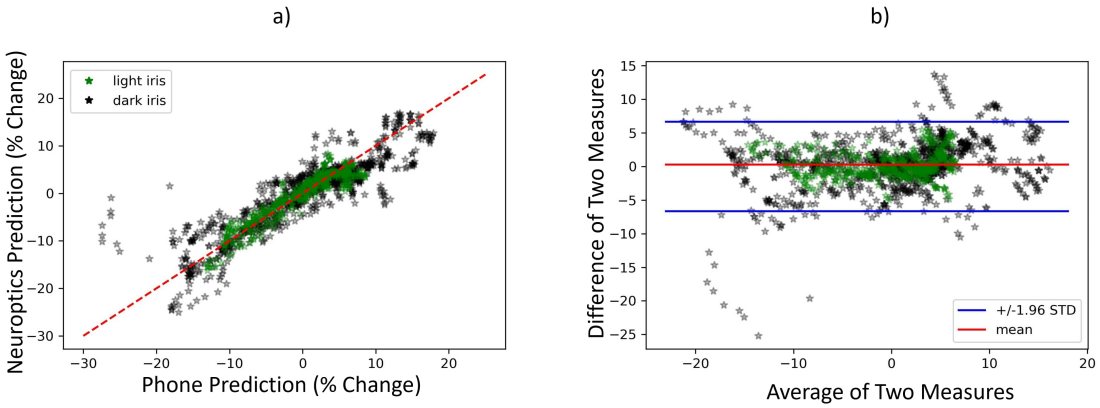


Figure 2.5: Additional regression and Bland-Altman analysis for the NIR pupillometry validation. These plots provide per-participant detail on the agreement between smartphone and clinical pupillometer measurements, supplementing the aggregate accuracy metrics reported in the main validation results.

camera has been phased out of most major smartphone lines, removed from the Pixel 5 onward, never exposed on iPhones. This limits the deployable population. The far-red approach in Section 2.4 was developed specifically to address this constraint.

2.4 Far-Red Spectrum Pupillometry for Universal RGB Cameras

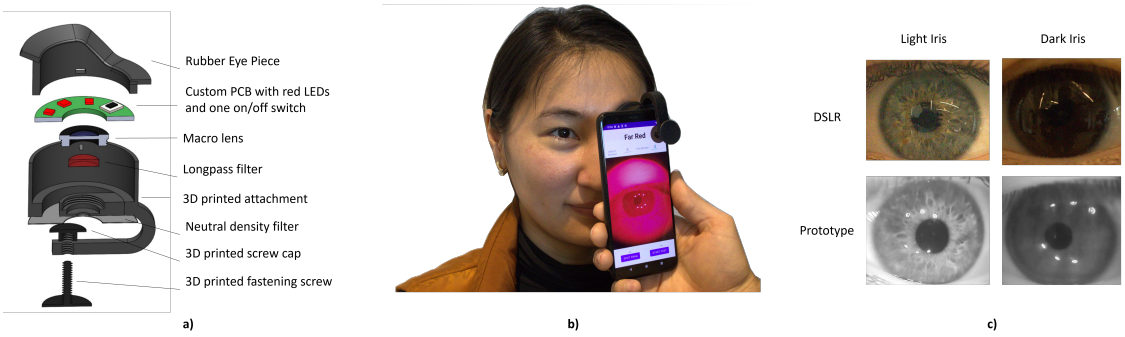


Figure 2.6: Overview of the far-red smartphone pupillometry system. The 630 nm longpass filter and 625 nm LED illumination restrict imaging to the far-red spectral window, where melanin absorption is substantially reduced, enabling high iris-to-pupil contrast on any RGB smartphone camera regardless of iris color.

2.4.1 Concept

The iris color problem in RGB cameras arises from broadband illumination that includes wavelengths at which melanin absorbs strongly. The NIR approach solves this by imaging outside the visible spectrum. The far-red approach solves it by imaging in the portion of the visible spectrum where melanin absorption is lowest [23].

The melanin absorption spectrum decreases monotonically across the visible range. In the far-red region (approximately 630–700 nm), absorption is substantially lower than in the green or blue regions. At these wavelengths, melanin in even the darkest brown iris reflects a meaningful fraction of incident light, creating measurable pupil-iris contrast. Filtering the illumination and camera image to the 630–700 nm band makes pupil-iris segmentation tractable on any RGB camera. Critically, this band falls within the passband of standard smartphone cameras (which are sensitive to approximately 700 nm before the built-in infrared cut filter attenuates) and within the red channel response of every major

smartphone [10].

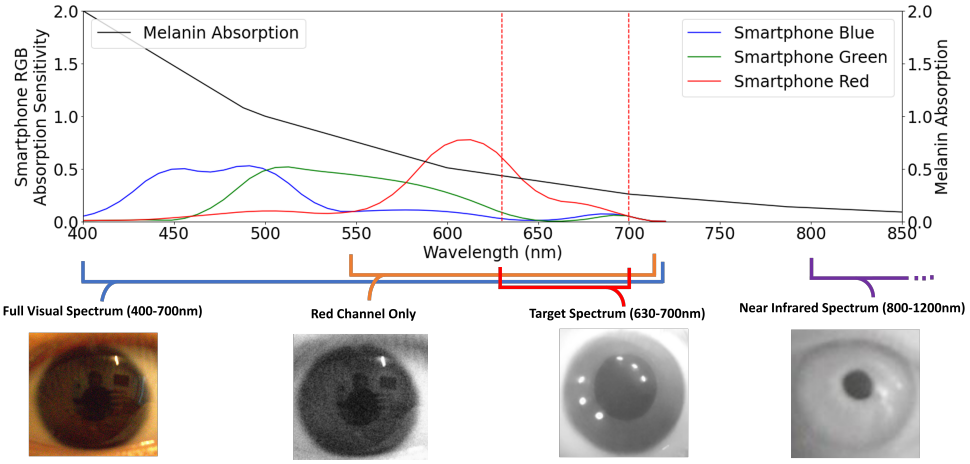


Figure 2.7: Melanin absorption spectrum across 400–900 nm. Absorption decreases monotonically across the visible range. The 630–700 nm far-red window (shaded) falls in a region of substantially reduced melanin absorption relative to the green and blue bands, making the iris appear lighter at these wavelengths and creating measurable pupil-iris contrast regardless of melanin content. At wavelengths above approximately 700 nm, the smartphone camera's built-in infrared cut filter attenuates the signal, defining the upper bound of the usable window.

2.4.2 Prototype Design

The far-red system is implemented as a compact smartphone attachment containing four components housed in a 3D-printed circular casing: a 630 nm longpass filter (Edmund Optics SCHOTT RG-630, 12.5 mm diameter), a custom PCB with six 625 nm LEDs (Marubeni 625SMT, spectral half-width 25 nm), a 15× macro lens, and a rubber eye cup. The attachment connects to any smartphone via a 3D-printed fastening screw with adjustable positioning to accommodate different camera locations across phone models.

The 630 nm longpass filter sits immediately in front of the smartphone camera, blocking all light below 630 nm. The smartphone's built-in infrared cut filter attenuates wavelengths above approximately 700 nm. The combined effect restricts imaging to the 630–700 nm far-red window. The six 625 nm LEDs create a controlled far-red illumination environment independent of ambient lighting. A rubber eye cup seals the attachment against the user's face, excluding ambient light and making the measurement environment-independent. A neutral density filter over the smartphone flashlight reduces its intensity to a

safe level for use as the PLR stimulus.

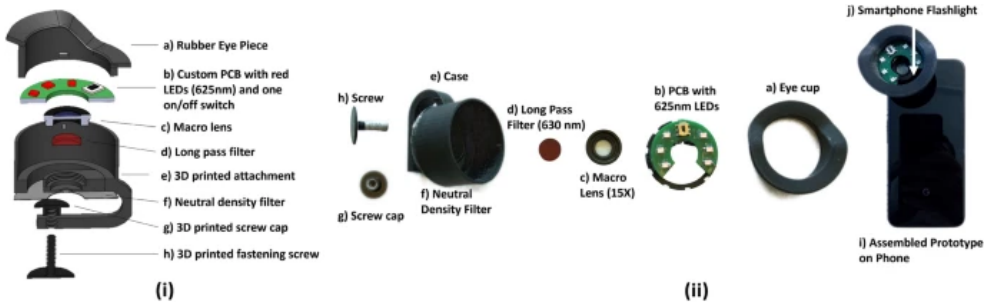


Figure 2.8: Far-red pupillometry prototype. (Left) Component diagram showing the 630 nm longpass filter, custom 625 nm LED PCB, 15× macro lens, and rubber eye cup housed in the 3D-printed casing. (Right) Photograph of the assembled attachment mounted on a smartphone, with the adjustable fastening screw positioned over the front camera.

2.4.3 Cross-Device Contrast Study

To characterize the pupil-iris contrast improvement across device models and iris colors, eleven participants were imaged on four smartphones (Google Pixel 4, Apple iPhone 12, Samsung Galaxy A53, Motorola Moto G Power 2022) with and without the far-red prototype. Contrast was quantified as the difference in mean pixel intensity between 15×15 pixel regions centered on the pupil and the iris, using the red channel. The contrast percent increase was:

$$\Delta C = \frac{C_{\text{far-red}} - C_{\text{full}}}{|C_{\text{full}}|} \tag{2.4}$$

The mean contrast increase across all participants and all smartphone models was 451%. For the seven dark brown iris participants, the mean increase was 624%, reflecting the disproportionate benefit for the population most poorly served by standard RGB imaging. The improvement was consistent across all four device models.

2.4.4 Accuracy Validation

The far-red prototype was validated against the Neuroptics PLR-3000. The smartphone recorded the left eye while the clinical device simultaneously recorded the right eye across eleven participants ($N = 12$ recruited; one excluded due to device error). The PLR

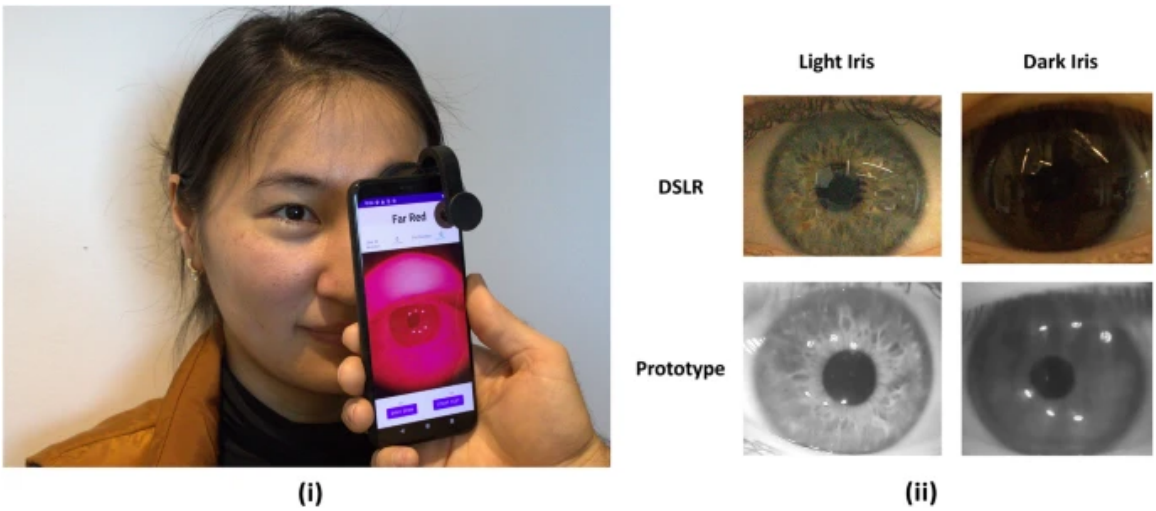


Figure 2.9: Side-by-side camera images of a dark iris under full-spectrum RGB illumination (left column) versus far-red illumination (right column), across four smartphone models (Google Pixel 4, Apple iPhone 12, Samsung Galaxy A53, Motorola Moto G Power 2022). Under full-spectrum illumination the iris-to-pupil boundary is nearly invisible; under far-red illumination the boundary becomes clearly defined, enabling accurate pupil segmentation regardless of device.

protocol consisted of a five-second recording: one second of no stimulus, one second of flashlight activation, three seconds of no stimulus. Far-red LED illumination was continuous throughout.

Comparison was performed using percent change from mean pupil size, a relative measurement appropriate for this prototype, which does not estimate absolute diameter. Across eleven participants, the system achieved a mean Pearson correlation of 0.93, a mean absolute error of 2.4%, and a standard deviation of 1.3%. Bland-Altman analysis showed no systematic bias. These results improve upon prior state-of-the-art smartphone pupillometry while extending performance to dark irises across four manufacturers [10].

2.4.5 Limitations of the Far-Red Approach

The far-red prototype measures relative, not absolute, pupil size. Adding absolute measurement would require a distance estimation component, either a fixed reference target in the eye cup or a stereo-vision approach analogous to the NIR system. The prototype also requires a physical attachment. As discussed in Chapter 4, the postable paradigm for

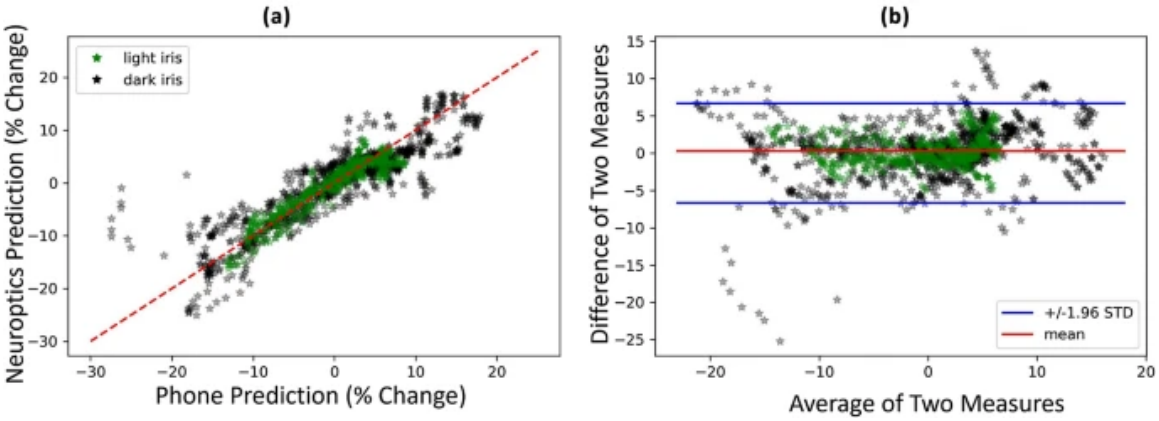


Figure 2.10: Far-red pupillometry accuracy validation against the Neuroptics PLR-3000. (Left) Linear regression of smartphone-measured percent pupil change versus clinical pupillometer measurement, across eleven participants. (Right) Bland-Altman plot showing the distribution of differences; the solid line indicates mean bias and dashed lines indicate 95% limits of agreement. No systematic bias is present. Mean Pearson correlation was 0.93; mean absolute error was 2.4%.

distributing such attachments, mailing a flat, self-assembled accessory to users, is directly applicable here.

2.5 Usability Study: Deployment to Older Adults

2.5.1 Motivation and Design

The populations most relevant to at-home pupillometry research are often the populations with the least experience using smartphones. A system that achieves high accuracy with young, technically experienced users in a laboratory has not solved the access problem if it fails for its intended population. To test real-world deployability, the NIR smartphone pupillometer was deployed to fifteen older adults (age 55+) recruited from an ongoing Alzheimer's cognition study at UC San Diego, approved by the Institutional Review Board (Project #210437X).

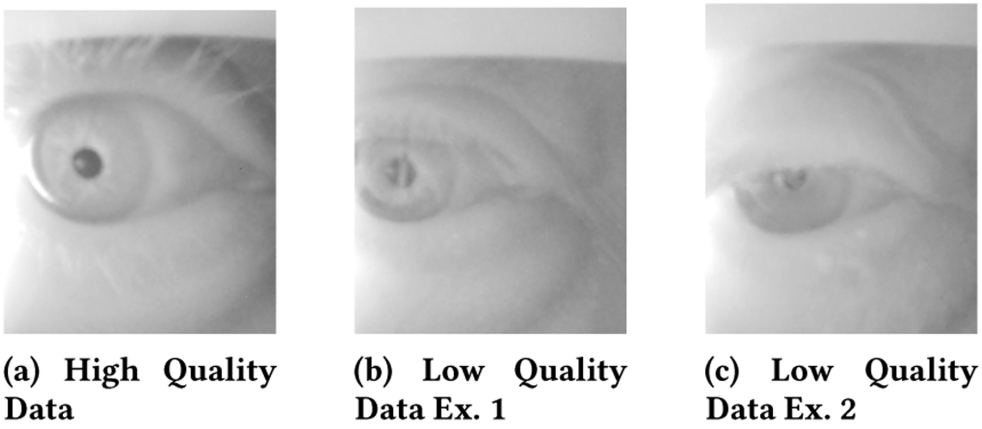
A Google Pixel 4 with the deployment application and a 3D-printed scope attachment was delivered to each participant's home. Each participant received a 20–30 minute video call instruction session with a researcher who had no control over the device. The testing application provided audio guidance through the digit span test. No in-person support was

provided.

2.5.2 Results

All fifteen participants successfully positioned the phone at their eye and completed at least one test. Eight produced data rated at least 80% usable, two produced data at least 50% usable, and five produced mostly unusable data. Usability ratings were based on the fraction of frames in which the pupil was clearly visible and unobstructed.

The most significant data quality problem was eyelid and eyelash occlusion of the pupil, which was not encountered in any laboratory validation study and was not anticipated in the initial design. The most common interaction problem was smartphone inexperience ($n = 7$), with the most difficult task being tap versus press interaction. Confusion about scope attachment ($n = 4$) and instruction misunderstanding ($n = 4$) were also common.



(a) High Quality Data **(b) Low Quality Data Ex. 1** **(c) Low Quality Data Ex. 2**

Figure 2.11: Representative frames from the older adult usability study. (Left) A high-quality frame from a participant who positioned the scope correctly, showing a clearly visible pupil suitable for automated detection. (Right) A frame exhibiting the most common failure mode: eyelash and eyelid occlusion of the pupil, which occurred in multiple participants and was not encountered in any laboratory validation session.

2.5.3 Design Implications

The eyelid occlusion finding motivates a scope attachment modification that changes the imaging angle: pointing the camera slightly upward from below the eye, achieved by inserting the phone upside down into the scope, reduces occlusion by capturing the lower

portion of the pupil. The smartphone inexperience findings informed simplified application design principles carried forward into the blood pressure application in Chapter 3: larger buttons, single-action screens, and explicit verbal instructions with built-in pauses.

2.6 Synthesis: Toward a Universal Smartphone Pupillometer

The NIR and far-red systems presented in this chapter address complementary aspects of the smartphone pupillometry problem. The NIR system achieves absolute pupil diameter measurements validated against a clinical gold standard, without any attachment, on devices that expose the NIR facial recognition camera. The far-red system achieves relative measurements with clinical-grade accuracy on any RGB smartphone, using a low-cost attachment compatible with every major device brand.

Together they define a practical deployment roadmap. For controlled research or clinical settings requiring absolute measurements and controlled device selection, the NIR approach is appropriate. For population-scale deployment where device uniformity cannot be assumed, the far-red approach provides clinically meaningful measurements on whatever smartphone the participant already owns.

Translational activity includes: a non-exclusive patent on far-red RGB pupillometry (licensed to Vitamin Greens, May 2024), the co-founding of Billion Labs Inc. to pursue FDA clearance, and the Best Demo Award at UbiComp 2025 for the ultra-low-cost pupillometry implementation.

2.7 Limitations and Future Work

Both validation studies used $N = 11$ participants, sufficient for proof-of-concept but not for regulatory submission. Larger, multi-site, demographically diverse studies are needed.

The MCI application analyzed an existing clinical dataset rather than smartphone-deployed data. Establishing the test-retest reliability of the smartphone systems, and validat-

ing them in a longitudinal study with known MCI status participants, are necessary future steps.

The NIR approach is limited to devices with accessible NIR cameras, a population that has been shrinking. Long-term scalability depends on hardware trends beyond the control of this research program.

Future work should also address eyelid and eyelash occlusion in older adults through automated occlusion detection and real-time quality assessment, and should formally validate the modified downward imaging angle for ptosis compensation described in Section 2.5.3.

Chapter 3

Cardiovascular Monitoring via Oscillometric Blood Pressure

3.1 The Hypertension Access Problem

Hypertension, persistently elevated blood pressure (BP), is the leading preventable cause of premature death and disability worldwide. Its downstream consequences include stroke, coronary artery disease, heart failure, chronic kidney disease, and accelerated cognitive decline. By 2025, hypertension is projected to affect 1.6 billion adults, a 60% increase from estimates at the turn of the century [1, 12]. Yet despite decades of awareness campaigns and effective pharmacological treatments, an estimated 46% of adults living with hypertension remain undiagnosed [12]. Undiagnosed hypertension cannot be treated.

The diagnostic gap is not primarily scientific. The blood pressure measurement itself, the oscillometric method employed by virtually all automated BP cuffs, has been well validated and standardized for decades. The AHA and AMA jointly recommend self-measured BP monitoring at home as a means of improving the diagnosis and management of hypertension [24]. The barrier is access. Traditional oscillometric arm cuffs cost between twenty and one hundred dollars, require correct cuff sizing, and depend on a degree of physical ability that may be limited in older or disabled individuals. More importantly, they are a separate device that must be purchased, stored, and actively used. The rate of sustained at-home BP monitoring in hypertensive populations remains low, particularly in lower-income communities where the prevalence of undiagnosed hypertension is highest.

The smartphone offers a compelling alternative platform. A downloadable BP monitoring application would require no additional purchase, no separate device, and no

specialized knowledge beyond what is needed to download and launch an app. It would be available to anyone who already owns a smartphone, a category that includes the majority of the global adult population, including many low- and middle-income countries where hypertension burden is high. If the measurement is accurate, calibration-free, and usable without instruction, the potential for widespread hypertension screening is substantial.

Realizing this potential, however, means rejecting the pulse-transit-time (PTT) and pulse-wave-analysis (PWA) methods common to consumer wearables: because these relative methods measure only the change in BP from an established baseline, they cannot produce a meaningful reading without prior cuff calibration, and are better characterized as BP change monitors than BP monitors. This chapter instead presents the first absolute oscillometric BP measurement on an unmodified smartphone, requiring no hardware modification, no physical attachment, and no individual calibration against a reference cuff measurement. The central technical contribution is Vibrometric Force Estimation (VFE): a method for measuring the force applied by a user's fingertip to the smartphone camera using only the smartphone's vibration motor and inertial measurement unit (IMU), two components present in essentially every modern smartphone [11]. Combined with camera photoplethysmography (PPG) as the blood volume sensing modality, VFE enables a complete oscillometric BP measurement pipeline in a downloadable software application. The chapter also contextualizes the VFE-based approach by describing BPClip, an alternative mechanical implementation of the same oscillometric principle, whose convergent accuracy strengthens the case for the oscillometric approach on smartphones [12].

3.2 Background: Oscillometry and the Two-Measurement Problem

3.2.1 The Oscillometric Method

The oscillometric method is the operating principle behind virtually all automated BP cuffs used in clinical and home settings. During a measurement, an inflatable cuff is placed around the arm and inflated above systolic BP, occluding the brachial artery. As

the cuff slowly deflates, blood flows through the artery intermittently, creating pressure oscillations within the cuff. The amplitude of these oscillations as a function of applied cuff pressure traces a bell-shaped curve known as the oscillogram. The peak of the oscillogram corresponds approximately to mean arterial pressure (MAP), and systolic and diastolic BP are estimated from the oscillogram shape using empirically derived algorithms [6].

The same principle can, in theory, be applied at a finger artery. The digital arteries of the index finger, specifically the transverse palmar arch branches accessible beneath the fingertip and the region overlying the fingernail bed, can be partially occluded by pressing a finger against a surface. If the applied force can be measured and converted to pressure, and if the blood volume oscillations at the artery can be simultaneously measured, the oscillometric method applies without a cuff. This is the principle underlying the finger oscillometric BP measurements described in this chapter.

The measurement therefore decomposes into two independent sub-problems:

1. *Blood volume sensing:* measuring the pulsatile variation in blood volume at the digital artery as a function of applied pressure. This is accomplished using camera PPG.

2. *Force sensing:* measuring the force applied by the user's fingertip to the smartphone camera, converted to pressure using an estimate of the fingertip contact area. This is the primary novel contribution of this chapter, accomplished via VFE.

3.2.2 Prior Work on Smartphone Blood Pressure

Prior approaches to smartphone BP measurement fall into two broad categories: relative methods and absolute methods. Relative methods based on pulse transit time (PTT) or pulse wave analysis (PWA) measure the change in BP from an established baseline, but cannot determine absolute BP without prior cuff calibration [25]. An individual who has never used a BP cuff cannot obtain a meaningful reading from a PTT-based system. These approaches are better characterized as BP change monitors than BP monitors, and they are not the approach pursued here.

Prior absolute methods based on finger oscillometry have been demonstrated, but all require hardware not universally present. Chandrasekhar *et al.* [26] demonstrated oscillometric finger pressing on an iPhone with a force-sensitive screen, a feature introduced in the iPhone 6S and subsequently removed. Chandrasekhar *et al.* [6] and Xuan *et al.* [12] demonstrated accurate oscillometric measurements using custom mechanical attachments that encode applied force as a measurable physical quantity. No prior work has demonstrated absolute oscillometric BP measurement on a completely unmodified smartphone without any special hardware component or attachment. The VibroBP system presented in this chapter is the first to do so.

3.2.3 Camera Photoplethysmography for Blood Volume Sensing

The PPG signal is measured using the smartphone's front camera with the display screen as a white light illumination source. With the user's index finger placed over the front camera, the screen illumination passes through the fingertip, and the camera records the variation in transmitted light intensity over time. Blood volume increases during systole, increasing light absorption and decreasing the camera signal; blood volume decreases during diastole, reducing absorption and increasing the signal.

As established in Appendix A, reliable PPG measurement across devices requires addressing the nonlinear tone mapping and sensor-specific zero light offset (ZLO) of the smartphone camera pipeline. For the oscillometric BP measurement, the critical requirement is signal consistency and low noise across the full range of applied forces rather than absolute amplitude calibration, since the oscillogram is constructed from pulse amplitude relative to its own maximum. The calibration approach in Appendix A is applied by setting the camera to linear tone mapping mode and locking white balance, exposure time, and frame rate, eliminating dynamic nonlinearities that would otherwise introduce pressure-dependent signal artifacts.

Skin tone considerations are also addressed. Individuals with darker skin tones have

somewhat lower PPG signal amplitude due to melanin absorption. Using the white screen as a broadband illumination source, rather than the narrowband LED flashlight, and calibrating camera exposure maintain adequate signal quality across Fitzpatrick Types II through V, as confirmed empirically by the absence of a statistically significant skin tone effect on BP measurement performance (Section 3.5.5).

3.3 Vibrometric Force Estimation

3.3.1 Physical Principle

VFE rests on a well-characterized mechanical phenomenon: when a vibrating object has an external force applied to it, the amplitude of its vibration decreases in proportion to the applied force. The relationship between applied force and vibration amplitude depends on the mechanical properties of the vibrating system, mass, stiffness, and damping, which are fixed for a given phone model and vary predictably across models.

In the VFE context, the phone itself is the vibrating object, driven by its vibration motor at maximum intensity. The IMU, rigidly mounted to the phone chassis, records the resulting oscillatory motion. When the user presses their index finger onto the front camera with increasing force, the finger applies compressive force and damping to the chassis, attenuating the vibration amplitude recorded by the IMU in proportion to the applied force. As shown in Figure 3.1, as applied force increases from zero to approximately 7 N, the oscillation amplitude of the IMU's y-axis acceleration decreases monotonically. This monotonic relationship is the basis for force estimation.

3.3.2 Signal Processing

The raw IMU data from three accelerometer axes and three gyroscope axes contains both the high-frequency vibration signal from the motor and lower-frequency motion noise. Applied force is encoded primarily in the amplitude of the vibration oscillation, which manifests as the rolling standard deviation of the signal within a short time window. As

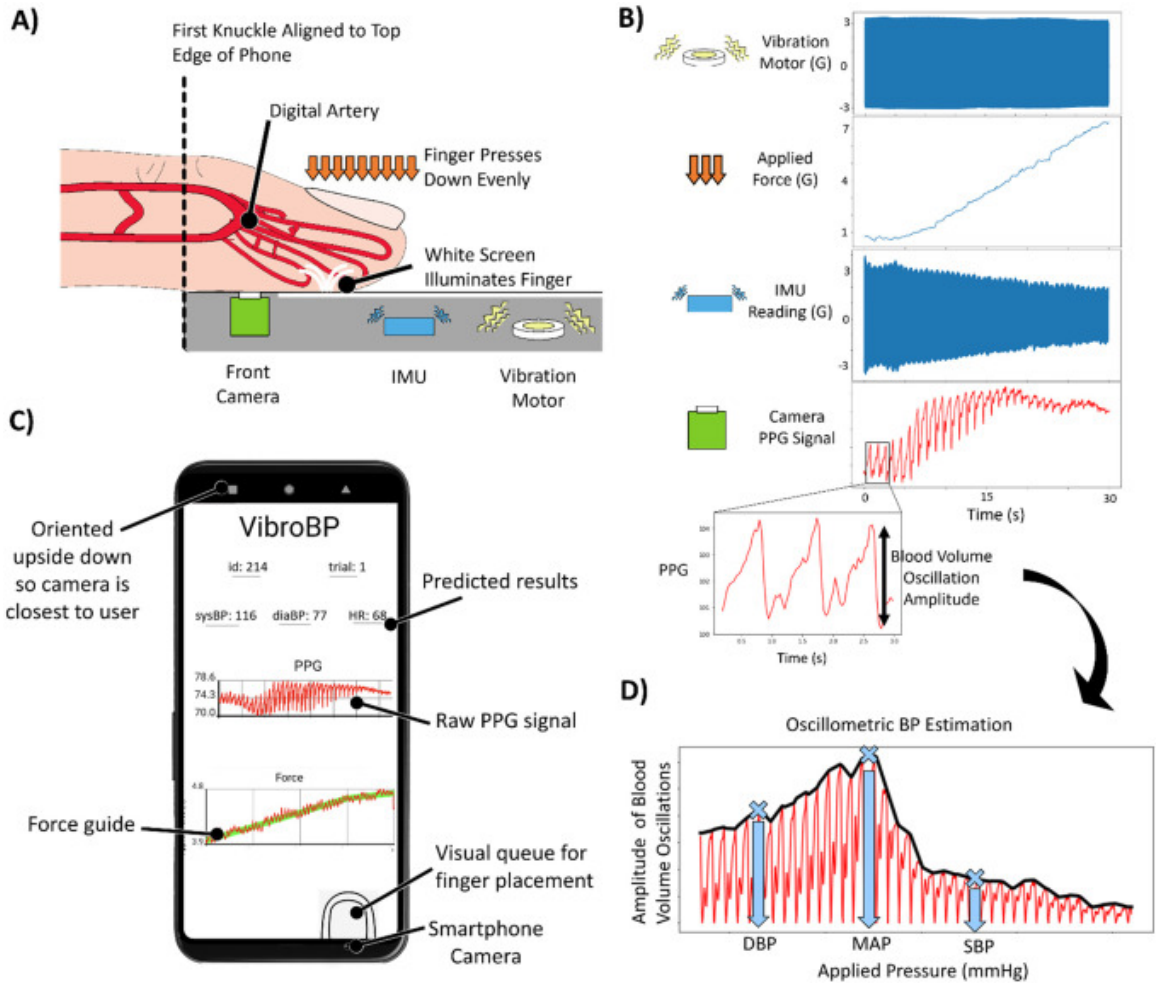


Figure 3.1: VibroBP system overview. (A) VFE schematic: the vibration motor drives the phone chassis, and fingertip pressure damps the resulting oscillation in proportion to applied force. (B) Representative IMU y-axis acceleration trace and simultaneously recorded reference force sensor trace during a ramp-up/ramp-down fingertip press, showing the inverse relationship between force and vibration amplitude. (C) Camera photoplethysmography (PPG) waveform recorded during the measurement. (D) Resulting oscillogram with DBP, MAP, and SBP feature points labeled. Adapted from Barry et al. *Scientific Reports* 2024, Fig. 1.

applied force increases and vibration amplitude decreases, the rolling standard deviation of the high-frequency vibration component decreases proportionally. The rolling standard deviation of each IMU axis thus serves as the feature vector for the force estimation model.

Features are selected by correlation with the reference force signal during calibration. The most informative features are typically the rolling standard deviation of the y-axis accelerometer, perpendicular to the phone face, in the direction of finger pressing, and one or more gyroscope axes sensitive to the rotational damping effect of the finger. Less informative axes are excluded to reduce noise.

3.3.3 Per-Device-Model Calibration

Because vibration motor characteristics and chassis resonance vary across phone models, VFE requires a per-model calibration procedure performed by the application developer, not the end user. During calibration, five participants apply a continuous range of forces (0 to approximately 8 N) with their index finger onto a calibrated force sensor (SingleTact 15 mm, 4.5 N) positioned over the front camera, while the phone vibrates and the IMU records. Three 40-second repetitions per participant yield approximately 10 minutes of calibration data per phone model.

A regression model is then trained for each phone model. The model type is selected by the linearity of the force-to-IMU relationship for the specific device: multivariate linear regression when the relationship is approximately linear, gradient boosting when it is more complex. The trained model is embedded in the application, producing real-time force predictions from the IMU signal during a user's BP measurement without any user-facing calibration.

3.3.4 Cross-Device Reliability

The calibration and force estimation procedure was performed on three smartphones representing different manufacturers, price points, and hardware configurations: the Google Pixel 4, Samsung Galaxy A53, and Motorola Moto G Power 2022. Performance was

evaluated using a 10% holdout from the calibration dataset.

Force estimation performance was consistent across all three devices. The Pixel 4 achieved MAE 0.77 N, standard deviation 0.90 N, bias -0.46 N, and $r = 0.92$. The Galaxy A53 achieved MAE 0.55 N, standard deviation 0.64 N, bias -0.18 N, and $r = 0.95$. The Moto G Power achieved MAE 0.39 N, standard deviation 0.53 N, bias 0.01 N, and $r = 0.96$. The average correlation coefficient across all three phones was 0.92, confirming that the vibration damping phenomenon is consistent and predictable across different phone designs.

3.3.5 Intra-Device Reliability

To assess the stability of a single per-model calibration across multiple units, the model trained on one Pixel 4 was tested on four additional Pixel 4 units: one new device, two refurbished, and one refurbished with heavy use and wear. MAEs were 0.97 N, 0.81 N, 0.71 N, and 1.45 N, respectively.

The heavily worn device performed notably worse (MAE 1.45 N), likely because vibration motor amplitude degrades with wear, shifting the force-to-IMU relationship. This is not a fundamental limitation: a baseline amplitude measurement at the start of each session could detect and correct for motor wear by shifting the model's operating point. For devices in normal operating condition, the per-model calibration transfers reliably across units, confirming that end users require no calibration.

3.4 Oscillometric BP Measurement Pipeline

3.4.1 Measurement Protocol

The BP measurement is performed as follows. The user holds the smartphone upside down so that the front camera is at the bottom edge, nearest to the hand. The user places their index finger on the front camera with the first knuckle aligned to the top edge of the phone. The vibration motor begins vibrating at maximum intensity, and the user is instructed to slowly and continuously increase the force applied by the index finger over

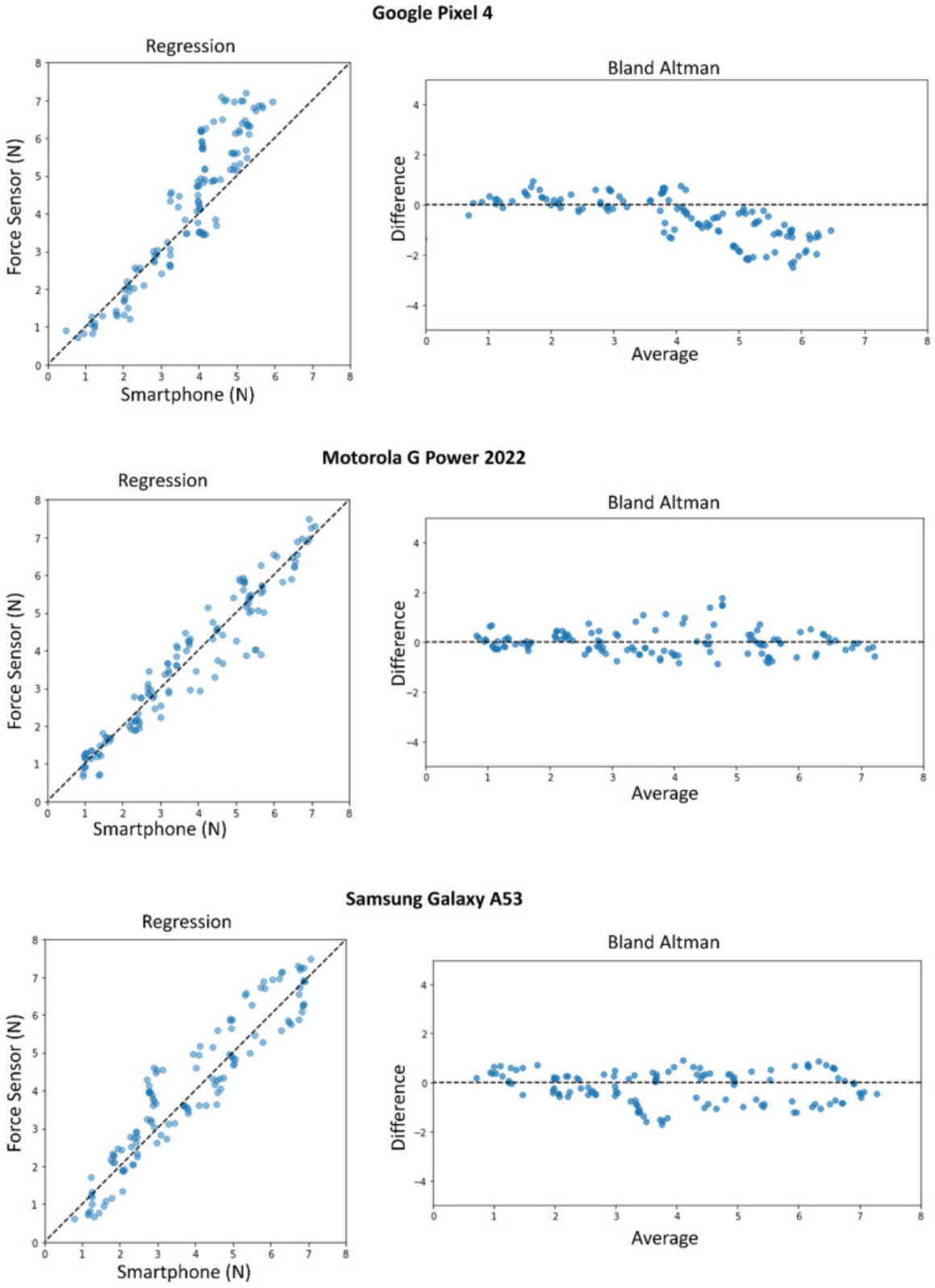


Figure 3.2: VFE cross-device force estimation validation. Regression plots (top row) and Bland-Altman plots (bottom row) for the Google Pixel 4, Samsung Galaxy A53, and Motorola Moto G Power 2022. Dashed lines in Bland-Altman plots indicate 95% limits of agreement. Mean Pearson correlation across all three devices was 0.92. Adapted from Barry et al. *Scientific Reports* 2024, Fig. 4.

approximately 30 seconds, following a visual force guide displayed on the screen. The white screen simultaneously illuminates the fingertip for PPG, the front camera records PPG, and the IMU records the vibration signal.

The visual force guide is a critical usability feature. Without real-time feedback, users tend to apply force too quickly or to maintain discrete levels rather than a smooth continuous ramp. The guide displays the current estimated force as a colored bar alongside a target ramp trajectory, enabling the user to match the desired force profile.

3.4.2 Signal Processing and Oscillogram Construction

The raw PPG signal is preprocessed to remove baseline drift using a high-pass filter with a cutoff below the heart rate range, then bandpass filtered to isolate the cardiac pulse component. For each cardiac cycle, the pulse amplitude is extracted as the peak-to-trough prominence of the filtered PPG waveform, producing a sequence of amplitudes, one per heartbeat, over the measurement.

The applied force from VFE is converted to estimated applied pressure at the digital artery using a fixed estimate of fingertip contact area (approximately 1.5 cm^2). Plotting pulse amplitude against applied pressure for each heartbeat produces the oscillogram. The oscillogram is fit to a skewed Gaussian to smooth the empirical data and enable robust feature extraction. As shown in Figure 3.4, the oscillogram peak shifts rightward as systolic BP increases, reflecting the higher MAP in hypertensive individuals, a physiologically expected result that validates the measurement principle.

3.4.3 Blood Pressure Feature Extraction and Model

Eleven features are extracted from the fitted oscillogram, drawing on concepts previously identified as informative for oscillometric BP estimation [6]: the applied pressure at maximum pulse amplitude; maximum and minimum amplitude gradients of the oscillogram and the applied pressures at which they occur; the maximum oscillogram amplitude; pulse amplitudes at the maximum and minimum gradients; and the applied pressures at the

maximum of the low-pass filtered PPG signal and at the maximum gradient of the low-pass filtered PPG signal.

These eleven oscillometric features are the only inputs to the BP prediction model. Heart rate, pulse shape, demographic variables, and skin tone are deliberately excluded. Including demographic inputs risks encoding demographic biases and limits applicability to users who provide those inputs. The model demonstrates that oscillometric shape features alone are sufficient for BP estimation, consistent with the operating principle of commercial oscillometric cuffs.

A LASSO regression model [27] is trained separately for systolic and diastolic BP. LASSO regularization reduces overfitting in the moderate-sized dataset and yields interpretable sparse feature weights. The model is trained and evaluated using leave-one-participant-out (LOO) cross-validation.

3.5 Blood Pressure Proof-of-Concept Study

3.5.1 Participants and Protocol

$N = 30$ participants were recruited under UC San Diego IRB protocol 804668. The population was 20% White, 67% Asian, and 13% Hispanic, with 50% male and 50% female. Mean age was 24.7 years ($\sigma = 4.3$). Mean resting systolic BP was 103 mmHg (range 89–132 mmHg) and mean resting diastolic BP was 70 mmHg (range 60–84 mmHg). Fitzpatrick skin types ranged from II to V. Of $N = 30$ recruited, $N = 6$ were excluded for the criteria described below, leaving $N = 24$ in the analysis.

Each participant performed one practice measurement (excluded from analysis) followed by three smartphone BP measurements. An FDA-approved oscillometric arm cuff (Omron BP7350) provided reference measurements in the sequence: cuff, smartphone $\times 3$, cuff. If both cuff readings were successful, they were averaged as the reference for all three smartphone measurements. $N = 23$ participants optionally completed an exercise phase (wall sit for approximately one minute), during which they simultaneously measured BP

with the cuff on one arm and the smartphone on the other, providing exercise-elevated BP data.

3.5.2 Measurement Exclusion

Individual measurements were excluded based on: cuff reference BP exceeding 155 mmHg (due to known finger-arm divergence at high exercise-induced pressures), undetectable or noisy pulse in the PPG signal, insufficient or inconsistently applied force, poor skewed Gaussian fit to the oscillogram, PPG saturation, and unchanging PPG pulse prominence. Of 133 total measurements (90 resting, 43 exercise), 65 were excluded, leaving 68 measurements from $N = 24$ participants (53 resting, 15 exercise). The high exclusion rate for exercise measurements (28 of 43) reflects the difficulty of performing the force application protocol during simultaneous physical exertion.

3.5.3 Blood Pressure Results

Systolic Blood Pressure

The LOO cross-validated LASSO model achieved MAE 9.21 mmHg for systolic BP, with mean error 0.09 ± 11.6 mmHg and Pearson $r = 0.65$. The reference systolic BP ranged from 89 to 147 mmHg with mean 108.6 ± 15.0 mmHg.

Diastolic Blood Pressure

For diastolic BP, MAE was 7.77 mmHg, mean error -0.24 ± 9.97 mmHg, and Pearson $r = 0.43$. Reference diastolic BP ranged from 60 to 107 mmHg with mean $74.6 \pm$ 10.7 mmHg.

Hypertension Screening

Classifying measurements as normotensive versus hypertensive (systolic BP ≥ 130 mmHg) per 2017 ACC/AHA guidelines [28], the smartphone system achieved AUC 0.92, sensitivity 0.86, and specificity 0.89 at the optimal operating threshold. This screening performance is sufficient to reliably flag individuals who warrant follow-up measurement

with a clinical device.

The limited diastolic correlation ($r = 0.43$) is a known limitation. Isolated diastolic hypertension, elevated diastolic BP with normal systolic BP, affects only approximately 3.2% of the hypertensive population [11]. For population hypertension screening, systolic BP alone correctly identifies hypertensive status in the vast majority of individuals. The diastolic limitation does not substantially diminish the clinical utility of the system for its primary screening application.

3.5.4 Oscillogram Group Analysis

As a physiological validation of the oscillometric mechanism, the mean oscillogram was computed separately for four systolic BP categories: below 100 mmHg, 100–120 mmHg, 120–140 mmHg, and above 140 mmHg. The oscillogram peak, which approximates MAP, shifts rightward as systolic BP increases, consistent with the expected physiological relationship between MAP and systolic BP. This group-level result confirms that the smartphone measurement captures genuine physiological variation rather than noise, even when individual-level prediction error is substantial.

3.5.5 Equity Analysis

A one-way ANOVA tested for statistically significant differences in BP measurement performance across ethnic groups and skin tone groups. For ethnicity (White, Hispanic, Asian), the group variation in systolic MAE was not statistically significant ($p = 0.22$). For Fitzpatrick skin type (II through V), the group variation was also not significant ($p = 0.86$). These results confirm that the measurement approach is equitable across the demographic groups represented in the study, consistent with the physics of the oscillometric method, which depends on mechanical force transmission and blood volume rather than optical tissue properties that vary with skin tone.

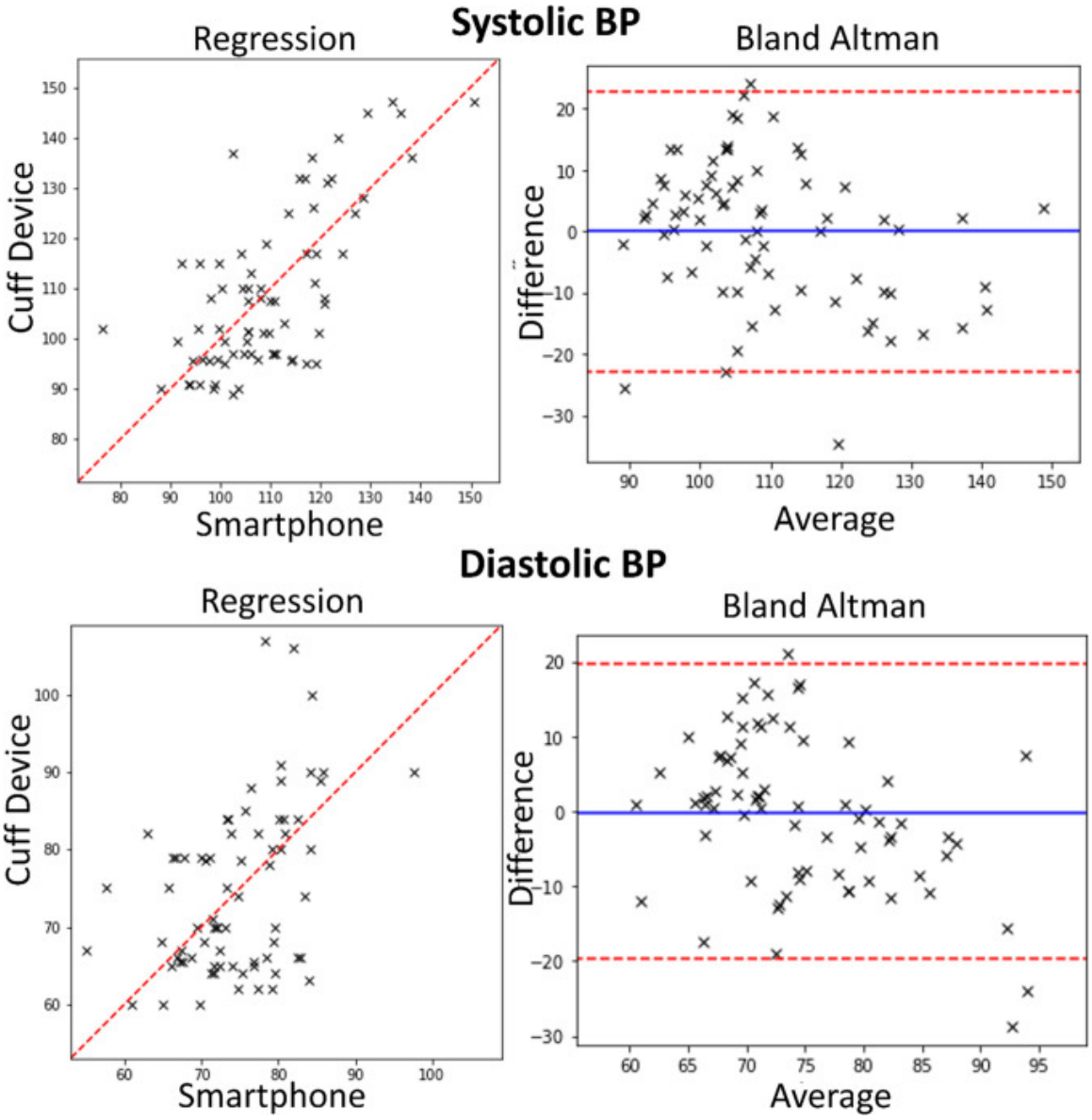


Figure 3.3: Blood pressure proof-of-concept validation. Regression plots (top row) and Bland-Altman plots (bottom row) for leave-one-out cross-validated systolic (left column) and diastolic (right column) BP prediction ($N = 24$ participants, 68 measurements). The systolic model achieved MAE 9.21 mmHg and $r = 0.65$; the diastolic model achieved MAE 7.77 mmHg and $r = 0.43$. Dashed lines in Bland-Altman plots indicate 95% limits of agreement. Adapted from Barry et al. *Scientific Reports* 2024, Fig. 2.

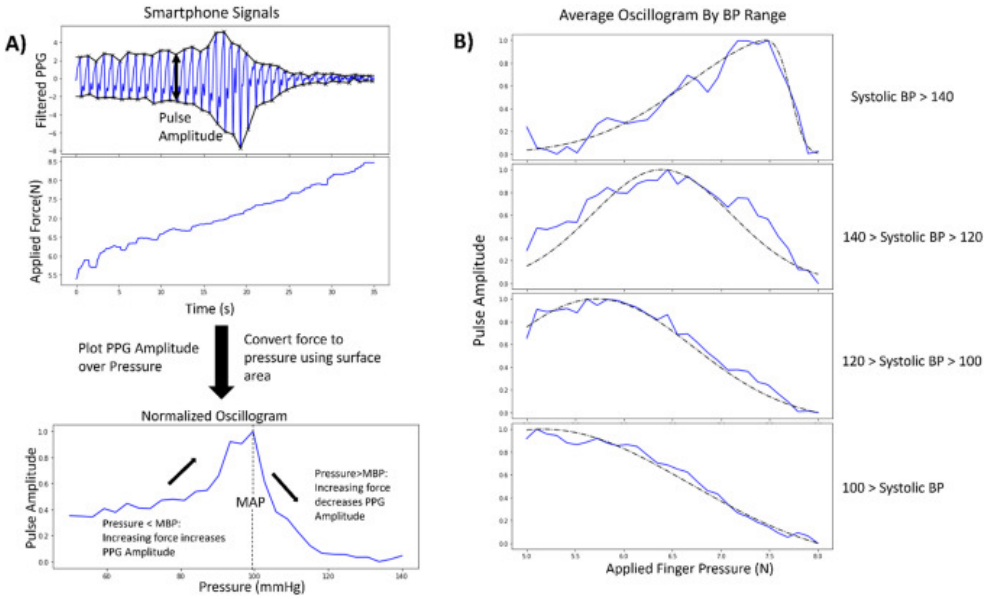


Figure 3.4: Mean oscillogram by systolic BP category (<100, 100–120, 120–140, >140 mmHg). Each curve shows pulse amplitude plotted against applied fingertip pressure. The oscillogram peak, which approximates mean arterial pressure, shifts rightward with increasing systolic BP, confirming that the smartphone measurement captures genuine physiological variation consistent with the oscillometric principle. Adapted from Barry et al. *Scientific Reports* 2024, Fig. 3B.

3.6 BPClip: Contextualizing an Alternative Implementation

3.6.1 Overview and Mechanism

BPClip [12] is an ultra-low-cost smartphone attachment that solves the two-measurement oscillometric problem through a purely mechanical approach. Rather than estimating force from an IMU signal, BPClip encodes applied force as the projection size of a movable pinhole imaged by the smartphone camera. As the user presses their index finger on the clip, a spring-loaded mechanism moves the pinhole closer to the camera lens, increasing the size of its projection on the camera sensor. The camera simultaneously measures projection size as a pressure proxy and brightness variation of the pinhole as the PPG signal, requiring no sensing modality beyond the camera and flashlight. The clip is entirely mechanical, no electronics, no battery, no IMU, with a material cost of $0.80 USD at a production volume of 1,000 units [12].

The iterative design process explored compressible urethane foam (rejected due

to hysteresis), 3D-printed rubber (rejected due to batch inconsistency), and a 3D-printed cantilever (rejected due to brittle fracture under cyclic loading) before converging on a metal spring mechanism, which satisfies the calibration, tolerance, and fatigue requirements for reliable repeated measurements [12]. Finger placement is enforced by physical flanges on either side of the finger slot, and the measurement protocol uses stepwise increasing force rather than a continuous ramp, since user testing showed discrete levels are easier to maintain consistently.

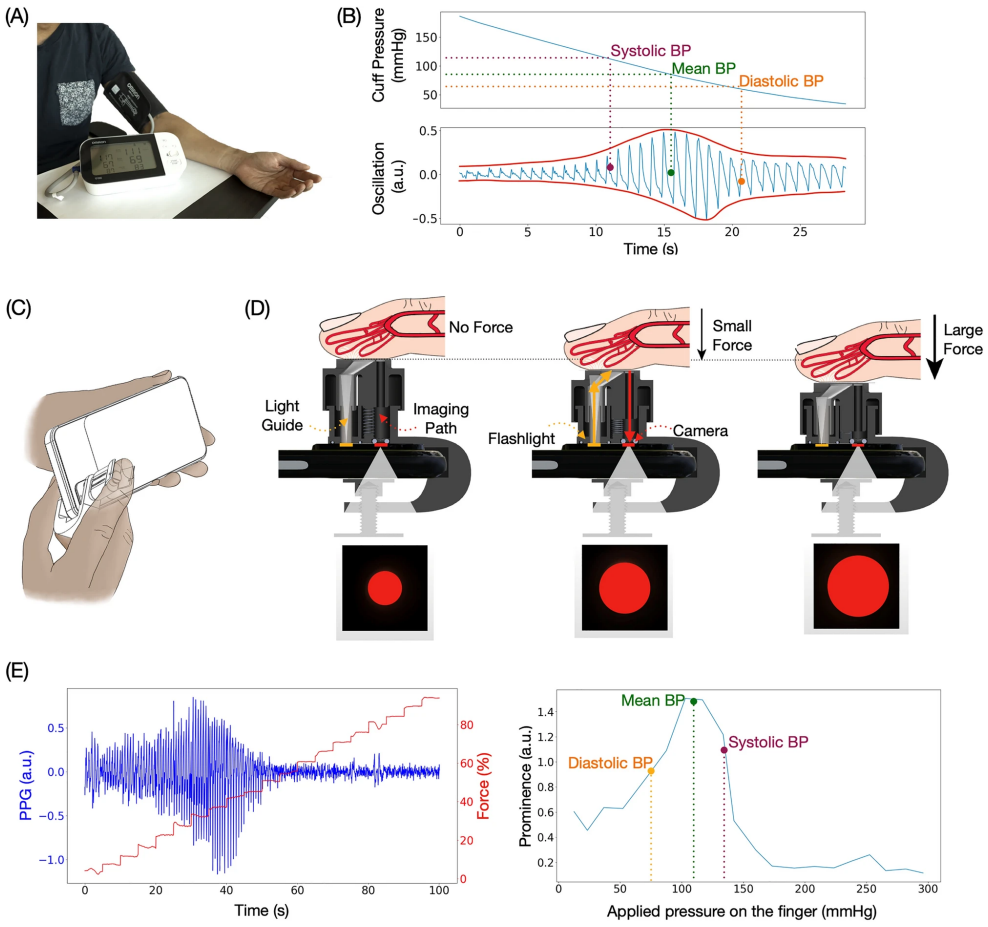


Figure 3.5: BPClip mechanism and design. (Left) Annotated diagram showing the metal spring, movable pinhole, finger slot with alignment flanges, and the optical path from pinhole to smartphone camera. As the user increases fingertip pressure, the spring compresses, moving the pinhole closer to the camera and enlarging its projection on the sensor. (Right) Photograph of the assembled BPClip attached to a smartphone, showing the clip's compact form factor and finger positioning. Adapted from Xuan et al. Scientific Reports 2023, Fig. 1.

3.6.2 Performance and Convergent Validation

BPClip was tested in a human subjects study of $N = 24$ participants (overlapping population with the VFE study), achieving MAE 8.72 mmHg for systolic BP and MAE 5.49 mmHg for diastolic BP. Subgroup analysis showed no significant performance difference across Asian (7.7 ± 4.2 mmHg systolic MAE), Hispanic (8.6 ± 5.2 mmHg), and White (10.2 ± 6.9 mmHg) participants.

The performance of BPClip and VibroBP are comparable: both achieve systolic MAE in the range of 8–10 mmHg and demonstrate no significant demographic performance difference. This convergence of two fundamentally different implementation strategies, software-only VFE via vibration damping versus passive mechanical force encoding via spring and pinhole, on similar accuracy levels is a meaningful result. It suggests that the accuracy of finger oscillometric BP measurement on smartphones is determined primarily by the oscillometric method itself applied at the finger, rather than by the specific technique used to estimate force. The oscillometric approach is robust to implementation details, and improvements to the signal processing and BP model will likely benefit both systems.

VibroBP and BPClip are also complementary in their deployment profiles. VibroBP requires no hardware beyond an unmodified smartphone, it is a purely software solution deliverable as a downloadable app to existing devices, making it optimal for population-scale screening. BPClip requires a physical clip, but that clip is mechanically passive, electronically inert, and inexpensive enough to be distributed at scale, including via the postal system in the manner described for postable accessories in Chapter 4. The two approaches serve different deployment contexts rather than competing for the same one.

3.7 The VibroBP Application

The VibroBP Android application was developed to guide users through the BP measurement and collect the necessary data. Two primary user-facing features are provided during a measurement. First, a real-time PPG signal plot before and during the measure-

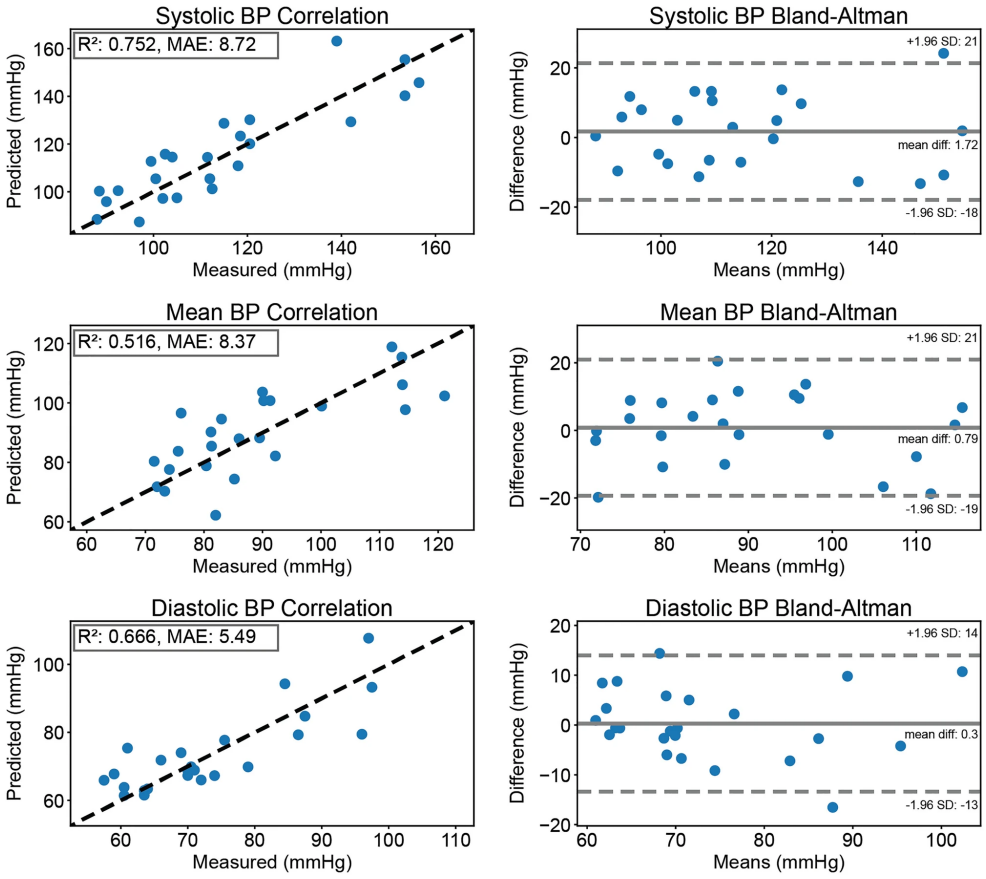


Figure 3.6: BPClip systolic BP validation. (Left) Regression plot of BPClip-measured versus reference systolic BP for $N = 24$ participants, with the identity line shown. (Right) Bland-Altman plot showing limits of agreement (dashed lines) and mean bias (solid line). BPClip achieved systolic MAE 8.72 mmHg, comparable to the software-only VibroBP system. Adapted from Xuan et al. Scientific Reports 2023, Fig. 2.

ment allows the user or researcher to verify adequate signal quality prior to initiating the measurement. Second, a real-time force guide during the measurement displays the current estimated applied force alongside a target ramp trajectory, enabling the user to match the desired force profile.

An optional PPG calibration step, implemented using the method of Xuan *et al.* [8], adjusts the camera exposure before the measurement to ensure the PPG signal is neither saturated nor too weak for the current user's finger optical properties. This step prevents the most common PPG signal quality failure mode without requiring any action from the user beyond briefly placing their finger on the camera.

The application supports both resting and exercise measurement modes. In exercise mode, the measurement is initiated during a wall sit to capture elevated BP. At the end of each measurement, the raw force and PPG signals are displayed within the application to allow the research staff to assess data quality.

3.8 Translational Pathway

The VibroBP proof-of-concept results establish feasibility but are not yet sufficient for clinical deployment. The path from the current system to a commercially available, FDA-cleared smartphone BP monitor is actively in progress on several fronts.

The system has been supported by an NIH POCTRN Program CAPCaT Innovation Award ($100k, 2024) and a Massachusetts AI and Technology Center A2Collective Pilot Award ($220k, 2023), both directed at developing oscillometric smartphone BP measurement toward clinical deployment. The system received the Best Demo Award at IEEE BSN 2025 and was a Stanford eWEAR Health Prize finalist ($1k award, 2025). A provisional patent on the oscillometric smartphone BP measurement method (attorney docket 09062-8487 US00) has been licensed to Billion Labs Inc., which was co-founded by the author and Professor Wang to pursue FDA clearance of this and related smartphone sensing technologies.

The regulatory pathway is as Software as a Medical Device (SaMD) under the

FDA's digital health framework. The current proof-of-concept study is not sufficient for a 510(k) submission, which requires validation against a cleared predicate device in a clinically representative population with an adequate BP range. Future studies are planned to expand to $N \geq 85$ participants spanning the full hypertensive spectrum, include older adults and individuals with relevant comorbidities, and test the system in unsupervised at-home conditions.

3.9 Limitations and Future Work

3.9.1 Measurement Performance

The systolic MAE of 9.21 mmHg falls short of the IEEE Standard for Wearable, Cuffless BP Devices (IEEE 1708-2014), which specifies MAE ≤ 5 mmHg and standard deviation ≤ 8 mmHg for clinical-grade measurements. Closing this gap is the central challenge for the next development phase.

The diastolic correlation ($r = 0.43$) is the most significant accuracy limitation. The oscillogram features most strongly associated with diastole correspond to lower applied pressures, at the beginning of the force ramp where force application technique is least consistent. Future work will explore lower initial force levels and diastolic-specific oscillogram features.

3.9.2 Usability

The 49% measurement exclusion rate is the most significant practical limitation. The majority of excluded measurements failed the force application criteria: the user applied force too quickly, inconsistently, or with the fingertip rather than the whole finger. The correct technique is not intuitive for users accustomed to touchscreen interactions. More prominent real-time feedback, haptic cues, and brief guided practice are likely to reduce the exclusion rate substantially.

The fingertip placement issue is particularly impactful. Pressing with the fingertip

positions the contact point above the camera, so the PPG signal reflects blood volume changes at the fingertip rather than beneath the fingernail bed overlying the digital artery. The resulting signal has lower pulse amplitude and poorer correspondence to arterial pressure. Resolving this without a physical guide, which would partially converge with the BPClip approach, is an open design challenge.

3.9.3 Population and Conditions

All proof-of-concept measurements were performed in supervised lab conditions with a predominantly young, healthy population. The mean systolic BP of 103 mmHg is well below the hypertensive threshold. Future studies must include participants with clinical hypertension, older adults, and individuals with comorbidities that may affect peripheral vascular compliance or fine motor control, including diabetic neuropathy, Parkinson's disease, and peripheral vascular disease.

3.9.4 Device Dependence

VFE requires a per-model calibration procedure and is sensitive to vibration motor wear. While the per-model calibration is designed for the application developer rather than the end user, it requires a small-scale data collection effort for each new phone model supported. A baseline amplitude normalization procedure that automatically corrects for motor wear without recalibration is a priority for future development.

Chapter 4

Smartphone-Augmented Functional Assessment for Musculoskeletal Health

4.1 The Information Discarded by Clinical Scoring

The 30-second chair stand test is administered in millions of clinical encounters annually. A patient holds the sides of a chair, stands and sits repeatedly for 30 seconds, and a clinician records the number of completed repetitions. That single integer is the outcome. The rest of the test is discarded [13].

This is a striking information loss. Each stand-to-sit cycle is a coordinated neuromuscular event requiring eccentric and concentric loading of the quadriceps, gluteals, and postural stabilizers. The velocity, smoothness, progressive fatigue, and inter-cycle variability visible across those 30 seconds reflect the very muscle quality and functional reserve that sarcopenia erodes. Yet standard clinical practice collapses this rich temporal signal into a scalar, discarding everything about *how* the person moved and retaining only *how many times* they did so.

The handgrip strength test presents a structurally similar problem from a different direction. Here the measurement itself is well-defined, peak grip force in Newtons or kilograms, obtained with a calibrated hydraulic dynamometer. The measurement is not informationally impoverished; it is access-impoverished. The dynamometer is clinic-bound. The majority of the global population at risk for sarcopenia will never have their handgrip strength measured because neither the device nor the clinical visit is accessible to them.

These two tests, handgrip strength and the 30-second chair stand, appear together in every major sarcopenia diagnostic framework [4, 29]. They represent the functional assess-

ment pillar of sarcopenia diagnosis, paired in those frameworks with a body composition measurement from dual-energy X-ray absorptiometry (DXA). DXA is itself expensive and clinic-bound, compounding the access problem. The result is a diagnostic infrastructure that is clinically well-validated but practically inaccessible for population-scale monitoring.

This chapter demonstrates that commodity smartphone sensing can address both limitations simultaneously, and in doing so illustrates both goals of the dissertation's central thesis. For handgrip strength, the smartphone replicates an established clinical measurement without the specialized device, demonstrating the access-expansion goal. For the chair stand test, the smartphone does something more: it recovers the movement information that standard scoring discards and translates it into body composition biomarkers that the traditional count cannot provide, demonstrating the biomarker discovery goal. Both contributions share a critical design property. The smartphone does not introduce a new clinical workflow. It accompanies tests that are *already being administered*, converting observational movement data into quantitative biomarkers as a passive, zero-burden upgrade to existing clinical practice.

4.2 Clinical Context: Sarcopenia and Functional Assessment

Sarcopenia is defined as the progressive, age-related loss of skeletal muscle mass, strength, and function. It affects 5–13% of adults aged 60–70 and up to 50% of those over 80, and carries well-documented associations with falls, fractures, disability, and mortality [4, 29]. Despite decades of research, the optimal diagnostic criteria remain unsettled, but the leading consensus bodies have converged on a point of significant practical importance: *muscle mass alone is insufficient for clinical diagnosis.*

The European Working Group on Sarcopenia in Older People (EWGSOP) requires both low muscle mass and low muscle function for diagnosis [29]. Muscle function is assessed via handgrip strength or chair stand performance, while severity is stratified using physical performance tests. The Asian Working Group for Sarcopenia (AWGS) adopts

a structurally similar framework, pairing a body composition measurement with either a strength or performance test [4]. Both groups explicitly acknowledge the barriers posed by the expense and equipment requirements of DXA. The Sarcopenia Definitions and Outcomes Consortium (SDOC) takes this reasoning to its logical conclusion, recommending a diagnosis based *entirely* on measures of muscular strength and physical performance, without any body composition measurement [4]. This reflects a growing clinical consensus that functional capacity is not merely a proxy for muscle mass, it is the more clinically meaningful target.

The diagnostic literature thus points toward a clear gap: the field needs assessments that capture functional muscle quality in a way that is accessible outside specialized clinical infrastructure. The 30-second chair stand test and handgrip strength test are already embedded in major diagnostic frameworks as exactly this kind of functional measure. What is missing is a means to extract more from them than their traditional scalar scores provide, and to make those measurements available beyond the clinical settings where specialized equipment currently confines them.

The current gold standard for body composition, DXA, is accurate but expensive, equipment-dependent, and confined to clinical or research settings. These limitations create real barriers to the early detection and longitudinal monitoring important for sarcopenia intervention. The repetition count from the 30s CST and the peak force from the dynamometer serve as coarse functional proxies precisely because better alternatives are inaccessible. The work in this chapter addresses both the information poverty of the count and the access poverty of the dynamometer.

4.3 Handgrip Strength: Replicating an Established Measurement

4.3.1 Clinical Significance and the Access Gap

Handgrip strength (HGS) is one of the most widely used and well-validated biomarkers in geriatric medicine. Its predictive validity extends beyond sarcopenia to encompass

cognitive decline, mobility limitation, functional status, and all-cause mortality [30]. The measurement is simple: the patient squeezes a calibrated dynamometer as hard as possible, and peak force is recorded in kilograms or Newtons. Its simplicity is both its clinical strength, it requires only seconds to administer and is reproducible across operators and settings, and its limitation. A Jamar hydraulic dynamometer, the clinical gold standard, costs several hundred dollars. Electronic dynamometers range from fifty to over one thousand dollars. Neither is present outside of clinical or research settings. For the majority of older adults living with or at risk of sarcopenia, handgrip strength has never been measured.

The smartphone addresses the access gap directly. If the force applied by a hand gripping a smartphone can be reliably measured using only the phone's existing sensors, the camera, the IMU, the screen, then handgrip strength becomes a measurement available to anyone who owns a smartphone. Unlike the chair stand augmentation work described in Section 4.4, the handgrip smartphone measurement does not aim to discover novel biomarkers beyond what the dynamometer measures. The goal is a lightweight proxy for that same construct: to approximate the muscular strength signal that hand grip strength testing captures, without a dedicated dynamometer, at the point of care.

4.3.2 Smartphone-Based HGS Measurement

The smartphone HGS system uses Vibrometric Force Estimation (VFE), the same underlying technique developed for the oscillometric blood pressure measurement in Chapter 3: applying force to a vibrating smartphone damps the amplitude of the phone's own vibration motor output as measured by its accelerometer and gyroscope, and that damping can be inverted to recover the applied force. For the grip strength application, the user grips the phone between the thumb and index finger while the vibration motor runs, and the resulting change in high-frequency accelerometer and gyroscope signal amplitude provides the basis for force estimation [14].

Because vibration damping characteristics differ across phone hardware, each smart-

phone model must first be profiled against a ground-truth reference. This work profiled a Google Pixel 4 using a 0.3 mm thick force sensor (SingleTact 15 mm, 100 N) positioned under the index finger; the user applied a range of forces guided by real-time visual feedback while the phone recorded synchronized six-channel IMU data (triaxial accelerometer and gyroscope). Forty-second profiling sessions captured varied force trajectories, including linear ramps, sinusoids, and square waves, to sample the full dynamic range of applied force.

Signal processing isolated the vibration-modulated component of each IMU channel with a 4th-order Butterworth bandpass filter centered at the vibration motor's dominant frequency (136 Hz, 10 Hz bandwidth), then computed a signal envelope as the sum of the upper envelope and the absolute value of the lower envelope. Cross-correlation between each IMU channel and the reference force signal determined the temporal alignment between sensor streams, and a derivative-based artifact correction step detected and interpolated over brief signal collapses caused by the vibration motor stopping and restarting. Two ridge regression models were trained: an absolute-force model using magnitude features from all six IMU channels, and a relative-force model using percentile-scaled accelerometer X-axis and gyroscope Y-axis signals normalized to the 5th–95th percentile range.

Model performance was assessed with 15-fold hold-one-out cross-validation across 209,603 six-channel IMU samples, with each fold corresponding to one profiling session held out for testing. Absolute force estimation achieved a mean absolute error of 1.88 lbs (SD 1.00) and RMSE of 2.10 lbs (SD 1.03) across folds. Relative force estimation, expressed as percentage of applied force magnitude, achieved a mean absolute error of 10.07% (SD 3.48%) and RMSE of 12.40% (SD 4.04%).

This measurement is a pinch-type proxy for grip strength, applied force between the thumb and index finger against the smartphone, rather than the standardized whole-hand, four-finger-and-palm measurement captured by a clinical hand dynamometer. This altered hand positioning was chosen deliberately to simplify how a user holds the phone and to reduce the variability that full-hand grip positions introduce into the sensor readings, at the

cost of departing from the exact clinical HGS metric. Realizing the access-expansion goal at scale requires only that the profiling procedure be repeated once per smartphone model; the resulting force model then applies unmodified to every identical device already in a user's hands.

4.3.3 HGS as the Replication Case Study

When viewed alongside the chair stand work in Section 4.4, handgrip strength occupies a distinct position in the dissertation's conceptual architecture. It is the clearest example of the access-expansion goal of the thesis: an established, validated, clinically meaningful measurement made available to populations currently excluded from it by device cost and institutional access. The measurement target is known, the gold standard is clear, and the contribution is deployment breadth rather than scientific novelty of the biomarker itself.

This distinction is not a limitation. The ability to replicate a clinical gold standard measurement on commodity hardware has direct and immediate clinical value. A patient who can check their handgrip strength at home, at the same frequency and with the same convenience as checking their step count, can participate in longitudinal sarcopenia monitoring that is currently unavailable outside clinical trials. The replication goal and the discovery goal are both necessary components of the dissertation's thesis; neither subsumes the other.

4.4 The 30-Second Chair Stand Test: Augmenting Beyond the Count

4.4.1 The Information Loss Argument

The central claim of the chair stand work is specific and falsifiable: the movement information discarded by standard 30-second chair stand test (30s CST) scoring is recoverable from a held smartphone and contains clinically meaningful signal that the repetition count does not provide [13].

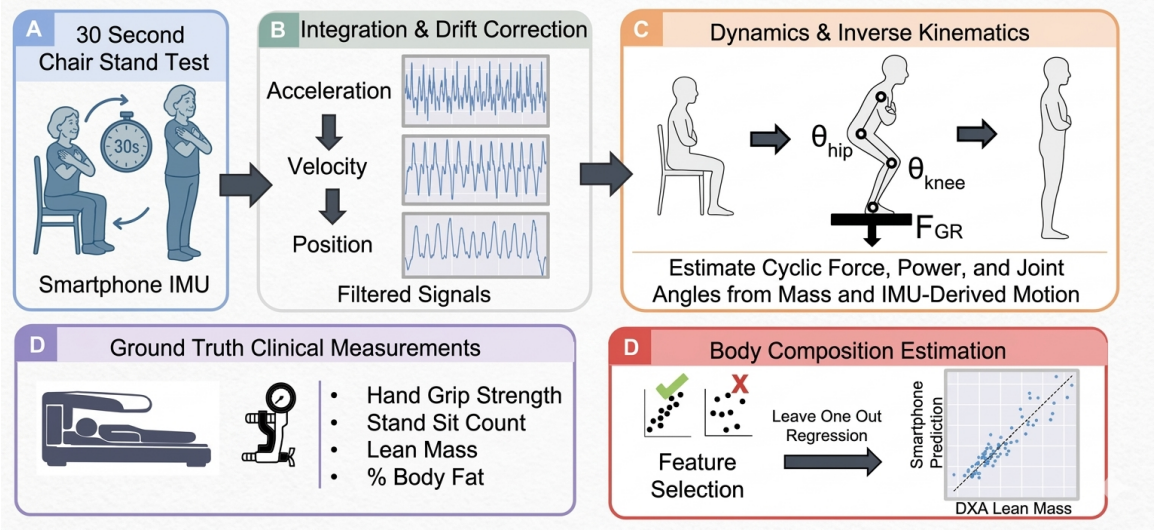


Figure 4.1: End-to-end 30-second chair stand study pipeline. (Left to right) A participant performs the 30-second chair stand test (30s CST) holding a Google Pixel 8 against their sternum; the raw IMU signal from the test is shown; six feature category icons represent the movement signal categories extracted (speed, power, smoothness, fatigue, kinematics, and cycle consistency); a DXA scanner provides ground-truth body composition outcomes; leave-one-out (LOO) ridge regression maps smartphone features to each outcome. Adapted from Barry et al. 2026 chair stand manuscript, Fig. 1.

Prior instrumented chair stand research has established that this discarded information is physiologically meaningful. Individuals with sarcopenia generate significantly lower ground reaction forces during the chair stand transition, and rate of force development is a particularly sensitive discriminator between sarcopenic and non-sarcopenic individuals [31]. Maximum knee extension angular velocity and maximum iliac elevation velocity discriminate between sarcopenic and non-sarcopenic individuals during chair rise [32]. Trunk flexion angle at seat-off, static stabilization time, and normalized chair stand power together explain a substantial proportion of variance in physical function and sarcopenia status [33]. These findings collectively establish that the 30s CST is not an impoverished test, it is an informationally rich test scored in an impoverished way.

The barrier to extracting this richer information has been the measurement instrument. Force plates cost tens of thousands of dollars and are fixed in laboratory settings. Multi-sensor IMU systems require body-worn hardware and expert fitting. Neither is compatible with the millions of clinical chair stand tests administered annually in community settings,

outpatient clinics, and research studies. A commodity smartphone held during the test changes this calculus entirely.

4.4.2 Study Design and Data Collection

Data were collected at the Exercise and Physical Activity Resource Center (EPARC) at UC San Diego under IRB approval (#811902). Eighty-five participants were recruited from individuals scheduled for clinical DXA scans at EPARC. Eligibility required English literacy, age 18 or older, and willingness to complete grip-strength and DXA assessments; individuals with neuromuscular disorders, conditions impairing hand function, recent upper-extremity injuries, or medications known to affect grip strength were excluded.

DXA scans were performed on a Hologic Horizon W or GE Lunar Prodigy system, providing reference measurements of total lean mass and percent body fat. Dominant and non-dominant handgrip strength were measured using a dynamometer. Demographic and anthropometric data including gender, height, and weight were also recorded. After excluding trials with missing IMU data and those failing internal consistency checks, the analytic sample comprised $N = 78$ participants. Participants were predominantly female (73.1%) with a mean age of 51.5 ± 16.4 years and mean BMI of 24.3 ± 4.5 kg/m^2. Mean 30-second chair stand count was 16.9 ± 5.5 repetitions (range 9–30).

Participants performed the 30s CST while holding a Google Pixel 8 against the chest with arms crossed, positioning the device near shoulder level. A custom application recorded triaxial acceleration and angular velocity at approximately 400 Hz throughout the task. Participants were instructed to stand fully and sit down as many times as possible within 30 seconds, consistent with standard clinical protocol [34].

Three regression outcomes were modeled against the IMU-derived features: total lean mass (lbs), percent body fat, and dominant handgrip strength.

4.4.3 Signal Processing and Cycle Detection

Raw triaxial acceleration channels were linearly interpolated to fill missing samples and zero-centered by subtracting the sample mean to remove DC offset. The sensor frame was reoriented by estimating the gravity vector from the median acceleration over the first 1.0 s of each trial and computing a rotation matrix aligning the vertical axis with a consistent world frame. A fourth-order Butterworth low-pass filter with a 2 Hz cutoff was applied to the reoriented signals to remove high-frequency noise while preserving chair stand transition dynamics.

Mean-subtracted acceleration was numerically integrated to obtain velocity. A 0.1 Hz low-pass filter was applied to the velocity signal to estimate and subtract integration drift, followed by a second integration to obtain position. Resultant (magnitude) acceleration, velocity, and position signals were computed from the three-axis signals for cycle detection and trial-level feature extraction.

Stand cycles were identified as peaks in the resultant position signal with a minimum prominence of 0.2. Rules were applied to retain the first and last peaks at trial boundaries. The number of detected peaks corresponds closely to the conventional 30s CST count, providing a direct bridge between the traditional score and the enriched feature set. For per-cycle analysis, a window spanning 0.4 s before to 1.2 s after each detected peak was extracted.

4.4.4 Feature Engineering: Six Movement Categories

The feature engineering process was motivated directly by the prior sensor-based chair stand literature reviewed in Section 4.4.1. The prior literature identified specific movement characteristics, ground reaction force, angular velocity, trunk kinematics, timing, as clinically meaningful, but measured each using specialized equipment. The contribution here is approximating all of them from a single held smartphone, organized into six interpretable movement categories.

Speed features characterize how quickly participants complete each stand and sit phase. Trial-level speed features include maximum resultant velocity, standard deviation and range of resultant velocity, 5th percentile velocity (a marker of sit-down speed), and time to complete the first five stands. Per-cycle speed features include mean rise time and mean sit time. Inverse-kinematics-derived speed features include peak and mean knee and hip angular velocities.

Power and impulse features capture the total force effort per repetition. Trial-level features include standard deviation and range of resultant acceleration and the 85th percentile of resultant acceleration as a proxy for peak force production. Per-cycle features include mean mechanical impulse (area under the absolute acceleration curve per cycle, scaled by body mass) and mean mechanical power (mean absolute product of acceleration and velocity per cycle). These features approximate the ground reaction force and rate of force development metrics identified as sarcopenia-sensitive by [31] and [35] from a shoulder-mounted smartphone rather than a force plate.

Smoothness features quantify movement coordination. Mean peak jerk (mean peak rate-of-change of acceleration per cycle, in m/s^3) marks abrupt, poorly coordinated transitions. Trial-level smoothness features include the number of velocity zero-crossings and the proportion of samples with resultant velocity below 0.2 m/s.

Fatigue features capture the within-test decline across the 30 seconds, information that is entirely invisible to the repetition count, which measures only the final total without recording how performance changed along the way. Five features were extracted from the inter-peak interval sequence and the per-cycle velocity sequence: fatigue delta (last minus first inter-peak interval), fatigue slope (linear slope of inter-peak intervals over cycles), fatigue ratio (mean interval in the final third divided by the first third), velocity trend (linear slope of per-cycle peak velocity), and velocity ratio (mean peak velocity in the final third divided by the first third). These features are unique to the smartphone approach: no scalar scoring system can capture within-test trajectory.

Kinematic features are derived from a planar two-link inverse kinematics model described in Section 4.4.5. Trial-level kinematic features include mean, standard deviation, minimum, maximum, and range of both knee and hip flexion angles. These approximate the angular velocity and trunk kinematic features identified as discriminating between sarcopenic and non-sarcopenic individuals by prior work using body-fixed inertial sensors [32, 36].

Cycle consistency features summarize repetition-level regularity. Mean stand asymmetry is the mean normalized absolute difference between rise time and sit time per cycle (0 = perfectly symmetric, 1 = entirely asymmetric). The total number of detected stand cycles is retained as a feature, providing a direct counterpart to the conventional count and enabling a formal comparison of count-alone versus enriched-feature model performance.

4.4.5 Inverse Kinematics Model

A planar two-link inverse kinematics model is used to estimate knee and hip joint angles from the IMU-derived shoulder position time series. The knee is treated as a fixed origin. The thigh segment (length $= 0.345\times$ height) connects the knee to the hip, and the torso segment (length $= 0.50\times$ height) connects the hip to the shoulder, where the phone is held. Segment lengths are derived from standing height using fixed anthropometric ratios. Shoulder position over time is obtained by double integration of vertical and anterior–posterior acceleration with drift correction.

For each time point, the knee-to-shoulder distance is computed from the estimated shoulder position. Knee and hip angles are then derived from the two-link geometry using the law of cosines. Angular velocities are obtained by differentiating the joint angle time series.

The model is a simplification of the complex multi-segment human body. It assumes sagittal-plane motion, a fixed knee position, and population-average segment length ratios that may not accurately represent individual participants. Sensitivity analyses perturbing segment length assumptions by $\pm 10\%$ and assumed phone placement height by ± 10 cm

demonstrated that IK-derived features are robust to these variations: perturbations of this magnitude changed mean estimated joint angles by less than a few degrees and produced minimal changes in LOO R^2 for IK-only feature models. Joint angle estimates should therefore be treated as movement-quality indices rather than anatomically precise angles.

4.4.6 Statistical Analysis

Predictive performance was evaluated under four feature-set conditions designed to isolate the contribution of IMU-derived movement information:

1. *IMU-only*: all features derived exclusively from the smartphone IMU signal. This is the primary condition of scientific interest, as it establishes what the phone contributes independently of body size information.

2. *All features*: IMU features combined with passive demographic predictors (height, weight, and BMI). Age was excluded from all predictive models.

3. *Count-only*: the 30s CST repetition count as the sole predictor, the conventional clinical score. This condition establishes the baseline against which the enriched IMU feature set is compared.

4. *Height + Weight*: height and weight only, without any test performance information. This condition isolates how much of each outcome is predictable from passive anthropometric measurements alone, without the movement or BMI information used by the all-features model.

For the all-features, count-only, and height-and-weight conditions, an ElasticNet regression was fitted with hyperparameters selected by 3-fold inner cross-validation (`ElasticNetCV`). Feature selection occurred inside each leave-one-out (LOO) fold, not before it, to prevent data leakage. The IMU-only condition is the exception: because its movement features are numerous and strongly collinear with no dominant predictor, ElasticNet selection is unstable at $N = 78$ and overfits, so that condition uses ridge regression

(`RidgeCV`) instead, which jointly shrinks the correlated movement features and recovers their shared signal. All feature values were standardized within each training fold using scaler parameters derived from the training data only. Reported metrics are LOO R^2, Pearson r, and mean absolute error (MAE).

To test whether the IMU-only condition produced significantly higher correlations than the count-only or height-and-weight baselines, Steiger's Z-test for dependent correlations with the Meng–Rosenthal–Rubin correction was applied. This test uses the LOO prediction vectors from two conditions alongside the observed outcomes; a two-tailed $p < 0.05$ was used as the significance threshold.

Feature importance was characterized using stability selection across 50 bootstrap resamples (75% subsampling without replacement). In each resample, an ElasticNet was fitted and the selected features (non-zero coefficients) were recorded. Stability selection frequency, the proportion of resamples in which a feature was selected, was computed per feature and per movement category. Features above the 0.5 threshold are considered reliably selected. Forward stepwise selection with plain linear regression (not ElasticNet) was additionally performed for each target to identify minimal sufficient feature sets, advancing one feature at a time based on marginal LOO R^2 gain.

4.4.7 Results

IMU Features Outperform the Conventional Count

The IMU-only condition achieved substantially higher predictive performance than the count-only baseline across all three regression targets. For total lean mass, the IMU-only condition achieved Pearson $R = 0.48$, compared to $|R| \leq 0.25$ for the count-only condition. For percent body fat, the IMU-only condition achieved $R = 0.50$ (95% CI: 0.33–0.65), compared to the same weak count-only baseline. For dominant handgrip strength, the IMU-only condition achieved $R = 0.52$. Steiger's Z-test confirmed that the IMU-only condition produced significantly higher correlations than the count-only baseline for all three outcomes

(percent body fat $p < 0.001$, dominant grip strength $p = 0.019$, lean mass $p = 0.007$).

The all-features condition achieved the highest overall performance (lean mass $R = 0.92$, percent body fat $R = 0.63$, dominant HGS $R = 0.85$). However, lean mass and grip strength results in the all-features condition should be interpreted with care: a height-and-weight-only baseline already reaches $R = 0.90$ for lean mass and $R = 0.77$ for grip strength, so adding movement and BMI features contributes comparatively little for these two outcomes. Percent body fat shows the opposite pattern: the height-and-weight baseline reaches only $R = 0.34$, well below both the IMU-only condition ($R = 0.50$) and the all-features condition ($R = 0.63$), indicating that for body fat the chair-stand movement signal carries body-composition information that height and weight alone do not provide. The IMU-only R is therefore the more meaningful evidence for the chapter's central claim, since it establishes the contribution of movement information independently of body size. The count-only model was the weakest condition for every outcome ($|R| \leq 0.25$), confirming that the repetition count alone discards the majority of the predictive signal present in the 30-second test.

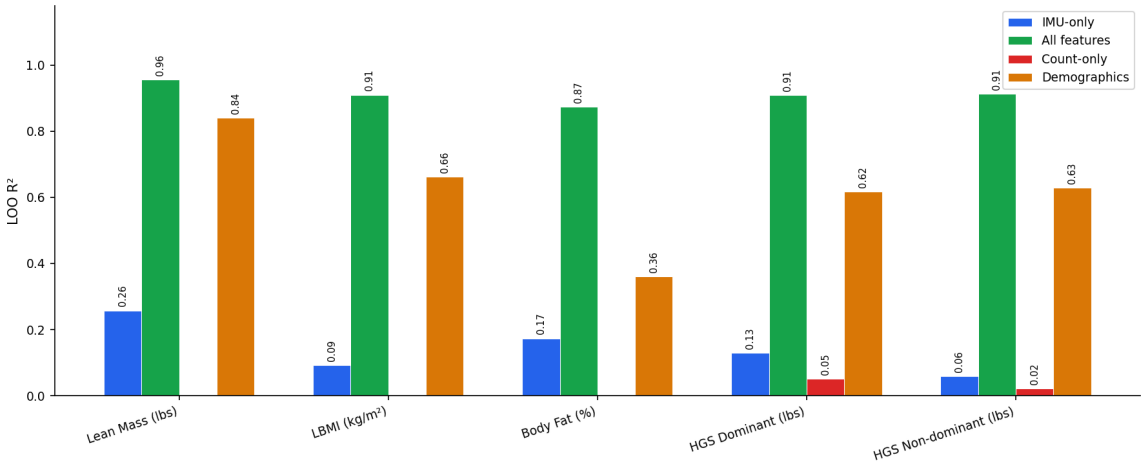


Figure 4.2: Leave-one-out R^2 grouped bar chart for four predictor conditions (IMU-only, all features, count-only, and height-and-weight) across all three regression targets (lean mass, percent body fat, and dominant handgrip strength). Error bars indicate 95% bootstrap confidence intervals. The IMU-only condition substantially outperforms count-only for all outcomes, demonstrating that the movement signal captured by the held smartphone contains clinically meaningful information beyond what the repetition count provides. Adapted from Barry et al. 2026 chair stand manuscript.

Movement-Specific Signal Independent of Demographics

A partial correlation analysis was conducted to assess whether the IMU signal captures information genuinely independent of what height, weight, and BMI already provide. Of the 33 curated IMU features, 15 retained a statistically significant partial correlation with lean mass after regressing out height, weight, and BMI ($p < 0.05$, two-sided Pearson); the corresponding counts were 12 for percent body fat and 19 for dominant grip strength. The reduction in correlation magnitude from raw to partial correlation was largest for the Power/Impulse feature group, consistent with those features partially tracking body-size and body-composition covariance in addition to genuine movement-specific physiology, suggesting that the surviving partial correlations reflect aspects of muscle quality or neuromuscular function not explained by body size alone.

This result is clinically significant. Features that retain predictive value after demographic adjustment are the ones that could eventually enable within-person longitudinal monitoring: tracking body composition change in the same individual over time without requiring DXA. This application, a passive, zero-burden body composition monitoring system embedded in a test the patient is already performing, is the longer-term clinical vision that the current feasibility results support.

Feature Importance and Minimal Sufficient Feature Sets

Stability selection (50 bootstrap resamples with 75Forward stepwise selection identified minimal sufficient feature sets of $k = 2$ features for lean mass ($R = 0.54$), $k = 4$ for percent body fat ($R = 0.60$), and $k = 1$ for dominant grip strength ($R = 0.60$), each capturing the majority of the predictive variance achievable with the full feature set. The features entering earliest in the stepwise sequence for each outcome are reported in the supplementary tables of [13]. The minimal feature set finding is practically important: it suggests that a future clinical deployment does not need to compute all features in real time but can focus on a small, interpretable subset.

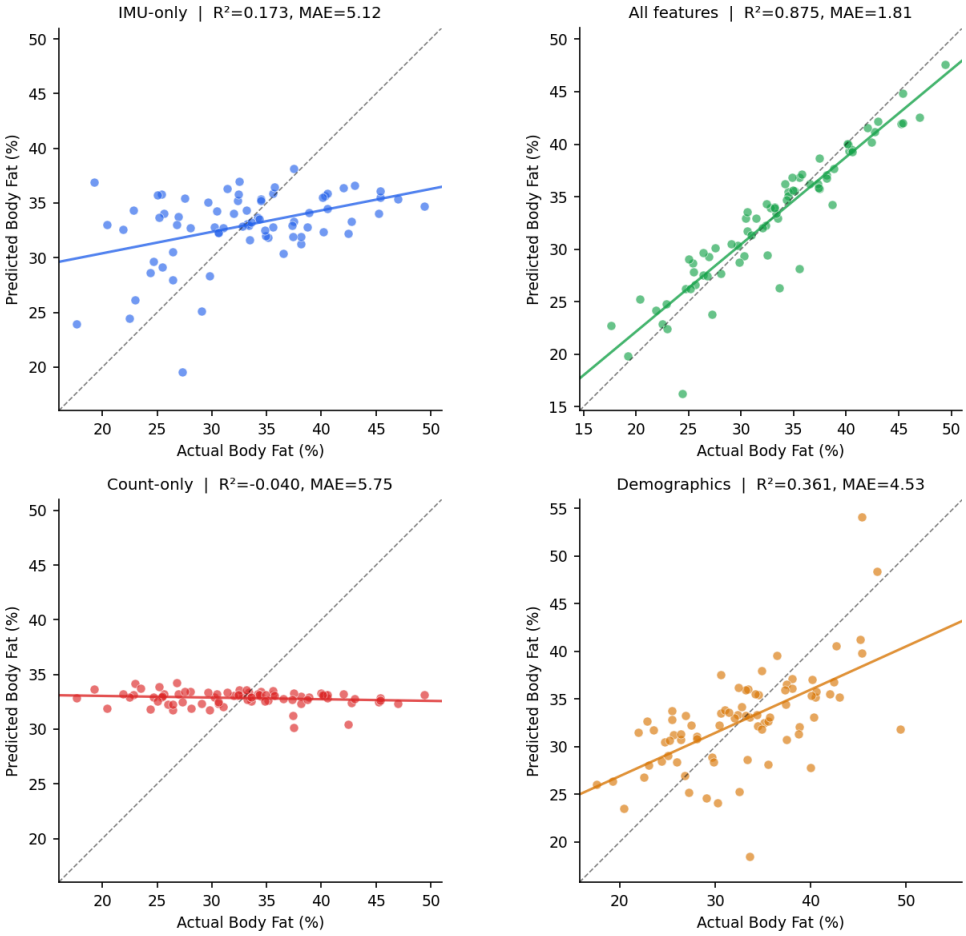


Figure 4.3: Leave-one-out predicted versus actual percent body fat under the IMU-only condition ($N = 78$ participants). Each point represents one participant; the identity line (dashed) and a linear fit (solid) are shown. Pearson r is annotated. The IMU-only model captures meaningful variance in percent body fat from movement features alone, without height, weight, or demographic inputs.

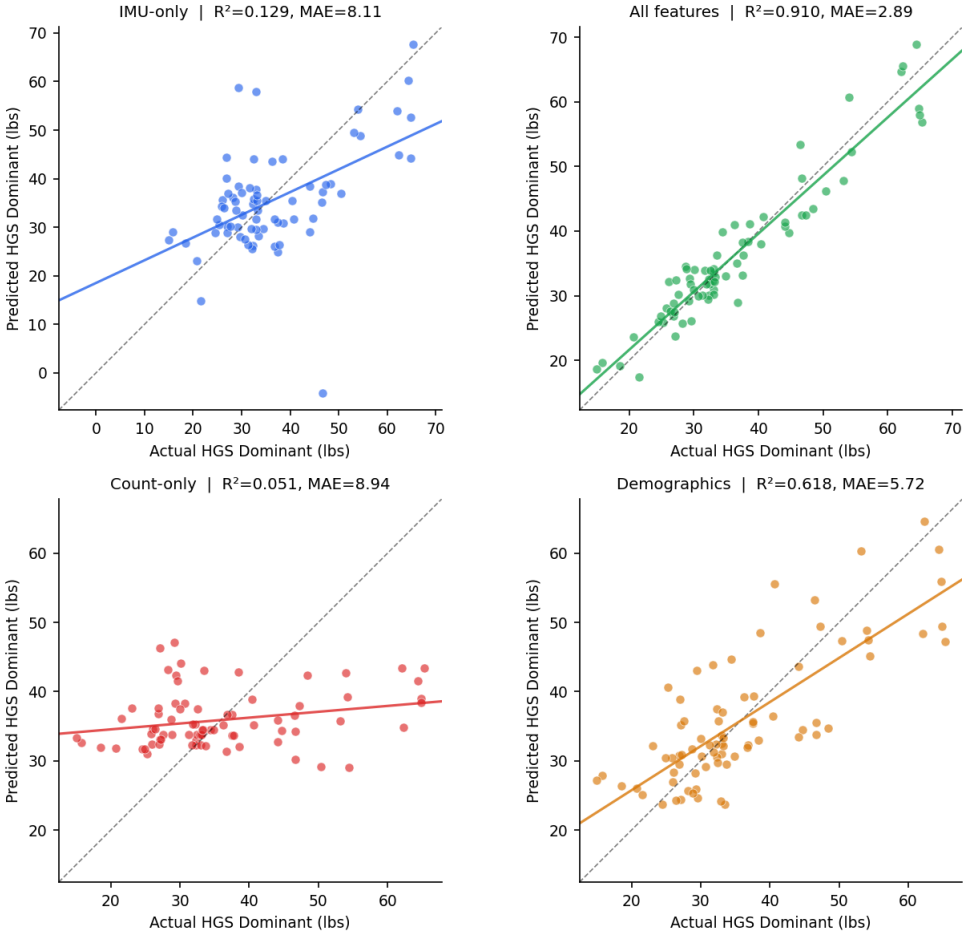


Figure 4.4: Leave-one-out predicted versus actual dominant handgrip strength (right hand) under the IMU-only condition. The chair stand movement features provide moderate predictive signal for handgrip strength, suggesting a shared neuromuscular substrate between lower-extremity functional strength and grip strength.

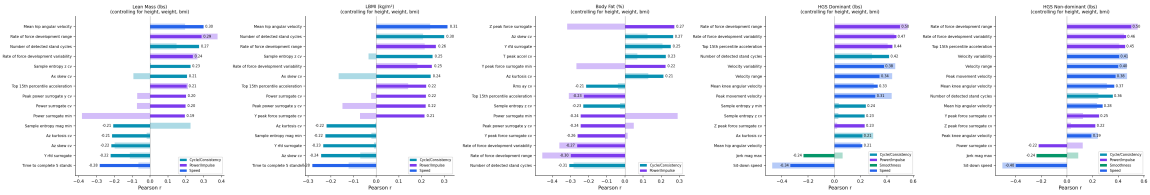


Figure 4.5: Raw Pearson *r* versus demographic-adjusted partial correlation coefficient for each IMU feature category, for total lean mass (left) and percent body fat (right). Features that retain large partial correlations after regressing out height, weight, and BMI reflect movement characteristics genuinely independent of body size, the subset most relevant for longitudinal within-person tracking.

Cohort Sarcopenia Distribution

The participants were largely normative with respect to published sarcopenia diagnostic thresholds. Under EWGSOP2 criteria, all participants (100LOO R^2 values therefore characterize model performance within a healthy distribution. The strong performance in this normative cohort is encouraging, but the results cannot be directly extrapolated to screening sensitivity or specificity at the sarcopenic boundary. This is acknowledged as the primary limitation requiring future study.

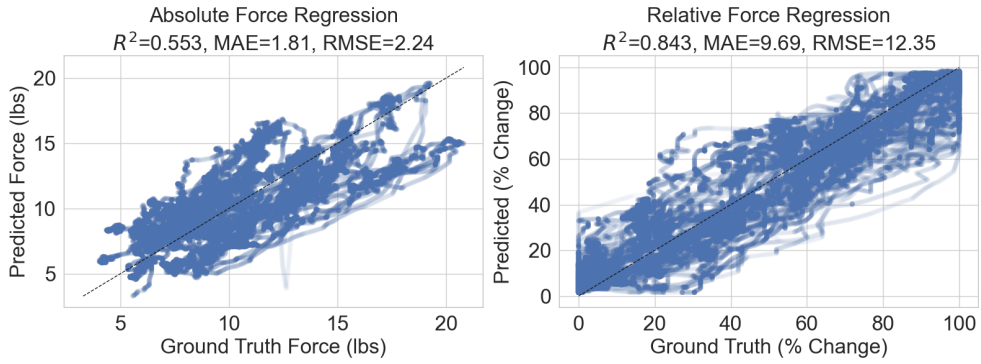


Figure 4.6: Smartphone-based pinch-force estimation validated against a reference force sensor (Google Pixel 4, $N = 209{,}603$ six-channel IMU samples, 15-fold hold-one-out cross-validation). (Left) Absolute force regression ($R^2 = 0.553$, MAE= 1.81 lbs, RMSE= 2.24 lbs). (Right) Relative force regression, expressed as percent of applied force magnitude ($R^2 = 0.843$, MAE= 9.69%, RMSE= 12.35%). Dashed line indicates identity; each point is one IMU sample across all 15 profiling sessions.

4.5 The Augmentation Principle

The two studies in this chapter, handgrip strength and the 30-second chair stand, illustrate complementary applications of a general principle: traditional clinical assessments are measurement instruments designed before ubiquitous sensing existed. Their scalar scores reflect the limitations of human observation in real-time clinical settings, not the limits of what the underlying physical tests could physiologically reveal. When a commodity sensor accompanies those tests, it does not displace clinical practice. It enriches it.

The distinction between the two studies is instructive. In the handgrip study, the

scalar score, peak grip force, is itself a good measurement. The limitation is not the score's information content but its distribution: the device that produces it is inaccessible. The smartphone's contribution is access replication. In the chair stand study, the scalar score, the repetition count, is genuinely informationally impoverished relative to the underlying test. The smartphone's contribution is not merely access replication but information recovery. The IMU-only features significantly outperform the count-only baseline, which would not be possible if the count were an adequate summary of the test's physiological signal. They also demonstrate independent predictive value beyond demographics, confirming that the signal is not simply recapitulating body size.

Together, the two studies span the range of what smartphone augmentation can contribute to traditional clinical assessment. At one end, the smartphone closes access gaps by replicating measurements that require clinic-bound devices. At the other, it discovers biomarkers that no element of the standard workflow was measuring at all. Both contributions are enabled by the same deployment principle: the smartphone rides along with tests that are already being administered, adding measurement depth without adding clinical burden.

This principle extends beyond the two tests examined here. The Timed Up and Go test, the Short Physical Performance Battery, gait speed assessment, and the six-minute walk test all discard rich temporal data in favor of a scalar summary. Each is a candidate for smartphone augmentation on the model established in this chapter. The movement signal that each test generates has been studied in specialized laboratory settings using force plates and multi-sensor systems; the contribution of a commodity smartphone would be to make that same signal accessible at the scale of ordinary clinical practice.

The Postable Spirometry work [3] extends this principle further, to the question of how even the physical attachment needed for some assessments can be distributed at population scale. A vortex whistle constructed from an oilboard postcard and assembled by the user, used with a smartphone microphone, achieves FVC estimation within 0.055 L mean

error, a pulmonary function measurement that previously required a clinic-bound spirometer costing between one and five thousand dollars. The *postable* paradigm, in which a flat, self-assembled accessory is mailed to the user, addresses the final distributional barrier: when even the attachment is a postcard, there is no hardware requirement that cannot be met through the postal system. This paradigm applies equally to the far-red pupillometry scope described in Chapter 2 and to the BPClip described in Chapter 3, representing a unifying deployment philosophy across the dissertation.

4.6 Toward At-Home Deployment

The studies presented in this chapter were conducted in supervised clinical settings with DXA ground truth. The translation to at-home, unsupervised deployment involves challenges that the current results illuminate without fully resolving.

For the chair stand test, the non-kinematic IMU features, speed, power, smoothness, fatigue, and cycle consistency, are less sensitive to exact phone placement than the inverse kinematics features, since they depend on the magnitude and temporal structure of the IMU signal rather than on the assumed geometric relationship between the phone and the body. These categories may provide a more robust foundation for at-home protocols where placement standardization is difficult to enforce. The kinematic features, which extract joint angles from the inverse kinematics model, require that the phone be held at approximately shoulder level during the test, a positioning requirement that can be communicated in a brief app instruction and verified by the IMU signal itself.

For handgrip strength, the at-home deployment challenge is different: the measurement requires the user to interact with the phone in a specific way, and the quality of that interaction determines measurement validity. The usability lessons from the older adult deployments in the pupillometry study (Chapter 2) and the blood pressure proof-of-concept (Chapter 3) apply here. Clear visual feedback, a practice measurement, and automatic quality checking are likely necessary components of any unsupervised deployment.

The ultimate vision for both measurements is a smartphone-based functional assessment battery, chair stand, handgrip, and spirometry, self-administered at home, without any clinical visit, and performed at the frequency needed to support longitudinal body composition monitoring. This vision is the natural extension of the measurements demonstrated in this chapter, and its realization will require the at-home usability studies that have not yet been conducted.

4.7 Limitations and Future Work

4.7.1 Cohort Representativeness

The most significant limitation of the chair stand study is the normative skew of the participant cohort. The EPARC DXA training school population from which participants were recruited skewed heavily toward healthy older adults without sarcopenia. Although the IMU features demonstrate clinically meaningful predictive value within this normative range, they have not been evaluated in a population with a sufficient proportion of sarcopenic individuals to characterize screening sensitivity and specificity. The critical next step is a study with pre-stratified recruitment targeting the sarcopenic range, with $N = 200$–300 participants sufficient to narrow confidence intervals on key performance metrics and to evaluate performance against EWGSOP2 and AWGS classification criteria.

4.7.2 Kinematic Model Validity

The two-link planar inverse kinematics model has not been validated against an optical motion capture system or a multi-sensor IMU ground truth. Joint angle estimates should be treated as movement-quality indices rather than anatomically precise angles. The sensitivity analyses reported in the supplementary materials of [13] demonstrate robustness to reasonable variations in segment length assumptions and phone placement height, but the accuracy of the model for participants with atypical body proportions or movement patterns is unknown. Concurrent validation against optical motion capture is a necessary

step before the kinematic features can be used for individual clinical decisions rather than population-level modeling.

4.7.3 Single-Visit Cross-Sectional Design

The current study does not establish whether the IMU features are reproducible across test occasions or whether they track changes in body composition over time. The clinically most valuable application of the system, monitoring the response to sarcopenia interventions such as resistance training or nutritional supplementation, without requiring repeated DXA scans, depends on the features being sensitive to longitudinal change. A longitudinal validation study is therefore the most important single future direction for this work.

4.7.4 DXA System Variability

Body composition measurements were obtained from two different DXA systems (Hologic Horizon W and GE Lunar Prodigy). Prior research has documented systematic measurement differences between these systems. For the purposes of this feasibility study, the potential cross-system error is not modeled, but any future study moving from feasibility toward clinical validation should use a single DXA system or apply published cross-calibration equations.

Chapter 5

Discussion and Conclusion

5.1 Revisiting the Thesis

This dissertation opened with a claim: that by repurposing universal smartphone sensors, deployable software provides opportunities for expanding access to established clinical measurements and facilitating the discovery of novel digital biomarkers, while prioritizing usability for older adults and other target populations. Three chapters of experimental work have now provided evidence for and against this claim across three distinct clinical sensing domains. This chapter evaluates the claim honestly, identifies the contributions that generalize across domains, and argues for what the accumulated evidence means for the future of mobile health.

The claim has three components. The first is that smartphone sensors can expand access to established clinical measurements. The second is that they can facilitate the discovery of novel digital biomarkers. The third is that these goals can be pursued while prioritizing usability for populations who have the most to gain and who face the greatest barriers. Each component is evaluated in turn.

5.1.1 Expanding Access to Established Clinical Measurements

The access-expansion claim receives its strongest support in the pupillometry work. The NIR-based smartphone pupillometer achieved absolute pupil diameter measurements with a median MAE of 0.27 mm against the Neuroptics PLR-2000, a clinical gold standard [9]. The far-red prototype achieved a mean absolute error of 2.4% relative change against the Neuroptics PLR-3000, with a mean Pearson correlation of 0.93 across eleven participants spanning four smartphone models [10]. These results are directly comparable to

what a clinical pupillometer produces, achieved on hardware that costs orders of magnitude less. The access gap for pupillometry is not fully closed, clinical-grade longitudinal validation studies remain to be conducted, but the existence of a technically viable path from consumer hardware to clinical accuracy is demonstrated.

The blood pressure work is more nuanced. As with VFE, this work rejects pulse-transit-time and pulse-wave-analysis methods, which report only a relative change in BP and cannot function without prior cuff calibration, in favor of replicating the oscillometric method's absolute, cuff-referenced measurement. The VibroBP system achieved systolic MAE of 9.21 mmHg and an AUC of 0.92 for hypertension screening [11]. This represents meaningful access expansion for population-level screening, where the clinical question is whether an individual's systolic BP is likely elevated and warrants follow-up. It does not yet represent access to the precision BP measurement that a clinician uses to titrate antihypertensive medication. The claim is accurate for the screening application and premature for the monitoring application. The BPClip convergent result (MAE 8.72/5.49 mmHg) [12] strengthens the case that the oscillometric method itself is viable at the fingertip; the engineering challenge that remains is closing the precision gap.

The handgrip strength work, whose results will be incorporated into this chapter when the manuscript is finalized, represents the clearest example of access replication: a measurement with well-established clinical validity, currently confined to a device costing hundreds of dollars, made available on a smartphone.

5.1.2 Facilitating Novel Digital Biomarker Discovery

The biomarker discovery claim is most directly addressed by the chair stand work, where the smartphone's contribution was not to replicate an existing measurement but to recover information that the standard clinical scoring discards entirely. The IMU-derived features, speed, power, smoothness, fatigue, kinematics, and cycle consistency, significantly outperformed the conventional repetition count for body composition prediction across three

regression targets [13]. The partial correlation analysis further established that movement-quality features retain independent predictive value after controlling for passive demographic variables, confirming that the signal captures aspects of neuromuscular function not encoded in body size alone.

The pupillometry MCI screening application illustrates both the promise and the current limits of biomarker discovery at the individual level. Group-level pupillary response features showed statistically significant associations with MCI classification, digit span dilation at the 3-digit load ($p < 0.001$) and 6-digit load ($p = 0.028$), and post-illumination PLR response at three seconds ($p = 0.03$). Individual-level classification, however, did not generalize beyond chance. This is not a failure of the biomarker concept; it is a finding about the study design. Individual-level detection of a slowly progressive condition requires longitudinal within-person change data, not cross-sectional group comparisons. The smartphone pupillometry platform described in Chapter 2 is the instrument that makes such longitudinal data collection feasible; the current finding motivates the study rather than negating it.

5.1.3 Prioritizing Usability for Target Populations

The usability goal receives the most mixed evaluation. The pupillometry deployment to fifteen older adults demonstrated that self-administered at-home measurement is possible, with eight of fifteen participants producing data rated at least 80% usable [9]. The eyelid and eyelash occlusion problem, which was not anticipated and not observed in any laboratory validation study, exemplifies the gap between laboratory and real-world performance that characterizes all three application domains. Usability is not a problem that can be solved once and assumed to stay solved; it must be re-evaluated each time the system is deployed to a new population in a new setting.

The blood pressure proof-of-concept study's 49% measurement exclusion rate is the most direct evidence that usability remains a central limitation. The force application

protocol is technically the most demanding user interaction of any system in this dissertation, and the results make clear that achieving the required force profile consistently and correctly is not intuitive without more scaffolding than the current visual guide provides. The BPClip's physical guides for finger placement represent one design response; future VFE-based iterations will need equivalent solutions.

The chair stand and handgrip studies were conducted in supervised settings and do not yet provide usability evidence for at-home deployment to older adults. This is the most significant gap between the current state of the work and the thesis claim for the musculoskeletal chapter.

5.2 Cross-Cutting Technical Contributions

Beyond the domain-specific contributions of each chapter, the body of work contains several technical contributions that generalize across sensing domains and represent intellectual advances independent of any single application.

5.2.1 Vibrometric Force Estimation as a General Sensing Primitive

VFE, using a smartphone's vibration motor and IMU to estimate the force applied to the phone's body, is presented in this dissertation in the context of oscillometric blood pressure measurement. But the technique is not inherently blood-pressure-specific. Any application that requires measuring the magnitude of a force applied to a smartphone without a dedicated force sensor is a potential application of VFE. Digital dynamometry, grip strength assessment, and even the detection of tapping or pressing gestures with quantitative force resolution are all potential extensions. VFE is the first demonstration that the vibration motor and IMU, present in essentially every smartphone, can be used as a force sensing system, and it establishes the physical principle and calibration methodology that future applications can build on.

5.2.2 Far-Red Spectral Filtering as a General Solution to Melanin-Confounded Optical Measurement

The far-red approach to pupillometry addresses a specific instance of a general problem: optical measurements that fail for individuals with high melanin concentration because melanin absorbs broadband visible light. The same physics that makes dark irises difficult to image in the visible spectrum affects other smartphone optical health measurements, including photoplethysmography at darker skin tones, colorimetric hemoglobin screening at the conjunctiva or fingernail bed, and jaundice detection at the skin. The far-red band, approximately 630–700 nm, where melanin absorption is substantially lower than in the full visible spectrum, is a physically principled solution to melanin interference that is accessible within the passband of standard RGB smartphone cameras. The methodology and the prototype design established in Chapter 2 constitute a reusable framework for any optical smartphone measurement that encounters melanin-dependent performance disparities.

5.2.3 The Postable Paradigm for Distributed Clinical Accessories

Several of the systems in this dissertation require a small physical attachment to achieve the necessary optical geometry, ambient light exclusion, or mechanical force encoding. The far-red pupillometry scope, the neutral density filter assembly, and the BPClip all represent cases where a passive, electronically inert attachment enables a measurement that the bare phone cannot achieve. The Postable Spirometry work [3] establishes a design and distribution paradigm for such attachments: construct them from flat sheet materials that can be cut by laser and mailed as a postcard, self-assembled by the user, and used once or discarded. When the attachment costs cents to manufacture, requires no electronics, and arrives through the postal system, the last remaining barrier to at-home deployment, the need for any physical accessory at all, is effectively eliminated. This paradigm applies to the far-red scope, to potential future BPClip iterations, and to any future smartphone sensing accessory where the measurement benefit justifies a minimal physical augmentation.

5.2.4 Camera PPG Calibration as Foundational Infrastructure

The camera PPG calibration methodology described in Appendix A is the most broadly applicable contribution in the dissertation. Smartphone camera photoplethysmography underlies the blood pressure measurement, and is relevant to pulse oximetry, hemoglobin estimation, and any other clinical application that depends on the ratio-of-ratios of optical absorption across color channels. The finding that default smartphone computational photography pipelines introduce systematic, device-specific, and lighting-dependent nonlinearities that corrupt amplitude information, and that these nonlinearities can be corrected by linear tone mapping and zero light offset calibration, applies to the entire class of smartphone optical biomarker applications. The calibration is performed once per phone model by the application developer, requires no action from the end user, and transfers across phones of the same model. It represents infrastructure that the mobile health research community can adopt across applications, and its absence from prior work is a likely contributor to the inconsistent cross-device performance reported in that literature.

5.3 The Equity Dimension

A recurring theme across the dissertation is the relationship between sensing accuracy and demographic equity. The pupillometry work is the clearest case: prior RGB-based smartphone pupillometers failed for up to 23% of dark-iris participants, a failure rate that is not random noise but a systematic performance disparity driven by the physics of melanin absorption [7]. The NIR and far-red approaches eliminate this disparity at the level of the optical signal rather than by training models on more diverse datasets or by post-hoc correction. The Student's t-test comparing light-iris and dark-iris performance in both systems found no statistically significant difference ($p = 0.55$ for NIR absolute error; no significant difference for far-red across four phone models). The contribution is equity by design, not equity by adjustment.

The blood pressure work demonstrates equity at the algorithm level: the LASSO

model was trained without demographic inputs, and the ANOVA testing for performance differences across ethnicity ($p = 0.22$) and Fitzpatrick skin type ($p = 0.86$) found no significant disparities. The oscillometric method, which depends on mechanical force transmission and blood volume rather than optical tissue properties, is inherently less sensitive to skin tone than optical methods.

The access dimension of equity is equally important, though harder to measure. The populations with the highest hypertension burden, the least access to pupillometry research, and the greatest risk of sarcopenia are disproportionately populations with less access to clinical infrastructure. A measurement that requires a clinical visit is not equitable even if its algorithm treats all skin tones equally. The deployment philosophy of this dissertation, software on existing hardware, with attachments that cost cents and arrive by mail, is an equity argument, not just an engineering choice.

What the dissertation does not yet demonstrate is equity at the deployment level for older adults in real-world conditions. The pupillometry older adult study provided the first evidence in this direction, but the eyelid occlusion problem, smartphone inexperience, and the usability challenges documented in that study make clear that equity of access requires more than equity of algorithm. Designing for older adults from the beginning, not as an afterthought, not as a secondary validation study, but as a first-class design constraint, is the commitment the results argue for.

5.4 Translational Pathway

Each system in this dissertation exists at a different stage of the translational pipeline, and honest accounting of that pipeline is necessary to contextualize both the achievements and the remaining gaps.

The pupillometry systems are furthest from clinical deployment in the traditional sense: neither the NIR nor the far-red system has been evaluated in a clinical population or in a longitudinal study, and neither has been through any regulatory process. Their

translational value at this stage is as research instruments that enable pupillometry studies at scales and frequencies that were not previously feasible, not as clinical diagnostic devices. The non-exclusive license of the far-red patent to Vitamin Greens (May 2024) and the Best Demo Award at UbiComp 2025 indicate commercial interest, but the path from research instrument to FDA-cleared medical device is long and the pupillometry systems are early in it.

The blood pressure system is actively on a translational trajectory. The NIH POC-TRN CAPCaT Innovation Award ($100k, 2024) and the Massachusetts AI and Technology Center A2Collective Pilot Award ($220k, 2023) are funding the next phase of development, which includes the larger-scale clinical study needed for regulatory submission. Billion Labs Inc. holds the provisional patent and is pursuing the FDA SaMD pathway. The Best Demo Award at IEEE BSN 2025 and the Stanford eWEAR Health Prize finalist recognition reflect the system's visibility in the medical device community. The regulatory question is not whether the oscillometric approach is physically valid, that is established, but whether the accuracy and usability can meet the IEEE 1708-2014 standard (MAE $\leq$ 5 mmHg, standard deviation $\leq$ 8 mmHg) at adequate scale and in adequate populations.

The chair stand and handgrip work are at the earliest translational stage: feasibility demonstrated, clinical validation yet to begin. Their translational value is currently as motivating evidence that the measurement direction is productive, rather than as systems approaching deployment. The DXA-validated predictive performance of IMU-derived chair stand features is a finding that will need replication in a sarcopenic population, longitudinal validation, and usability testing before it can inform clinical practice.

Across all three domains, the translational bottleneck is the same: the gap between proof-of-concept performance in a young, healthy, supervised laboratory population and validated performance in the target population, older adults, individuals with the condition being measured, in unsupervised at-home conditions. This is the canonical challenge of mobile health translation, and this dissertation provides evidence that the technical sensing

barriers are not the primary obstacle.

5.5 Honest Limitations of the Body of Work

A dissertation that does not identify its own limitations honestly is incomplete. The following limitations apply across the body of work and are not fully addressed within it.

Sample sizes. The pupillometry validation studies used $N = 11$ participants each. The blood pressure proof-of-concept used $N = 24$ participants with 68 usable measurements. The chair stand study used $N = 78$ participants, which is more substantial, but the sarcopenic range is underrepresented. None of these sample sizes is sufficient for regulatory submission or for establishing population-level screening performance metrics.

Population homogeneity. All studies were conducted primarily with young, healthy, English-speaking adults recruited through university-affiliated facilities. The one exception, the pupillometry older adult usability study, surfaced usability problems that had not appeared in laboratory work. This pattern is predictable: laboratory populations are not representative of deployment populations. The findings are evidence of feasibility, not of generalized performance.

Supervised conditions. Every quantitative performance result in this dissertation was obtained in a supervised laboratory or clinical setting, with researcher oversight and, in most cases, practice measurements to familiarize participants with the protocol. Unsupervised at-home performance is unknown for all three systems. The high exclusion rate in the blood pressure study (49%) was obtained even in supervised conditions, suggesting that unsupervised performance could be substantially worse without significant usability improvements.

iOS exclusion. The pupillometry and blood pressure applications were developed for Android using the Camera2 API, which provides the camera control necessary for linear tone mapping and locked exposure settings. The Apple AVFoundation framework does not expose equivalent controls for video recording. This means that the measurement systems

do not currently function on iPhones, which account for approximately half of the global smartphone market in high-income countries. The iOS limitation is a real constraint on the access-expansion claim and is not resolved within the scope of this work.

The validation-deployment gap. Validation against a clinical gold standard in a controlled setting establishes that the measurement can be accurate, not that it will be accurate in deployment. The deployment gap, the difference in performance between controlled validation and real-world use, is the single largest unknown in the translational picture. Closing it requires the at-home usability studies that are the most important next step for all three application domains.

5.6 Future Directions

The findings of this dissertation motivate a set of research directions that are both natural extensions of the current work and represent meaningful advances in their own right.

Longitudinal at-home studies. The most important single next step across all three domains is deployment of the measurement systems to clinical populations in unsupervised at-home conditions, with longitudinal follow-up sufficient to establish test-retest reliability and to detect clinically meaningful change over time. For pupillometry, this means a multi-week deployment to older adults with known MCI status. For blood pressure, this means a months-long study with hypertensive participants measuring daily or weekly. For the chair stand, this means a study in which sarcopenic participants perform the test before and after a resistance training intervention and the IMU features are tested for sensitivity to body composition change.

Closing the iOS gap. The development of camera calibration and controlled PPG measurement approaches that function within the constraints of iOS AVFoundation is a high-priority engineering goal. One potential path is the use of RAW image capture, which bypasses tone mapping in iOS for still images; whether RAW video capture or equivalent frame-by-frame control can be achieved for the frame rates needed in clinical

PPG measurement is an open technical question.

Paper-based and colorimetric extensions. The postable paradigm and the far-red spectral filtering approach together suggest a natural extension toward paper-based diagnostic strips that are illuminated by the smartphone flashlight and read by the smartphone camera. Urine creatinine and albumin-to-creatinine ratio for chronic kidney disease screening, hemoglobin concentration from a fingerprick blood spot, and cholesterol from a lateral flow strip are all targets where a smartphone-readable paper strip would provide access to a clinically meaningful measurement at essentially zero per-test cost. The optical calibration work in Appendix A is direct technical preparation for this class of measurement.

Integration across modalities. Each measurement in this dissertation was developed and validated independently. A clinically meaningful application may require multiple measurements within a single assessment session. A smartphone-based sarcopenia screening session that includes chair stand, handgrip, and spirometry, all performed within five minutes, without any clinical visit, is technically feasible given the systems demonstrated here. A future integrated system could also combine pupillometry with cognitive battery tests to create a richer neurological assessment profile. The technical infrastructure for such integration is largely present; what is required is usability engineering for multi-test sessions and clinical validation of composite assessments.

Standardization and regulatory frameworks. The smartphone medical sensing field currently lacks standardized validation protocols, performance benchmarks, and regulatory pathways that are specific to software-only or minimal-attachment medical sensing on commodity hardware. The tenets proposed in the introduction of this dissertation, universal hardware anchoring, gold standard validation, equity-first design, and deployment-centered development, are proposed as design principles, not regulatory standards. The development of such standards, in collaboration with regulatory bodies, clinical professional societies, and device manufacturers, is a community-level challenge that the results of this dissertation help motivate.

5.7 Closing Statement

The smartphone in a person's pocket already contains the sensors needed to measure their pupil, their blood pressure, and their functional strength. This dissertation demonstrates that accessing those measurements does not require new hardware. It requires new uses of the hardware we already carry, combined with new ways of thinking about what counts as a clinical instrument, who counts as the target user, and what counts as adequate validation.

The vision articulated in the introduction, an app store of medical sensors, in which a developer can write an application that transforms existing devices into clinical instruments, has not been fully realized by the work in this dissertation. The technical demonstrations here are proofs of concept, not deployed systems. The regulatory pathways are open, not cleared. The populations most likely to benefit have been studied in small numbers, under supervision, in settings that do not represent their lives.

But the dissertation does establish something important: the barrier is not the hardware. The sensors are sufficient. The physics works. The access-expansion goal is achievable. The biomarker discovery goal is productive. The usability goal is tractable, even if it has not yet been achieved at scale.

The smartphone in a person's pocket has always been, in principle, a medical instrument. The work required to make that principle a reliable clinical reality is the work that follows from this dissertation.

Appendix A

Camera Photoplethysmography Calibration

A.1 Motivation and Background

Camera photoplethysmography (PPG) uses a smartphone camera to detect the cardiac pulse by recording the variation in light absorbed by the fingertip or other tissue over time. In principle, this is straightforward: more blood in the tissue during systole means more light absorbed and a lower camera signal; less blood during diastole means less absorption and a higher signal. In practice, the relationship between the true optical signal and the value recorded by the camera is distorted by two systematic nonlinearities present in the image signal processing (ISP) pipelines of virtually all modern smartphones.

These nonlinearities do not merely add noise to the PPG signal. They introduce systematic, device-specific biases that are sensitive to the ambient light environment and to the brightness of the scene being imaged. For the oscillometric blood pressure measurement described in Chapter 3, the consequence is that the pulse amplitude recorded at different levels of applied fingertip pressure—which determines the shape of the oscillogram—is corrupted by ISP-dependent artifacts that cannot be separated from the true physiological signal without calibration. This appendix describes the two nonlinearities, the calibration procedures used to correct them, and their impact on PPG signal quality.

The calibration methodology was developed and validated by Xuan, Barry *et al.* [8], and the description here follows that work closely. The calibration is applied as a preparatory step before every blood pressure measurement in Chapter 3 and is the foundational infrastructure for any smartphone optical health sensing application that requires accurate amplitude information.

A.2 The Two ISP Nonlinearities

A.2.1 Tone Mapping

The first nonlinearity is tone mapping. The raw sensor output of a smartphone camera has a linear relationship with the number of photons striking each pixel: double the light, double the pixel value. Smartphone ISP pipelines apply a nonlinear tone mapping curve to the raw sensor data before the image is made available to applications, converting the linear sensor response to a perceptually pleasing, contrast-enhanced output. The tone mapping curve is typically a sigmoid-like function that compresses both shadows and highlights while boosting midtone contrast.

For photographic purposes, tone mapping is beneficial. For PPG measurement, it is destructive. The PPG signal is a small fractional change ($< 1\%$) in the overall brightness of the fingertip image. The derivative of a nonlinear tone mapping curve at any given brightness level determines how the PPG signal amplitude is scaled. Because the derivative varies with brightness, the effective gain applied to the PPG signal changes as the fingertip brightness changes— which is precisely what happens during an oscillometric blood pressure measurement as the user presses harder and the tissue is increasingly occluded. The result is that the observed pulse amplitude at high applied pressure is scaled differently than the pulse amplitude at low applied pressure, distorting the oscillogram in a way that cannot be predicted or corrected without knowledge of the tone mapping curve.

A.2.2 Zero Light Offset

The second nonlinearity is the zero light offset (ZLO). In an ideal camera, the pixel value recorded when no light reaches the sensor is zero. In real sensors, a fixed positive offset is added by the readout electronics, so that the recorded value at zero light is a small but nonzero positive number. This offset is subtracted in the ISP pipeline when the camera is operating in its default mode; however, the subtracted value is specific to the sensor hardware and operating temperature, and is not necessarily equal to the true ZLO for a given

device.

For PPG measurement, the ZLO introduces a systematic error in the DC amplitude of the recorded signal. Because PPG signal ratios across color channels are used in some applications to estimate tissue oxygenation or to reduce motion artifact, a device-specific ZLO that differs between the red, green, and blue channels introduces cross-channel biases that cannot be corrected by a global brightness calibration. The ZLO is stable within a measurement session but varies across devices and across firmware versions on the same device.

A.3 Calibration Procedure

The calibration procedure corrects both nonlinearities in a single protocol that is performed by the application developer once per phone model. The procedure requires the phone whose camera is to be calibrated and a neutral density filter set.

A.3.1 Tone Mapping Linearization

To characterize the tone mapping curve, a series of images is captured with the camera aimed at a uniform white surface under controlled illumination, with the camera exposure time varied across a range from the minimum available exposure to the maximum. For each exposure time, the mean pixel value across the image is recorded for each color channel. The resulting curve—mean pixel value as a function of exposure time—encodes the ISP tone mapping function. If the tone mapping were linear, this curve would be a straight line through the origin with slope proportional to the sensor gain; deviations from linearity indicate the tone mapping nonlinearity.

The tone mapping curve is then used to select a fixed exposure time within the linear region of the curve— typically in the lower portion of the sensor's dynamic range, where the tone mapping is approximately linear for most devices. For the PPG measurement, the camera is locked to this fixed exposure time by setting the Camera2 API to manual exposure

control. With a fixed exposure in the linear region, the relationship between the true optical signal and the recorded pixel value is approximately linear, and changes in pixel value over time accurately reflect changes in the amount of light absorbed by the fingertip.

A.3.2 Zero Light Offset Measurement

To measure the ZLO, a series of images is captured with the camera lens completely covered (no light reaching the sensor). The mean pixel value across these dark images, for each color channel, is the ZLO. This value is stored as a device-specific calibration constant. During PPG measurement, the ZLO is subtracted from each frame before the PPG signal is extracted, ensuring that the baseline pixel value when there is no light signal is zero rather than the device-specific offset.

A.3.3 Pre-Measurement Exposure Calibration

The tone mapping linearization procedure described above identifies a fixed exposure time that works in the linear region for average imaging conditions. However, the appropriate exposure also depends on the brightness of the tissue being imaged, which varies with individual skin tone, tissue thickness, and the camera-to-tissue geometry. A fixed exposure that is appropriate for one individual may result in signal saturation (too bright) or noise dominance (too dim) for another.

To address this, a brief adaptive exposure calibration is performed at the beginning of each PPG measurement session. With the user's finger placed over the front camera, the application captures a short sequence of frames at multiple exposure times and selects the exposure that places the mean fingertip brightness in the target range—neither saturated nor below the noise floor. This exposure time is then used for the entire measurement session. The adaptive calibration adds approximately three seconds to the start of the measurement and requires no action from the user beyond placing their finger on the camera.

A.4 Validation and Impact on PPG Quality

The impact of the calibration on PPG signal quality was evaluated in a study comparing calibrated and uncalibrated PPG signals across multiple smartphone models and participants spanning a range of Fitzpatrick skin types [8]. The key metrics were the signal-to-noise ratio (SNR) of the PPG waveform, the linearity of the pulse amplitude across different simulated brightness conditions, and the cross-device consistency of the PPG amplitude for a given physiological condition.

Without calibration, the pulse amplitude recorded at higher simulated applied pressure (lower transmitted light through the fingertip) was systematically underestimated relative to the pulse amplitude at lower applied pressure, due to the compressive tone mapping in the low-brightness region. This distortion corrupted the oscillogram shape and degraded blood pressure prediction performance. With calibration applied, the pulse amplitude varied linearly with the true physiological signal across the full range of applied pressures, producing oscillograms with the expected bell-shaped profile and significantly improved blood pressure prediction performance.

Cross-device consistency also improved substantially with calibration. Uncalibrated PPG amplitude varied by a factor of approximately three across the smartphone models tested, due to differences in tone mapping curves and ZLO values. After calibration, the variation was reduced to approximately 15%, indicating that the calibrated signal is largely independent of device-specific ISP behavior.

For participants with darker skin tones (Fitzpatrick Types IV–VI), the uncalibrated system required higher exposure times that placed the signal in the nonlinear region of the tone mapping curve for some devices, producing corrupted oscillograms. After calibration and adaptive exposure selection, signal quality for darker skin tone participants was comparable to lighter skin tone participants on all devices tested, confirming the equity benefit of the calibration approach.

A.5 Implementation Notes

The calibration is implemented in the VibroBP Android application using the Camera2 API. The tone mapping linearization and ZLO measurement are performed during application development and stored as device-specific constants embedded in the application. The per-session adaptive exposure calibration runs automatically at the start of each measurement. Users see only a brief prompt to place their finger on the camera; the calibration logic runs transparently in the background.

For researchers or developers wishing to apply this calibration to new applications or new phone models, the full calibration protocol and the code used to characterize the tone mapping curve are available in the supplementary materials of [8]. The calibration is designed to generalize to any Camera2-compatible Android device; iOS compatibility requires adaptation to the AVFoundation API and RAW capture modes, which do not provide equivalent exposure control for video capture.

Appendix B

IRB Protocols and Study Demographics

This appendix consolidates the institutional review board (IRB) approvals and participant demographic summaries for the human subjects studies reported in Chapters 2, 3, and 4. All studies were conducted in accordance with the Declaration of Helsinki. Informed written consent was obtained from all participants prior to data collection.

Some fine-grained demographic subcategories reported below are marked N/R (not tabulated in the original study records available for this appendix); readers seeking this additional detail may contact the author.

B.1 Chapter 2: Pupillometry Studies

B.1.1 NIR Validation and Far-Red Contrast Studies

The NIR pupillometry validation study and the far-red cross-device contrast study were conducted under UC San Diego IRB approval. Eleven participants were recruited for the NIR validation study (five blue iris, one hazel, three brown, two dark brown); twelve were recruited for the far-red contrast and accuracy validation (one excluded due to equipment error, leaving $N = 11$ in the analysis). Both studies recruited young adults from the UCSD campus community with no known ocular pathology and normal or corrected-to-normal visual acuity. Participants were excluded for: active ocular infection, history of iris-altering ocular surgery, use of medications known to affect pupil size (including anticholinergics, opioids, and certain antidepressants), and inability to remain still for the duration of the test.

Table B.1: Pupillometry validation study participant demographics.

	NIR Study	Far-Red Study
N (analysed)	11	11
Age, mean ± SD (years)	N/R	
Sex (% female)	N/R%	N/R%
Iris color		
Blue	5	N/R
Hazel	1	N/R
Brown	3	N/R
Dark brown	2	N/R
Fitzpatrick skin type		
Types I–II	N/R	N/R
Types III–IV	N/R	N/R
Types V–VI	N/R	N/R

B.1.2 Older Adult Usability Deployment

The older adult usability study was conducted under UC San Diego IRB protocol #210437X. Fifteen participants aged 55 and older were recruited from an ongoing Alzheimer's cognition study at UC San Diego. Inclusion criteria required English literacy, age 55 or older, and willingness to participate in a remote video-call-guided self-administration session. Exclusion criteria included active ocular infection, known severe visual impairment uncorrectable by glasses, and inability to use a touchscreen smartphone.

Table B.2: Older adult usability study participant demographics.

	Usability Study ($N = 15$)
Age, mean ± SD (years)	N/R
Age range (years)	N/R–N/R
Sex (% female)	N/R%
Fitzpatrick type, mean ± SD	N/R
Smartphone experience	
Comfortable	N/R
Some experience	N/R
Limited experience	N/R

B.1.3 MCI Biomarker Analysis: VETSA Dataset

The MCI biomarker analysis used existing data from the Vietnam Era Twin Study of Aging (VETSA), a longitudinal study conducted at UC San Diego with long-standing IRB approval. The VETSA dataset used in the Chapter 2 analysis contained pupillary response recordings and neuropsychological MCI classifications from approximately $N = 900$ male participants. Participants were classified under the Jak/Bondi neuropsychological criteria as cognitively normal (CN), non-amnestic single-domain MCI (NA S-MCI), amnestic single-domain MCI (A S-MCI), or multi-domain MCI (M-MCI). No new data collection was conducted for this analysis; secondary use of existing data was covered by the standing VETSA IRB approval. Full VETSA demographic information is available in the primary VETSA publications.

B.2 Chapter 3: Blood Pressure Study

The blood pressure proof-of-concept study was conducted under UC San Diego IRB protocol 804668. Thirty participants were recruited from the UCSD campus community; six were excluded from the analysis (one unable to perform valid force measurements, one with low pulse prominence, one with poor oscillogram fit, three meeting a mixture of criteria), leaving $N = 24$ in the analysis. Inclusion criteria required age 18–65, ability to use a touchscreen smartphone, and no known cardiovascular disease, arrhythmia, or peripheral vascular disease. Exclusion criteria included pregnancy, active use of vasoactive medications, Raynaud's phenomenon, and any condition causing chronic upper-extremity pain.

B.3 Chapter 4: Musculoskeletal Studies

B.3.1 30-Second Chair Stand Study

The chair stand study was conducted at the Exercise and Physical Activity Resource Center (EPARC) at UC San Diego under IRB approval #811902. Eighty-five participants

Table B.3: Blood pressure study participant demographics ($N = 24$).

Characteristic	Value
Age, mean $\pm$ SD (years)	24.7 ± 4.3
Sex	50% male, 50% female
Ethnicity	
White	20%
Asian	67%
Hispanic	13%
Fitzpatrick skin type	
Type II	N/R%
Type III	N/R%
Type IV	N/R%
Type V	N/R%
Resting systolic BP, mean $\pm$ SD (mmHg)	103.0 ± 15.0
Resting diastolic BP, mean $\pm$ SD (mmHg)	70.0 ± 10.7
Systolic BP range (mmHg)	89–147
Exercise measurements (N)	23 consented; 15 included

were recruited from individuals scheduled for clinical DXA scans at EPARC; after excluding trials with missing IMU data or failing internal consistency checks, $N = 78$ were included in the analysis. Inclusion criteria required English literacy, age 18 or older, and willingness to complete grip strength and DXA assessments. Exclusion criteria included neuromuscular disorders affecting lower extremity function, conditions impairing hand function, recent lower-extremity injuries (within six months), and use of medications known to affect muscle performance.

B.3.2 Handgrip Strength Study

Unlike the pupillometry, blood pressure, and chair stand studies described above, the smartphone handgrip strength (HGS) system was developed and validated through benchtop instrumentation rather than a clinical human-subjects study. The Google Pixel 4 vibration-damping force model was profiled against a calibrated benchtop reference force sensor (SingleTact, 15 mm, 100 N) rather than a clinical dynamometer or a cohort of enrolled participants; force was applied directly to the sensor-instrumented phone across fifteen

Table B.4: Chair stand study participant demographics ($N = 78$).

Characteristic	Value
Age, mean ± SD (years)	51.5 ± 16.4
Sex (% female)	73.1%
BMI, mean ± SD (kg/m^2)	24.3 ± 4.5
Height, mean ± SD (cm)	N/R
Weight, mean ± SD (kg)	N/R
30s CST count, mean ± SD	16.9 ± 5.5
30s CST count range	9–30
Total lean mass, mean ± SD (lbs)	N/R
LBMI, mean ± SD (kg/m^2)	N/R
Percent body fat, mean ± SD (%)	N/R
Dominant HGS, mean ± SD (kg)	N/R
Non-dominant HGS, mean ± SD (kg)	N/R
DXA system used	
Hologic Horizon W	N/R participants
GE Lunar Prodigy	N/R participants
EWGSOP2 classification	
Normal muscle mass (LBMI $\geq$ cutoff)	100%
Low muscle mass (LBMI $<$ cutoff)	0%
Probable sarcopenia (low LBMI + low HGS)	0%

40-second profiling sessions spanning a range of force trajectories (linear ramps, sinusoids, and square waves) to characterize the device's vibration-damping response [14]. Because this procedure did not involve identifiable human-subjects data collection, it did not require separate IRB review. Readers seeking further protocol detail may contact the author.

Appendix C

Signal Processing Code Documentation

This appendix documents the key signal processing pipelines implemented for the three application chapters. The goal is to provide sufficient algorithmic detail to enable independent replication of the core analyses. Full source code for the Android applications and the Python analysis pipelines is available through the Wang Lab GitHub repository and through the supplementary materials of the corresponding publications.

C.1 Pupillometry Processing Pipeline

C.1.1 NIR Pupil Detection and Distance Estimation

The NIR pupillometry pipeline consists of four sequential stages: frame acquisition, pupil detection, distance estimation, and diameter calculation.

Frame acquisition. NIR video is recorded at 30 fps using the Android Camera2 API with the YCbCr_420_888 format. A simultaneous RGB still image is captured at the beginning of the recording for distance calibration. The proximity sensor is polled at 10 Hz; if the reading exceeds a threshold indicating the phone has been moved more than 5 cm from the eye, recording is paused.

Pupil detection. Each NIR frame is passed to the Eivazi *et al.* convolutional neural network [22], which outputs the pupil center $(x^*_{\mathrm{NIR}}, y^*_{\mathrm{NIR}})$ and radius r_{px} in pixels. The network was pretrained on the OpenEDS and LPW datasets and runs at approximately 15 fps on the Pixel 4's neural processing unit. The RGB still image is processed by the same network to extract x^*_{RGB}.

Distance estimation. The horizontal disparity between normalized pupil centers is:

$$D = \left| \frac{x^*_{\text{RGB}}}{W_{\text{RGB}}} - \frac{x^*_{\text{NIR}}}{W_{\text{NIR}}} \right| \tag{C.1}$$

The camera-to-eye distance is $Z = K/D$, where K is a device-specific calibration constant determined empirically on a 3D-printed stair structure.

Diameter calculation. The absolute pupil diameter is:

$$d = 2\,r_{\text{px}} \times \frac{2Z\tan(\theta/2)}{W_{\text{NIR}}} \tag{C.2}$$

where θ is the NIR camera's horizontal field of view. Blink-contaminated frames (detected by a sudden drop in detected radius $> 30\%$ between consecutive frames) are linearly interpolated over.

C.1.2 Far-Red PPG-Based Pupil Tracking

The far-red pupillometry pipeline follows the same neural network detection step, applied to red-channel frames from the smartphone RGB camera captured through the 630 nm longpass filter. Pupil diameter is reported as percent change from the mean pupil size across the recording:

$$\Delta d(t) = \frac{d_{\text{px}}(t) - \bar{d}_{\text{px}}}{\bar{d}_{\text{px}}} \times 100\% \tag{C.3}$$

No distance estimation is performed; the output is a relative trace comparable to the Neuroptics device's normalized PLR output. Downsampling from 60 fps to 30 fps uses a triangular anti-aliasing filter before comparison with the reference device.

C.2 Blood Pressure Processing Pipeline

C.2.1 PPG Signal Extraction

After camera calibration (Appendix A), raw frames are saved as 8-bit YUV. The green channel of the fingertip region of interest (ROI, approximately 50×50 pixels centered on the camera) is averaged per frame to produce the raw PPG time series $s(t)$ at the camera frame rate (30 fps default).

Preprocessing applies a 4th-order Butterworth bandpass filter (0.5–4.0 Hz) to isolate the cardiac frequency band. The high-pass cutoff removes DC offset and slow baseline drift; the low-pass cutoff attenuates high-frequency noise above the maximum physiological heart rate. Pulse peaks are detected using a modified Pan-Tompkins algorithm adapted for fingertip PPG: a moving window peak detector with a minimum inter-peak interval of 0.35 s (corresponding to a maximum heart rate of 171 bpm) and a minimum prominence threshold of 1.5 times the local noise floor.

C.2.2 Vibrometric Force Estimation

IMU data are recorded at 400 Hz using the Android `SensorManager` with `TYPE_ACCELEROMETER` and `TYPE_GYROSCOPE` listeners. Synchronization with the camera frames is achieved by recording the system timestamp of each IMU sample and each camera frame and aligning by timestamp after collection.

The rolling standard deviation of each IMU axis is computed over a 100 ms window (40 samples at 400 Hz) at 30 Hz output rate. Feature selection identifies the axes most correlated with the reference force sensor during the developer calibration procedure. For the Pixel 4, the selected features are the y-axis accelerometer rolling standard deviation (perpendicular to the phone face) and the z-axis gyroscope rolling standard deviation. For the Galaxy A53 and Moto G Power, device-specific feature subsets were selected by the same procedure.

The force estimation model is a multivariate linear regression (Pixel 4, A53) or

gradient boosting regressor (Moto G Power), trained using scikit-learn. Training used the developer calibration data (five participants, three 40-second sessions each), with the SingleTact force sensor output (sampled at 200 Hz, downsampled to 30 Hz) as the target. Model coefficients are embedded as constants in the Android application.

C.2.3 Oscillogram Construction and Feature Extraction

The pulse amplitude sequence A_i (one value per heartbeat) and the simultaneously estimated applied pressure P_i (in mmHg, converted from Newtons using $P = F/A_{\text{finger}}$ with $A_{\text{finger}} = 1.5$ cm^2) define the empirical oscillogram. A skewed Gaussian is fit to (P_i, A_i) using the `scipy.optimize.curve_fit` function with the functional form:

$$A(P) = A_{\max} \exp\left[-\frac{(P-\mu)^2}{2\sigma^2}\right] \left(1 + \operatorname{erf}\left[\frac{\alpha(P-\mu)}{\sigma\sqrt{2}}\right]\right) \tag{C.4}$$

where μ, σ, α, and $A_{\max}$ are free parameters. The eleven oscillometric features described in Section 3.4.3 are extracted from this fit.

The LASSO regression models are trained and evaluated in Python using `sklearn.linear_model.Lasso` with α selected by 5-fold cross-validation on the training set within each LOO fold.

C.3 Chair Stand Processing Pipeline

C.3.1 IMU Preprocessing

Raw IMU data are recorded by the Google Pixel 8 at approximately 400 Hz using the Android Sensor API. Data are exported as CSV files containing timestamp, triaxial accelerometer, and triaxial gyroscope readings.

Processing applies the following sequence:

1. *Linear interpolation* to a uniform 400 Hz grid, filling gaps of up to 5 samples.
2. *Gravity estimation*: median of the first 1.0 s of accelerometer data gives the gravity

vector; a rotation matrix is computed to align the vertical axis with $[0, 0, 9.81]$ m/s^2.

3. *Butterworth low-pass filter*: 4th-order, 2 Hz cutoff, applied to all axes after gravity reorientation.

4. *Numerical integration*: trapezoidal integration of the vertical acceleration to obtain velocity; 0.1 Hz low-pass filter applied to velocity to estimate and subtract integration drift. Second integration yields position.

5. *Resultant computation*: magnitude of the 3-axis acceleration, velocity, and position vectors.

C.3.2 Cycle Detection

Stand peaks are detected in the resultant position signal using `scipy.signal.find_peaks` with minimum prominence 0.2 and minimum inter-peak interval 0.8 s. Boundary conditions retain the first and last detected peak even if they fall within 0.4 s of the start or end of the trial. Per-cycle windows span $[-0.4, +1.2]$ s around each peak.

C.3.3 Feature Computation

All 37 features across the six movement categories (speed, power/impulse, smoothness, fatigue, kinematics, cycle consistency) are computed in Python using NumPy, SciPy, and a custom inverse kinematics module. The IK module implements the two-link planar chain described in Section 4.4.5 using participant height (recorded at DXA visit) and the fixed anthropometric ratios $\ell_{\text{thigh}} = 0.345\,h$ and $\ell_{\text{torso}} = 0.50\,h$.

C.3.4 Model Training and Evaluation

ElasticNet models are fit using `sklearn.linear_model.ElasticNetCV` with 3-fold inner cross-validation, `l1_ratio` in $[0.1, 0.5, 0.7, 0.9, 0.95, 1.0]$, and `n_alphas = 100`. Feature standardization uses `sklearn.preprocessing.StandardScaler` fit on the training fold only. All feature selection (via ElasticNet's ℓ_1 penalty) occurs inside each LOO fold.

Reported metrics (LOO R^2, Pearson r, MAE) are computed on the out-of-fold predictions only.

Stability selection uses `sklearn.linear_model.Lasso` across 50 bootstrap resamples (75% subsampling, without replacement) with α fixed at the mean cross-validated value from ElasticNet training. Forward stepwise selection uses ordinary least squares (`sklearn.linear_model.LinearRegression`) with LOO R^2 as the selection criterion.

References

[1] K. T. Mills, J. D. Bundy, T. N. Kelly, J. E. Reed, P. M. Kearney, K. Reynolds, J. Chen, and J. He. “Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries”. en. In: *Circulation* 134.6 (Aug. 2016), pp. 441–450.

[2] M. D. M. Larson and M. M. Behrends. “Portable Infrared Pupillometry”. In: *Anesthesia and Analgesia* 120 (June 2015).

[3] A. J. Fascetti, A. Pestonjamasp, C. Barry, and E. J. Wang. “Postable Spirometry: User-Assembled Vortex Whistle Spirometer”. In: *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Orlando, FL, USA: IEEE, July 15, 2024.

[4] S. Von Haehling, J. E. Morley, and S. D. Anker. “An overview of sarcopenia: facts and numbers on prevalence and clinical impact”. In: *Journal of Cachexia, Sarcopenia and Muscle* 1.2 (Dec. 2010), pp. 129–133.

[5] A. Mariakakis, J. Baudin, E. Whitmire, V. Mehta, M. A. Banks, A. Law, L. McGrath, and S. N. Patel. “PupilScreen: Using Smartphones to Assess Traumatic Brain Injury”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1.3 (2017), p. 81.

[6] A. Chandrasekhar, C.-S. Kim, M. Naji, K. Natarajan, J.-O. Hahn, and R. Mukkamala. “Smartphone-based blood pressure monitoring via the oscillometric finger-pressing method”. en. In: *Science Translational Medicine* 10.431 (Mar. 2018), eaap8674.

[7] A. E. Neice, C. Fowler, R. A. Jaffe, and J. G. Brock-Utne. “Feasibility study of a smartphone pupillometer and evaluation of its accuracy”. In: *Journal of Clinical Monitoring and Computing* (2020).

[8] Y. Xuan, C. Barry, N. Antipa, and E. J. Wang. “A calibration method for smartphone camera photophlethysmography”. In: *Frontiers in Digital Health* 5 (2023).

[9] C. Barry, J. de Souza, Y. Xuan, J. Holden, E. Granholm, and E. J. Wang. “At-Home Pupillometry using Smartphone Facial Identification Cameras”. In: *CHI Conference on Human Factors in Computing Systems*. CHI ’22: CHI Conference on Human Factors in Computing Systems. New Orleans LA USA: ACM, Apr. 29, 2022, pp. 1–12.

[10] C. Barry and E. J. Wang. “Racially fair pupillometry measurements for RGB smartphone cameras using the far red spectrum”. In: *Scientific Reports* 13.1 (2023), p. 13841.

[11] C. Barry, Y. Xuan, A. Fascetti, A. Moore, and E. J. Wang. “Oscillometric blood pressure measurements on smartphones using vibrometric force estimation”. In: *Scientific Reports* 14.1 (Oct. 31, 2024), p. 26206.

[12] Y. Xuan, C. Barry, J. De Souza, J. H. Wen, N. Antipa, A. A. Moore, and E. J. Wang. “Ultra-low-cost mechanical smartphone attachment for no-calibration blood pressure measurement”. In: *Scientific Reports* 13.1 (May 29, 2023), p. 8105.

[13] C. Barry, E. J. Wang, and D. Wing. “Smartphone Instrumented 30-Second Chair Stand Recovers Body Composition and Strength Biomarkers”. Manuscript in preparation. 2026.

[14] C. Barry and E. J. Wang. “Smartphone-Based Hand Grip Strength Measurement”. Manuscript in preparation. 2026.

[15] S. Mathôt. “Pupillometry: Psychology, Physiology, and Function”. In: *Journal of Cognition* (Feb. 2018).

[16] P. van der Wel and H. van Steenbergen. “Pupil dilation as an index of effort in cognitive control tasks: A review”. In: *Psychonomic Bulletin and Review* (2018).

[17] G. Micieli, C. Tassorelli, E. Martignoni, C. Pacchetti, P. Bruggi, M. Magri, and G. Nappi. “Disordered pupil reactivity in Parkinson’s disease”. In: *Clinical Autonomic Research* (1991), pp. 55–58.

[18] R. LF, R. EA, W. JK, and G. MF. “Pupillary responses to a cognitive effort task in schizophrenia”. In: *Schizophrenia Research* (Mar. 2018).

[19] E. L. Granholm, M. S. Panizzon, J. A. Elman, A. J. Jak, R. L. Hauger, M. W. Bondi, M. J. Lyons, C. E. Franz, and W. S. Kremen. “Pupillary Responses as a Biomarker of Early Risk for Alzheimer’s Disease”. In: *Journal of Alzheimer’s Disease* 56 (2017).

[20] K. WS, P. MS, E. JA, G. EL, A. OA, G. N. Dale AM, G. DE, L. MW, L. MJ, N. MC, R. CA, W. N, and F. CE. “Pupillary dilation responses as a midlife indicator of risk for Alzheimer’s disease: association with Alzheimer’s disease polygenic risk”. In: *Neurobiology of Aging* (2019).

[21] E. JA, P. MS, H. D. Jr, E. LT, G. EL, F.-N. C, L. MJ, M. LK, F. CE, D. AM, and K. WS. “Task-evoked pupil dilation and BOLD variance as indicators of locus coeruleus dysfunction. Cortex”. In: *Cortex* (2017).

[22] S. Eivazi, T. Santini, A. Keshavarzi, T. Kübler, and A. Mazzei. “Improving real-time CNN-based pupil detection through domain-specific data augmentation”. In: *In Proceedings of the 11th ACM Symposium on Eye Tracking Research and Applications*. ETRA ’19. New York, NY, USA: Association for Computing Machinery, 2019.

[23] N. Kollias. “The Spectroscopy of Human Melanin Pigmentation”. In: (Jan. 1995).

[24] D. Shimbo, N. T. Artinian, J. N. Basile, L. R. Krakoff, K. L. Margolis, M. K. Rakotz, G. Wozniak, and On behalf of the American Heart Association and the American Medical Association. “Self-Measured Blood Pressure Monitoring at Home: A Joint Policy Statement From the American Heart Association and American Medical Association”. en. In: *Circulation* 142.4 (July 2020).

[25] R. Mukkamala, J.-O. Hahn, O. T. Inan, L. K. Mestha, C.-S. Kim, H. Toreyin, and S. Kyal. “Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice”. In: *IEEE Transactions on Biomedical Engineering* 62.8 (Aug. 2015), pp. 1879–1901.

[26] A. Chandrasekhar, K. Natarajan, M. Yavarimanesh, and R. Mukkamala. “An iPhone Application for Blood Pressure Monitoring via the Oscillometric Finger Pressing Method”. en. In: *Scientific Reports* 8.1 (Dec. 2018), p. 13136.

[27] R. Tibshirani. “Regression shrinkage and selection via the lasso: a retrospective: Regression Shrinkage and Selection via the Lasso”. en. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73.3 (June 2011), pp. 273–282.

[28] P. K. Whelton, R. M. Carey, W. S. Aronow, D. E. Casey, K. J. Collins, C. Dennison Himmelfarb, S. M. DePalma, S. Gidding, K. A. Jamerson, D. W. Jones, E. J. MacLaughlin, P. Muntner, B. Ovbiagele, S. C. Smith, C. C. Spencer, R. S. Stafford,

S. J. Taler, R. J. Thomas, K. A. Williams, J. D. Williamson, and J. T. Wright. “2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines”. In: *Hypertension* 71.6 (June 2018).

[29] A. J. Cruz-Jentoft, J. P. Baeyens, J. M. Bauer, Y. Boirie, T. Cederholm, F. Landi, F. C. Martin, J.-P. Michel, Y. Rolland, S. M. Schneider, E. Topinková, M. Vandewoude, and M. Zamboni. “Sarcopenia: European consensus on definition and diagnosis”. In: *Age and Ageing* 39.4 (July 1, 2010), pp. 412–423.

[30] J. M. Rijk, P. R. Roos, L. Deckx, M. Van Den Akker, and F. Buntinx. “Prognostic value of handgrip strength in people aged 60 years and older: A systematic review and meta-analysis”. In: *Geriatrics & Gerontology International* 16.1 (Jan. 2016), pp. 5–20.

[31] T. Kera, H. Kawai, J. Takahashi, H. Hirano, Y. Watanabe, Y. Fujiwara, K. Ihara, H. Kim, and S. Obuchi. “Development of a screening formula for sarcopenia using ground reaction force during sit-to-stand motion”. In: *Gait & Posture* 93 (Mar. 2022), pp. 177–182.

[32] A. Kanai, S. Kiyama, H. Goto, H. Tomita, A. Tanaka, M. Kunimi, T. Okada, and T. Nakai. “Use of the sit-to-stand task to evaluate motor function of older adults using telemetry”. In: *BMC Geriatrics* 16.1 (Dec. 2016), p. 121.

[33] L. Meulemans, J. Alcazar, L. M. Alegre, S. Dalle, K. Koppo, J. Seghers, C. Delecluse, and E. Van Roie. “Sensor- and equation-based sit-to-stand power: The effect of age and functional limitations”. In: *Experimental Gerontology* 179 (Aug. 2023), p. 112255.

[34] J. M. Guralnik, E. M. Simonsick, L. Ferrucci, R. J. Glynn, L. F. Berkman, D. G. Blazer, P. A. Scherr, and R. B. Wallace. “A Short Physical Performance Battery Assessing Lower Extremity Function: Association With Self-Reported Disability and Prediction of Mortality and Nursing Home Admission”. In: *Journal of Gerontology* 49.2 (Mar. 1, 1994), pp. M85–M94.

[35] K. Ota, J. Yoon, K. Shibuya, M. Miyashita, T. Miyama, and T. Okura. “GROUND REACTION FORCE IN SIT-TO-STAND MOVEMENT PREDICT SARCOPENIA”. In: *Innovation in Aging* 8 (Supplement_1 Dec. 31, 2024), pp. 1248–1248.

[36] A. Zijlstra, J. H. Goosen, C. C. Verheyen, and W. Zijlstra. “A body-fixed-sensor based analysis of compensatory trunk movements during unconstrained walking”. In: *Gait & Posture* 27.1 (Jan. 2008), pp. 164–167.