Mobile Health Sensing for Widespread Disease Screening and Chronic Management

Edward Jay Wang

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Shwetak Patel, Chair

Robert Darling

Anind Dey

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

Abstract

Mobile Health Sensing for Widespread Disease Screening and Chronic Management

Edward Jay Wang

Chair of the Supervisory Committee:
Professor Shwetak Patel
Allen School of Computer Science & Engineering
Department of Electrical & Computer Engineering

Medical care today revolves mainly around the clinic, with devices and care designed with accuracy as the main metric of success. Many medical tests can be accomplished with just a small vial of blood, and in the ICU we can probe a variety of physiological parameters continuously. In the search for accuracy, accessibility, however, has often taken a backseat. The inability to spread medical sensing out of high-resource regions leads to many medical scenarios that are difficult to care for. A promising approach to increasing people's access to medical sensing and care in diverse situation is to leverage the computing and sensing infrastructure that has been laid out through the adoption of smartphones in the world. I explore using the smartphone as the platform to build a variety of Software as a Medical Device (SaMD), creating solutions for (1) total hemoglobin measurement, (2) personal blood pressure tracking, and (3) Intraocular pressure. In this dissertation, I show that SaMDs enabled by smartphoneś built-in sensors has the potential for widespread disease screening and chronic disease and health management through balancing trade-offs between optimizing the sensing solution of the biomarker and addressing practical limitations of designing on real world smartphone devices.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Thank you so much to everyone. Reaching this milestone in my life would certainly not have been possible without the support, the guidance, the gripes we've shared over drinks (you know who you are), and most importantly the good times I've shared with all those around me who made this a fun and memorable journey. Doing a PhD was not really in my life plans, but I dare say, the decision to work on this PhD (and finish it) will probably be the thing that will define me more than anything else in my life. So thank you all for helping me define who I am.

Working through this PhD has been a journey for sure. One that I would have given up on many times if it wasn't for the support of my wife Chelsea. Her love and support pushed me through every doubtful moment. Her understanding allowed me to be selfish during stressful times. Her interest in what I actually work on made me a much better communicator of my own work. Her willingness to help me in my work made it possible for me to always have a helping hand whenever I needed one. Her love for travel made it particularly fun to go to conferences as we got to see the world together. And when it was time to finish up and find a job, her selflessness to let me go into a job search not knowing where we might end up was truly invaluable. This PhD degree belongs to me just as much as it belongs to her.

To my mentors, thank you all for giving me the resources and guidance to help me through this process and develop my own independence as a researcher and innovator. Thank you to all my professors at Harvey Mudd who taught me how to learn to learn. It is without a doubt the most useful skill that's empowered me to take on the PhD. To Shwetak, it's been great working with you and having a front row seat observing how you're pushing the world of technology and healthcare to change. To Anind, you've pushed me to examine my research direction deeper than anyone else. To Prof O, thank you for looping me into helping design that Biosignal processing class! It showed me a whole new world. To Mayank, Gabe, and Sidhant, thanks for showing me the ropes and being my role model.

To the colleagues and collaborators I've worked with over the years, thank you all so much for helping me make my work better. Many project would not have been possible if it stayed in the confines of our lab. All the data collection that was needed would have been impossible if I didn't get significant amount of help from everyone. Thanks Alex for taking my bomb looking apparatus everywhere snooping for EMI.

To those who let me mentor them, especially Junyi, William, Rajneil, Zhao, and Parker, I also want to thank you for putting your trust in me. You all gave me a chance to truly realize how much I enjoy working with students. I hope that I didn't mess up too many of your lives. I promise though that I've gotten better at it. So the next student who let me mentor them, I'm sure your pain and suffering will help them get better advise.

And I certainly cannot finish this acknowledgement without thanking my parents for having provided me with the opportunity of being educated in America. The way I learn, the way I think, the way I work has all been defined by their decision to bring our family here. The education here gave me the foundation to think creatively and challenge established knowledge. They never doubt my pursuits, they never pressure me with "finish up already", and they always express support for whatever I tell them I am working on. They believe in me, and that means a lot.

Chapter 1

WIDESPREAD DISEASE SCREENING AND MANAGEMENT WITH A SMARTPHONE APP

Current medical diagnostic devices are typically designed for a direct one-to-one physiological metric to measurement device pairing or require laboratory access. That means to monitor all the different physical phenomenon of the human body, a large variety of special medical equipment is needed. Ultimately, this means that the medical sensing infrastructure becomes more concentrated into high-resource regions and cannot be easily spread out. The inability to spread medical sensing out of high-resource regions leads to many medical scenarios that are difficult to care for. First is population screening in low resource or remote regions. In many cases, because the medical capability is centralized, people living in remote regions cannot be reached for screening. Second is at home monitoring. Because the cost of owning a medical device is typically above individual means, it becomes costly and often infeasible for at home care. In both of these cases, the fundamental issue is one of scale. Because not everyone needs to own a medical device and different people may need to monitor a different physiological phenomenon, the cost of each medical device can only be so cheap. To this day, only a few medical devices have been made to reach mass population and generally accessible in high-income countries. These include body thermometers, blood pressure monitors, and glucometers.

In recent years, researchers have demonstrated that smartphone sensors hold the potential to be repurposed for medical sensing.The uniqueness of this solution centers around the use of sensors that are **built-in** to a smartphone. This concept purposely constrains the sensing solution to sensor systems that are common to smartphones or has the potential of being common given the cost, size, and general use cases beyond medical sensing. By building the essential capabilities of the sensing needed for the medical end-usecase with hardware that is already widespread, the vision is that medical devices can be created out of smartphones with a software download. This concept follows the idea around Software as a Medical Device (SaMD). SaMDs are medical devices, as defined by

the International Medical Device Regulators Forum (IMDRF), is "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device." [31]

The issue of scale is addressed by leveraging the fact that smartphones are already made and sold in high volumes for diverse usecases that the general population wants and needs. Smartphones have already been shown to be able to use its own built-in sensors to perform physiological measurements such as hemoglobin [79, 80, 27], bilirubin level [20, 53], lung function [45, 24], and eye health [52]. By being able to build medical sensing right out of sensors already present in common smartphones, we are presenting a potential future where makeshift medical devices can be made by simply downloading an app onto phones already in people's pockets.

1.1 Dissertation Goals, Contributions, and Organization

In this dissertation, I have applied this concept specifically to addressing the global need for anemia screening and monitoring with HemaApp, chronic blood pressure management with Seismo, and eye pressure screening with iPressure. Through these different projects, I will illustrate various sensing solutions with different combinations of cameras, microphone, inertial sensor, and light emitting devices to achieve the measurement necessary to capture the underlying biomarker of interest. Central to the dissertation is a nuanced discussion about limitations of phone hardware, firmware, and software I encountered in the various explorations and identify solutions such as alternative bio-physics phenomenon than standard sensing solutions, as in the case for HemaApp, and the use of simple add-ons to augment the phone for iPressure.

The goal of this dissertation is to provide support for the thesis statement:

> SaMDs enabled by **smartphone's built-in sensors** has the potential for **widespread disease screening and chronic disease and health management** through balancing **trade-offs** between optimizing the **sensing solution** of the biomarker and addressing **practical limitations** of designing on real world smartphone devices.

Each work has been carried out in various stages. The document is organized by splitting each project. Each project is presented in a somewhat chronologically linear fashion. For clarity, the related work will be split as well. There will be a related work section that describes the prior-arts

on smartphone based medical sensing, but specific related work that are relevant for each project will be grouped per project. For each project, we will start with initial feasibility studies, system improvements based on the feasibility study, and system validation. The aim of this dissertation is to act as a documentation of the development of medical sensing apps to provide insights for engineers and researchers who want to apply the concept for similar solution spaces. The final section will distill the comprehensive learning culminated from the various projects, pointing out the properties of the current smartphone devices that particularly worked well in enabling these solutions, drawing out limitations, and suggesting future directions.

The document is organized as follows:

- Chapter 2 describes prior-arts on smartphone medical sensing including work that uses external sensors and built-in sensors.

- Chapter 3 presents the background, initial feasibility study, and system design of HemaApp. This section is based on the work published in Ubicomp ’16 and is a stand-alone end-to-end implementation and study of the HemaApp concept. This initial exploration used external lighting to provide a broad examination of the concept without the constraint of the "phone only" concept.

- Chapter 4 describes the extension from the feasibility work where I re-examined the software of the initial system and developed an app-only version of HemaApp. This work is published in EMBC ’17. Included in this chapter is also a separate characterization study of the effects of oxygenation state on hemoglobin estimation.

- Chapter 5 documents the first international study performed in various regions of Peru. The chapter focuses on the lessons learned. The majority of the chapter discusses the usability of the concept in the target users.

- Chapter 6 presents the motivation and background for the development of Seismo, a phone based blood pressure monitoring system.

- Chapter 7 presents the initial feasibility study of Seismo. The feasibility study was a series of developmental trials that involved multiple instantiations of smartphone-based blood pressure measurement app.

- Chapter 8 presents the system design and evaluation of the Seismo app that was developed based on the initial feasibility study. This work was published in CHI '18.

- Chapter 9 presents the design and early stage evaluation of iPressure.

- Chapter 10 will serve as a closing summary of the dissertation that gathers together all the learning in the development of HemaApp, Seismo, and iPressure and present a set of insights to help guide future endeavors in smartphone medical sensing, ranging from the development process, technological challenges, and usability findings.

1.2 Overview of Dissertation Scope

The focus of this dissertation is on the development and testing of smartphone-based medical sensing technology. These project sits within a larger research scope during my PhD studies which explored broadly the concept of bringing clinical sensing outside the clinic through utilizing existing mobiles devices like the smartphone and exploring next generation mobile devices like wearable technologies. The following section quickly goes over various domains and the work I developed as part of my PhD. Although each project is mainly stand-alone in terms of technology, together they form a complete picture of how I view the role and potential of mobile devices in changing clinical sensing.

1.2.1 Widespread Coverage

One of the biggest limitation to clinical monitoring is the lack of distribution of medical measurement equipment. Many medical equipment are expensive, thus often restricted to high-resource or easily-accessible, centralized areas. This leads to the inability to perform medical diagnostics in many remote regions, low-income regions, at-home care scenarios, and emergency situations. One of the key aspects of my work is to utilize widespread computing resources established by the

prevalence of smartphones to perform medically relevant physiological measurements. The utility of out-of-clinic widespread medical diagnostics capability are two pronged. First is in the form of more prevalent screening of potential diseases. This is particularly useful for epidemic level diseases that requires a case-finding tool to diagnose who is actual sick and provide the most effective treatment plan to individuals who are actually sick or an indication that an individual may need the more expensive, clinic based testing. Second is in the form of chronic disease management. Diseases such as hypertension is important to frequently monitor the changes over time in everyday life rather than a more yearly based health checkup. Smartphones already have a variety of sensors, and by utilizing these sensors in new ways, I have been able to show the ability to use phone camera to measure hemoglobin as in the case of HemaApp, use camera with accelerometer to track blood pressure changes, and use camera with a small passive augmentation to measure internal eye pressure. This solution is inherently advantageous because of the reuse of technological resources that are being widely produced, disseminated, connected, and supported.

HemaApp - Noninvasive hemoglobin measurement using a smartphone camera and flash LED. The camera measures the change in spectral absorption distribution to determine the concentration of hemoglobin in the blood. The user places their finger over the camera and flash LED, allowing the camera to capture the blood movement in the finger. The blood movement caused by the heart beat provides a signal source for non-invasively extracting the absorption contribution by blood against the baseline absorption of skin, muscle, and bone.

Seismo - Blood pressure monitoring using a smartphone camera and accelerometer. An alternative method to measuring blood pressure besides the use of a blood pressure cuff is a method called pulse transit time (PTT). PTT is the time it takes for the pressure pulse generated by the heart to travel a set distance on the artery. Seismo, utilizes the smartphone accelerometer to pick up the acceleration created by the aortic valve opening as a timing marker for the genesis of the pressure pulse. With the phone camera, the ensuing pressure pulse is captured down stream as it arrives at the finger tip. By relating these two timings, the PTT is measured using only a smartphone.

iPressure - Internal eye pressure measurement using a small hardware augmentation to smartphone camera. The hardware augmentation is a 5 gram clear acrylic weight nested in a black tube. This attachment is secured on the phone camera such that the camera can see through the acrylic weight. A yellow flourescene analgesic is applied to the eye. By placing the 5 gram acrylic weight

on the eye directly, the eye is depressed by the weight, a technique called applanation. The camera captures the amount of depression which can be used to estimate the internal pressure of the eye.

1.2.2 Continuous Monitoring

The current clinical paradigm relies on active measurement taking using a medical device. May it be blood pressure measurement, glucose monitoring, blood testing, oral hygiene, etc, our medical diagnostics rely on physiological metrics that are a single measurement in time. However, with the advent of consumer wearable devices, more and more attention is being put on the possibility of doing more continuous monitoring of physiological metrics. The first of such is continuous heart rate monitoring, overall activity level, and sleep quality through movement analysis. In my research, I have worked on a variety of aspects to push the boundaries of continuous monitoring, ranging from new physiological monitoring techniques to solutions to charge wearable devices in a passive way.

Project Glabella is a device that integrates multiple optical sensors into a pair of glasses to measure blood pressure continuously using PTT. The PTT is measured using optical sensors placed strategically along the frame of a pair of glasses. When the user wears the glasses, the optical sensor captures the arrival time of the pulse at different parts of the facial arteries. Such a solution can be integrated into any head-mounted wearable device.

CASPER is a capacitive through body charging solution that can charge and power devices worn by the user without being taken off. It works by using the human body as a conduit for 13.56 MHz AC signal, and strategically designing the environment around the user to promote AC coupling of the wearable device and the charging station. This is accomplished by embedding the AC charging station into environments the user often come into contact with, such as furniture and clothing. When the user put on their jacket, drives, sits at a coffee shop, or goes to bed, these charging enabled "stations" deliver AC power through the body and charges the wearable device that are worn on the user. Such a solution can provide a new design space of low power sensors that are distributed on the body, each performing a specific health monitoring task. One such use case is wound monitoring. The current paradigm for wearable health monitors focus on single location, high power devices that tries to infer general health information. However, in cases like wound monitoring, the most effective mechanism is a direct, local measurement. With CASPER, it

is possible for a gauze pad with wound monitoring technology to be charged as the person goes to bed, and for the wound monitor to then operate for the rest of the day checking up on the wound's pH and wetness, to track healing progress and healing environment.

1.2.3 Contextual Understanding

Beyond direct physiological monitoring, one of the aspects that is needed for a full understanding of the human health is the activity and contextual information the person is in. In some ways, this is the outside influence and the physiological metric is the response. Being able to capture information about what a person is doing through out the day can ultimately provide us opportunities to understand how our behaviors are related to our health, or even improve our sensing capability by using our current context to adjust the sensing. In this domain, my work has broadly explored novel sensing solutions to for wearable devices using magnetic sensing.

MagnifiSense is a wrist-worn magnetic sensing technique to capture what a person is doing by detecting the electronic environment they are current in the presence of. Most electronics have a unique electromagnetic radiation due to the internal electronics such as motors, power switches, processors, heating elements, etc. By knowing what someone is using, for example turned on the stove, it is possible to infer someone is cooking. If someone is driving vs riding a bus vs biking (the lack of EMI), the EMI footprint is different, providing information about commute patterns. The EMI based classification of activity patterns provides a rich set of information while maintaining a fairly low power requirement as compared to vision based systems, and only requires a wrist-worn device versus needing to instrument the environment.

The variety of work I explored throughout the PhD along side my main endeavor in building the HemaApp system has been an essential part of how I formed my approach to this project. The work on Seismo, while a fully different use case and validation approach, was a similar base application. The software development that occurred for Seismo provided me with more insight into how to adjust the original HemaApp system that relied on external lighting into the new version that uses only the flash LED. Much of the early work I did in MagnifiSense formed my understanding of signal processing and machine learning solutions, which prepared me for analyzing the data in this project. The other projects provide me with a big picture view of how my work fits into the

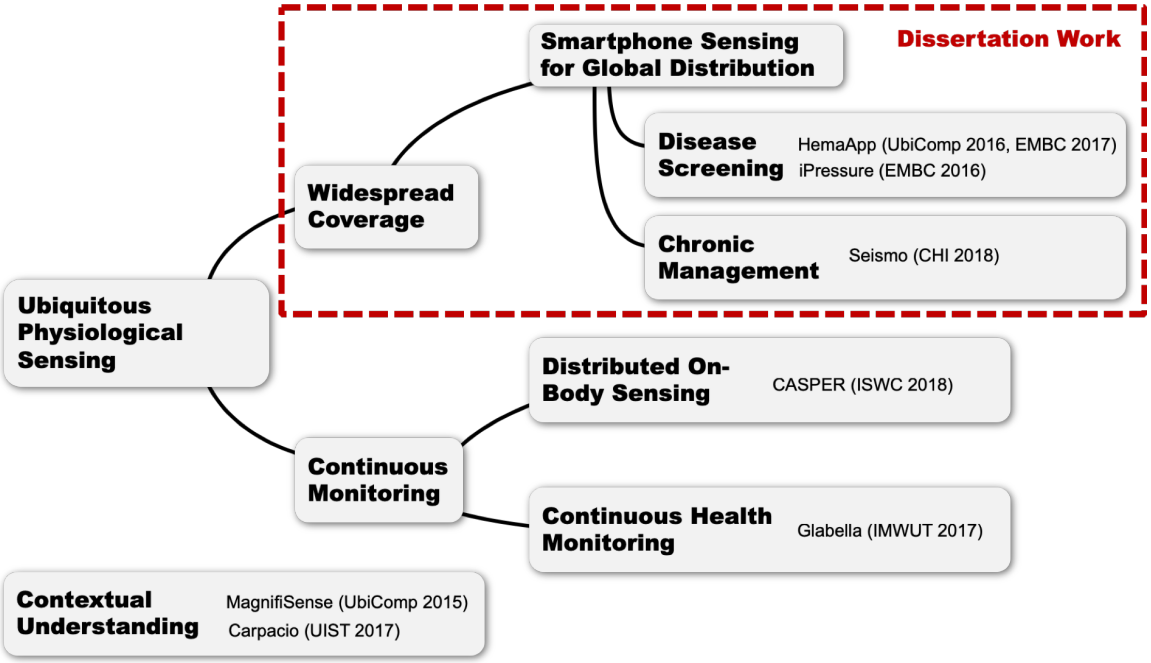

Figure 1.1: Scope of This Dissertation

potential ecosystem of ubiquitous health technologies, giving me the preparation to form a proposal of HemaApp that will ultimately coalesce towards a medically relevant technology.

Chapter 2

BACKGROUND AND RELATED WORK

The focus of this dissertation is the development of smartphone-based SaMD solutions in order to improve the access to widespread for disease screening and chronic disease management. The purpose of this section is to focus on prior art in the use of smartphone for performing medical sensing and will leave specific background and related work for the different condition and measurement systems in the individual chapters to follow.

2.1 Mobile health

The umbrella term of mobile health refers broadly to the use of smartphone in healthcare. The purpose of this section is to categorize the different uses of smartphone as related to health **sensing**. This is actually in contrast to the common referral to the term mobile health. In the field today, smartphones have already been integrated into the work flow of many community health workers. Open Data Kit is a good example of a toolkit meant to establish an infrastructure to utilize the phone's sensing capability, device usability, and internet connectivity to facilitate the work flow of global community health workers [13, 26, 14]. Such framework does not, however, focus on the use of mobile sensors to perform medical sensing akin to a medical device. This type of mobile health is, in spirit, more similar to the class of mobile health solutions that connects a medical device attachment to a phone. The purpose of the phone is to relay the information captured by the attachment to a centralized information repository, in the form of personal data logging or telemedicine. This use of the smartphone in the healthcare system, although extremely useful, is not the focus of this dissertation.

In my work, I am focusing on the use of smartphone sensors directly to perform physiological measurement and analysis. In essence, the smartphone itself becomes the medical device. The advantage to this solution over standard medical devices (with or without smartphone connection) is the potential for easier dissemination of the medical diagnostic capability. Since the sensing

relies on hardware that is already present in a smartphone, the medical device is created when the user downloads the app that measures a specific physiological metric. Within this approach to mobile health are two main distinctions. First is the use of passive augmentations and second is with only the smartphone sensors. Passive augmentations are defined as objects or devices that are mechanical and in theory easy to mass produce. These passive augmentations are needed because the phenomenon needed to be measured cannot be achieved with just the smartphone sensor. The defining characteristic, however, is that these passive augmentations will likely never be included in a common smartphone.

For example, microscope attachments have largely been explored as a way to perform in the field blood testing using the phone's camera [61]. Computer vision algorithms can leverage these magnified images of blood slides to perform cell counting and cell identification. These attachments generally are a set of lenses that a relatively cheap to make and are attached onto the camera using either a clip or a phone stand that also holds the specimen in place. It is reasonable to assume that microscopes will likely not be integrated into phones.

Bilicam [20], developed by de Greef et al., uses a smartphone camera to measure the blood bilirubin level to screen for neonatal jaundice. This is performed by capturing a picture of the baby's skin to perform a color-based analysis to estimate bilirubin level. In order to correct for ambient lighting, Bilicam relies on a color correction card, similar to the McBeth color checker for photography. Biliscreen [53], by Mariakakis et al., uses a similar concept for measuring bilirubin levels in adults. Instead of a photo of the skin, Biliscreen relies on measuring the color of the sclera, or the whites of the eye. Similar to Bilicam, Biliscreen uses a color correction card.

In this vein, my work on a system called iPressure [52] combines the smartphone camera with a clear acrylic weight attachment to measure a person's intraocular pressure. This system replicates other intraocular pressure measurement devices which applanates (flattens) the eye with the known weight to infer the internal pressure of the eyes. To augment this measurement, a standard fluorescene eye drop is applied to amplify the applanation surface. In this case, the attachment weight and fluorescene will never be a part of a standard phone. However, the attachment is purely passive and fluorescene can readily be purchased.

The second category, which will be the main focus of this dissertation, aims to enable medical sensing using only the phone. As compared to solutions that require passive augmentations, this

solution is a true "one download away," and supports the concept of ease of dissemination more readily. In the following section, you will notice that I include examples where the research prototype was not built directly out of a phone, but rather a stand-alone prototype that either attaches to the phone or does not use a phone at all. I include these examples because the sensor technology may already be in most phones but lacks the necessary software control to achieve the proposed solution. I will, however, mainly focus on solutions that are implemented and tested on an actual smartphone.

2.2 Medical Sensing with a Smartphone

Medical sensing with a smartphone can be differentiated into two types: Active and passive. Active sensing involves user interaction during a measurement. For example, the user puts their hand on the camera, blows at the phone, or put the phone on their chest. Passive sensing is when the phone monitors the user in the background without explicitly requesting the user to engage in a measurement. In this dissertation I do not make any arguments about whether one is better than the other, but will focus on active sensing, which is the category my work resides in.

2.2.1 Active

Active sensing is similar to how medical devices are designed today. When a measurement needs to be taken, a prescribed action is required. A variety of work have utilized this concept to perform physiological measurements with a smartphone.

The most common measurement, one that has already made it to main stream applications, is to measure heart rate with the phone camera [18]. When the finger is placed over the camera and flash LED, the camera captures the color of the tissue and blood in the finger. When the heart beats, some amount of blood is pushed into the finger. This change in blood quantity in the finger leads to a change in the light reflected into the camera, which can be used to calculate when a pulse happens. This phenomenon is fundamentally similar to the way HemaApp works and I will go into more detail on this in the following chapter. Heart rate can also be captured using the inertial measurement unit (IMU) [59]. This work demonstrates that when the phone is placed on the chest, the microvibration caused by the valve movement of the heart leads to a measurable signal.

Similarly, this can be measured with a microphone placed on the chest [36]. A work by myself, called Seismo, also explores this topic by utilizing the smartphone accelerometer to capture the heart beat and the camera to capture the pulse.

Chandrasekhar et al. demonstrates a novel oscillometric finger-pressing method on a optical sensors instrumented with a pressure sensor. Their method utilizes the same theory as standard brachial cuff oscillometric BP measurement, but instead uses the digital artery pressure to estimate brachial pressure. Although their prototype uses a standalone sensor system placed on a phone, reminiscent of other external medical device that connects to the phone, their implementation takes into consideration a form factor and usability that would likely work on a smartphone. Various smartphones are equipped with pressure sensitive screens and as demonstrated by many prior work, cameras are capable of capturing a similar pulse waveform. The configuration and sensing technique can be reasonably considered to be feasible without any external hardware.

Various research groups have shown initial feasibility studies or technical limitations in performing pulse oximetry on a smartphone camera. Karlen et al. have explored the design challenges for camera oximetry on mobile phones [25, 34]. They presented the theoretical relationship of the SaO2 calculated using different wavelengths of light based on known absorption coefficients of hemoglobin. Their work showed that if a phone camera's IR filter can be removed, the use of 660nm and 950 nm would be ideal, but if no modification is made to the camera, 460nm and 660nm can be used instead. In a separate work, Karlen et al. also presented automated solutions for detecting finger placement on the camera lens and regions of interest in the camera feed. Our system draws on the wavelength recommendation from these works. Although we recognize that automated detection of signal quality is important for end-use scenarios, we did not implement these for our system due to the controlled nature of our experiment, which guarantees proper placement of the finger and low movement.

Larson et al. [46]. demonstrated the use of the microphone to estimate lung function in a spirometry-like effort with their application called SpiroSmart. Using only the microphone of the smartphone, Spirosmart captures the sound of the user's exhalation effort, estimating the air flow speed through analysis of the frequency content of the exhalation.

Instead of direct sensor based measurements, researchers have also explored ways to infer physiological metrics through other means. Mariakakis et al. [54], for example, developed a blood

alcohol estimation system based on user performance on a set of cognitive acuity tests on the smartphone to capture information like response time to on-screen challenges.

2.2.2 Passive

Compared with the active sensing solution, passive solution is constraint to the everyday "natural" uses of the smartphone to capture health information about the user.

Hernandez et al. has demonstrated the ability for phones in a user's pocket to measure heart rate using the accelerometer [28]. Their technique relies on a similar phenomenon as Mohamed et al. introduced earlier, but instead of a direct measurement at the chest, the phone in the pocket captures the peripheral vibrations caused by the proximal cardiac ballistocardiography.

Sleep tracking is another popular domain with a phone in the background solution. Phones are placed on the bed or next to the bed at night, and the sensors on the phone capture activity throughout the night to infer sleep information about the user. Downloadable applications today uses the phone's IMU to capture movement information on the bed to infer sleep stage of the user, information that is used to perform smart alarm features or logging of sleep quality. Apps also exist that uses the microphone by the bed to capture snoring activity. Nandakumar et al.'s ApneaApp uses the smartphone's microphone and speaker to detect breathing motion, which is then related to apnea, stopping of breathing, events [63].

Beyond analog sensors like cameras, IMUs, and microphones, researchers have also demonstrated the utility of phone use information as a method to infer health information. Jacques et al. has shown that phone use information, like when a person uses a phone (middle of the night), has the potential for detecting depression [32, 33]. Bae et al. demonstrated that tracking metrics like activity level, screen duration, time between key presses, etc can be used to detect potential high-risk alcohol consumption episodes for just-in-time interventions [4].

Chapter 3

HEMAAPP: NONINVASIVE BLOOD SCREENING OF HEMOGLOBIN USING SMARTPHONE CAMERAS

HemaApp is a design of a noninvasive hemoglobin measurement solution that relies on the use of a smartphone camera. The development of this work started in April 2015 with the first publication on the feasibility of this concept at Ubicomp 2016 [79]. In this initial work, the system did not actually rely solely on the hardware of the phone. Instead, I had explored the use of a variety of external lighting sources to establish an overview of what light works well and whether the concept even works. The following chapter is a reproduction of the 2016 publication modified to fit within the context of this dissertation.

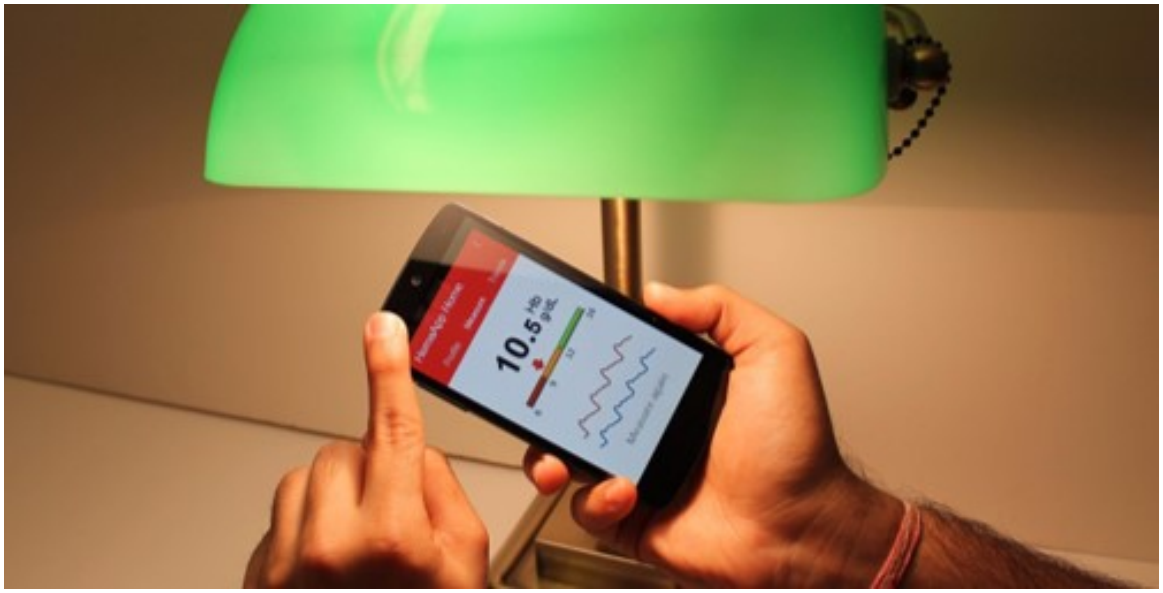

Figure 3.1: HemaApp is a smartphone application that noninvasively estimates blood hemoglobin concentration using a smartphone camera. Analysis of the color of the blood in a user's finger yields an estimate of the user's hemoglobin level. We evaluated the system using the smartphone's LED flash and incandescent light bulbs as illuminating sources.

Figure 3.2: Three embodiments of HemaApp in increasing hardware augmentation.

By leveraging the absorption properties of hemoglobin and blood plasma at multiple wavelengths of light, HemaApp, can measure hemoglobin concentrations through a method called photoplethysmography (PPG) combined with hemochromatography. PPG is an optical measurement of the blood flow, in our case measured at the finger. Hemochromatography is an analysis of the coloration of the blood. By combining these two techniques, HemaApp can extract the coloration of the blood from the finger optically, a method that is employed to reduce the effect of skintone. This is achieved by using the RGB camera with different light sources illuminating the fingertip.

Three different hardware embodiments that vary the level of hardware augmentation necessary were evaluated (Figure 3.2): (1) white flash + infrared emitter, (2) white flash + infrared emitter + incandescent lamp, and (3) white flash + custom infrared led array. Most smartphone cameras have a white LED, and many smartphones have started to introduce IR emitters for autofocus, potentially making the first embodiment augmentation free. The second embodiment requires a low overhead modification by introducing an incandescent lamp for an additional IR spectrum. The third embodiment requires the most hardware augmentation by introducing a custom array of LEDs as a phone case.

3.1 Background and Related Work

3.2 Anemia

When someone is anemic, their blood has a reduced number of red blood cells (RBC). The central function of RBCs is performed by the protein molecule hemoglobin, an iron-containing molecule,

that has the ability to bind to oxygen. As the RBCs circulate with blood throughout the body, oxygen is delivered to various parts of the body, allowing proper physiological function. As a result, if the number of RBCs is reduced, the oxygen delivery capability is also reduced. The most common symptom of anemia is fatigue due to this decrease in oxygen carrying capacity. If the evolution of anemia is slow, however, symptoms are often hard to detect. This is because of the body's homeostatic adaptation such as increasing tissue's oxygen extraction from the baseline of 25% up to about 60%, allowing the body to maintain proper oxygenation [40]. Symptoms might only surface during exertions in this case where the need for oxygen increases beyond the extraction compensatory response. Acute or severe anemia typically have more pronounced symptoms, and can lead to lethargy, confusion, and other potentially life-threatening complications.

3.2.1 Defining Anemia

The strict definition of anemia is a state of reduced absolute number of RBCs in circulation. For clinical practice, however, this definition is not practical because RBC counting is not cost-effective and general unavailable. In practice anemia is defined as the a reduction in one or more of the following RBC measurements: Hemoglobin concentration (HGB), hematocrit (HCT), and RBC count.

Hemoglobin Concentration (HGB) - HGB is the concentration of hemoglobin molecules, the oxygen-carrying molecule, in whole blood. It is typically measured in grams per volume, often in terms of g/dL (for deciliter = 100 milliliter) or g/L.

Hematocrit (HCT) - HCT is the percent packed spun volume of intact RBC to volume blood. This measured by spinning the blood sample in a centrifuge and measuring the height of packed RBC.

RBC Count - RBC count is the number of RBCs in a specific volume of whole blood, expressed as millions of cells per μL of whole blood.

In many cases, either HGB or HCT is used for diagnostic purposes due to availability of these testing solutions. For the purpose of this dissertation, I will refer to HGB as the RBC metric for anemia levels because our system ultimately measures HGB and not HCT because it performs a color measurement. To determine whether a person is anemic depends on many factors like age,

sex, ethnicity, and sometimes life-style. For this reason, clinically it is important to consider the various factors when determining if the RBC measurement is considered below normal levels for this person.

HGB is elevated in the first month of life at >14g/dL, and rapidly declines to about 11g/dL at six to nine weeks of age [41, 58]. Between three to six months, anemia detected usually suggests a hemoglobinopathy because nutritional anemia is unlikely in term infants. The hemoglobin level increases as the infant grows up and typically to about 14 g/dL for female and 15 g/dL for male [48]. As a person ages past 60 years old, their hemoglobin level tends to then decrease. Table 3.1 documents a normal range of hemoglobin level for male and female of Caucasian and African-American population in America of each age group. Anemia is often defined as having a value of more than two standard deviations below the mean [10].

3.2.2 Causes of Anemia

As mentioned previously, the purpose of having a proper concentration of hemoglobin is to enable oxygen delivery. To maintain a proper concentration of working and effective RBC, the body has to have a balanced production, maintenance, and breakdown of RBC [48]. A disruption in any of these processes can lead to a reduction in healthy RBC count. There are a variety of causes and forms of anemia. This review will cover some of the most common causes, but is not meant to act as a comprehensive list.

- **Iron-Deficiency** - Iron-deficiency is the most common cause of anemia globally, at about 1 in 5 people in the world [38, 85]. Because the hemoglobin protein molecule requires iron to be produced, if a person lacks iron, the production of hemoglobin will be reduced. Iron-deficiency is most often a result of malnutrition, which is the main reason that iron-deficiency anemia is most prevalent in low-resource regions of the world. Increased iron intake through diet, supplement, or infusion are treatments for iron-deficiency anemia.

- **Sickle Cell Disease (SCD)** - SCD is a genetic blood disorder that results in a mutation of hemoglobin production [67]. Healthy RBCs are shaped like a flat disk. Sickle mutation in the hemoglobin protein leads to a change in RBC into a sickle-like shape, changing the hemo-

Table 3.1: Normal Values of Hemoglobin for Different Age, Sex, and Ethnicity Groups

Caucasian				African American			
Age (Years)	Sex	Hemoglobin (g/dL)		Age (Years)	Sex	Hemoglobin (g/dL)	
		50th percentile	Lower Limit			50th percentile	Lower Limit
<1		12.5	11	<1		12	11
2 or 3		12.6	11	2 or 3		12	11
4 to 6		12.9	11.7	4 to 6		12.5	11
7 to 10		13.5	12	7 to 10		12.7	11.2
11 to 14	Female	13.7	12.3	11 to 14	Female	12.9	10.6
	Male	14.3	12.6		Male	13.6	11.8
15 to 18	Female	13.7	11.5	15 to 18	Female	12.8	10.7
	Male	15.4	13.7		Male	14.9	12.9
20 to 49	Female	13.6	12.2	20 to 49	Female	12.9	11.5
50 +	Female	13.6	12.2	50 +	Female	13	11.5
20 to 59	Male	15.2	13.7	20 to 59	Male	14.6	12.9
60 +	Male	14.9	13.2	60 +	Male	14.3	12.7

Age 1-18 data reproduced from Copyright © 2018 UpToDate which references: Brugnara et al. [12], Cembrowski et al. [16], and Baker and Greer [6].

Age 20+ data reproduced from: Beutler E and Waalen J [9].Copyright © 2006 The American Society of Hematology.

dynamic properties both inter-cellular and in general circulation. The clinical manifestation of SCD is variable, with a variety of mechanisms that causes acute anemia and pain crisis. The change in hemodynamics causes excessive hemolysis (rupture of RBC), vaso-occlusion (reduction in circulation due to sickle RBC adherence to the endothelium) [19], splenic sequestration (trapping of RBC in the spleen) [5, 21], and severe anemia during aplastic crisis (cessation in RBC production in the bone marrow) [73]. Patients with SCD are chronically anemic and are often treated with either hydroxyurea to induce production of non-mutating fetal hemoglobin or regular transfusion of whole blood.

- **Thalassemia** - Thalassemia is a genetic blood disorder that results in abnormal production of hemoglobin. There are many forms of Thalassemia, classified as the imbalance of alpha- or beta-globin production, leading to unpaired alpha- or beta-chain to precipitate and causing destruction of RBC precursors in the bone marrow. This leads to ineffective erythropoiesis (production of RBC) and hemolysis, both of which contributes to chronic anemia [57].

- **Malaria** - Malarial infection occurs when the malaria parasite enters the blood through infected mosquito bites. The parasite infects the RBC and at the end of the infection cycle, ruptures the RBC and the released merozoites from the infected RBC cause fever and other manifestations of malaria [86]. This rupturing of RBC causes hemolytic anemia.

- **Hookworm Infection** - Hookworm infections' main impact is on nutritional status [7]. Hookworm infection causes blood loss due to lacerating capillaries and also by the hookworm ingesting extravasated blood. This daily loss of blood, iron, and albumin leads to anemia of malnutrition.

3.3 Hemoglobin Measurements

Clinically adopted methods to measure hemoglobin are currently restricted to blood tests. In a lab test, a technician draws about 3mL of blood for a complete blood count (CBC). In a CBC, information about red blood cells (RBC), white blood cells (WBC), and platelets are measured electronically. The hemoglobin concentration is measured optically after the blood is mixed with a chemical agent that changes the solution's density proportionally to the concentration of hemoglobin. For

tests that only require a hemoglobin concentration, such as those for screening during blood donations, a point of care (PoC) device is often used instead of the more time consuming CBC. An example of a PoC is the HemoCue device. The HemoCue requires a finger prick to draw a small amount of blood. A microcuvette then draws the blood and mixes it with a chemical reagent on a test panel. This test panel is read by the HemoCue device and a result can be determined in about a minute. This method has been shown to have a rank order of 0.89 at a mean accuracy of ±0.5 g/dL when compared to results from a CBC test [23, 78].

Most recently, noninvasive measurement of the hemoglobin concentration through optical measurement of the blood at the fingertip has been developed. Alam et al. has created a finger probe with 6 LEDs that cover multiple wavelengths in the red to IR spectrum (630, 660, 680, 770, 880 and 1300nm) to measure the hemoglobin to water ratio in blood plasma [1, 3]. Kraitl et al. conducted a 41 people study using a three LED (670, 810, 1300nm) finger probe, and found similar success. Timm et al. based their technique off of Kraitl et al.'s system to create a device called OxyTrue Hb®. A commercially available system developed by Masimo is similar to Alam et al.'s system, but instead incorporates 7 LEDs and uses a proprietary algorithm [51]. Validation of this system has shown a mean accuracy of ± 1.0 g/dL [51, 69]. We draw from these existing works for noninvasive hemoglobinometry in our design of the lighting sources we incorporated in our study. The aim for our system is to provide a similar level of accuracy, but reusing the hardware of a smartphone to make the system more easily deployable.

HemaApp is the first noninvasive hemoglobin measurement system that leveraged the camera of the smartphone to perform similar measurements as the Masimo Pronto and other finger probe systems. Since the Ubicomp 2016 and EMBC 2017 publication of HemaApp, other works have explored the use of smartphone camera to analyze hemoglobin concentration [27]. Hasan et al.'s work demonstrates a different approach to the noninvasive measurement with a smartphone. Instead of a pulse extraction based system, their system uses a more statistical based, histogram analysis of the measurement at the finger. Their method and results present an interesting and novel take on the problem and in my proposed analysis, I plan to explore more closely how their system can contribute to the future improvements of my own solution.

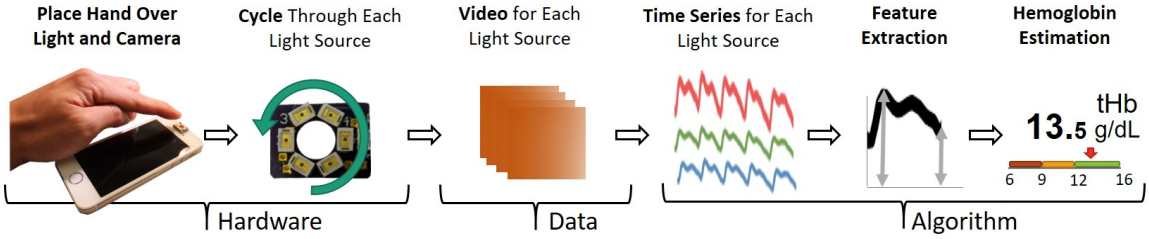

Figure 3.3: System overview of HemaApp, which consists of the phone and lighting hardware, data, and algorithm. The HemaApp application communicates with an LED attachment. After a user places their finger over the LED and camera, multiple light sources are cycled through. A video is recorded for each light source. The algorithm then extracts the R, G, and B time series waveform for each video by averaging each RGB channel independently for each frame. The algorithm then extracts machine learning features including peak and trough measurements for each light source, and also interaction terms between light sources. Finally, a SVM based regression is applied to estimate the hemoglobin concentration for the user.

3.4 System

Figure 3.3 shows the HemaApp system, consisting of hardware components that record data which are processed by an algorithm to calculate the hemoglobin concentration. The following hardware section describes the choice of light wavelengths and the phone application that controls the lights and records the light reflected from the finger to the camera. The data obtained by the phone app are a series of videos for each light source. The algorithm section describes the process of converting the series of video recordings to RGB time series waveform and the physical intuition behind the machine learning features used in the hemoglobin estimation.

3.4.1 Hardware

Due to the limited access to low level hardware control of the phone, we choose to develop a hardware add-on to provide the lights for hemachrome analysis. However, unlike other noninvasive optical hemoglobin measurement systems, we limit the lights we choose to ones that can be commonly found on smartphones today and in typical hardware stores. Furthermore, we make no modifications

to the phone's camera itself, unlike prior work's use of phone cameras for pulse oximetry, which requires the removal of the IR filter of the camera. Instead, we rely on the front facing camera, which tends to have weaker IR cut off than the back facing cameras, making it more suitable for our application.

Wavelength Selection

We address a major limitation of using a smartphone camera for doing hemoglobin measurement; a lack in sensitivity to wavelengths above 1000nm. First, we use only IR absorption below 1000nm. Water begins to have a response above 940nm and has an initial relative maximum at 970nm. Second, we not only measure the water content in the plasma, but also the proteins that make up about 10% of the blood plasma by volume as a proxy for capturing the plasma volume to compare against hemoglobin concentration.

This is accomplished by leveraging the blue absorption of the plasma. By illuminating the finger with a white LED (which contains a strong blue component) and an IR light at 970nm, the system is capable of capturing the plasma response. We also included another IR LED (880nm) to help capture the different absorption between the various forms of hemoglobin. Both the 970 and 880nm LED are in the range of IR autofocus LEDs equipped in the current generation of smartphones. Our IR LEDs are sourced from the Marubeni SMC series.

Even though some new smartphones are beginning to be equipped with IR LEDs, most smartphones only have a white LED. To address this issue, we look towards a commonly available IR source, the incandescent light bulb. Incandescent light bulbs typically have strong IR light emission in the NIR range. Our system uses a candelabra 6W incandescent light bulb that can be found in hardware stores.

The LEDs are placed in a ring around the phone camera, while the incandescent light bulb was placed about 3 inches above the finger.

Phone App

For the HemaApp prototype, we use a Nexus 5 with an unmodified front-facing camera. For typical optical hemoglobin measurement, each lighting condition is cycled through at a high rate repeatedly

to measure the response at each wavelength pseudo-simultaneously. In this way, the lighting conditions are all measured for the same pulse. However, due to the limited sampling rate of the camera and the inability to synchronize the lighting circuit with the camera's frame refresh, it is not feasible to measure all wavelengths of light at the same time without considerable access to low level hardware control. We discuss the tradeoffs of multiplexing and series testing in the discussion and make suggestions on how multiplexing can be achieved based on our own experiments. For our prototype, each lighting condition shines continuously for 15 seconds and cycles to the next light with an assumption that hemoglobin concentration and average blood flow does not change significantly during the course of the test, which takes a few minutes. The mobile app is built on the Android platform with the Android Camera 2 API, which allows for full control over the white balancing and exposure. This is important because infrared is generally considered an unwanted spectral response that camera applications will detect and rebalance settings to avoid. Prior work has found that certain white balancing settings, such as the "incandescent mode" can be used to avoid the rebalancing [34]. The Camera 2 API allows for full control over the exposure, white balance, and sensor sensitivity. The hardware gains are manually set for each RGB channel using presets that are empirically found for each lighting condition such that the three channels reported a similar level of light. This is necessary because the red channel typically has a much stronger response due to the red blood under the white and incandescent lights. If left to a flat white balance, the auto exposure will be set to the red channel, leaving the G and B channels highly underexposed. Finally, exposure is set using the camera API's auto-exposure settings. Once the image is auto-exposed, the exposure is locked and a 15-second video is recorded for each of the lighting conditions sequentially.

3.4.2 Data

Each light source is cycled through one after another and a 15 second RGB video is recorded for each light source. The exposure, frame rate, white balance gain, and ISO settings are recorded for calibration. At an average resting heart rate of about 75 beats per minute (BPM), around 15-20 pulses were captured for each light source.

3.4.3 Algorithm

The algorithm section first explains hemachrome analysis, which is the blood color analysis we base our algorithmic implementation on. We then detail the two stages of the implementation. First, the video processing step extracts the pulsatile signal recorded in each video. The second stage is a feature generation step that involves combining the intensity values extracted from each wavelength's video. These features are then used with an SVM to train regression models based on the ground truth blood test hemoglobin.

Hemachrome Analysis

Hemachrome analysis is the study of blood coloration to analyze the components in the blood. HemaApp aims to measure the concentration of hemoglobin as compared to the concentration of plasma in the blood. The Beer-Lambert law states that the absorption of light is proportional to the concentration and thickness of the medium, given by:

$$I_{measured} = I_0 e^{-\alpha[C]d} \tag{3.1}$$

where I_0 is the incident light intensity, α is the absorption coefficient, $[C]$ is the concentration, and d is the thickness of the medium that the light travels through. When the finger is illuminated with a single wavelength of light, the measured intensity $I_{measured}$ represents the absorption due to tissues, hemoglobin, and plasma:

$$I_{measured,\lambda} = I_{0,\lambda} e^{-d(\alpha_{tissue,\lambda}[tissue]+\alpha_{Hb,\lambda}[Hb]+\alpha_{plasma,\lambda}[Plasma])} \tag{3.2}$$

where λ is the wavelength of the incident light. To obtain the ratio of [Hb] and [Plasma], it is necessary to eliminate the attenuation of the intensity signal due to finger tissue. This is accomplished by measuring the temporal change of the measured intensity as the thickness of the arteries oscillate with respect to the heartbeat.

The change in arterial thickness Δd affects only the path length for Hb and Plasma. By measuring the ratio of the maximum and minimum intensity of the light received, the effect of the tissue is

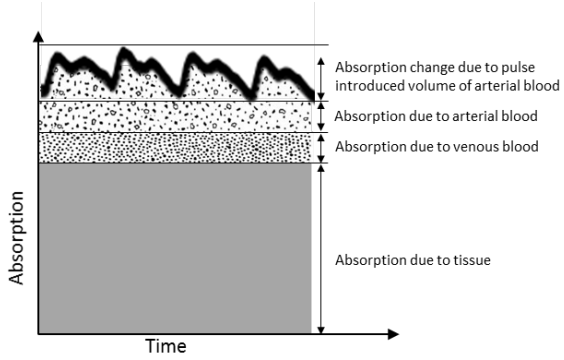

Figure 3.4: Light absorbed by living tissue. Adapted Web09[26]. The absorption of light changes due to the change in volume of blood when the heart pulses.

removed:

$$\frac{I_{peak,\lambda}}{I_{trough,\lambda}} = e^{\Delta d(\alpha_{Hb,\lambda}[Hb]+\alpha_{plasma,\lambda}[Plasma])} \tag{3.3}$$

Where the ratio of intensities can then be expressed as:

$$I_{R,\lambda} = ln\left(\frac{I_{peak,\lambda}}{I_{trough,\lambda}}\right) = \alpha_{Hb,\lambda}[Hb]\Delta d + \alpha_{plasma,\lambda}[Plasma]\Delta d \tag{3.4}$$

The measured ratio of maximum and minimum values of intensity thus provide a measure of absorption due to the different components of blood. In theory, a system can use empirically measured absorption coefficients for each compound at a specific wavelength of light to predict the hemoglobin concentration using Equation 4. By measuring the response at multiple wavelengths of light, multiple $I_{R,\lambda}$ can be calculated. Ratios of $I_{R,\lambda}$ across wavelengths over determine the ambiguity of Δd allowing an estimate of [Hb] to be made. However, factors such as the distribution of the emitter, sensitivity of the sensor, and complex reflection properties of tissue make this infeasible. In our case, the incandescent and white LED are very broadband emitters that produce measurements that are poorly modeled by Equation 4. Using machine learned regressions to estimate the $\alpha_{Hb,\lambda}$ and $\alpha_{Plasma,\lambda}$ for each lighting source, we overcome our reliance on predetermined absorption coefficients at specific wavelengths for blood.

3.4.4 Video Processing

For each lighting condition, we extract the intensity of the peak and troughs for each pulse. The center section, measuring half the width and half the height of the image, is cropped and an average intensity for each channel is calculated. This is done because the image from the center of the image is the most consistent and stable. The signal processing technique that extracts I_R

1. High pass filter with a cut off at 0.5Hz. This is to remove the fluctuations due to breathing, which are around 0.2-0.3 Hz.

2. Apply an FFT on the filtered waveform and extract the dominant peak. The dominant peak gives an estimate of the heart rate.

3. Using the heart rate as a threshold between successive peaks, false peaks due to the dicrotic notch and the diastolic peak are avoided. The threshold is set at $\frac{3}{4}$ of the heart rate to avoid missing actual peaks.

4. Map the index of peaks and troughs onto the unfiltered signal. We need the original magnitudes of the peaks and troughs to calculate $I_R = ln\left(\frac{I_{peak}}{I_{trough}}\right)$

5. Calculate I_R for each peak

Feature Selection

In order to estimate the absorption coefficients at the broadband wavelengths used by our system, the HemaApp system cannot rely on the basic derivation of multi-wavelength hemachrome analysis. The following is a list of the features. The first three features are derived directly from the Beer-Lambert equations in hemachrome analysis. The fourth and fifth features aim to capture nonlinear interactions between wavelengths.

$I_{peak} = I_{DC}$: The baseline intensity due to tissue.

$I_{peak} - I_{trough} = I_{AC}$: The amplitude of the pulsatile absorption.

$ln\left(\frac{I_{peak,\lambda}}{I_{trough,\lambda}}\right) = I_R$: The adjusted amplitude of absorption that eliminates the effect of tissue.

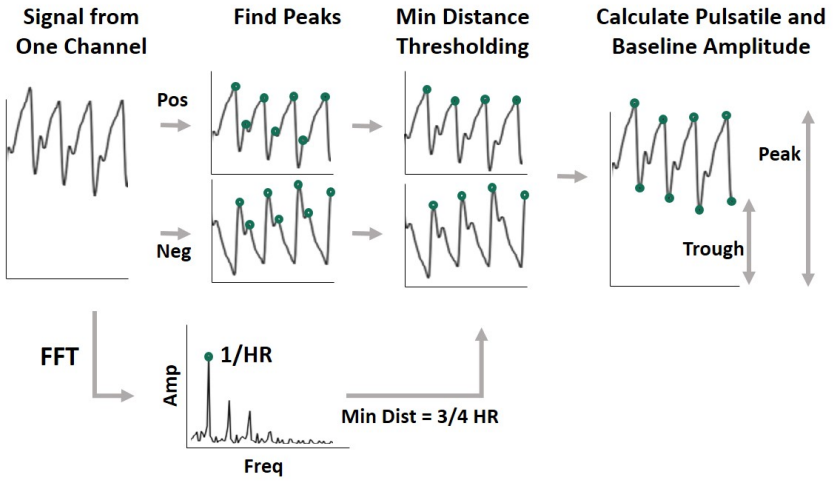

Figure 3.5: To calculate the absorption change due to blood, a peak detection algorithm is applied on the temporal signal. An FFT is used to estimate the heart rate. Using the estimated heart rate as the minimum spacing between two peaks, the peaks and troughs of the heart beat signal is extracted.

$I_{R,AC}(\lambda_i,\lambda_k) = I_{AC,\lambda_i}/I_{AC,\lambda_k}$: A pairwise ratio of pulsatile absorptions between all wavelengths.
$I_{R,ACDC}(\lambda_i,\lambda_k) = |\frac{I_{R,\lambda_i}-I_{R,\lambda_k}}{I_{DC,\lambda_i}-I_{DC,\lambda_k}}|$: Absorption difference across wavelengths, baseline adjusted.

Regression

A separate SVM regression (SVR) model is trained for each embodiment of our proposed system: (1) white + 970 nm LED, (2) white + 970 nm LED + incandescent light, (3) white + 970 nm + 880 nm LED + incandescent light. We will refer to each embodiment as EMB#. The regressions are made based on ground truth values obtained from a blood test. During development, we find that a linear regression do not produce as good of a result as an SVR, but provides good insight into feature significance. As such, the linear regression is used to help in the feature selection process, as it is easier to interpret the resultant model, but the SVR is used to produce a model for evaluation by employing the features chosen through the linear regression. Table 3.2 displays the feature list for the three embodiments:

The SVM model is tested using a leave-one-subject out validation. The training is done using the MATLAB implementation of SVM regression with a Gaussian kernel with default parameters,

Table 3.2: Features used for training in each embodiment.

EMB1	**EMB2**	**EMB3**
$I_{AC}(WhR)$	$I_{R,AC}(InR, WhB)$	$I_{R,AC}(InR, 880)$
$I_{AC}(WhB)$	$I_{R,AC}(InR, 970)$	$I_{R,AC}(InB, WhR)$
$I_{AC}(970)$	$I_{R,AC}(InB, WhB)$	$I_{R,AC}(WhR, WhB)$
$I_{R,AC}(WhR, WhB)$	$I_{R,ACDC}(InR, InB)$	$I_{R,ACDC}(InR, InB)$
$I_{R,ACDC}(WhR, 970)$	$I_{R,ACDC}(InR, WhB)$	$I_{R,ACDC}(InR, 970)$
	$I_{R,ACDC}(WhB, 970)$	$I_{R,ACDC}(WhR, 880)$
		$I_{R,ACDC}(Whb, 880)$

to avoid overfitting due to parameter tuning.

3.5 Data Collection & Validation

To evaluate and inform the design of HemaApp, we conduct a clinical study with three groups of people: healthy students and staff of the university, in-patients at a children's cancer and transfusion clinic, and in-patients at an adult cancer and bone marrow transplant clinic. Data collections at these sites provide a diverse dataset paired with ground-truth hemoglobin concentration from CBC tests. We note that typical initial feasibility studies for hemoglobin measurement systems involve 10 – 60 people and typically cover population ranges of 8 – 16g/dL hemoglobin concentration [42, 70]. In order to recruit patients from beyond these ranges, a system must be validated enough for hospital use during surgical procedures, as people who are well below the 8 g/dL range are likely in need of a transfusion or are already in surgery. These initial feasibility studies are then followed by a series of larger follow up clinical validations in more diverse populations and repeated measurements per person during surgeries such as urologic procedures where hemoglobin concentrations change dramatically [51, 70, 77]. We base our study design to reflect the feasibility studies of these previous works. Our study include 31 patients in a range of 8.3 g/dL to 15.8 g/dL.

Table 3.3: Demographic information of subjects.

Participant Demographics (N = 31)	
Age (years)	6 – 77 (μ=31, σ=17.5)
Hemoglobin (g/dL)	8.3 – 15.8(μ=12.1, σ=2.2)
Gender Ratio	15 male:16 female
Reported Ethnicity (n, %)	
East Asian:	7
South East Asian/Indian:	6
Latino:	3
White:	14
Mixed:	1

Each patient's data set include a series of videos measuring the absorption change under multiple wavelengths of light. We collect videos within 24 hours of the ground truth CBC blood draw to ensure that hemoglobin measures are as accurate as possible. Within a day, hemoglobin concentration is typically stable within 0.5 g/dL [11]. Patients who have hemoglobin transfusion or heavy bleeding between the study and the blood draw were excluded.

The in-patient clinics are centered on cancer patients, in particular leukemia and bone marrow transplant patients, because these patients tend to be chronically anemic. As such, these patients often have CBCs done as part of their clinical care. By including both children and adult clinics, the study cover a wide range of age and hemoglobin variations, allowing us to validate on different groups of people. Our study population include about 1/3 pale skintone, 1/3 sepia – light brown skintone, and 1/3 dark brown skintone. Due to a low population of people with black skintones in the city we conduct our study, we are not able to include those with very dark pigmentation in our study. However, it should be noted that pigmentation on the underside of the hand tends to be much lighter in comparison to the backside of the hand, as such, it is likely that the effects of very dark pigmentation is covered well enough by our inclusion of dark brown skintones. In our follow up

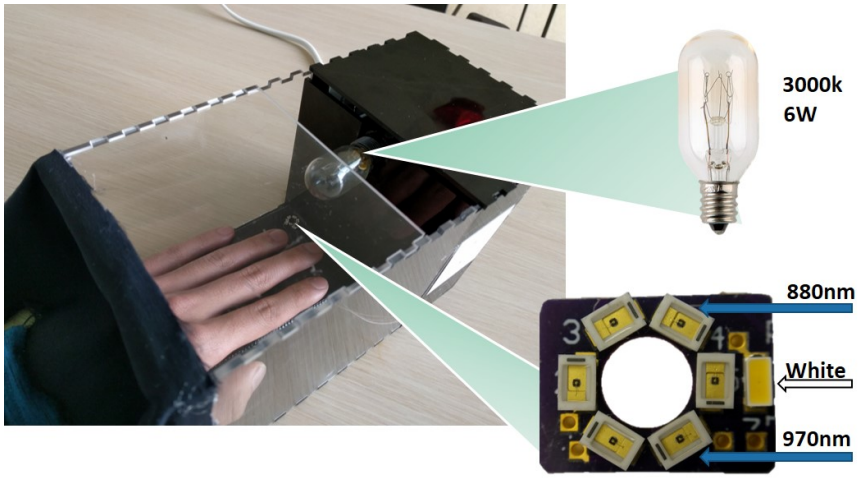

Figure 3.6: The experimental set up included an acrylic box housing a 6W incandescent light bulb of 3000k color temperature and a Nexus 5 smartphone with an LED array attached to its front facing camera. The array is equipped with white, 880nm, and 970nm LEDs. The participant places their finger over the camera while covering the LEDs.

study, we will be expanding to other hospitals nationally and internationally to have a wider spread of pigmentation to validate these claims.

3.5.1 Data Collection Procedure

In order to evaluate the effects of camera hardware and lighting conditions, we built a setup that allows us to efficiently cycle through all the combinations for every subject in our clinical validation. The setup is an acrylic box that contains a Bluetooth-enabled microcontroller that controls each of the light sources. The top of the box has a 6W incandescent light and a white piece of card stock with a hole cut in the middle. A Nexus 5 smartphone is placed in the box with the camera pointing up to the ceiling of the box. The LED circuit is then placed over the camera. The box is clear except for the portion holding the electronics, which lets ambient light shine through. The box also has a black cover that is used to block out ambient light.

The subject places the fingertip of the ring finger on their non-dominant hand on the camera. The subject is asked to sit still and not speak during the test to reduce movement. Each lighting

condition is then cycled through in the following order: incandescent, white, 970 nm, 880 nm.

The study consists of taking a series of videos of the participant's finger under various lighting conditions. Tests are done during the day, with no particular control over the ambient lighting conditions. An optical hemoglobin measurement is obtained using the FDA cleared Masimo Pronto 7 right before the recordings with the HemaApp system. The CBC blood test is used as ground-truth data and the optical Hb is used as a source of comparison to a specialized noninvasive device.

3.6 Results

3.6.1 Estimating Hemoglobin Levels

The top half of Figure 7 shows the predictions for each HemaApp embodiment comparing to the ground truth CBC, calculated using a leave-one-subject out cross validation. The bottom half of Figure 7 shows the corresponding modified Bland-Altman plot, where the residuals (HemaApp – CBC) are plotted against the CBC hemoglobin. For comparison, the results of the Masimo Pronto are included in each plot.

Results: HemaApp's hemoglobin estimations correlate with the CBC's predictions with a rank order correlation of 0.69, 0.74, and 0.82 with a mean error of 1.56 g/dL, 1.44 g/dL, and 1.26 g/dL respectively for each embodiment. The results of the Pronto are also compared to the CBC, which yields a rank order correlation of 0.81 with a mean error of 1.28 g/dL.

Implications: The improvement from EMB1 to EMB2 shows that simply supplementing the smartphone with an incandescent light both improves the correlation and decreases the error significantly. With the addition of an extra IR source, thus including both a 970 nm and 880 nm IR LED, the performance of HemaApp is comparable to the results of the Masimo Pronto.

A Wilcoxon signed rank test and an F-test of the residual variances fails to show statistically significant differences between the EMB3 predictions, Pronto's Hb predictions, and ground truth CBC values ($p>0.05$). An N-way ANOVA on the residual magnitude (|HemaApp EMB3 – CBC|) also did not reveal statistically significant effects on the residual magnitude due to age and race ($p>0.05$).

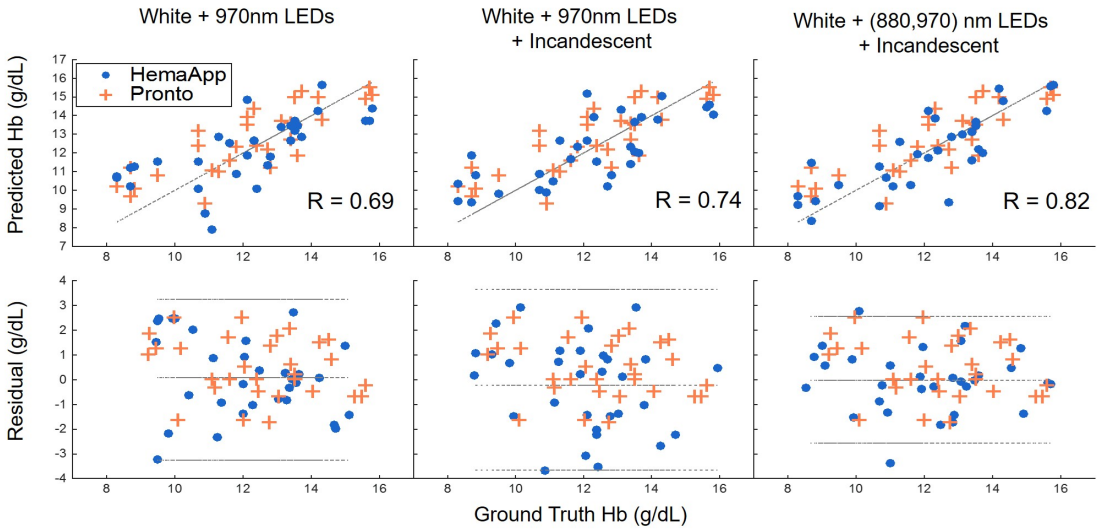

Figure 3.7: (Top) A comparison between predicted hemoglobin levels and ground truth hemoglobin level. (Bottom) A Bland-Altman plot showing residuals of predicted hemoglobin level against the ground truth hemoglobin. The ±1.96SD is shown for each plot. Embodiments 1, 2, and 3 shown from left to right. Blue represents the HemaApp results and orange represents the Pronto results.

3.6.2 Screening for Anemia

To evaluate how well HemaApp can be used as a tool for screening for anemia, the results of the regression are classified into two groups: anemic and normal. The classification is based on the average expected hemoglobin for each age group and gender. Figure 8 shows the classification results plotted against age. A different reference line is shown for male and female above the age of 16 as female tends to have lower hemoglobin than males. Those classified correctly are shown in gray. Higher sensitivity signifies better ability to detect the presence of anemia. Higher specificity signifies better ability to detect normal levels of hemoglobin. Table 3 breaks down the results by sensitivity and specificity for each embodiment and the Pronto. The 31 participants included 14 with anemia. The Pronto did not produce a hemoglobin measurement for one of the anemic participant due to low signal quality, thus the sensitivity results are for 13.

Results: HemaApp's classification results for each embodiment have higher sensitivity (78.6%,

85.7%, and 85.7%) as compared to the Masimo Pronto (69.3%). In comparison, the Pronto performed better in specificity, 88.2% compared to HemaApp (70.6%, 70.6%, and 76.5%).

Implications: The results are in favor of HemaApp as a useful screener for anemia. This can be seen in the sensitivity of each embodiment being near or above 80%. The Pronto on the other hand performed better in terms of specificity, at just under 90%. Comparing HemaApp to Pronto, the two systems are calibrated differently. HemaApp is tuned to be more cautious in order to reduce missing patients who are anemic, but at the cost of lower specificity. We believe this is a safer choice for use cases such as community health workers screening in the field, as missing an anemic diagnosis is a higher risk. In comparison, the Pronto's higher specificity is useful for situations such as in home monitoring. In situations where the patient is measuring frequently, having too many false alarms will result in unnecessary patient burden and hospitals getting frequent calls from unnecessarily worried patients.

Table 3.4: Anemia classification results for each embodiment.

	Sensitivity	Specificity
EMB 1	78.6% (11/14)	70.6% (12/17)
EMB 2	85.7% (12/14)	70.6% (12/17)
EMB 3	85.7% (12/14)	76.5% (13/17)
Pronto	69.3% (9/13)	88.2% (15/17)

3.7 Discussion

Our analysis of HemaApp shows that with some augmentation to the current smartphone hardware, our smartphone-based hemoglobin system compares favorably with the Pronto predictions. HemaApp cannot replace the CBC blood test, but can be used as an effective screener to determine whether further blood testing is necessary. The availability of the smartphone makes HemaApp a good candidate for at-home monitoring of chronically anemic patients, and its low cost can help equip community health workers in low resource areas screen for iron deficient anemia. The findings

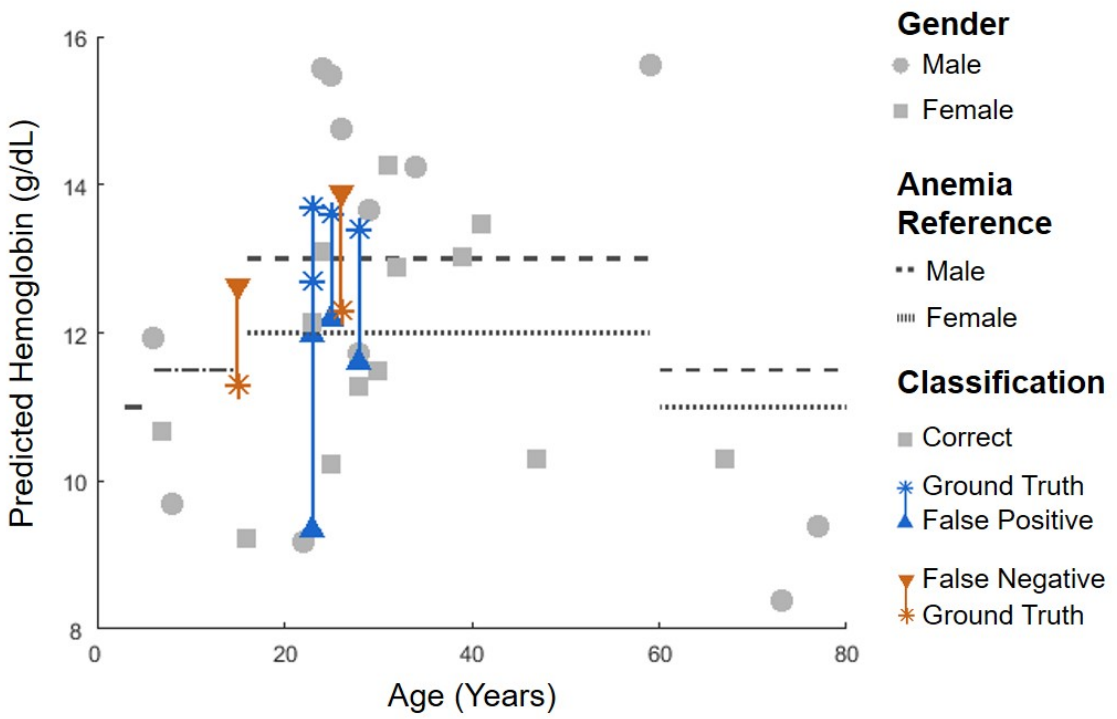

Figure 3.8: Anemia classification of HemaApp using typical hemoglobin reference ranges for each age and gender groups. Colored lines map incorrectly classified points to ground truth classification and hemoglobin predictions.

show that by using another ubiquitous item, the incandescent light bulb, the accuracy of HemaApp can be significantly improved. Although such light bulbs are beginning to be phased out in developed countries for more efficient alternatives, they are still easily obtainable as decorative lights in typical hardware stores.

3.7.1 Limitations & Next Steps

The current results are based on data collected on a Nexus 5 device and using only one brand of incandescent light bulb. As such, for the system to be more accessible, it needs to function on various types of devices and supplemental lighting. Different brands and models of phone use different cameras, lenses, and filters. The strength of the IR filter will strongly influence how strong

the IR LEDs need to be for penetration. The necessary adjustments based on these factors need to be investigated before the results of HemaApp can be generalized to different hardwares.

The incandescent light's properties also present another variable for future exploration. The current light bulb is a 3000K bulb, which places its peak wavelength at about 1000 nm, making it ideal for the detection of water. Another variable is the effect of the age of the light bulb. As light bulbs become less efficient, their emitted spectrum changes. It is possible that in order for the incandescent light bulb method to remain accurate over time, a calibration stage would be necessary after some time. Another issue with using an incandescent light bulb is ambient lighting. When using the data collected when ambient light is allowed into the experimental setup, the results of the best case EMB3 dropped to R = 0.60 and a mean error of ±1.72 g/dL. Although this does not limit the use of EMB2 and 3 in a dark room, it does suggest a user would have to make sure to use a lamp shade to reduce large fluctuations due to changes in ambient lighting. Further experiments will have to be done to determine to what degree ambient light exposure affects the results. We believe using a brighter light bulb and placing the hand closer to the lamp will reduce this problem significantly, limiting the effect of ambient light.

3.7.2 Limited Hemoglobin Range

The current data collection focuses on a range of 8-16 g/dL, which is sufficient for covering certain levels of anemia and normal populations; however, this means HemaApp cannot be used for those who are so gravely anemic that they need immediate transfusion or those with elevated hemoglobin. Our current data collection has come across a few patients in these extreme ranges, but we will need a more targeted recruitment to have enough data to develop the system to work in these ranges. To find people with elevated hemoglobin, we will need to expand our clinical study to include pulmonary clinics with patients that have high hemoglobin due to advanced pulmonary ailments. For obtaining extremely low hemoglobin, our collaborators have suggested taking data at midnight, when most of their transfusion patients are admitted and their hemoglobin measures as low as 4-5 g/dL. For this to work, our system needs to become much quicker, more robust, and self-contained such that running our study does not require the participation of the patient and fast enough that we do not interfere with the patient's clinical care.

3.7.3 Effects of Hemoglobinopathy

Only patients with normal red blood cell traits were recruited in our initial study, but we plan to expand our evaluation to include different blood disorders (e.g., sickle cell, thalassemia, and hemolyzing patients). Each of these disorders can potentially cause different optical absorption variations that may or may not be observable by our current measurement technique. Further studies of the optical properties of different hemoglobin types will have to be done through laboratory experiments. On the other hand, patients with hemolysis have ruptured red blood cells that results in free floating hemoglobin in the blood plasma. In the measurement of hemoglobin concentration, it is important to distinguish between hemolyzed and non-hemolyzed hemoglobin. In a laboratory test, the blood is first separated into the plasma and intact red blood cell components before hemoglobin concentration is determined.

3.7.4 Ensuring Data Quality

A major hurdle to making the system truly deployable is in making a system that can ensure data quality. All the data collection is conducted either by the development team or trained data collection assistants who have experience using prototype medical devices. The interface produces a real-time visualization of the camera feed to ensure that the patient has fully covered the camera/LEDs and the waveform being recorded is void of excessive movement.

To enforce proper finger placement and reduce movement, a finger cuff can be designed to center the finger over the camera. A snug fit for such a cuff would be ideal, but it cannot be so tight that the finger loses circulation. Furthermore, the finger cuff will also function as a cover to block the ambient light. A small window can then be integrated into the top of the cuff to be opened for the incandescent light.

An automated signal quality detector will also be necessary for detecting whether a signal is stable enough for analysis. The usual components in the signal consists of the pulse signal and the breathing signal, both of which are typically periodic between 0.3 – 2 Hz. An automated system can analyze for sudden DC shifts caused by the finger shifting, resulting in a non-periodic signal alteration. This signal quality detector then determines whether a segment of collected data is useable for analysis, prompting the user to perform the data collection again if it failed.

Chapter 4

REMOVING EXTERNAL HARDWARE

The initial development of HemaApp involved the use of external lighting. However, as noted in the limitations of the initial work, the system needs to be more stand-alone and easier to deploy. The focus of the initial development was to determine feasibility and identify whether certain lighting can ultimately work. In re-examining the initial approach to the work, I noted that it is likely possible to improve the feature strength by eliminating the reliance on measuring each wavelength sequentially. Since most of the strongest features required a cross wavelength normalization, the sequential measurement introduces a noise in the feature generation due to the inter-heartbeat variation in pulse amplitude. In revisiting the Android Camera API control, I found that it is possible to manually adjust the gains of individual color channels, providing an opportunity to capture individual wavelengths of vastly different amplitudes at the same time. In this section, I will go into detail the implementation of this new standalone system, which I will refer to as HemaApp V.2, and a performance analysis in a small validation study.

4.1 Phone Hardware and App

An RGB camera has three broadband color channels, giving a relatively even response to the entire optical spectrum. These cameras typically respond to light from the near infrared spectrum; however, to reduce glare from the infrared spectrum, manufacturers typically insert an IR filter, limiting the camera to visible spectrums. The absorption property of the blood is about two orders of magnitude higher in the blue and green wavelengths than it is in the red wavelengths. This means that under the same exposure settings, either the red channel will completely dominate and the green and blue channels will be unobservable, or the red channel will be clipped for the green and blue channels to be measureable. In HemaApp V.1, I had implemented a single channel measurement for each wavelength, turning on only the red, green, or blue at one time. That was satisfactory in the initial implementation as I was cycling through narrow band LEDs one at a time. However, this

leads to an inherent disadvantage of relying on the assumption that the pulse size does not change from beat to beat. Instead, if we can capture the different wavelengths at once, we can remove such an assumption. Prior work in using the smartphone camera and flash to perform pulse oximetry has suggested the use of white balancing techniques such as incandescent light mode [42]. In this mode, the blue channel is amplified to alleviate the issue slightly. We take this concept further by manually setting each color channel gain individually to adjust the relative exposures between channels. The current system is built on the Huawei Nexus 6p with an Android version 6 operating system. The Android 6.0+ uses the new Camera 2.0 API. This new API gives developers fine grain control of the individual amplification gain for each color channel. The white balance gain set for HemaApp V.2 is R = 0.25 | G = 0.85 | B = 3.00, where a gain of 0.0 turns off the color channel and a gain of 4.0 was determined to be the maximum. These values were determined empirically with 20 subjects of varying skin tones to determine the signal quality for each color channel. In this experiment, the amplitude of the fluctuation channel due to blood absorption was measured for each color channel under various gain settings. The above setting gave the most balanced fluctuation across channels while still maintaining a low noise that occurs at high gain settings. Although the maximum achievable gain is 4.0, this results in very noisy measurements. The frame rate is set at 30 fps and an auto exposure routine is used to set the integration to properly expose the camera once the finger is placed over the flash and camera.

A blue with yellow phosphorus fluoresced white LED is used in the development of HemaApp V.2. This type of LED generates white light using a blue LED coated with a yellow phosphorous coating which fluoresces a yellow light under illumination by the blue light. The resulting mixture creates a white light with a strong narrowband blue light mixed with a broadband light between 500-700nm. This type of white LED is more commonly used in a smartphone than a tri-color white LED which composes of one red, one green, and one blue LED, due to the quality of the white light produced. Although optical hemoglobin level is typically measured using a combination of visible and near infrared at 1300nm to estimate plasma volume, due to the limitation of smartphone CMOS cameras and the white LED, only absorption changes between 400 – 700nm can be observed. The system, instead, relies on the blue absorption to estimate plasma volume. Blue wavelengths can act as a proxy due to the bilirubin in the plasma, which has strong blue wavelength absorption. Note that the bilirubin concentration in the blood would thus likely affect the accuracy of the system.

HemaApp has thus far been tested in populations that are unlikely to have high levels of bilirubin.

4.2 Algorithm

This section will describe the algorithm that calculates a hemoglobin value from a video collected from the camera. The majority of the algorithm remains the same up to the feature generation. As there are only three wavelengths, there are limited number of combinations. The purpose of the initial validation is to check for whether there are feature significance, but avoid any overfitting. However, through the field deployment, a larger dataset should allow an introduction of machine learning based regression. By observing the different features extracted from the pulse signal, it was found that $I_{AC}(G,B) and R_{R,B}$, in particular the summation of the two, as the feature that best tracks hemoglobin concentration. A linear regression between $I_{AC}(G,B)+I_{AC}(R,B)$ and the ground truth generates a model that converts PPG to hemoglobin levels.

4.3 In-Lab Validation

To evaluate our system, we conducted a 32 participant study that measured the hemoglobin level using two systems: HemaApp V.2 and the Masimo Pronto 7, with participants ranging from 18 to 35 years old and having various skin tones. The performance of Masimo Pronto 7 has been evaluated in several independent clinical trials, demonstrating a mean error of about 1 g/dL against a blood test CBC golden standard. Although the performance of our system against the Pronto 7 does not directly represent the performance of our system against a blood test, the result of this test provides an estimate of how well the system would perform compared to a blood test.

For each test, a participant first had their hemoglobin measured using the Masimo Pronto 7. Next, they were measured by the HemaApp three times. Each recording was one minute long. The participant was asked to remove their hand between each measurement.

Accuracy

For each one minute measurement, the average $I_{AC}(G,B)+I_{AC}(R,B)$ is calculated. For each subject, three measurements were made and the median value of $I_{AC}(G,B)+I_{AC}(R,B)$ is used in the linear regression. We noted that in our 32 subjects, five subjects deviated from the regression significantly. We treated them as outliers in the regression, and fit the linear regression based on the

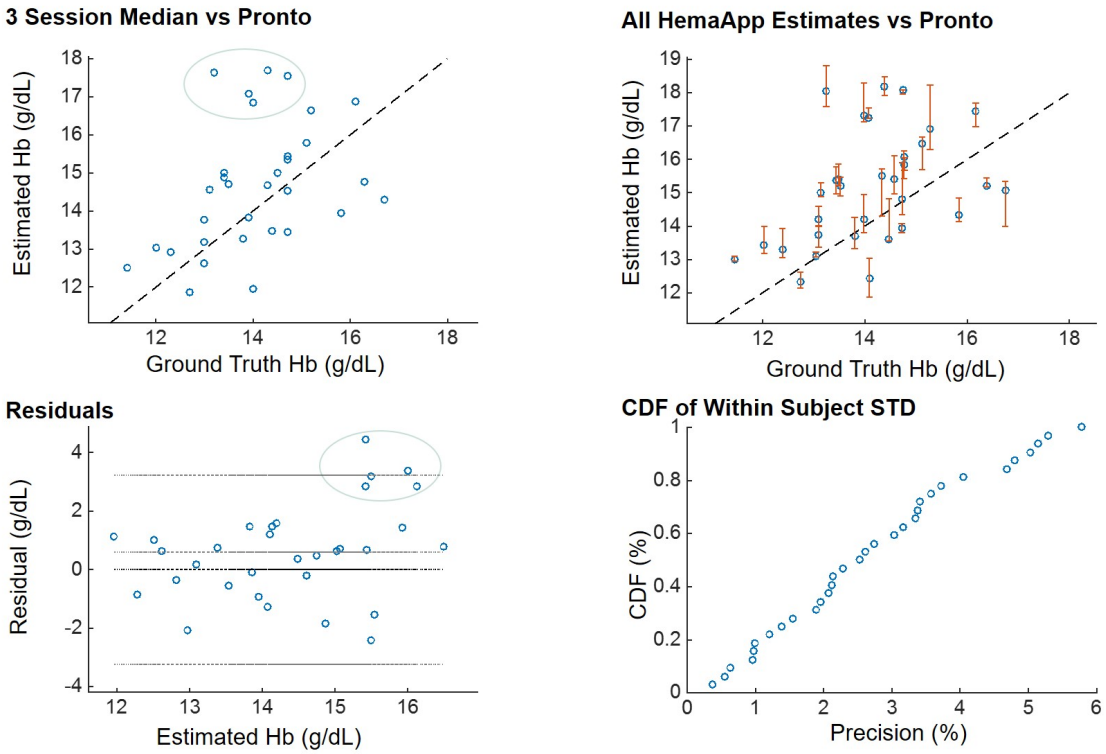

Figure 4.1: (Left-top) Linear regression results showing the HemaApp estimation of hemoglobin vs the Masimo Pronto's estimation. (Left-Bottom) Bland-Altman plot showing the HemaApp estimation's residual against the Pronto.(Right-Top) The linear regression shown with the median value for each subject and error bars indicating the two other readings. (Right-Bottom) The CDF of the precision of each subject's readings.

27 remaining data points instead. When considering the outliers, R = 0.42 and the RMSE = 1.93. When outliers are removed, R = 0.62 while the RMSE = 1.27. For comparison, Pronto's RMSE = 1.1 [72].

Precision Looking at the repeatability of the measurement for each subject, we separate the three Hb estimate as independent measurements and compare it to the Pronto measurement. The repeatability, calculated as the standard deviation divided by the mean of three measurements, are all under 6%, with 50th percentile of the CDF at 2.5% and 75th percentile at 3.6%. The Masimo Pronto's reported precision is 4.1% in a study where the device is used in two consecutive measurements [72].

Chapter 5

IN-FIELD USABILITY STUDY

5.1 Field Study in Peru

The initial validation of the phone camera and LED system provided the feasibility needed to carry forward with a field deployment using the new portable system. Through collaboration relationships established with Peruvian NGO Red Innova, I was able to setup a deployment study in Peru, focusing on answering two questions: 1) What is the performance of HemaApp in a target anemia screening population, and 2) What is the user experience of HemaApp when used by community health workers in actual anemia screening practice. The study was conducted in conjunction with a rural community anemia screening campaign in Indiana, Peru, a community in the Amazonian Jungle. Red Innova operates anemia screening campaigns, during which, community health workers from their Lima and Iquitos office transport finger prick based hemoglobin measurement equipment (Hemocue) to the communities and host a 2-3 day adhoc clinic to screen for anemia, provide micro-nutrient supplements, and general health education around anemia and sanitation practices.

Just before deploying for the study with Red Innova, I had organized a trial run with another NGO operating in Arequipa, Peru, which ran a similar version of the rural community screening. In this trial, I did not travel to the study site to observe the study and, instead, had trained one of the study coordinators who took the setup with her to the study site. In a conference call after the deployment, I received the feedback that it was unclear whether the subject's finger was in the right place and that it was confusing for the staff to determine whether they were doing the procedure right. In analyzing the data, I noticed two main issues: 1) the finger was often misaligned, and 2) even when the finger covered the camera, the subject's fingers were pressed so hard that the blood does not enter the finger. To alleviate this, I adjusted the setup to include a finger holder made by cutting two slots in a black phone case and slipping a black cloth through the slots to make an adjustable finger holder. The purpose of the finger holder is to help provide alignment and appropriate pressure on the finger. In this way, to make a measurement, first the staff has the subject

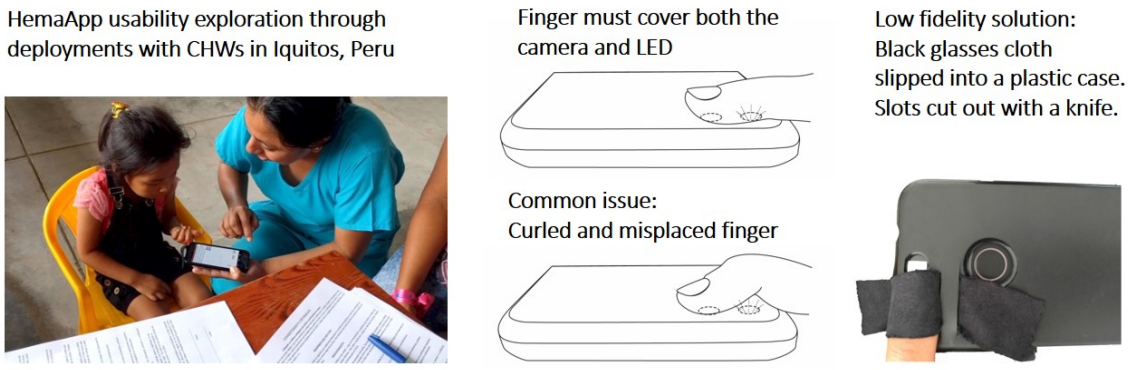

Figure 5.1: Data collection in Peru. Finger not placed on the right place on the camera is a common issue. To solve this, a low fidelity solution of a black cloth weaved through a phone case was devised.

put their index finger into the black cloth holder. Then they cinch down the cloth. And then turn the phone around such that the staff can see the screen to check the quality of the signal and hit start. Through training, I explain to the study staff how to identify quality signal and signal that is of poor quality such as when the subject is pressing too hard.

The following sections will provide the details of each study location, insights gained from discussions with local staffs, researcher observation in person, and data review.

5.2 Field Study Procedure

We mainly used the same procedure for each study location, only adjusting slightly depending on the needs of the conjoint studies as we either needed to streamline our data collection protocol to the existing clinical procedure or study. The core procedure is as follows:

1. Introduce the HemaApp study to the subject. Explain to them that this will not change their clinical procedure and is part of a research aiming to make hemoglobin measurements more noninvasive. Since many subjects are minors, the consent process was mainly conducted with the subject's parents. The procedures were approved by an ethics review board in Peru.

2. The HemaApp measurement is conducted before the reference measurement. In the clinical

office and home study, the reference measurements are done using a venapuncture while the community screening campaign relied on a finger-prick Hemocue measurement.

3. The staff requests the subject to put his or her index finger over the camera, which is under the black cloth holder. Once the finger is in place, the black cloth is cinched down on the finger. The phone is then turned around to face both the subject and the staff.

4. The staff checks the pulse signal to make sure the finger is placed correctly and the signal is coming through properly. If the signal looks to be incorrect, the staff adjusts the subject's finger either by having them loosen their grip or adjust the placement of the finger, whichever is appropriate.

5. Once the signal looks appropriate, the staff starts the measurement. The measurement lasts 15 seconds. During which time, the subject is asked to remain still and not move their finger position, not speak, and remain still.

6. Once the measurement is complete, the subject's hemoglobin is measured using the reference measurement.

Figure 5.2: Data collection in Peru. Me working with a children to obtain a HemaApp reading.

5.3 Insights from Field Deployment

As the study's main focus was on development of the algorithm, the study was not designed in a formal ethnographic style of observational study. However, to the best of my abilities, I recorded personal observations and interviewed the study staff during and after the study to gain insights on staff experience, subject experience, issues with the procedure, and how all of this can ultimately lead to better user experience to promote better quality data capture for future studies.

The main issue we found at all the sites is many children were scared during the procedure and often resulted in either drawing away from the measurement phone or crying when the parent or staff tried to keep the child's hand over the phone. Even the slightest level of fear would result in a finger curling response, which is a common pediatric fear response. This causes issues with the HemaApp measurement. A proper measurement requires the finger to be placed steadily over the camera. Curling causes both a misalignment of the finger and also excessive pressure that reduces perfusion at the finger tip. I noted that the staff did not fully understand the issue of this at first as they were use to typical blood draw based hemoglobin measurements. Blood draws does not rely on the subject being still. As long as blood is pulled out, the procedure is considered successful. As a result, the community health workers are used to grabbing the child's hand forcefully and swiftly using the finger prick device to get blood out of the finger. No matter how much resistance is met does not affect the procedure. However, for HemaApp, this type of subject interaction is counter productive. Note that many of these observations are subject to cultural influence. Most of these children have previously attended anemia screening clinics where they have experienced fairly forceful finger pricks and an overall stressful experience. Due to these past experiences, they are more likely to associate the entire experience as stressful even if our procedure does not involve any puncturing and pain.

To mitigate the issues caused by general fear of participating in medical procedures, the staff and myself came up with a variety of compensatory solutions to try to reduce the stress on the children subjects. First, the immediate solution that each of the staff gravitated towards was to try to tell the children about the measurement and how it's just a camera. This seems to have worked only when they first show the front of the phone, but when the phone is turned around and they ask for the children's finger, many children immediately started crying. We were unsure why that is, but

it may be due to unfamiliarity with the device and associating to how the procedure to how most finger pricking procedure occurs. Some children calms down once the phone is turned back around and they see the moving signal and begins to pay attention to the situation and the lack of pain. However, often times the procedure do not go beyond trying to position the finger on the camera or even if the finger is on the camera, because the fear has already set in, the subject is unable to stay still during the measurement, resulting in an unusable measurement.

Interestingly, we found that when children sit in their own chair, and not on their parent's lap, almost none of the children started crying nor initiated fear responses. Although we did not extensively study this, we found that this seemed independent of age. On the first day of the trial, we found that most of the children, even the older ones were quite scared. But on the second day when we began to instruct the staff to actively have parents let their children sit in their own chair, the entire procedure became much smoother. In future procedures, we will incorporate this as part of the procedure.

5.4 Issues in Data Analysis

In the Amazon portion of the study, we were able to capture 83 measurements that passed the quality check. Applying a similar linear regression analysis in the HemaApp V.2 study, I found a much weaker regression. Given the issues I found in the data collection and the resulting reduction in signal quality, the poor regression is not surprising.

First, the data quality is the major issue. Even the data that made it past the manual data quality checking, it is unclear, based on in-person observation, how well the measurements were actually made. However, this is likely not the main issue.

Second, and more importantly, the baseline measurements were made using a Hemocue device for the study as the rural campaign could not perform CBCs. Hemocue measurement, in it of it self, has a fairly large margin of error, with reports between 0.5 to 2.3 g/dL. As we are treating this as both a training and testing data set, the high margin of error of the Hemocue baseline leads to issues of poor ground truth. Instead, it would be ideal to collect data based on CBC measurements as the baseline and re-test the estimation of the measurements made in this field study to examine the accuracy.

5.5 Summary of Learning

The study in Peru, although did not provide me with the data to study the performance of HemaApp further, did provide invaluable information through observation of usability to improve the overall protocol. First, the procedure involving placing the subject's finger in the cuff and then turning the phone over to observe the screen was difficult. Instead, it will be easier to simply not use the screen as a feedback mechanism and let the phone be placed screen side down so that the subject can put their finger on the camera directly. A useful future development would be an automated algorithm that can then detect the quality of the signal to provide feedback to the subject staff and automatically begin the measurement. Second, it is important to make sure subjects are calm. This is particularly difficult with young subjects. Instead of having parents hold the child, anxiety can potentially be managed by first giving more independence to the children by allowing them sit in their own chair and placing the phone on the table and letting the children put their finger over the camera themselves. Third, although we did not try this in the field, some of the community health workers suspect the look and feel of the device may contribute to the general scariness of the unknown procedure. They suggested that instead of a pure black cloth, a cloth with black on one side and a brighter color on the other side can be used to make the device look more welcoming.

Chapter 6

SEISMO: BLOOD PRESSURE MONITORING USING BUILT-IN SMARTPHONE SENSORS

6.1 Blood Pressure

High blood pressure (BP), also known as hypertension, is a medical condition in which the blood exerts high force on the artery walls as it circulates through the body. According to the WHO [81], hypertension causes 7.5 million deaths per year, about 12.8% of all deaths worldwide, and in 2008, around one billion people had hypertension. Frequent monitoring of BP at home has been show to improve the management of hypertension and effectiveness of treatments over infrequent in-clinic monitoring [2]. Currently, at home monitoring of BP relies on automated oscillometry-based BP arm cuffs for daily measurements. Although this way of automated way of measuring blood pressure enable out of clinic monitoring, they are cumbersome and inconvenient [76, 71].Due to these drawbacks, patients typically only perform arm cuff-based measurements once a day, and thus cannot capture the fluctuations in BP that occurs throughout the day due to exercise, activities, and other stressors.

Advances in mobile technologies now present an opportunity to have always available health monitoring capabilities as more and more people carry a smartphone around with them. Incorporating BP monitoring using only the built-in sensors on all smartphones has the potential of enabling unobtrusive and convenient BP monitoring at home and frequently throughout the day. As a software-only solution leveraging commodity devices that are already disseminated widely, it would be possible to enable higher adoption, and as a result improve the population awareness of their cardiovascular risk.

In this work, we develop and evaluate *Seismo*, a BP measurement technique using the existing sensors on the smartphone. Seismo uses pulse transit time (PTT) – the time taken by the heart's pulse to propagate between two arterial sites – which is inversely related to BP [37, 82, 88]. In particular, Seismo tracks the time when the blood is ejected from the heart as the aortic valve opens

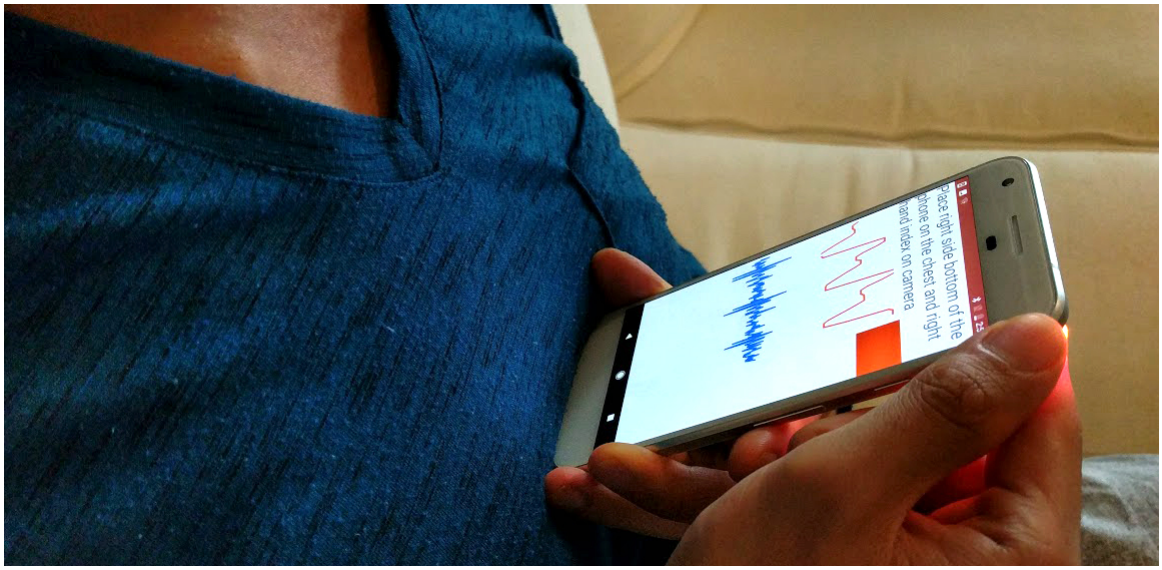

Figure 6.1: Measuring blood pressure using Seismo, a smartphone application that uses the built-in accelerometer and camera to calculate pulse transit time.

and when the pulse arrives at the fingertip. To perform this, Seismo relies on Seismocardiography (SCG), which uses the vibration caused by the movement of the blood and valve activities as the heart beats, allowing for accurate measurement of aortic valve opening time. The SCG is captured using the phone's accelerometer pressed against the chest (Figure 6.1). In this position, the user holds the phone with their finger covering the camera, which then captures the photoplethysmogram (PPG) at the finger, thus measuring the pulse as it arrives. This technique conveniently captures both the proximal (close to the heart) and distal (away from the heart) timing all from one device, without the need for any supplemental hardware. Additionally, PTT-based techniques can measure beat-to-beat BP, thus it can more reliably measure short-term BP changes (such as post-exercise), which are difficult to measure using cuff-based devices. One major distinction between Seismo and previous solutions that enable smartphone blood pressure tracking without additional hardware is the use of accelerometer to capture SCG as a proximal timing. Other work has mainly focused on using the sound created by the heart, otherwise known as phonocardiography (PCG). However, the fundamental limitation of using PCG as a proximal timing is that the sound being captured is actually created by the *closing* of the heart valves rather than the *opening*, thus not an ideal reference

for when the blood is actually ejected. Although prior work have demonstrated the use of SCG and PPG to reliably capture PTT for measuring BP, they use custom hardware with ultra low-noise accelerometers [15] to resolve the signal, while we focus on using only off-the-shelf commodity smartphones.

In coming down to the above design, I performed a series of prototype pilots. In this initial feasibility work, I developed test applications that implemented a variety of sensing modalities to explore the many instantiations of PTT measurements that can performed with a smartphone. This was a very breadth first search, with two main branches: (1) a two camera mode and (2) a microphone and camera combination. The two camera mode is used to capture PPG from two distal sites and PTT captured using the time difference of arrival at each site. The microphone and camera combination implements the PCG measurement as a proximal timing and PPG from the finger. As mentioned above, PCG is not an ideal measure, but the feasibility study was performed prior to the team learning about the nuanced difference between PCG and SCG signals. The properties of the signal and general usability finding are still mostly relevant. Through the series of pilot studies and multiple development iterations, we created the Seismo application, which uses an accelerometer in the phone to capture SCG and the camera to capture finger pulse arrival time.

To measure the effectiveness of Seismo, we collected data from nine participants in four lab-sessions each. Participants biked at multiple intensities during each session to raise their blood pressure. For each participant, the first session BP measurements are used for calibration and learning. Based on which, the subsequent three sessions of blood pressure measurements are estimated. We note that Seismo struggles to capture reliable SCG and PPG signals from two of the nine participants. One participant has very shaky hands during the measurements, which is accentuated when holding the phone in the measurement position. The other participant has a comparatively weak heart beat, while also having more muscle and fat tissues between the accelerometer and the heart, which significantly reduces the transduction of the signal. Of the seven participants whom Seismo is able to capture clean and extractable signals, Seismo achieves Pearson correlation coefficient scores of 0.20-0.77 (μ=0.57, σ=0.15) and RMSE of 3.3-9.2 mmHg (μ=0.52, σ=2.0) for each participant.This is comparable to prior work [39, 87] that uses custom hardware to measure PTT, demonstrating the feasibility of using commodity smartphones to capture SCG and PPG for BP estimation, while we also identify the limitations of our system.

6.2 Background and Related Work

6.2.1 Blood Pressure

Blood pressure is the pressure which the circulating blood exert on the walls of blood vessels. During each heart beat, blood pressure reaches a maximum, systolic, and minimum, diastolic, pressures. Typically, blood pressure, when addressed to without further context, is referring to systemic blood pressure in the central arteries and measured in millimeters of mercury (mmHg) above surrounding atmospheric pressure. Blood pressure is desirably between 90-129 mmHg for systolic blood pressure and 60-84 mmHg for diastolic blood pressure. This pressure is needed to maintain circulation of the blood through the entire body, and is influenced by cardiac output, total peripheral resistance, and arterial stiffness.

Fluctuations occur depending on situation, emotion, disease state, digestion, physical activity, and with the time of day (cite). The circadian cycle influences blood pressure, with blood pressure typically being higher in the morning and lower at night(cite). There is evidence that show nighttime blood pressure is a stronger predictor of cardiovascular outcome of patients (cite).Diet also influences blood pressure. A high sodium diet tends to increase the blood volume in a person, thus increasing cardiac output thereby increasing blood pressure. Age and sex also affect blood pressure. Arteries tend to stiffen as adults age, which is thought to be the reason why elderly tend to have higher than normal blood pressure.

6.2.2 Blood pressure measurement

Blood pressure is directly measured through invasive insertion of a catheter tube and pressure transducer into the artery. This is typically only done in intensive care units and is considered the most accurate method for measuring blood pressure. However, invasive measurement using an arterial line catheter is not feasible for most use cases. Instead, blood pressure is more commonly measured at upper arm's brachial artery as a proxy for estimating the blood pressure in the central artery. The more common way to measure blood pressure is using a sphygmomanometer, which uses an inflatable cuff worn on the arm to collapse and then release an artery in a controlled manner and a manometer to measure the pressure. Although it is less reliable than an intra-arterial catheter, it is non-invasive and accepted by the medical community when the sphygmomanometer has a grade

A/A rating (cite). Several easy-to-use electronic home sphygmomanometer devices are available as well.

6.2.3 PTT vs PAT

Compared to cuff-based BP measurement approaches, PTT-based techniques provide continuous beat-to-beat BP measurements. A typical method for obtaining PTT [37, 62, 64, 83] is to use the R-wave of the ECG to mark the timing of the genesis of the pulse, and measure the pulse at the periphery (typically at the fingertip) using PPG. However, the ECG R-wave is not a reliable timing marker for the genesis of the pulse [15, 88, 62], as it marks the electrical stimulation that would then trigger the start of a heartbeat, when the blood leaves the heart.

The time between the occurrence of the R-wave and the actual aortic valve opening (AO) is called the pre-ejection period (PEP). In reality, systems using a combination of ECG and PPG to measure PTT are actually measuring Pulse Arrival Time (PAT=PEP+PTT) to estimate BP, by wrongly assuming PEP to be constant. Previous work has shown that PAT is not an accurate measure to estimate BP [15, 88], as PEP is connected to the nervous system activity and is variable up to tens of milliseconds. Instead, prior work have shown that a better way to resolve PTT is to capture the mechanical vibration at the chest caused by the series of muscular contractions during a heart beat, known as seismocardiography (SCG) [30]. The advantage of SCG is that the signal captures the valve movements of the heart, showing the mitral valve closure time (MC) and aortic valve opening time (AO).

In this work, similar to [15, 39], we rely on SCG instead of ECG to capture the proximal timing of aortic valve opening in order to measure accurate PTT value (not PAT) for BP estimation (Figure 6.2). PTT-based BP estimation approaches suffer from a few limitations. PTT can only estimate relative changes in blood pressure, *i.e.*, an absolute measure of PTT cannot be extrapolated to an absolute measure of BP. Also, the relationship between PTT and BP is dependent on the physiological properties of an individual. Both of these limitations can be addressed by performing per user calibration, and is an active area of research. Various methods have been proposed to determine the calibration between PTT and BP [62]. The most common is to perform a set of interventions that would perturb an individual's blood pressure. Exercises such as running and biking

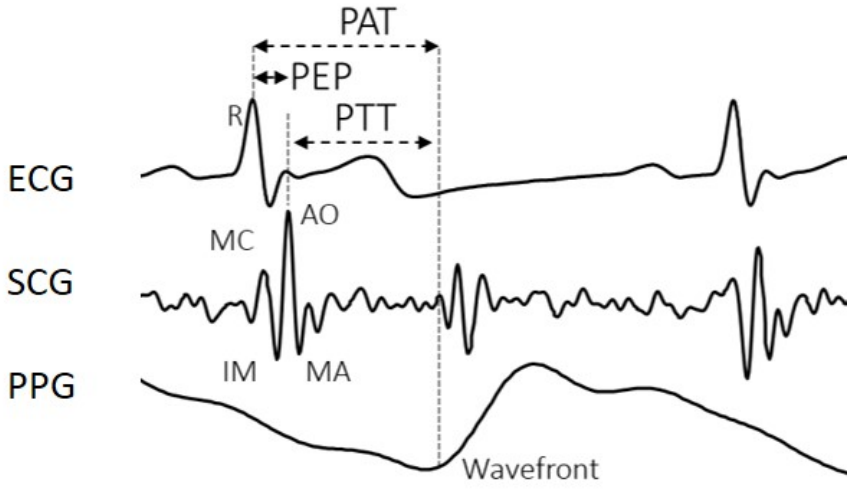

Figure 6.2: PTT using SCG and PPG as measured by the Seismo system on a Google Pixel phone. ECG based timing results in the inclusion of PEP, and thus measures PAT.

are effective methods to drastically increase the blood pressure. Other methods include postural changes, valsalva maneuvers, and cold pressor [62]. Finally, another limitation of PTT is that it is most strongly influenced by diastolic blood pressure. As such, in our evaluation we focused only testing the correlation between PTT and diastolic blood pressure. However, because diastolic blood pressure changes and systolic blood pressure changes are often strongly correlated, PTT remains a highly useful for monitoring blood pressure fluctuation.

6.2.4 Measuring BP using Wearable Devices

Researchers have proposed PTT-based wearable devices to continuously sense BP throughout the day [15, 22, 29, 82]. Glabella is a wearable glasses [29] with three optical sensors for PPG to measure PTT on different arterial sites on the face. Fortino and Giampa [22] use a combination of a ring and a wrist band with optical sensors for two separate PPG measurements, while SeismoWatch [15] use a wristwatch form factor with optical and IMU sensor to calculate PTT. SeismoWatch [15] shows the usage of SCG and PPG to accurately measure PTT and predict BP. However, it uses custom wearable hardware with ultralow-noise accelerometer (ADXL354) and IR photodiode arrays

to measure SCG and PPG, in order to measure high resolution SCG and PPG values. Previous work has shown that SCG can be accurately measured using smartphone's built-in accelerometer [43, 60], though only examined more measuring heart rate and heart rate variability, not pulse transit time. In our work, we explore measuring SCG from the phone's accelerometer by placing the phone over the chest, and measuring PPG at the fingertip from the same phone's camera (Figure 6.1), to determine PTT and estimate BP while the user is seated and with measurements taken over clothing. This is of significance because the measurement of PTT necessitates better capturing of distinct time characteristics of aortic valve opening time over just identifying the presence of a heart beat.

6.2.5 Measuring BP using Smartphones

As wearable devices are potentially cumbersome and are not widely adopted, researchers have explored exploiting the ubiquity of smartphones to measure PTT and predict BP [17, 49, 65, 35]. Most of the phone-based BP techniques [17, 35] using a combination of the smartphone's microphone to record the heart sounds created by the closure of the mitral and aortic valves, called phonocardiogram (PCG), and the smartphone's camera to determine the heart beat at the fingertip. As the smartphone microphone is sensitive to noise, a stethoscope-based hardware is usually attached to enhance the smartphone sensing capability [65, 17]. Dias *et al.* [35] has demonstrated the use of the commodity smartphone without any attachments to capture PCG. However, as mentioned previously, PCG does not provide the correct time to determine PTT, as PCG captures the closing of the valves sound, not the opening. There is currently a gap in using built-in smartphone sensors to accurately capture PTT, which our work addresses through the use of accelerometer and camera paired PTT measurement.

Chapter 7

PHYSIOLOGICAL SITE TESTING FOR PTT MEASUREMENTS

In our development of the Seismo system, we did a broad exploration of the many different ways of capturing PTT using smartphone sensors. One of the key work that was carried out was implementing a number of methods to measure pulse signals on a smartphone that were demonstrated in prior work [50, 17, 35, 43, 60]. We found that many of the configurations presented in related work was indeed usable in a lab setting where a trained staff is present for obtaining good data, but when implemented as an app meant for use by individual users, many configurations became infeasible. The following is a thorough documentation of the various pilot studies and design consideration that went into the final design of Seismo.

Using a smartphone's built-in sensor, we implemented three ways to measure the pulse to determine the pulse transit time: camera to measure PPG, microphone for PCG, and accelerometer to measure SCG. In our final configuration, we chose to use the camera based PPG and accelerometer based SCG. The choice between using SCG and PCG came about from personal communication with other researchers in the field of PTT measurement. The literature has identified that SCG is a more reliable measure of the aortic valve opening timing. However, prior to making this decision, we had also spent considerable time examining the PCG measurement, and thus include that exploration in this section. PCG measurements, although may not be as suitable for PTT monitoring, is still a useful signal for capturing cardiac activity. We group the following in the methods in which the sensors are combined to capture PTT by combining two of the signal sources. First is using two PPG sources on different parts of the body, referred to as the PPG/PPG based system. The second is using a PPG source and a SCG source.

7.1 PPG/PPG based systems

Using two PPG measurements at different distances from the heart provides a PTT measurement that can be used to estimate changes in blood pressure. The best location to measure PPG is the

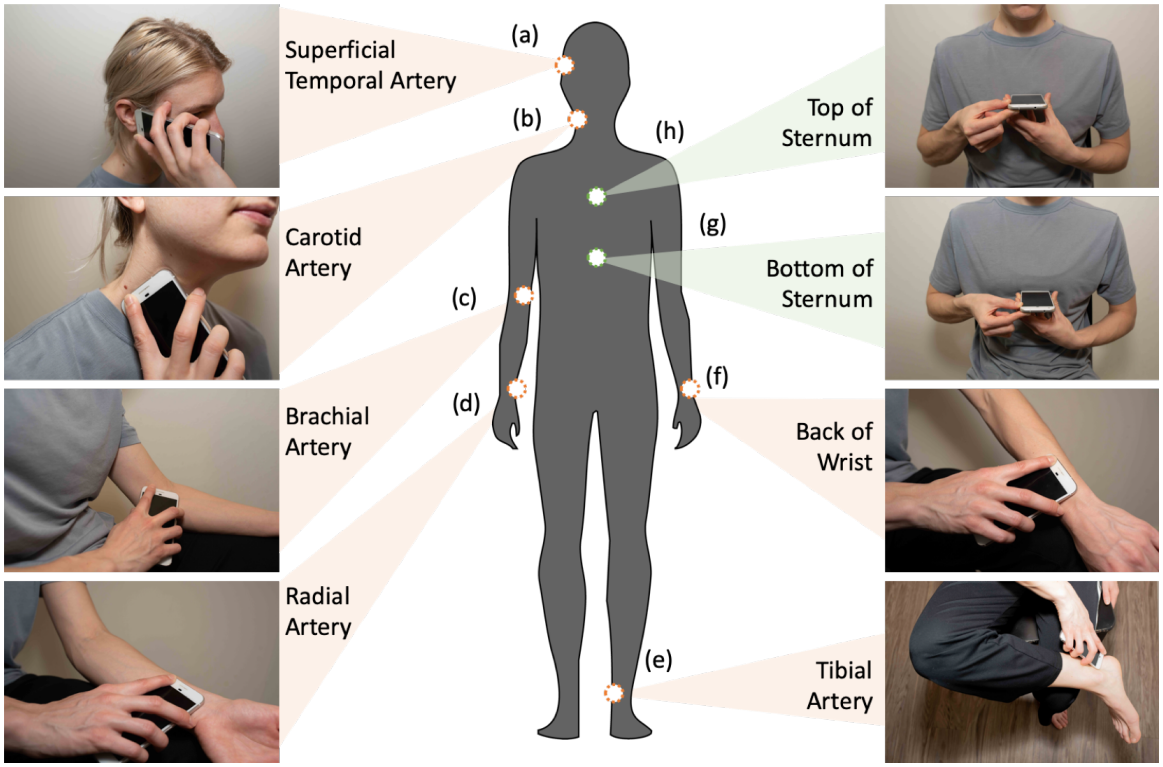

Figure 7.1: Different body sites to record PPG and PCG.

finger. The blood is typically significantly perfused, the location is easily accessible, and the hand is already operating the phone. The challenge is to determine another site on the body to measure the second PPG. In various prior work in both mobile health research and clinical health research, various second site have been proposed. PPG can be measured at various locations on the body where a main artery is directly under the skin or in the peripherals where blood is significantly perfused, meaning blood is distributed to the capillaries. Arteries that can be used to measure PPG from include the brachial artery (at the elbow) the superficial auricular artery (in front of the ear), carotid artery (at the neck), and tibial artery. Body peripherals with significant perfusion include the fingers, toes, forehead, earlobe, and the wrist. The different sites are shown on figure 7.1. However, no PPG/PPG based system have been implemented on a standalone smartphone application format, so it is unclear which of these pairings to measure the two PPG locations using the smartphone's dual cameras is feasible.

To experiment with measuring from the various sites, we implemented a smartphone app on the Nexus 6P Android phone that can simultaneously access both the front and back facing camera and operate both at 30 Frames Per Second. This is made possible with the use of the Android Camera2 API and a smartphone with a GPU with multiple Image Sensor Processor (ISP), which allows it to stream and process more than one camera feed at once. The Nexus 6P uses the Adreno 430 GPU We tested the same app on multiple phones and noted that the dual camera access was capable on the LG G5, using the Adreno 530 GPU, and the LG Nexus 5X, using the Adreno 418 GPU. This app, however, did not operate on the Google Pixel from 2017, even though it uses the same Adreno 530 GPU as the LG G5. We suspect this to be a firmware difference designed to limit the application developers from accessing both cameras at once to reduce potential instabilities in apps.

To use the app, the user places one finger over the front facing camera and places the back facing camera against another part of the body shown in Figure 7.1. Because the back facing camera is the only one that has a LED flash sitting beside it, it is suitable for measuring reflective PPG. Aside from the toe and the earlobe, all the other PPG sites require the use of reflective PPG measurement due to the inability to shine light through the tissue from the other side of the body.

To test the app, we conducted a series of design iterations and pilot studies with three different groups of users: 1) students within the same research laboratory between the age of 20 to 30 (n=10), 2) academic students and staffs between the age of 20 to 50 (n=30), and 3) senior residents at a nursing home between the age of 70 to 85 (n=10). We aimed to understand if the signal quality obtained from the PPG is suitable for PTT measurement and if it was possible for the users to operate the app. Through testing multiple versions of the app with these three groups of users, we found four major issues with the use of PPG/PPG based blood pressure measurement with a smartphone: difficulty positioning the phone due to a lack of visual feedback during the use of the app, signal quality measured from some locations are too inconsistent, access to certain body locations introduces awkward postures and inconvenience, and flaws in camera synchronization using stock phone firmware.

7.1.1 Lack of Visual Feedback

One major limitation we identified is that many of the PPG measurement locations result in the inability to get visual feedback from the smartphone screen when performing the measurement. Specifically, when the phone is placed on arteries not on the extremities like the carotid by the neck site (b) and the superficial temporal artery by the ear site (a), the user cannot see the screen. Placed against the neck, the screen is below the user's chin. While against the ear, the screen points away from the user in order for the back facing camera to aim towards the face. Although we did not formally try measuring the PPG from the forehead with users in groups (2) and (3), it is clear that this form factor suffers from the same limitation.

In prior work [50], the use of the artery near the ear is viable because the data collection is assisted by the use of a laptop that visualizes the data in real time. However, in a end-user application, it is unreasonable to assume that the user would be able to stream the data or mirror the screen on another computer screen. In our pilot study with groups (2) and (3), we used a laptop screen to visualize the data in real-time to assist in placing the camera to gather PPG. We found that due to the lack of clear anatomical marker that is tacitly easy to identify, it is extremely difficult to place the phone's camera to capture the PPG on first try. Only with the help of the mirrored screen were we able to place the phone in such a way that we can capture PPG that is usable for analysis. With group (1), a few of the developers of the app were able to learn over a course of using the app and with visual feedback with the mirroring for a few weeks to learn to successfully place the camera in the same location that had the best pulse without visual feedback. This potentially suggests that with enough practice and some feedback, users can learn to place the phone properly without feedback. From a end-user standpoint, this seems infeasible.

Because visual user interface is not possible for these PPG measurement locations, we developed a few prototypes using audio feedback. One prototype modulated a single tone sound output by increasing the volume and decreasing the volume with respect to the amplitude of the PPG. We tried the interface with group (1), who were researchers familiar with the concept of PPG. Of the five people who tried using the audio feedback interface, all found it potentially possible to learn to use, but did not feel that it was easy without significant training. The issue a few of the participants found is a lack of mental model for what a good PPG "sounds" like.

Another audio feedback system we tested is an automated feedback system that tells users when the phone has been properly placed. We wizard-of-oz'd an automated signal quality detector system to mimic a system that can provide audio feedback for users when they reached a proper position. The users in group (2) were asked to place the phone on their superficial temporal artery. The users were shown an example of where to roughly place the phone along with an explanation to place the top corner of the phone against the front of ear. The user makes an initial attempt at placing the phone. Looking at a mirrored screen feed of the PPG signal, the staff determines whether the signal quality is good. The staff then tells the participant if the signal is good or bad, and asks the participant to move it. In the case where the participant successfully placed the phone by the ear, asking the participant to move the phone a little up or a little down a few times generally led to usable signal. The major limitation we found is that due to the lack of anatomical markings that clearly constrain the user, there is not a clear strategy for guiding the user in finding the pulse unless they were already very close. If the user placed it over their hair, from the PPG, it just looks like a lack of signal. If they placed the phone too far down, it takes more than five minutes to slowly move the camera around before the camera simply gets overheated and is too painful to keep using the app. Furthermore, users typically did not understand they cannot just keep moving the phone around. PPG is a signal that takes at least a few heart beats to make sure the signal looks correct. The user actually needs to move the phone, wait for about 5 seconds before a decision can be made. This was almost universally unclear to the user unless we explain why the 5 seconds are needed.

From our learning in working with the lack of visual feedback with our prototype app, we concluded that performing PPG measurement at sites that do not allow for visual feedback is not feasible without significant re-examination of the feedback provided to the user. Audio feedback interface has potential and further research in this form of interface could potentially improve the usability. However, through testing our wizard-of-oz'd automated system with untrained users, we found that the superficial temporal artery and the carotid artery lacked a clear anatomical position and even with a feedback system that tells the user the position is correct, it takes too long to find the position.

7.1.2 Poor Signal Quality

Although all of the PPG sites shown in Figure 7.1 exhibit the possibility of measuring a signal, the quality of the signal differs depending on the site. The quality of measurement is affected in a few ways. For arteries, it is particularly difficult to measure the same signal each time. This is because the camera must be placed directly above the artery. In testing with groups (1-3), we found that two issues arise when the camera is not placed directly above the artery. First, for locations that are difficult to palpate, or feel, the artery, it is easy to completely miss the artery, leading to no signal. This is particularly true for the brachial artery located at the elbow which is difficult to locate. For arteries that are very close to the surface, a different problem exist. For the carotid and superficial auricular artery, if the camera is placed off to the side of the artery and instead the flash is placed over the artery, the PPG that is measured becomes inverted. The exact reason for this is not clear, but is suspected that the arteries can sometimes pulsate strong enough for a person that the surface of the skin displaces in response to the pulse, leading to a change in the contact of the skin with the camera. We found this to be more common for younger people who have a strong pulse. Another way the quality of the signal is affected is a lack of a clear peak in the PPG for a timing reference. This was true for the signal measured at the forehead and the back of the wrist where the blood vessels feeding the skin area originates from multiple arterial sources. Although the PPG's pulsatile nature is clearly measurable and the heart rate can be measured, the peak is blurred and the peak from beat to beat shifts by at least 10-20 ms. This is particularly limiting for the wrist measurement when used to compare with the finger because of the extremely short distance from the wrist to the finger. Given that the change in PTT averages to be about 10ms for every 10mmHg of difference from the heart to the finger tip, the short distance between the wrist and finger would necessitate about 1ms resolution for every 10mmHg of change.

7.1.3 Inconvenient and Difficult Posture

A few of the PPG sites that have been used in prior work for PTT measurement involves measurements from areas that are not easily accessible. This is particularly true for measurements on the legs and feet. The femoral artery is accessible only when the user either rolls up their pants up to their hip level or remove their pants. This might only be possible in a limited number of scenarios

such as at home or in the changing room, thus would significantly limit the usability. Furthermore, placing a phone between the thigh is inconvenient and is likely not socially acceptable. The toes introduces a potentially attractive location to measure PPG. The toes is just as well perfused as the finger, have a clear anatomical marking to place the sensor, and is of the furthest distance from the finger tip giving the best resolution for PTT. The toes and finger pairing has been investigated in clinical settings to show good measurement of the PTT for measuring arterial stiffness. However, this pairing is only usable in a clinical setup where the patient has two separate sensors that can instrument the finger and toes in a natural position, typically laying down with the hand and feet outstretched. This is not possible with the smartphone camera setup. The user would have to reach down to the feet or pull the feet up. In this posture, the blood pressure is affected and the pulse transit time is also affected due to the significant bending of arteries leading to the toes, leading to inconsistent measurements time to time.

Through our prototype of the dual camera setup on an actual smartphone, we found that the PPG/PPG based measurement has a multitude of usability limitations. Many of the PPG pairing have been demonstrated in prototypes or clinical measurements of PTT systems to be effective, but when actually implemented on a smartphone, the limitations become clear.

7.2 SCG/PPG based systems

In the prior work, although there have been explorations of using smartphone accelerometer to capture SCG[43, 60], there has not been a system that combines the use of SCG and PPG captured on the phone. So in our initial exploration, our goal was (1) test out the sensing of SCG on a phone, and (2) develop an application that combines the use of SCG and PPG.

7.2.1 Contact Position with the User

To measure the SCG from the heart, the phone needs to be placed on the left side of the chest, close to the heart. Depending on the position the SCG is recorded from, the shape of the SCG signal may differ. Placing the phone lower down the chest positions, it is closer to the Mitrial and Tricuspid valves. Whereas placing the phone at higher positions, it is closer to the Aortic and Pulmonic valves. We noted that although the bottom of the sternum typically produced the cleanest signal,

this position was not always possible. First, the position the phone needs to be placed is often where the bra frame is positioned under the breast. Second is even if the user removes their bra, the phone has to be held in a position that will likely be covered by the breast, creating two issues: the user can't see the screen for visual feedback and the breast resting on the phone dampens the vibration signal. For the female subjects, we found that it was possible to obtain a SCG signal from the upper chest. This position, however, results in a different shape of the SCG.

We also tried doing the measurement using either the base of the phone or the face of the phone in contact with the body. We noted that using the base of the phone worked far better. This is contrary to the setup used by Mohamed et al. [60], which describes using the phone placed directly on the chest with the phone face. We note two differences. HeartSense uses the gyroscope to detect the rolling motion that the chest undergoes. By placing the phone on the chest while the user is laying down and letting the phone rest freely, the phone is free to rotate, which may be why the signal works particularly well. For our system, we aimed for the user to perform the measurement sitting up as that is the standard for blood pressure monitoring. However, when seated, the phone cannot rest on the chest without being held. In this situation, we found the signal to not transduce as reliably, likely because the hand fixes the phone in a way that does not allow the natural rotation to be transduced. We did, however, find that the accelerometer had a more reliable signal in this posture. One avenue we did not fully explore was the use of gyroscope signal in combination of accelerometer, which may work well.

We found another issue with placing the phone on the chest, unintended screen clicks. This was true for both using the based of the phone and for the face of the phone. For the face of the phone, in order for the user to be able to place their finger on the back camera for the PPG measurement, the screen must face the body. Depending on the shirt thickness, there are sometimes false touches. It was very awkward for the user to flip the phone around and stick their finger on to the back camera in that configuration. For the base of the phone, the false screen touches was an issue for female users. Depending on the user's breast size, when the phone was placed on the lower part of the sternum, false screen touches would occur. We were able to solve the false screen touch issue by implementing a touch blocker that intercepts touches and ignore it. And since the on screen UI is no longer available, we implemented a start trigger using the volume button.

7.2.2 Sensor Synchronization

In our initial prototype, we had assumed that by triggering the start of the measurement one line of code apart for the camera and accelerometer, the signals would be more or less aligned. However, this was far from true. In hindsight, this is maybe obvious. At the API level, the app simply can only trigger an asynchronous command and the camera system goes through a series of setup of framerate, exposure settings, etc. The first frame asynchronously returns once all that is complete. This is similar for the accelerometer. There does not exist synchronization between the two data streams, so by capturing the two streams independently, the end result is completely unsynchronized. Furthermore, the lag of starting the sensor capture and actual first data point is inconsistent measurement by measurement. We found that although the average difference was around 165 ms between the camera and accelerometer, the difference of the lag time can be plus or minus 30 ms or so from measurement to measurement.

7.3 PCG Recording with Smartphone Microphone

A note worth mentioning is in our initial prototype, we also explored the use of microphone to capture phonocardiogram (PCG). Similar to the PPG/PPG based exploration, we implemented an Android phone based app that accesses the microphone and the back facing camera simultaneously. We tested the app on four phones with different variants on their microphone placement and design: LG Nexus 5, LG Nexus 5x, Huawei Nexus 6P, and the Google Pixel. The Nexus 5 has the microphone placed on the bottom edge of the phone with sound inlet made up of multiple pinholes. The Nexus 5X has the microphone placed on the bottom edge of the phone with a pinhole inlet to the microphone. The Nexus 6P has the microphone placed on the bottom front of the phone under the slotted opening below the screen. The Pixel also has the microphone placed on the bottom edge of the phone, but has a slotted opening that is about a centimeter wide. The different configurations of these three phones allowed us to examine different advantages and disadvantages of each design.

7.3.1 Contact Position with the User's Chest

For all PCG measurements, the microphone has to be placed on the chest very close to the heart. The heart is situated towards the middle of the chest a little towards the left (Figure 7.1 g). To best

measure the ejection timing of the heart, the microphone should be placed close to the bottom half and slightly left of the chest to capture the S1 sound. We found through testing with groups (1-3) that depending on the position of the microphone, this introduces a problem. Specifically, the Nexus 6P's front facing microphone introduces difficulty in accessing the chest in this location. In order for the front facing microphone to be placed flush on the chest, the entire screen has to be placed to the chest. This introduced two issues. The first is that the user can't see the screen. As compared to the PPG/PPG based system where exact positioning is very crucial, PCG measurement is more lenient, but this still results in difficulty in gaining proper feedback. Furthermore, by placing the entire screen surface against the chest, the configuration introduces a large area of contact with the skin. This large contact introduces a source of sound created by the friction against the skin. In group 1 and 2 where most of the participants were young and had a strong heart beat, this effect was not as significant. However, in group 3 where the participants were much older, the sound produced by the heart was much weaker, and sometimes the sound from the friction would significantly over power the heart sound. This was further compounded because the older participants tended to have a bit of shake in their hands, leading to a much more significant movement of the phone. For the Nexus 5, Nexus 5X, and the Pixel, the bottom edge placement has a significantly smaller footprint for contact on the skin and leaves the screen easily visible to provide feedback to the user (Figure B). This is similar to the phone placement as [17, 35].

7.3.2 Enclosure Shape Affects Sound Quality

The medical stethoscope amplifies the heart sound effectively by using a large surface area in contact with the skin, the origin of the sound, and channeling the sound through a sound chamber into a smaller opening into the microphone. This sound chamber increases the amplitude of the sound by compressing the sound wave from the source into a higher pressure sound wave, increasing the amplitude of the measurement. The same effect happens on the smartphone microphone depending on the enclosure of the microphone. For the Pixel and the Nexus 6P, the microphone is placed under a large slot, mimicking the sound chamber of a stethoscope when the slotted opening is pressed firmly against the chest making a tight seal. In contrast, the Nexus 5 and Nexus 5x does not create the same sound chamber effect due to the use of pinholes. As a result, the PCG recording is much

more amplified on the Pixel and Nexus 6P. This is important for the participants in group 3, whose heart sound tended to be quieter and hard to capture unless amplified significantly by the sound chamber.

Chapter 8

SEISMO IMPLEMENTATION AND EVALUATION

Seismo implements the SCG and PPG based PTT measurement method with the built-in sensors of a smartphone. To take a measurement, a user needs to hold the phone in one hand, and press it onto the left side of the chest in one of the two locations: (1) below the pectoral muscles on the sternum, (2) above the pectoral muscles and below the clavicles.

These two positions provide proper ergonomics and access to the heart vibration, and also are less obstructed by body fat compared to other portions of the chest. The user can try placing the phone in both the positions, and our system Seismo will determine the subject-specific location depending on the quality of the signal received. The SCG morphology changes depending on the position of the recording device (Figure 8.2). Thus once the chest position is determined, the user needs to place the phone at the same position thereafter. To measure PPG, the user places their index finger of the other hand over the phone's back camera, covering both the camera and flash. In this position (Figure 6.1), the SCG is measured using the accelerometer Y-axis data (along the height of the phone) and the PPG is measured using the camera data.

8.1 Implementation

The prototype of Seismo is implemented on the Google Pixel phone, which has the BST-BMI160 inertial measurement unit with a stock setting of 250uG/LSB and 400Hz sampling rate. Although it is technically possible to improve the resolution to 61uG/LSB and sampling rate to 2000Hz, we found that the stock setting is good enough for our use case. The phone's back camera operates at a standard frame rate of 30 fps.

The following sections describe the signal processing of the captured SCG and PPG signals in this setup, and the calculation of PTT from the processed signals.

8.1.1 Synchronization

The Android SDK does not guarantee perfect synchronization between different sensor subsystems, hence we need to synchronize the data collected from the camera and the accelerometer subsystems. For the synchronization, the user places the phone on a flat surface for each use, and the phone's speaker plays a series of three beeps (Figure 8.1). The vibrations caused by the beeps is captured by both the microphone (which is synchronized with the camera subsystem) and the z-axis of the accelerometer (axis perpendicular to the screen).Using a matching filter of the expected vibration pattern, the timing of the beeps was extracted from both the microphone signal and the accelerometer signal. The timing delay is then used to shift the accelerometer signal to synchronize with that of the microphone signal based on the beep timing. We found this timing to be about 170$\pm$15ms.

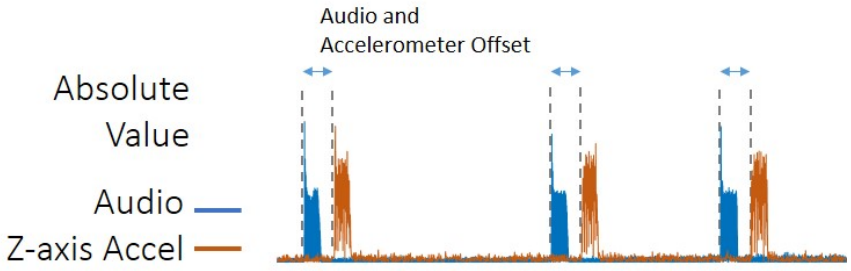

Figure 8.1: A series of three beeps is played through the phone's speaker to synchronize the camera and accelerometer subsystems.

8.1.2 PTT Calculation

To calculate PTT, we need to measure the time difference between the aortic valve opening (AO) and the pulse reaching the finger. The SCG signal from the phone's accelerometer provides the AO, while the PPG from the camera provides the pulse timing at the finger. PPG has a much higher SNR, compared to SCG. Thus, we first extract the timing of the pulse arriving at the fingertip using the camera data, and use that to search for the position of the AO relative to it. From the camera data, for each heartbeat, we need to find the foot of the PPG waveform, which denotes the end of diastole. First we calculate the heart rate by taking an FFT of the PPG, and retrieving the dominant

frequency. The heart rate is then used to predetermine the minimum spacing between successive pulses. To calculate the actual timing of the pulse wavefront, we use the second derivative of the PPG signal, called acceleration photoplethysmogram (APG), instead of the original PPG waveform. The reason for this is the peak of the APG accentuates when the pulse shows up at the finger. By using the heart rate as the minimum spacing, we then apply a standard peak detection algorithm to capture the peak of the APG.

The Y-axis of the accelerometer records SCG at 400Hz. The accelerometer captures the vibration induced by the mechanical motion of the heart when it beats. The SCG is the low-frequency (0.8-20Hz) component of this vibration. Figure 8.3 shows the SCG waveform captured by the accelerometer after bandpass filtering. The SCG waveform has a series of characteristic points resulting from the different phases of the heartbeat. A typical SCG captures the mitral valve closure (MC), isovolumetric moment (IM), the aortic valve opening (AO), and maximal acceleration (MA) (Figure 8.2). As discussed before, the major advantage of SCG is that it can capture AO, which cannot be captured by the sound-based phonocardiogram (PCG) as it measures the high-frequency (30-150Hz) component caused by the closure of the mitral valve. The PCG is sometime coupled to the accelerometer as the accelerometer acts as a contact microphone. To ensure that the timing of the AO is not affected by the PCG, a bandpass filter from 0.8-20Hz is applied to the raw accelerometer signal to remove breathing artifacts below 0.8 Hz and the sound of the PCG signal .

We need to find the timing of the AO point for every beat. Some characteristics of SCGs include: (1) the AO point happens before the PPG signal begins to rise, and (2) the AO point happens after the IM, which typically has the strongest negative amplitude signal before the arrival of the distal pulse. Hence we can use the PPG waveform extracted in the previous step to aid in finding AO point, which is expected to be between two successive APG peaks. We extract all the positive and negative peaks from the SCG prior to each APG peak, the duration being determined by the heart rate. We designate the negative peak with the highest prominence as the IM point. The prominence of a peak measures how much it stands out due to its intrinsic height and its location relative to other peaks. However, for certain SCG morphologies, in particular those recorded closer to the sternum, the MA point can have a similar prominence as the IM point (Figure 8.2). To take care of such cases, the algorithm then compares the negative peak ahead of the most prominent negative peak to determine if the most prominent peak is IM or MA. If the peak extracted is from MA, then the peak

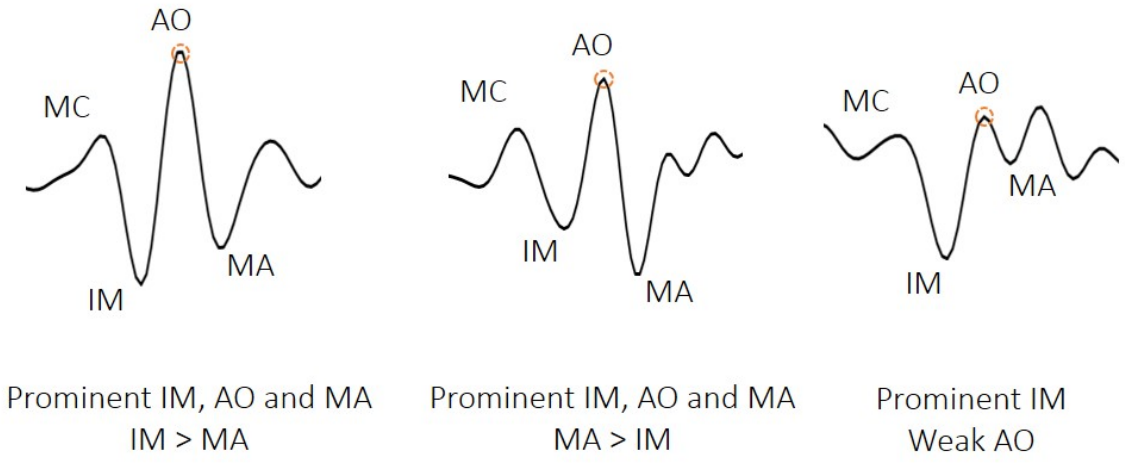

Figure 8.2: The SCG is produced by the consecutive mechanical movements during a heartbeat (MC, IM, AO, then MA). However, the relative amplitude of each movement changes depending on the person and position of the accelerometer placement.

ahead of it would also have a high prominence. In this case, we compare the prominence of the two peaks and if the ratio of their prominence is above 90%, a threshold that we empirically determined to work well, the algorithm chooses the peak ahead of the most prominent peak as the actual IM. Finally, since the AO point necessarily occurs after the IM, the positive peak following immediately after IM is chosen to be the AO. Finally, the time difference is calculated for each AO and APG peak pair, with the median taken as the PTT.

8.1.3 Blood Pressure Calculation

To calculate blood pressure from the PTT, an individualized calibration for each user needs to be created. In our evaluation, we use the data from the first day of the intervention study for calibration between PTT and BP, and test how well it performs for the remaining three days. A least-square fit is used to fit the PTT measured with Seismo and the ground-truth blood pressure measured by a cuff-based sphygmomanometer, using the Moens-Kortweg equation (Figure 8.3) to estimate the subject-specific coefficients K_1 and K_2.

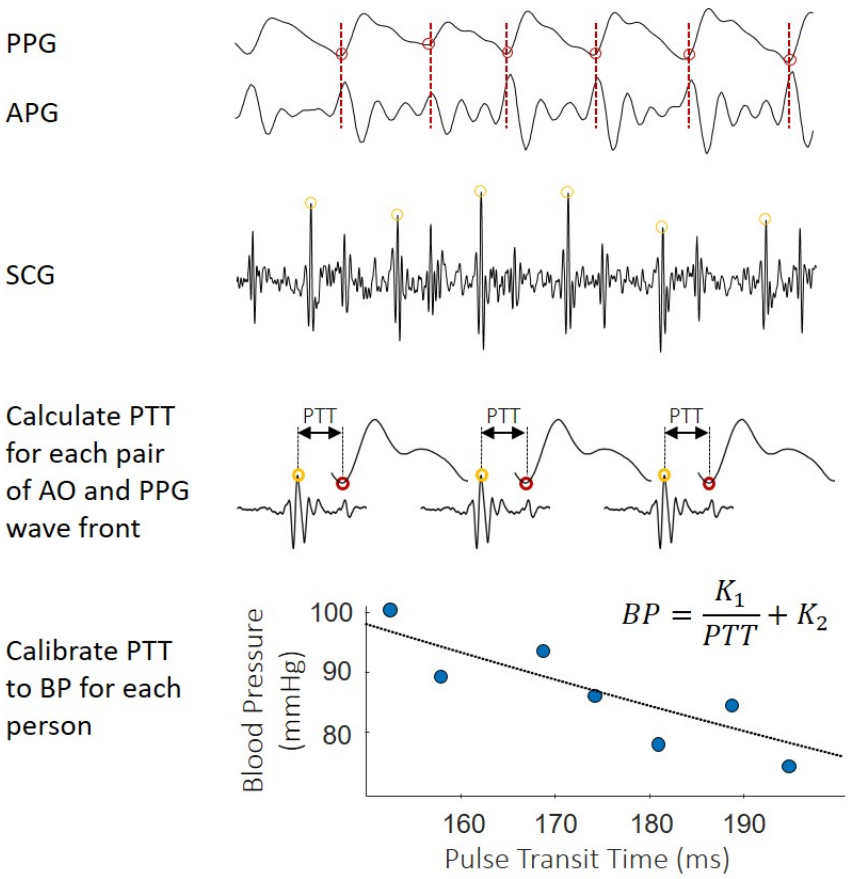

Figure 8.3: PPG and SCG signals are used to measure the PTT. The maximal point of acceleration of the PPG (APG) is compared with the AO point of the SCG. To convert PTT to BP, an individualized calibration of PTT to BP is generated based on reference recording with a blood pressure cuff.

8.2 Evaluation: Blood Pressure Estimation

To validate the performance of Seismo, we recruited nine participants (five female) to perform a series of stationary bike tasks. All participants were recruited from within the author's institution, with an age range from 26 to 66 years (μ=44, σ=17). Participants were paid 50 USD/session and all protocols are approved by our institution's IRB.

Participants are asked to ride a stationary bike five times, for 2-minutes each. The bike riding intensity varies from 50%, 75%, 100%, 75%, and 50% of the participants' maximum speed. The

Table 8.1: Participant statistics and results for the 4-day bike study

ID	Age	Sex	Weight [lb]	Height [in]	Day 1 DP [mmHg]	All Days DP [mmHg]	RMSE [mmHg]	R
1	27	F	130	67	81 - 96	75 - 101	5.2	0.63
2	61	F	134	65	92 - 114	91 - 114	5.8	0.52
3	26	F	101	64	68 - 79	68 - 84	3.3	0.68
4	27	M	145	70	84 - 89	75 - 91	3.3	0.20
5	56	M	190	72	90 - 97	80 - 107	4.9	0.59
6	55	F	114	64	62 - 85	62 - 88	4.4	0.77
7	66	M	162	72	79 - 92	79 - 112	9.2	0.45
μ	44		139	68			5.2	0.55
σ	17		30	4			2.0	0.19

participant's blood pressure is first taken at rest. Then the participant performs the first 2-minute bike ride at 50% of the maximum speed. Just after the bike ride, their blood pressure is taken again. The participant repeats this until they finish all five rounds of biking. At the end, the participant is asked to rest for two minutes, and their blood pressure is measured again after the rest. In total, the blood pressure and PTT are measured 7 times (once before biking, 5 times after biking at each intensity, and once after two minutes of rest).

The blood pressure is measured simultaneously using Seismo and with a standard blood pressure cuff device (Microlife bp3na1-1x) as ground truth. Both the measurements are started at the same time. The Seismo measurement is taken with the right hand index finger over the camera, while holding the phone at the sternum, or upper chest just below the clavicle (Figure 6.1). Each Seismo measurement takes 30 seconds. The blood pressure cuff is placed over the left upper arm. The blood pressure cuff measurement typically takes between 30-60 seconds depending on the participant's blood pressure, and provides a measure of systolic and diastolic blood pressure.

The study was conducted in a quiet office, so external vibration is not evaluated by our study. However, during development, we observed negligible impact of surrounding noise, such as speaking, passing cars, and gym activities.

8.2.1 Results

All the participants completed all four days of the bike study. Of the nine participant, two participants are removed from analysis as the data collected from them is inconsistent. One old male participant suffers from severe hand tremor, and as SCG is of similar frequency to his hand tremor, most AO peaks are masked. Another female participant exhibited poor SCG signal both at the sternum and at the clavicle. We suspect this is due to low acceleration caused by dampening by fat and muscle tissues.

After removing the two participants, we have a total of n=196 (= 7 participants x 4 sessions x 7 measurements/session) blood pressure and PTT measurement pairs for analysis. The average diastolic blood pressure, and average blood pressure fluctuation per participant observed in the study is μ=88$\pm$11.5 mmHg. Individual breakdown is shown in Table 8.1.

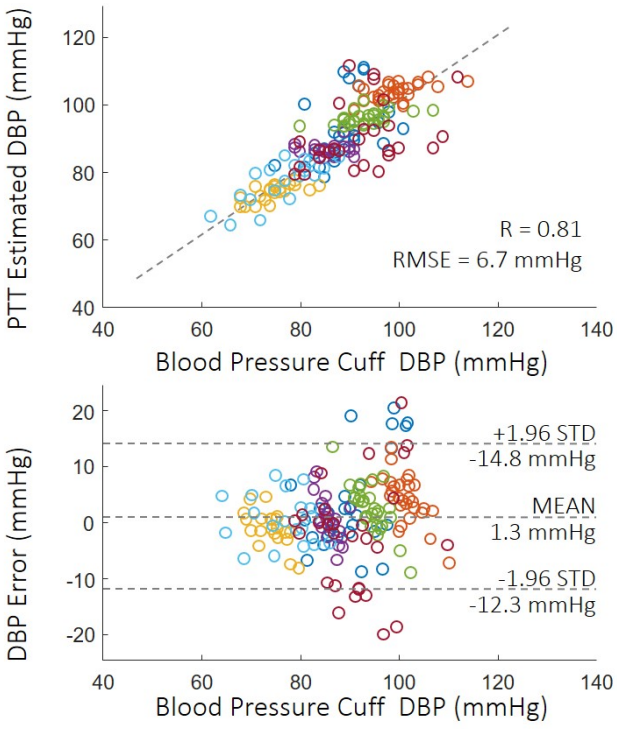

Figure 8.4: Correlation and Bland-Altman plot between DBP estimated using individual calibration and reference blood pressure measurements of the seven participants included in analysis.

For each participant, the first days' PTT measurement by Seismo and the blood pressure reference measured by the cuff are used for calibration. With this calibration, the subject's blood pressures for the rest of the sessions are estimated from the PTTs measured and compared against the reference cuff-based BP measurements. The calibration performed in our work is limited to the mechanism of BP perturbation caused by exercise, and our work does not include other vasomotor tone modulation induced by cold pressor or stress.

We analyze the relationship between the pulse transit time and diastolic blood pressure (DBP). This is because the foot of the PPG signal (or peak of the APG signal) measures the arrival time of diastole and not systole. Correlation of pulse transit time with systolic blood pressure (SBP) comes from an associated increase of SBP with DBP, but is not a direct relationship. For this reason, we focus only on the correlation of DBP with PTT. The average Pearson correlation coefficient across all participants is 0.55 and the average RMSE is 5.2 mmHg across participants, with R between 0.20-0.77 and RMSE between 3.3-9.2 mmHg.

Figure 8.4 shows the calibrated PTT and DBP for all participants. When all participants' estimated blood pressure using the individual calibration is correlated against the reference blood pressure measurement, we find the Pearson correlation coefficient = 0.81 with an RMSE = 6.7mmHg. As seen in the CDF, 67% of absolute errors fall under 5 mmHg, 87% fall under 10 mmHg, and 94% fall under 15 mmHg (Figure 8.5), with the maximum being 21.3 mmHg.

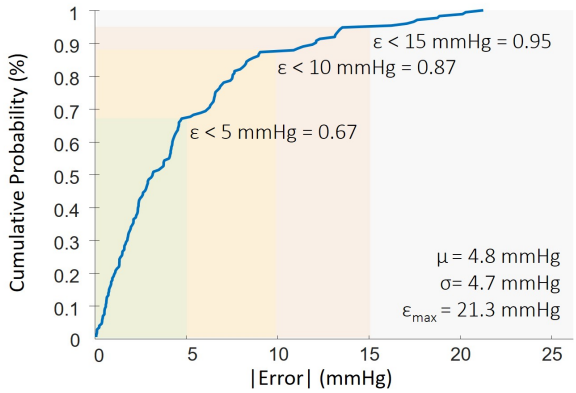

Figure 8.5: CDF of the absolute error between the estimated DBP from Seismo to the reference DBP from the Microlife blood pressure cuff.

8.3 Discussion

8.3.1 Limitations

One of the main limitation of the evaluation is the reference blood pressure. Although the Microlife bp3na1-1x used in the evaluation is rated A/A under the BHS Standard, it is not designed for measuring rapidly changing blood pressure. In comparison, other works [15, 39] that have demonstrated high correlation between PTT and BP during BP perturbation studies instead use a reference system that measures blood pressure continuously, such as a ccNexfin finger cuff (Edwards Lifesciences, Irvine, CA). Such a system can more reliably capture the fast-changing blood pressure after exercise, but it is very expensive and specialized. It is likely that some of the disparity between our results and results from prior work [15, 39] can be attributed to using a different system to measure ground truth. In fact, the participant with the lowest correlation, Subject-4, has sessions where his/her reference blood pressure remain unchanged even after intense biking. When PTT of Subject-4 is measured, it reduces as expected. Figure 8.6 shows Subject-3 which has the expected trend of both – the reference blood pressure and the inverse of pulse transit time – increasing due to exercise, while for Subject-4, the inverse of PTT is following the expected trend, however the reference blood pressure does not. If we remove Subject-4, the group correlation improves to 0.61.

We compare our system with prior work, using either ECG or SCG as a reference. Although widely used as a method for noninvasive blood pressure monitoring, ECG-based pulse arrival time measurements have been shown to poorly correlate with diastolic blood pressure due to the pre-ejection period [88, 15]. In contrast, techniques that can thus capture the actual timing of aortic valve opening can accurately determine the pulse transit time and thus correlate with blood pressure more reliably. Previous studies using ECG and PPG positioned at the finger/wrist found correlations between PAT and DBP to be 0.40 ± 0.35 [44], 0.26-0.57 [56], and 0.41 [8]. In comparison, our average correlation between PTT and DBP is comparatively high at 0.55 (or 0.61 when removing the outlier Subject-4).

When compared to other methods that accurately capture the aortic opening time, our correlation results are comparable to 0.57 ± 0.13 by [87], 0.62 ± 0.16 by [39], but is comparatively lower than 0.84 ± 0.09 by [15]. Although our system demonstrates a slightly lower accuracy compared with state-of-the-art in prior work, a main contribution of our work is in showing that a smartphone's

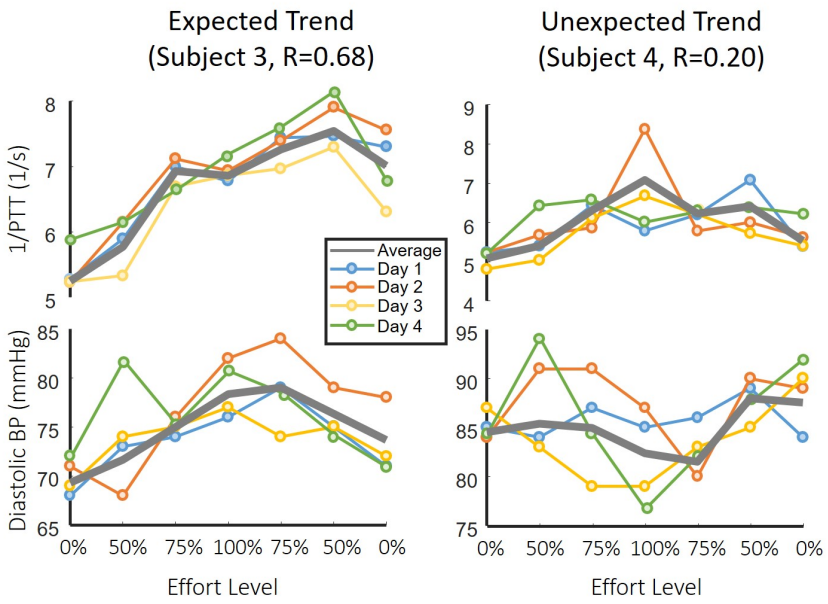

Figure 8.6: Left: Subject-3 with an R=0.68 exhibiting the expected pattern in both reference blood pressure and PTT changes. Right: Subject-4 with an R=0.20 exhibiting only the PTT showing the expected trend across multiple sessions.

built-in SCG and PPG measurement can be used in measuring pulse transit time for tracking blood pressure. Compared with SeismoWatch [15], which uses a ultralow-noise accelerometer (ADXL354 with a noise floor $20\mu g/\sqrt{Hz}$, instead of using a dedicated hardware system, Seismo relies on a built-in smartphone accelerometer (BST-BMI160 with a noise floor of $180\mu g/\sqrt{Hz}$), which is a typical sensor for commodity phone hardware. To enable built-in sensors to perform PTT measurements, we came up with a synchronization technique that uses the built-in speaker that is always available on the phone as well.

Of the nine participants in the study, two participants were taken out of the analysis due to poor signal quality. One participant's hand shook while performing the measurement, and the motion artifact often over-shadowed the SCG. The current prototype has the person holding the phone with the screen facing them. This position is ideal since in actual use, it is more suitable for users to be able to see the screen while operating the app. However, for users whose hand shakes, holding the

phone in a floating position is difficult. Instead, if the phone were held to the chest with the screen against the chest, the hand would no longer be floating, making the position more stable.

Another participant was removed from the analysis due to low signal strength caused by thicker tissue mass between the heart and the phone's accelerometer. This demonstrates a limitation of the sensor. As the noise floor of the sensor is relatively high compared to that of a low noise accelerometer used by [15], with a low measurable heartbeat acceleration possibly caused by damping by fat tissues, our system was not able to resolve such a signal. This was particularly the case when the participant was at rest. During the high intensity biking efforts, however the participant's increase in cardiac output did allow for a stronger SCG signal, and the sensor was able to reliably detect the SCG.

8.3.2 Incorporating Smartphone Blood Pressure Monitoring to Personal Health Tracking Applications

By only using the built-in sensors of the smartphone, Seismo can monitor blood pressure without any modifications to a smartphone with a camera and accelerometer. We envision such a smartphone based PTT measurement can be useful in a variety of scenarios depending on the availability of resources.

Supplement to At-Home Blood Pressure Cuff

A smartphone solution to measure blood pressure can be a good supplement for patients who are already monitoring their blood pressure with the cuff. Although a person can easily keep a cuff around at home and monitor once or twice a day to gain knowledge of their general blood pressure trends, it is hard to carry a cuff around as they go out of their house or when they travel. Having a smartphone based method can certainly help in such a scenario. Moreover, Seismo can be use for blood pressure monitoring for pre- and post-exercise. Here blood pressure change pre- and post-workout can even act as an additional measure of workout quality beyond heart-rate tracking. Using Seismo, the user can also track their blood pressure recovery time after a cardio workout. One of the advantage to a pulse transit time based blood pressure monitoring is the ability to track blood pressure continuously, something that can not be tracked even if the user had brought a cuff with

them to the gym.

Cardiovascular Risk Screening

In order for PTT-based blood pressure monitor to measure blood pressure accurately, the user needs to calibrate it. As such, for users who do not already have access to a cuff, Seismo would not be able to readily translate the measured PTT to blood pressure. However, in this case, Seismo can still offer insight into the person's cardiovascular health. The relationship between PTT and blood pressure also depends on arterial stiffness. Arterial stiffening occurs naturally as people age, however excessive stiffening usually suggests some form of health risks, such as plaque build-up. PTT has been shown to be a potent index for monitoring arterial stiffness [89] by taking into account arm length, gender, and age. In well-equipped hospitals, an ECG and PPG system is often used to monitor PTT to screen for arterial stiffening. However, this is not available in low resource areas. Using Seismo, reliable PTT measurements can be made available through phones that are already available in these areas to enable arterial stiffness screening.

Chapter 9

IPRESSURE: A SMARTPHONE-BASED SYSTEM FOR ASSESSING INTRAOCULAR PRESSURE

9.1 Intraocular Pressure

Intraocular pressure (IOP) is the innate fluid pressure within the eye. IOP is maintained by the trabecular meshwork, which manages the leakage of the aqueous humor in the anterior chamber of the eye. The typical IOP of humans ranges from 7-21 mmHg with a mean of approximately 16 mmHg. Elevated IOP is an important risk factor for glaucoma, a progressive optic neuropathy that can lead to visual field defects or eventual blindness. A study carried out by Quigley and Broman in 2006 [66] predicts that the global population affected by glaucoma will reach 80 million by 2020; it further postulates that half of the people living with glaucoma are unaware that they have the disease, which can largely be attributed to a lack of resources or incentive for IOP assessment. Glaucoma also imposes a significant burden on the US healthcare system, costing roughly $3 billion USD and over 10 million visits to physicians per year [68].

Tonometry is the diagnostic procedure performed by eye care professionals for measuring IOP. The clinical gold standard for measuring IOP is Goldmann Applanation Tonometry (GAT) [75]. Applanation tonometry in general relies on Goldmann's observation that "the pressure in a sphere filled with liquid and surrounded by an infinitely thin membrane is measured by the counterpressure which just flattens the membrane" [74], also known as the Imbert-Fick law. In GAT, a form of fixed-area tonometry, a topical anesthetic with a fluorescein dye is placed on the eye. When the dye mixes with the tears and the eye is fluoresced with a cobalt blue light, the dye appears as a brighter yellowish green. A split optical prism is then pressed against the eye, resulting in two semicircles. The ophthalmologist adjusts the force exerted by the prism until the semicircles align on opposite ends, indicating that the area of the applanation surface has reached a predetermined quantity. That force measurement is mapped to an IOP value using a clinically validated lookup table [74]. The complement to fixed-area tonometry is fixed-force tonometry. Instead of measuring the

Figure 9.1: iPressure emulates fixed-force applanation tonometry using the hardware adapter pictured above. The clear acrylic cylinder is allowed to move freely within the black casing so that its mass provides a constant force on the patient's eye.

force required to make an applanation surface of known area, a cylinder of known mass is allowed to rest on the eye without any external forces and the area of the applanation surface is mapped to an IOP value.

Tonometry requires either a trained eye care professional or access to dedicated medical devices. These constraints make tonometry difficult in low-resource environments. Smartphones, on the other hand, have seen a rapid uptake all over the world and contain a myriad of sensors that can be used for mobile health applications. In this paper, we propose a smartphone-based system that allows for minimally trained individuals to perform IOP assessments on other individuals. Rather than requiring precision from specialized hardware or a trained professional, the precision of this system is placed within the smartphone. The user attaches a low-cost smartphone adapter that we have developed to emulate fixed-force tonometry. While the patient lies supine, the cylinder inside the instrument is rested on the patient's eye, allowing the smartphone's camera to automatically detect and measure the applanation surface, from which the patient's IOP may be inferred.

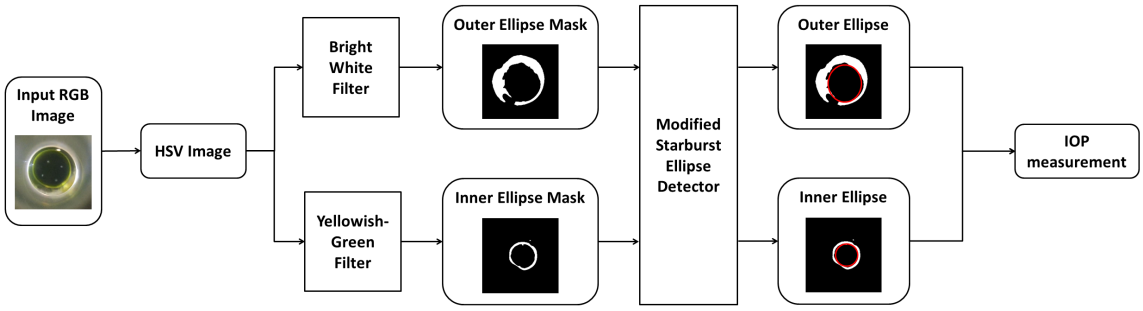

Figure 9.2: The steps taken to estimate intraocular pressure from an RGB image of the applanation procedure. After converting the image into the HSV space, masks are defined for the clear acrylic cylinder's base (outer ellipse) and the applanation surface (inner ellipse) using color and intensity features as filters. Ellipses are detected on the insides of those masks and then mapped to absolute measurements given the 8 mm diameter of the acrylic cylinder. The diameter of the applanation surface is then mapped to the patient's estimated IOP.

9.2 iPressure System and Usage

9.2.1 Hardware

The hardware adapter is shown in Figure 9.1 attached to an iPhone case. The most important part of the hardware is the clear acrylic cylinder inside the black casing. The acrylic cylinder is allowed to move freely within the casing, but has notches to ensure that it does not fall out of the adapter. The acrylic cylinder has a diameter of 8 mm and a height of 63 mm. The 8 mm diameter was chosen such that it would capture a fairly large circle from a low eye pressure without being too difficult to use on patients with small palpebral fissures. The height of 63 mm was chosen for two reasons. (1) If the applanation surface is placed too closely to the smartphone's camera, the resulting video becomes difficult to focus and the edges become blurry. (2) This combination of diameter, height, and material leads to a mass of 5.0 g, a mass for which the conversion from applanation surface diameter to IOP has already been clinically validated for human eyes [74, 84]. Although conversion tables for larger masses have been produced, studies have shown that the weight of the tonometer

induces an increased pressure due to the displacement of aqueous humor during applanation [84].

The black casing itself is designed such that the acrylic cylinder is optimally positioned in front of the smartphone's back-facing camera. Not only does this positioning include the alignment of the acrylic cylinder with the camera, but also the distance between the base of the acrylic cylinder and the camera. The black casing also blocks out ambient lighting to prevent any extraneous reflections from appearing in the acrylic cylinder.

To enhance the visibility of the acrylic cylinder in the camera, the edge of the cylinder's bottom surface is frosted. To emphasize the applanation surface in the camera, an external LED is mounted on the casing; in the future, the smartphone's flash could be redirected to the outside of the acrylic cylinder via a short fiber optic cable. When this lighting is reflected off of fluorescein dye, it shines as a bright yellowish green.

9.2.2 Performing the Assessment

Before receiving the assessment, the patient assumes a supine position. The user conducting the test administers a topical anesthetic with fluorescein dye (Fluorescein sodium 0.25%/Proparacaine 0.5%) to the patient's eye. The user then holds the smartphone over the patient's eye such that only the weight of the acrylic cylinder is applied to it. This means that the smartphone should be as flat as possible (i.e., parallel to the ground) and the user should not apply any extra force on the smartphone (i.e., pressing down). The flatness of the smartphone is measured with the smartphone's accelerometer, operating as a sort of bubble level.

The weight of the acrylic cylinder creates an elliptical applanation surface with a yellowish green outline when the LED is shone on the patient's eye. The smartphone's camera records the applanation of the eye. The frames from the resulting video are then processed using computer vision to give a real-time estimate of the patient's IOP.

9.2.3 Video Analysis

Figure 9.2 outlines the algorithm used to extract an IOP measurement from an RGB image. The overall goal of the algorithm is to detect two ellipses: the base of the clear acrylic cylinder (outer ellipse) and the applanation surface (inner ellipse). Since the diameter of the acrylic cylinder is

known, the applanation surface can be assigned an absolute measurement by using the cylinder as a reference. Both of the ellipses should be relatively circular; however, the acrylic cylinder may appear slightly elliptical if the hardware adapter is improperly mounted, and the applanation surface may be elliptical if the patient has significant astigmatism or corneal surface irregularities.

As shown on the far left of Figure 9.2, the edge of the acrylic cylinder appears bright and white, and the edge of the applanation surface appears as a dimmer yellowish green. By filtering the image according to intensity and color information, binary masks can be produced to select the outlines of the circles. This information is most intuitively recovered from the image after it is converted into the HSV space. The mask for the inner ellipse bounds the hue between 15-45%, the saturation between 35-100%, and the value 15-100%. Together, these thresholds encode the greenish yellow that appears due to the fluorescein. The mask for the outer ellipse is simpler, thresholding the saturation between 0-20% and the value between 25-100%. Both masks are smoothed using morphological filtering operations to create contiguous contours.

Each of the masks will have some non-uniform thickness due to the application of the dye and extraneous reflections in the cylinder. The diameters of interest correspond to the innermost edges of these masks. Standard circle detection methods would either discover many overlapping circles or none at all, depending on the evenness of the masks. Even worse, only part of the applanation surface may be visible if it overlaps with the sclera, which makes it more difficult to see the fluorescein dye. For these reasons, we apply an adaptation of the pupil contour detection algorithm (Figure 9.3) used by Li et al. in their Starburst work [47]. The ellipse detection starts by assuming the center of the ellipses given that the position of the clear acrylic in the camera's view is known. The algorithm then steps radially at 20 evenly spaced angles until an edge is reached in the mask (illustrated with fewer angles for clarity). This assumes that there are no contours that appear within the mask, which can happen for the applanation surface if the fluorescein pools in the patient's eye. Since extra blobs appear in the middle of the mask due to the distribution of the fluorescein, the radial steps start from a fixed distance just below the minimum expected radius to prevent them from stopping short.

Most of the detected edge points should belong to the desired ellipse, but some may still belong to artifacts along the edge of the contour. The original Starburst algorithm accounts for noisy ellipse points by fitting random subsets of points to ellipses and selecting the ellipse that minimizes the number of outliers. In the case of applanation, there is almost always a clean arc that appears in the

image. Instead of randomized subsets of points, as used by Li et al., the proposed system fits contiguous subsets of points (three-quarters of the entire circumference) to ellipses. Although we noted earlier that the base of the acrylic cylinder and the applanation surface may appear elliptical, the ellipses should be relatively rounded. If the percent difference between an ellipse's major and minor axes is greater than 10%, it is automatically rejected. Amongst the rest of the ellipses produced by the different subsets of edge points, the ellipse that best fits the data according to Euclidean distance is selected.

The ellipses recovered from the two masks are then translated into circles with a radius equal to the average of the ellipses' axes. Given that the diameter of the clear acrylic cylinder is 8 mm, the absolute measurement of the applanation surface can be recovered by using the cylinder as a reference. Every time the user performs an applanation, a time series of diameter measurements is produced. The measurements of interest occur when the data is most stable since that is when the weight of the acrylic cylinder should be resting on the eye; therefore, the system combines diameter measurements by taking a mean over the measurements within a standard deviation of 0.25 mm over the course of 0.5 s. The final diameter measurement is mapped to an IOP value using a clinically validated lookup table, such as the one published by Adolph Posner [74].

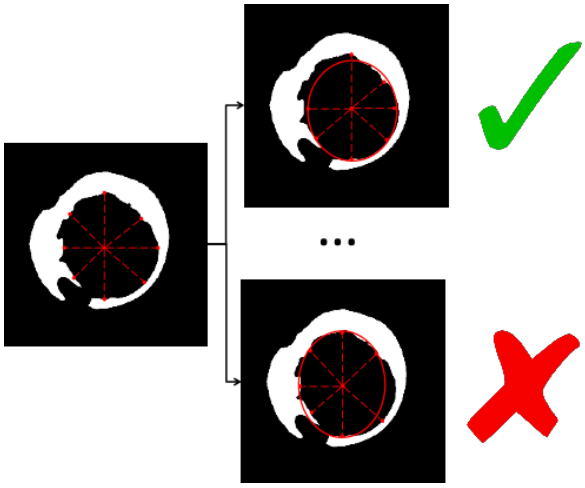

Figure 9.3: A variation of the Starburst technique by Li et al. [47]is used to estimate the innermost ellipse from a binary mask. After candidate points are selected from the inside, contiguous subsets of points are tested with least-squares ellipse fitting until the most circular is found.

9.3 Results

9.3.1 Data Collection

Given the invasive nature of contact applanation tonometry, a feasibility evaluation has been performed on two freshly enucleated ex vivo porcine eyes before deploying the system to living human patients. Although there is not a clinically validated table that maps applanation surface diameter to IOP for animal eyes, the Imbert-Fick law still applies. The porcine eyes were inserted into a clay mold such that the iris was horizontal to the ground, as if a patient were supine. The anterior chambers of the eyes were cannulated and the IOP was artificially varied by the height of a saline-filled reservoir. A topical anesthetic with a fluorescein dye (Fluorescein sodium 0.25%/Proparacaine 0.5%) was applied to the eyes to assist in imaging the applanation surface. Tonometry measurements were obtained three times at every 5 mmHg between 15 and 40 mmHg, leading to a total of 32 measurements. Below 15 mmHg, the applanation surface's edge begins to overlap with the edge of the acrylic cylinder, making it difficult to separate the two. The upper bound exceeds the limits of diagnostic significance for elevated IOP.

9.3.2 Comparison to the Imbert-Fick Law

Although other methods of tonometry are available as a point of comparison for validation, they all have two drawbacks for our validation: (1) they require manual inspection (e.g., Schiøtz or Maklakov tonometers) and/or (2) they are calibrated specifically for in vivo human eyes (e.g., GATs). Instead of a comparison to a possibly inaccurate ground truth, we validate our findings against what is expected from the Imbert-Fick law:

$$P = \frac{F}{A} = \frac{mg}{\frac{1}{4}\pi d^2} \tag{9.1}$$

where P is the intraocular pressure, m is the mass applied to the eye, g is the acceleration due to gravity, and d is the diameter of the applanation surface. Although corrections have been proposed by ophthalmologists to account for properties like the coefficient of ocular rigidity and corneal curvature (e.g., [84]), these models break down for ex vivo eyes. The measurements from each of the eyes were independently fit to the Imbert-Fick law using non-linear fitting. The mass parameter was

used as the unknown parameter; if the data were to follow the Imbert-Fick law without deviation, the parameter resulting in the best fit would be 5g, the mass of the acrylic cylinder.

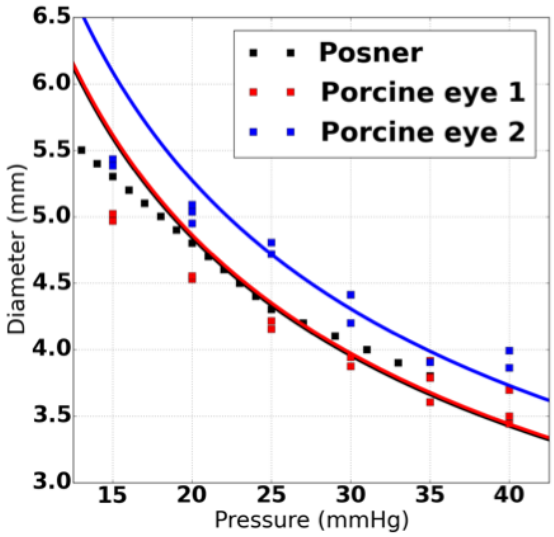

Figure 9.4: The data recorded from the smartphone system and fit to the physical model expected from the Imbert-Fick law. The two curves lead to coefficients of determination of 0.89 and 0.88.

Figure 9.4 shows the models that were fit to the two datasets. The clinical measurements validated by Adolph Posner [74] are also included as a point of reference. Even though Posner only dealt with human eyes, the shape of the data should be similar to that of the ex vivo porcine eyes. When compared to the Imbert-Fick law, Posner's data shares a coefficient of determination (R^2) 0.95 and the estimated mass according to the optimal fit is 5.01g, showing that it follows the model very closely. The fit for the first porcine eye leads to a coefficient of determination of 0.89 and an estimated mass of 5.04g. The fit for the second eye does not obey the Imbert-Fick law as well; it results in a lower coefficient of determination of 0.88 and an estimated mass of 5.94g. The regressions overestimate low pressures and underestimates high pressures in all cases, a fact that has been observed by clinicians for other forms of tonometry as well [8]. The most promising observation is that there is a statistically significant separation between the diameters for 20 mmHg and 30 mmHg, the boundary that clinicians consider for the diagnosis of elevated IOP.

9.4 Discussion

The Imbert-Fick law assumes that the eye's membrane is perfectly elastic, flexible, and infinitely thin; prior work has shown that these assumptions do not hold for the cornea of the eye. For example, DF Woodhouse [84] notes that the volumetric displacement of aqueous humor induces an increase in the IOP, so he derives a correction term that takes into account the corneal curvature and rigidity. We plan to apply Woodhouse's model once evaluations are performed on cadaveric and in vivo human eyes. It is very important that the entire weight of the acrylic cylinder is the only force applied to the eye during applanation. The smartphone's accelerometer is used to ensure that the cylinder's weight is applied perpendicular to the eye, assuming that the user is supine. However, it is possible that the user can add an additional force by pushing down on the cylinder. When this happens, the cylinder slides into the black casing past a specific depth. This could be solved using a contact sensor, but would require an active sensing mechanism Instead, the next iteration of the hardware design has holes along the edge of the casing at the threshold depth. When the acrylic cylinder is pushed past the holes, additional light is shone into it. This can be detected through the smartphone's camera without additional sensing hardware, making it a more desirable design.

Another planned refinement in the next hardware iteration is the use of a macro lens, which is used in photography when a camera must focus on a small subject at close range. The primary reason that the acrylic cylinder used for applanation is 63mm tall is to ensure that the applanation surface remains in focus during the assessment. With a macro lens, the applanation surface can be much closer to the smartphone camera. Assuming a constant diameter, reducing the height of the acrylic cylinder would decrease its mass, but the cylinder could be made with a denser acrylic or metal rings could be attached to the base to add weight.

Chapter 10

SMARTPHONE AS THE PLATFORM FOR UBIQUITOUS HEALTH SENSING

In this dissertation, I have presented a collection of novel approaches in the use of smartphone sensors to perform medically relevant measurements. Much of the focus of the work has been an exploration towards making the system work. Along the way, however, I have met a variety of limitations in the developmental process due to limitations of hardware/firmware/software within and across devices, created workarounds that address limitations, and also identified what limitations may need to be solved at a more fundamental level. Throughout the chapters for each project, I mention these points in context of the project. In this chapter, my goal is to draw out in one place the various findings under a more centralized view point. This chapter is put together through nuanced reflections based on the many lessons learned from my own work and having been in close contact with my colleagues Alex Mariakakis and Mayank Goel throughout their development of various mobile health technologies BiliScreen [53] and SpiroSmart [45, 24]. Partially adapted from our article in IEEE Pervasive Computing, Challenges in Realizing Smartphone-based Health Sensing [55], the purpose of this chapter is to serve both as discussion driven by talking frankly about the limitations of the technology as faced during our own development process and present a critical yet forward looking direction for this line of research.

10.1 Limitations of Smartphone as the Platform

A central theme to this research space has been the use of smartphone as the platform to develop the various health sensing solutions. We use sensors already embedded in the smartphone sensor suite and aim to enable various health sensing solutions. In this solution space, the explicit design boundary we impose is to use the capabilities inherent to a "smartphone." And therein lies a flaw in this design specification. What is "inherent to a smartphone?"

As a research topic, the fundamental purpose in developing these novel sensing solutions using smartphones is to demonstrate the potential enabling solutions that can leverage this form factor

and the typical sensors for clinically impactful applications. In developing the different applications presented in this dissertation, I encountered many challenges that were created due to this wide spectrum, some of which were honestly quite surprising.

10.1.1 Unimplemented Hardware Control

Many of the challenges for the development process was hooked directly to making such apps work on the phone. The mode in which we intended to use the phone sensors is atypical and not use cases directly supported by phone manufacturers or OS developers. As a result, even though conceptually, what we want to do with the sensors is possible based on sensor specifications, what is actually possible depends largely on decisions made by what manufactures of phones and/or operating systems decided as use cases that app developers should be developing. Although we like to present the phone as a multi-functional sensor package, it is quite different than having a sensor package. The hardware level specification is quite different than the actual sensor capability at the level of software SDK access. The following are some examples:

Camera Setting Control For HemaApp, it is crucial to be able to control the red, green, and blue color channels independently. This is because the sensors are exposed at the same time, but the absorption property of blood is drastically different across the color spectrum. Blood absorbs blue and green significantly more than it does red. As a result, if all channels are recording at the same exposure, which is typical, either the red channel will be overexposed and saturated or the blue and green channel will be underexposed and the signal to noise ratio is too poor. The only OS system that provided this level of control is the Full Hardware Control mode of Android Camera2API. At Legacy or Limited control modes, the Android Camera2API or standard Camera API do not have such control. iOS 8 and onward also provide individual gain control, however, the maximum difference can only have a gain of 3x, compared to Android which provides more than 24x gain. Using the 3x gain in an early prototype, I found the resulting signal to noise ratio to still be very high. For Android, in legacy mode, which is the mode for the cheaper phones, the white balance, focus, and exposure settings all cannot be controlled. As such, the app will have to use automatic settings. For limited mode, the phone typically has focus and exposure control, but does not have control over the white balance settings and thus no individual gain control.

LED Brightness Control LED brightness can easily be controlled using a scheme called Pulse Width Modulation (PWM). As it turns out, it is uncommon for phones to have LED brightness control for the flash LED. It is either full on or off. Brightness control is extremely useful, though not absolutely necessary for use cases that typically use this LED. For flash light apps, the LED on full brightness is perfect. For both HemaApp and Seismo, however, because of the amount of time it takes to perform a reading, the LED can often get too hot at full on. This depends on the phone and how many measurements the phone has made already. We found that in Peru where we were performing measurements almost one after another, the phone heated up significantly. To alleviate this, we had to wipe the phone down with alcohol wipes after each reading and take the phone out of its case. We found that during the exploration of the dual camera version of Seismo that because it takes a while to find the location of the superficial temporal artery, the phone got way too hot. In other work where the eye is directly exposed to the bright LED light, it would also be advantageous to be able to dim the light down to a more acceptable brightness.

Recording Multiple Sensor Streams Many of the high-end phones have multiples of each sensor on the phone. In many cases during my development of Seismo I attempted to access either two camera streams or two microphone streams at the same time. In terms of two audio streams at the same time, this looked to be do-able out of the box. As it turns out, the stereo recording made available at the SDK level is not always true stereo, but rather a duplicate of one audio stream on both channels. If one were to want true access to a different microphone for each audio stream captured, it was necessary to root access the phone and alter a hidden hardware configuration file to route each channel to the desired audio channel. Something I found out through hacker forums. For accessing two cameras at the same time, for the case where I wanted to capture two PPG signals at the same time, this was more of a combination of limitations. First for the phone hardware to be able to make this happen, the GPU on the phone has to have multiple Image Signal Processors (ISP). In many of the phones during my early development of Seismo, such as the Nexus 6p, LG G5, Nexus 5x, there were at least 2 ISPs so this was possible. However, interestingly, when the Pixel 1 came out, I noticed that my app no longer worked even though the GPU clearly had enough ISPs on board. My speculation is that this is caused by a few selfie apps that came on the market that utilized the simultaneous 2 camera capture function. These apps allowed people to take a photo of the scenery and a cropped image of themselves with a selfie, a photo in photo. Since this was an

unintended use case, these apps were not superbly stable; admittedly similar to my own dual camera Seismo implementation.

Sensor Synchronization For all instantiations of Seismo, cross sensor synchronization is crucial. For the dual camera version, we needed the camera frames to be synchronized, which is not guaranteed. For the camera and microphone version, the camera and microphone being on the same sensor stream was synchronized based on the graphical processing unit. However, through empirical testing, we noticed that there is usually a lag between the video and sound up to two video frames. It appeared that this lag is consistent recording to recording but not necessarily the same across phone models. For the camera and accelerometer version of Seismo, there is no inherent synchronization between the two sensing systems. Synchronization is certainly possible to implement through firmware integration, however, at the software level, it is not. Without firmware implementation, the best solution I came up with is to use physical time points that transduces measurable signals in both sensor domains (visual for camera, audio for microphone, vibration for accelerometer). For the dual camera, we tried using a tapping motion of the phone on a hard object. The physical jolt gave a distinct and sharp signal in both cameras. For accelerometer and camera, we took advantage of the inherent synchrony of the microphone and camera and performed the synchronization through a buzz tone created by the speaker. The speaker buzz created both a measurable pitch and vibration that gave a clean synchronization point.

10.1.2 Hardware Differences Across Devices

Most of the phones on the market have a lot in common. A camera in the back and maybe a camera in the front. A microphone towards the bottom of the phone and a microphone on the back. A white LED next to the back camera. An accelerometer and gyroscope in the phone. A speaker. There are, however, variations that separate the cheap phones and the more full featured smartphones. In the high-end phones today, we are seeing capacitive sensors, a skintone LED, IR Time-of-Flight sensor, a second camera next to the back camera, and structured light sensor for facial recognition. In much of my exploration, I aimed my solution at using only sensors that are available in all phones: camera, accelerometer, a single microphone. However, even though the sensors exist, variations still exist tremendously between phone models.

One of the differences come from the actual hardware quality. Cameras and microphones are available, however, the specifications are different. For cameras, this is typically resolution, dynamic range, and frame rate. For microphones, this is typically the frequency response and frequency bandwidth. For inertial measurement, this is typically the sensitivity range, resolution, and sampling frequency.

Physical design of the phone also affects the actual sensor system as well. For the camera, the shape of the camera lens and the LED and camera relative positioning are commonly different. For microphones, the position of the microphone varies from being at the bottom edge or at the bottom face of the phone. Furthermore, the microphone is typically housed under some opening that can either be a pin hole, a grating, or wide opening. The opening shape affects the dynamics of how sound is transduced. For example, with the wide opening, when the microphone is pressed against the chest, the sound of the heart beat is far better transduced. Something that was some what unexpected, but had significant effect nonetheless was the color of the phone. For HemaApp, because the measurement mode is one that is reflection based, the color of the phone actually affects significantly the measured response. When we used the Google Pixel phone, the exposure was significantly different for a white Pixel and a black Pixel, with the white Pixel reflecting a lot more light into the camera, leading to much higher exposure.

One of the interesting phenomenons is the prevalence of one-off implementations of sensing capabilities in models created and sold commercially. One of the common claims in the research space is that if the implementation is made possible on a phone that can be bought, then it is possible on "commodity devices". The idea is that even if it is a newer model, it is likely that in the future, more phones will incorporate such capabilities and thus become widely implementable. However, we have actually found this to often not be the case. This is particularly the case for the camera system during the development of this dissertation. During the development of the dual camera implementation of Seismo, we found a few phones that had incorporated a front facing camera paired with a front facing flash. I presume this was incorporated for a brighter selfie-image. This would have been particularly useful for our use case since when the the front facing camera is pressed against the facial artery, without a flash LED, the signal is typically too weak to capture. However, the front facing LED was only incorporated in a few off-shoot models of HTC and Samsung phones. Unfortunately for us, this solution never caught on.

In a similar vein, pressure sensitive screens seems to also be facing a similar fate. Introduced in 2017 in the iPhone and poised as one of the next mainstay sensors in the smartphone sensor suite. Two years later, it seems to not have found a home in more than just the iPhone line, and even within the iPhone line the continued integration of the sensor seems to be dwindling. Many of these decisions are strongly driven by economic factors. And with most sensors that are almost universal to all phones being either core to the function of the phone (i.e. touch screen, microphone and speaker), a big value add (i.e. camera, gps), or multi-functional and application enabling (i.e. accelerometer and gyroscope). Even the magnetometer, which is often bundled with the accelerometer and gyroscope as a 9dof IMU is often omitted in the low-end models. The increased accuracy is not truly necessary for the functions performed and is only useful for compassing in navigation. A function that is arguably none crucial. For the pressure sensitive screens, I speculate that application developers simply did not find compelling use cases beyond the hard press to reveal menus, which is very similar to the interaction of press and hold.

Companies have also considered directly implementing sensors meant for health monitoring. This can be seen in the line of Samsung Galaxy flagship phones which includes a pulse oximeter sensor. The inclusion of this sensor has persisted through multiple generations from 2014 to present (2019) and might continue for at least a few more generations. However, this sensor has not been included in any of the non-flagship phones or none Samsung phones even after 5 generations, a factor likely due to patents and limited use cases.

10.2 The Inaccessibility of Smartphone Health Sensing

A more nuanced examination of the limitations of smartphones as a platform begs the question of whether enabling medical sensing through the use of smartphone sensors is indeed increasing access to health services in a truly widespread manner as motivated by the original intent.

10.2.1 Socio-Economic Disparity

A different way of evaluating the differences in the capabilities of smartphones, not in the view of hardware/firmware/software, is looking at the socio-economic boundaries in which the different kinds of phones are adopted. As mentioned previously, smartphones are different model to model.

Much of the major difference is often driven by the cost of the phone. Smartphones have clearly become another class of everyday item and not just a singular product. Being a class of item means that there is a fundamental capability and specification of the item, and that there are pricing tiers that define the features of each level of the item. As an analogy, cars are also a class of item. A car fundamentally can seat a driver, be steered, and move; this is what even the cheapest car should be expected to do. The higher the tier, the more features. The lowest tier likely does not have heated seats, for example. There is a distinct difference between low-tier phones (<$150USD) versus mid-tier phones ($150USD-$300USD) versus high-tier phones ($300-$700USD) versus flagship phones (>$700USD).

In trying to transfer HemaApp to low-tier and mid-tier phones, the app no longer worked due to the lack of gain control in the color channels, which means that the fundamental sensing solution has to be completely different. Similarly, for Seismo, the accelerometer needs to be sampled at above 100Hz to reliably resolve the seismocardiogram. In this case, at lower sampling rates of 50Hz, typical in low-end phones, the Nyquist frequency is too low for a signal that has components above 40Hz. A major focus of smartphone health sensing is on the possibility of widespread health monitoring because everyone has a smartphone. However, what is in everyone's pocket is different, and that difference comes from how much the user can afford. From a use case stand point, it may be possible to project that a single high-end phone can be procured for performing screening in a village, but it is unreasonable to assume every phone owned by individuals living in the village can be used for personal monitoring. Smartphone-based health sensing solutions often focus on enabling healthcare for low-to-medium income resource situations and communities. To truly reflect that motivation, there should be more of a clear identification of whether or not the solution's requirements are truly representative of the projected use case.

If the specification to enable a solution requires a "commodity" device that has the newest sensor or one-off implementation, it may not be a good solution. This is particularly true of implementing the solution on a flagship device with a unique sensor. Even if the sensor is on a phone currently sold on the market, it does not mean the solution will transfer to other systems as the sensor on the flagship phone is very likely never to be transferred to future generations. Similarly for high-tier devices, the sensors found in these phones may or may not stand the test of time in terms of transfer to future cheaper models. The accelerometer sampling rate is only increased in the higher-

tier phones because it provides a more seamless game play experience. The microphone sampling rate of 48kHz is not present in many mid-tier phones and not present at all in low-tier phones even though it has been a standard for higher-tier model for years. This limits much of the ultra-sound applications utilizing 17-20kHz, for example doppler sensing for breathing detection [63].

Although it is difficult to predict what smartphones in the different tires will feature in the future, I think it is possible to at least speculate how feature adoption will occur in the near term of 5 to 10 years given the last decade of smartphone progression. The flagship phones are a test bed of new features like interaction sensors, screen technologies, camera technologies, sensor sensitivity and performance, etc. The flagship is expected to change every year incorporating the newest sensors to introduce some new experience. Many of these features are completely experimental and likely integrated to test market interest and user review. As new features are tested out with flagship phones each year, some features become defining features that make for what should be in a "full featured" smartphone. The most recent example is likely the dual camera depth mapping to provide computational depth of field portrait photos. Prior to it was the IR time-of-flight focusing, capacitive finger print sensor, multi-tone flash , multiple microphones, capacitive touch screens, front facing camera etc. These features are adopted to high-tier phones relatively quickly within a year or two of introduction in the flagships depending on previous years' reception. Over time, the "standard" features of high-tier phones are still relatively transient year to year, albeit not as mercurial as the flagships. Even some of the examples I mentioned above, like the multi-tone LED, seems to also be losing steam in adoption. The adoption for mid-tier smartphones is more delayed, and seems to be based on a secondary filter from the high-end phones, with an integration timeline more in the range of 3-5 years since introduction. In the above list of features commonly found in high-end phones today, only the capacitive touch screen and multiple microphones can be found universally at this tier. For low-end smartphones, the integration timeline is more akin to 5-10 years. Looking at smartphones priced under $150, we do indeed find phones that have updated wireless connectivity technology but with features similar to the first iPhone. Under $50, the smartphones resemble feature phones with a blend of touchscreen technology. The higher-tier phones don't simply become cheaper over time, instead, what happens after a period of sales, models are slightly discounted until the stocks are all sold to make room for the next generation the year after. As such, the kind of phone features a person has access to is more or less locked to a tier. As such, even though access to

smartphone is becoming more and more universal, we as developers of phone based health sensing solutions should be mindful of the implications of assuming features will one day be adopted and our solution will be ubiquitous.

10.2.2 Anatomical Differences

An important consideration for medical device design is the demographic for which the system is designed to be used by. This is a relatively common issue. For example, blood pressure cuffs typically have two different sizes. Pediatric pulse oximeters have a different wrap design to accommodate different finger size and are sometimes put on the foot. Chest straps for respiration monitoring do not work well for obese patients, so instead a nasal cannula is often used in these cases. Medical device designs often have two factors. The sensor technology and the interface technology. In the example cases for blood pressure cuff and pulse oximeter, the interface technology is adjusted based on the end user. In the case of respiration, the sensing technology is adjusted. For medical device design, the designer controls both technologies and ultimately creates the solution that fits well for the end user.

In the solution space for smartphone health sensing, the control of interface technology is particularly limited and can potentially limit equal access to these solutions. In developing Seismo, we found many examples of the lack of control over interface technology limiting the types of users the solution can accommodate. The first issue we found is the solution does not work well for female users because the phone needs to be placed in the sternum area. This position has two issues. First, the position the phone needs to be placed is often where the bra frame is positioned under the breast. Second is even if the user removes their bra, the phone has to be held in a position that will likely be covered by the breast, creating two issues: the user can't see the screen for visual feedback and the breast resting on the phone dampens the vibration signal. The second issue is also an issue we found for obese patients, which has significantly more fat tissue between the heart and the phone. For the female subjects, we found that it was possible to obtain a seismocardiogram signal from the upper chest. However, in changing the position, we also noticed that the morphology of the signal changed and it was still not as reliable. A different issue was found for more senior subjects who have significant hand tremors. Because the phone is held to the chest by the user, when subjects

have a tremor, both the signals measured by the accelerometer and the camera were affected and resulted in poor signal quality.

In contrast, the sensing technology can perform unequally depending on the anatomical differences. A sensing limitation of HemaApp is the lack of spectral response for >1000nm optical sensing capability of the CMOS camera. This limits the ability to measure blood's water content. To compensate for this issue, our solution relied on using the blue absorption of blood plasma as a proxy for blood volume. Although this appears to work well in our initial study, the assumption may not hold in a more diverse patient population. Depending on the underlying condition that is causing anemia for the patient, the patient may have significant hemolyzation of their blood. This results in elevated bilirubin levels. In this case, the amount of blue wavelength that is absorbed will increase and be misrepresented as plasma volume increase.

For traditional medical devices where both aspects can be controlled, it may be possible to develop the solution for one population and translate it over by making adjustments. Because of the limitation in control for the sensing and interface technology, it is important to consider the end-user when developing smartphone health sensing solutions, because one solution may not transfer at all to a different end-user.

10.2.3 Limited Physical Affordance

A main challenge that came up during the exploration of phone based medical measurements is that there is a mismatch between putting the phone in a position that can capture the necessary signal and how easy it is to operate the phone app. A big issue is visual feedback. As identified during the dual camera exploration of Seismo, we found that there are postures of holding the phone and measuring of facial areas that prevent visual access to the phone screen. A similar issue exist for HemaApp as well. To use the back camera for measurement, the phone either needs to be placed with the screen down or the user has to hold the finger to the camera while the phone is held up, thus cannot rest the finger on the phone or see where the finger is placed. Another issue is the ease of controlling the phone during measurement. If the phone is placed with the screen down, then it is not possible to click the on-screen button UI. Alternatively, for Seismo, when the phone is held against the chest, one hand holds the phone and the other is the finger placed on the phone. This

leaves no hand to click the screen. In these cases, we found that using the volume buttons as physical buttons worked pretty well. We also found that since the phone is being held in weird positions, it is easy to touch the screen and initiate unintended actions like quitting out of the program. We were able to implement a screen blocker that mitigated this issue. All of these issues can be summarized as a physical affordance problem. The phone was designed assuming the user holds the phone in their hand or placed on the table, and the main interaction is to the screen. The only time the user holds the phone to their body, is to talk to the microphone and listen to the other speaker. With that mental model, it makes total sense that the phone has a screen that can be viewed when placed in front of the user. The phone was never meant to be press and shoved onto every inch of the body besides on the user's ear with no expectation of looking at the screen in that case.

10.2.4 Limited Interpretability for Untrained Users

The acceleration of hypochondria due to information available on the Internet, also known as cyberchondria, is likely to be exacerbated by ubiquitous medical testing. Doctors receive years of training on how to apply Bayesian reasoning when accounting for a test result in the diagnostic process. This procedure requires calculating the patient's pretest probability of having the condition and then updating that probability according to the accuracy and result of the test. Calculating the prior probability requires knowing the prevalence of the condition and the specific risk factors that may increase a person's likelihood of having the condition, such as family history and environmental factors. Updating to a post-test probability given a positive test result entails calculating the positive predictive value (PPV) of a test and how often people with a positive test result actually have the medical condition. Calculating PPV requires knowing the test's sensitivity (true positive rate), the test's specificity (true negative rate), and the prevalence of the condition:

$$PPV = \frac{sensitivity \times prevalence}{sensitivity \times prevalence + (1 - specificity) \times (1 - prevalence)} \tag{10.1}$$

This calculation does not always lead to intuitive results. A test with a sensitivity and specificity of 80% for a disease that occurs in 15% of the population will have a PPV of 41.3%. A similar test for a disease that occurs in 5% of the population will have a PPV of only 17.4%. In both cases, the test performs worse than random chance despite having a seemingly decent accuracy.

Test results are never black-and-white; all models have uncertainty bounds that complicate decisions. For example, the current state of HemaApp has a RMSE of 1.2 g/dL. This is reasonable for severe anemia screening scenario screening for people around 7-10 g/dL with a typical range of 7 - 16 g/dL. For use-cases that are very borderline, for example blood donation centers where a potential donor who is determined to be borderline anemic of 11 g/dL is asked to take iron pills until their hemoglobin level raises above 12 g/dL, our system would be poorly suited.

We have yet to tackle this challenge ourselves, but we foresee significant contributions that can be made by researchers. If smartphone-based health sensing apps are going to be freely distributed to the general public rather than prescribed and supervised by trained physicians, the routine of estimating a post-test probability should be as automated as possible. Apps should be able to calculate a pre-test probability by collecting risk factor information. Family history and demographic data can be explicitly recorded through digital forms. Sensors can also be used to infer risk factors. As an example, GPS data could reveal that a person is at a higher risk of a lung condition because of poor local air quality.

The weighing scale provides an interesting study of how important the presentation of results can be to the decision-making process. All scales have uncertainty, yet people tend to fixate on the number they see. Weight is also a function of how much the person is wearing and how much they ate and drank before the measurement. Kay et al. found that people often forget these factors, leading to stress over negligible weight changes. One scale design they suggest graphically emulates a traditional analog scale with exaggerated needle movement to reflect uncertainty. Kay et al. also propose an “always-on” scale design that accounts for daily variance and incorporates information through low burden question prompts so that measurements can be automatically adjusted closer to their true value. Researchers in the machine learning community have actually trained models that learn how different clinical measurements vary over time to help clinicians identify high-risk patients. The same models could be used to help users extrapolate reasonable trends in their data if they feel the need to do so.

10.3 Moving Forward with Smartphone as the Platform

It is reasonable to assume that smartphones as we know them today will continue to evolve year to year. However, I think it is unreasonable to assume that sensors and capabilities we see in high-end or flagship phones will necessarily move down to mid and low tier phones over time. It is important to keep that fact in mind. With all that said, the proliferation of smartphones in the world still presents a great opportunity towards improving access to health sensing.

10.3.1 Sensor-variation Tolerance

As mentioned many times now, sensors are different between phones. Where the sensor is placed, the sensitivity of the sensor, how the sensor is tuned across the bandwidth, etc. The sensors are typically cheap and that means even for the same phone and sensor, because the manufacturing tolerance is low, the sensor may not be exactly the same. There are, however, phenomenon that are more tolerant to the differences between, say, one camera to another. For example, shapes are well preserved. In the case of iPressure, the phone camera is responsible for capturing the shape of the eye under applanation. It is reasonable to assume that this sensing solution will be well preserved across devices. In contrast, ratiometric measurements are more affected. For example, solutions that depend on color measurement like HemaApp is highly subject to sensor variations. Differences in white balance, lighting condition, and sensor tuning can create big differences.

Phenomenon that are frequency dependent will be affected by the tier of the phone. Currently, looking at mid-range phones, it is reasonable to assume cameras of <25 FPS, microphone sampling of 16kHz, and IMU sampling of 200Hz. How these sampling frequency will improve is unclear, and it may be more suitable to target designing for these specifications if the intended use case is for low-resource communities where even mid-range phones may not be extremely prevalent. This can be said about the utility of exploring work-around solutions in the process of researching in this solution space. Work-arounds that can be used by mid-tier phones are much more impactful than one-off work arounds that is very device specific. The camera and accelerometer synchronization technique introduced in Seismo would be a relatively good example of a work around that is translatable to most phones.

10.3.2 Low-Fidelity Augmentation

Many of the solutions focus on using the phone's own sensor capabilities. However, in many cases, it is beneficial to enhance the phone's capability through external supplements. For the most part, in my exploration I focused on passive and mechanical supplements. For example, in HemaApp, the phone case for finger placement is an example of such a supplementation. Although one might argue that any such supplement nullifies the advantage of a software solution for deployment, I believe that there is a level of deployability that can make such supplementation reasonable. For the finger holder case, the case is easily acquired and a black cloth is not difficult to come by.

There are also cases where it may not be possible to use ultralow fidelity add-ons. For the example of iPressure, the add-on device is much more custom and likely not DIY-able. It is possible that such a solution can be part of a fairly cheap kit that has a number of such physical add-ons that can expand the capabilities of the phone. A kit might be created, such as for specific use cases like a "eye exam in a box" for low resource clinics, or just a few general purpose physical interfaces that helps block out light, provide color reference, direct sound waves, or press against the body more ergonomically. The success of this line of solution would come down to whether the add-ons are truly low cost in manufacturing and a major advantage if it is usable across multiple use cases. Another important consideration in the design of such a kit is the reuse of manufacturing. Beyond the electronic component reuse, I think another important aspect to mass production is the hardware enclosure should also be explicitly designed to use physical artifacts that are massively reproduced. For example, a commonly useful measurement is an image/video capture of the eye or skin patch without ambient light. This can be accomplished by some form of a conical cup with a deformable ring, with the potential of a lens. This is a relatively common place item when we start to think about single lens cameras and the eye cup of the view finder. A sensible direction in the design of these cheap medical phone add-ons is to think critically of the physical enclosure and interfaces that are already massively produced and manufactured that can reduced the cost of custom manufacturing.

10.3.3 Long-term Tracking in the Background

One major advantage of the phone that I have so far not talked about, but nonetheless should be noted, is the phone's special role of often being close to the user. In the next section, I'm going to

discuss how we might want to move forward in designing a more suitable platform for health sensing. The role of what I will propose focuses on creating a solution that can better support developing cheap and ubiquitous medical diagnostic and monitoring capabilities for active measurements, akin to the solutions I have presented in this dissertation. However, certain instantiation of the proposed solution may or may not always be with the user if it is not implemented on to a phone. I believe the smartphone, in these cases, will play a strong supporting role to such a new class of solution by being the agent for long-term tracking of the user.

General mobility and sleep patterns, for example, are important factors that play a role in many clinical end usecases. Phone sensors are well suited for monitoring in the background. An example where an interplay of phone sensing and a cheap medical sensor can operate in synergy is the monitoring of sickle cell disease patients. Patient living with SCD are at risk of acute anemia due to breakdown of mutated red blood cells. A major improvement for disease management for sickle cell disease is to provide a cheap solution for daily monitoring of hemoglobin, a solution space of HemaApp. However, in conjunction to hemoglobin deficiency, pain induced lowering of mobility is often a co-existing condition. Due to the mutation of red blood cells, blood flow for SCD patient is often impeded and the decreased flow often leads to pain episodes. The pain typically leads to decrease mobility, and the decreased mobility then leads to decreased blood flow. This vicious cycle leads to acute pain crisis episodes. In conjunction to monitoring hemoglobin regularly, monitoring the patient's general mobility and sleep time motion can add a layer of information important for understanding the state of the patient.

This relates to the previous point about interpretability of mobile health sensing. The goal for at home monitoring of SCD patients is to prevent cases of extreme anemia brought about by these acute crisis. In this case, the goal of the monitor is to flag potential crisis situations that need the patient to be admitted for treatment. What the phone can provide is a more complete prior to the measurement provided by the cheap medical monitor. By only monitoring hemoglobin levels, it may either instill unwarranted confidence or false alarms. Using changes in mobility pattern as a prior could help in the robustness of such a monitoring solution.

10.3.4 Explicitly Design a Platform for Software as a Medical Device

The difference between the phone based solution versus designing a pieced together medical device is not fully about sensor reuse. Rather, the core difference is the sensors of the phone only gets us part of the way to the final measurement. An image of the eye. A color measurement. An accelerometer that can measure the vibration. At a meta level, the smartphone provides a measurement of some physical phenomenon of the body that can then be processed via software to perform biomarker estimations. For medical sensor like a pulse oximeter, the measurement and processing fully encapsulates the end biomarker and has no room for reuse beyond designing for specific applications. This is fundamentally different because the phone's collection of sensors gives rise to new biomarkers that can be measured based on the collective creativity of those who are working closer to the clinical end points and have a better sense for the actual needs and use cases.

The phone, however, was never intended to be the hardware platform to enable what is sometimes referred to as a Software as a Medical Device (SaMD). This can be seen in the multitude of limitations I drew out from the different explorations made through my work and also a review of the related work in this domain. There are certainly circumstances in which the current phone is already well suited for the solution, mainly solutions where the end measurement is less reliant on controlled environments. And for these scenarios, I believe the current phone infrastructure can indeed be a good solution. However, I believe a promising direction of research will be to explore from the ground up a platform that will enable software as a medical device in a more robust, flexible, and reliable manner. Much of the work so far has focused on using the phone as the central hardware platform and add-on where needed. Such as a lens attachment, color correction card, finger cover. However, much of the limitation of the phone at the core is still that the phone has many other requirements such as telemetry, productivity, entertainment, etc. Such a hardware platform should be standalone and aim to facilitate capturing the data necessary for an expandable list of health/medical measurements, while standardizing the hardware/firmware requirements to provide a uniform specification or tiered specification that can be queried. In this way, a developer of a SaMD can have a notion of the hardware interface that their solution is deployed through SDK level query of the hardware support, which can include both sensor and enclosure configuration. The SDK should support abstractions closer to the level of physiological measurements rather than sen-

sor measurements. Instead of just being able to request a camera feed, it would be more powerful to be able to request an image of an eye measured at specific exposure wavelengths captured at a specific framerate. The SDK should also provide a quality check of the data return. The software developer can be guaranteed that what they get back is color corrected for the sensor, is actually an eye, and potentially other meta data such as position of the eye. In this way, the focus of the software developer is one that is of "how can I use an image of the eye to address a clinical end usecase" and not "what do I need to do to make sure this photo is useful."

It is certainly open ended as to how such a system will take form, but the ultimate vision of such a system is that it will be a blueprint specification for a platform dedicated for SaMD. It might take shape as a weirdly shaped mobile health smartphone with added affordances to interface nicely with the body, a peripheral device that could use a smartphone as its computation, or as a completely standalone device. My vision is that instead of a thermoeter in a first aid kit, a low cost device for SaMD will be in every first aid kit. I believe the shape it takes is potentially irrelevant, but rather the goal for this new direction is define explicitly what the right specifications should be to enable widespread ability to develop medical devices via software.

10.4 Conclusion

As a researcher in this topic of smartphone-based health sensing, my dissertation has been aimed at advancing the state-of-the art of this form of health sensing solution space. My main solution space has been to push the limits of what we can do with the smartphone's camera, demonstrating that it is feasible to use cameras that are commonly integrated in smartphones to measure hemoglobin concentration. Utilizing multiple sensors in the phone, such as combining cameras with accelerometers and microphones, I then enabled blood pressure monitoring. With passive augmentation to the phone, I then demonstrated the capability of eye pressure measurement. The premise of the work focuses on working within the confines of what is "already integrated on phones" as a design criteria. However, I have found limitations to our work's motivation of "make it work on a phone and it has the potential of becoming ubiquitous."

Just because a capability exist in one phone or that the app can be implemented on a phone that was bought off the shelf, it does not mean it has the potential of being ubiquitous. The "Smart-

phone" is not a singular item that has a clear specification. Smartphone is really a class of item that differ from product to product. A major factor to the difference between actual products is often driven by cost. With one of the central theme being to enable low-to-middle income communities to access healthcare services in limited resource situations, it is important to consider the long-term applicability of our research in this practical frame.

The importance of working directly within the confines of what currently exists in phones, however, should not be brushed over. Although we cannot necessarily guarantee long-term applicability and feasibility for true ubiquitiy, I was able to draw out a nuanced discussion of the limitation of our current implementation of smartphones towards health sensing in any price tier because of the deep exploration within this design constraint. When we refer to user interfaces for smartphones today, we focus on how a user interacts with the in-screen digital content. Almost none of the use-cases focuses on interfaces that require the interfacing with the human body to capture physiological measurements. This lack of physical interfacing with the human body creates a major challenge for adapting smartphones we have today for measuring for medical and health metrics. Smartphones as a class of item hold the potential to help deliver health and medical monitoring to the masses as the platform for Software as a Medical Device. However, to make this a reality beyond research ideas or one-off implementations, the next step is to explicitly work towards designing our platform to have health and medical monitoring as a central use case. This includes a blueprint for physical affordances, cross-device transferability, and data formatting that have higher-level abstractions that more closely represent physiological measurements rather than sensor data. The ubiquity of this solution space hinges on shifting the central definition of the "smartphone" to include health and medical sensing as a use case no less important than making a call or connecting to the internet.

BIBLIOGRAPHY

[1] Omar Abdallah, Mohammed Natsheh, Kawther Abo Alam, Qasem Qananwah, Ahrned a Nabulsi, and Armin Bolz. Concentrations of Hemoglobin Fractions Calculation Using Modified Lambert-Beer Law and Solving of an Ill-Posed System of Equations. *Spie*, 7715(1):1–8, 2010.

[2] Rajiv Agarwal, Jennifer E. Bills, Tyler J.W. Hecht, and Robert P. Light. Role of home blood pressure monitoring in overcoming therapeutic inertia and improving hypertension control. *Hypertension*, 57(1):29–38, 2010.

[3] Kawther Abo Alam. *Fuzzy logic hemoglobin sensors*. PhD thesis, Karlsruhe Institute of Technology, 2011.

[4] Sangwon Bae, Tammy Chung, Denzil Ferreira, Anind K. Dey, and Brian Suffoletto. Mobile phone sensors and supervised machine learning to identify alcohol use events in young adults: Implications for just-in-time adaptive interventions. *Addictive Behaviors*, 83(July 2017):42–47, 2018.

[5] R Bainbridge, D R Higgs, G H Maude, and G R Serjeant. Clinical presentation of homozygous sickle cell disease. *The Journal of pediatrics*, 106(6):881–885, 6 1985.

[6] R. D. Baker and F. R. Greer. Diagnosis and Prevention of Iron Deficiency and Iron-Deficiency Anemia in Infants and Young Children (0-3 Years of Age). *Pediatrics*, 126(5):1040–1050, 2010.

[7] Sarah M Bartsch, Peter J Hotez, Lindsey Asti, Kristina M Zapf, Maria Elena Bottazzi, David J Diemert, and Bruce Y Lee. The Global Economic and Health Burden of Human Hookworm Infection. *PLoS neglected tropical diseases*, 10(9):e0004922, 9 2016.

[8] A Bernjak and A Stefanovska. Pulse transit times to the capillary bed evaluated by laser doppler flowmetry., Mar 2009.

[9] E Beutler and J Waalen. The definition of anemia: what is the lower limit of normal of the blood hemoglobin concentration? *Blood*, 107(5):1747–1750, 2006.

[10] Ernest Beutler and Jill Waalen. The definition of anemia : what is the lower limit of normal of the blood hemoglobin concentration ? *Blood*, 107(5):1747–1750, 2006.

[11] A. Brown and A. L. Goodall. Normal variations in blood haemoglobin concentration. *The Journal of Physiology*, 104(4):404–407, 1946.

[12] C Brugnara, FA Oski, and DG Nathan. Diagnostic approach to the anemic patient. In SH Orkin, DE Fisher, D Ginsburg, and L Thomas, editors, *Nathan and Oski's Hematology and Oncology of Infancy and Childhood*, page 293. WB Saunders, 8th ed edition, 2015.

[13] Waylon Brunette, Rita Sodt, Rohit Chaudhri, Mayank Goel, Michael Falcone, Jaylen Van Orden, and Gaetano Borriello. Open data kit sensors: a sensor integration framework for android at the application-level. *Proceedings of the 10th international conference on Mobile systems, applications, and services - MobiSys '12*, page 351, 2012.

[14] Waylon Brunette, Samuel Sudar, Mitchell Sundt, Clarice Larson, Jeffrey Beorse, and Richard Anderson. Open Data Kit 2.0. *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services - MobiSys '17*, pages 440–452, 2017.

[15] Andrew M. Carek, Jordan Conant, Anirudh Joshi, Hyolim Kang, and Omer T. Inan. Seismowatch: Wearable cuffless blood pressure monitoring using pulse transit time. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3):40:1–40:16, September 2017.

[16] GS Cembrowski, J Chan, and C Cheng. NHANES 1999-2000 data used to create comprehensive health-associated race-, sex- and age-stratified pediatric reference intervals for the Coulter MAXM. *Laboratory Hematology*, (10):245, 2004.

[17] V. Chandrasekaran, R. Dantu, S. Jonnada, S. Thiyagaraja, and K. P. Subbu. Cuffless differential blood pressure estimation using smart phones. *IEEE Transactions on Biomedical Engineering*, 60(4):1080–1089, April 2013.

[18] Thomas Coppetti, Andreas Brauchlin, Simon Müggler, Adrian Attinger-Toller, Christian Templin, Felix Schönrath, Jens Hellermann, Thomas F. Lüscher, Patric Biaggi, and Christophe A. Wyss. Accuracy of smartphone apps for heart rate measurement. *European Journal of Preventive Cardiology*, 24(12):1287–1293, 2017.

[19] Carlton Dampier, B N Yamaja Setty, Barry Eggleston, Darcy Brodecki, Patricia O'neal, and Marie Stuart. Vaso-occlusion in children with sickle cell disease: clinical characteristics and biologic correlates. *Journal of pediatric hematology/oncology*, 26(12):785–790, 12 2004.

[20] Lilian de Greef, Mayank Goel, Min Joon Seo, Eric C Larson, James W. Stout, James a Taylor, and Shwetak N Patel. Bilicam: Using Mobile Phones to Monitor Newborn Jaundice. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing - UbiComp '14 Adjunct*, pages 331–342, 2014.

[21] A M Emond, R Collis, D Darvill, D R Higgs, G H Maude, and G R Serjeant. Acute splenic sequestration in homozygous sickle cell disease: natural history and management. *The Journal of pediatrics*, 107(2):201–206, 8 1985.

[22] G. Fortino and V. Giampa. Ppg-based methods for non invasive and continuous blood pressure measurement: an overview and development issues in body sensor networks. In *IEEE International Workshop on Medical Measurements and Applications*, pages 10–13, April 2010.

[23] H. Gehring, C. Hornberger, L. Dibbelt, A. Roth-Isgkeit, K. Gerlach, J. Schumacher, and P. Schmucker. Accuracy of Point-of-Care-Testing (POCT) for Determining Hemoglobin Concentrations. *Acta Anaesthesiologica Scandinavica*, 46(8):980–986, 2002.

[24] Mayank Goel, Elliot Saba, Maia Stiber, Eric Whitmire, Josh Fromm, Larson Eric C., Gaetano Borriello, and Shwetak N. Patel. SpiroCall : Measuring Lung Function over a Phone Call. In *CHI '16*, 2016.

[25] Mathew J. Gregoski, Martina Mueller, Alexey Vertegel, Aleksey Shaporev, Brenda B. Jackson, Ronja M. Frenzel, Sara M. Sprehn, and Frank a. Treiber. Development and Validation of a Smartphone Heart Rate Acquisition Application for Health Promotion and Wellness Telehealth Applications. *International Journal of Telemedicine and Applications*, 2012, 2012.

[26] Carl Hartung, Yaw Anokwa, Waylon Brunette, Adam Lerer, Clint Tseng, and Gaetano Borriello. Open Data Kit: Tools to Build Information Services for Developing Regions. *Proceedings of the International Conference on Information and Communication Technologies and Development*, pages 1–11, 2010.

[27] Md Kamrul Hasan, Munirul Haque, Nazmus Sakib, Richard Love, and Sheikh I. Ahamed. Smartphone-based Human Hemoglobin Level Measurement Analyzing Pixel Intensity of a Fingertip Video on Different Color Spaces. *Smart Health*, 5(May 2017):26–39, 2018.

[28] Javier Hernandez, Daniel J. McDuff, and Rosalind W. Picard. Biophone: Physiology monitoring from peripheral smartphone motions. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2015-Novem:7180–7183, 2015.

[29] Christian Holz and Edward J. Wang. Glabella: Continuously sensing blood pressure behavior using an unobtrusive wearable device. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(3):58:1–58:23, September 2017.

[30] O. T. Inan, P. F. Migeotte, K. S. Park, M. Etemadi, K. Tavakolian, R. Casanella, J. Zanetti, J. Tank, I. Funtova, G. K. Prisk, and M. Di Rienzo. Ballistocardiography and seismocardiography: A review of recent advances. *IEEE Journal of Biomedical and Health Informatics*, 19(4):1414–1427, July 2015.

[31] International Medical Device Regulatory Forum. IMDRF Software as a Medical Device (SaMD). Technical report, 2013.

[32] Natasha Jaques, Sara Taylor, Akane Sano, and Rosalind Picard. Multi-task , Multi-Kernel Learning for Estimating Individual Wellbeing. *Multimodal Machine Learning Workshop in conjunction with NIPS*, pages 1–7, 2015.

[33] Natasha Jaques, Sara Taylor, Akane Sano, and Rosalind Picard. Predicting Tomorrow's Mood, Health, and Stress Level using Personalized Multitask Learning and Domain Adaptation Ognjen (Oggi) Rudovic. *Journal of Machine Learning Research*, 66:17–33, 2017.

[34] E. Jonathan and Martin Leahy. Investigating a Smartphone Imaging Unit for Photoplethysmography. *Physiological measurement*, 31(11):79–83, 2010.

[35] A.Dias Junior, Srinivasan Murali, Francisco Rincon, and David Atienza. Estimation of blood pressure and pulse transit time using your smartphone. In *Euromicro Conference on Digital System Design (DSD)*, pages 173–180, 2015.

[36] Alair Dias Junior, Srinivasan Murali, Francisco Rincon, and David Atienza. Estimation of Blood Pressure and Pulse Transit Time Using Your Smartphone. *18th IEEE/Euromicro Conference on Digital System Design (DSD 2015)*, (1):10–18, 2015.

[37] M. Kachuee, M. M. Kiani, H. Mohammadzade, and M. Shabany. Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. In *International Symposium on Circuits and Systems (ISCAS)*, pages 1006–1009, May 2015.

[38] Nicholas J Kassebaum, Rashmi Jasrasaria, Mohsen Naghavi, Sarah K Wulf, Nicole Johns, Rafael Lozano, Mathilda Regan, David Weatherall, David P Chou, Thomas P Eisele, Seth R Flaxman, Rachel L Pullan, Simon J Brooker, and Christopher J L Murray. A systematic analysis of global anemia burden from 1990 to 2010. *Blood*, 123(5):615–624, 1 2014.

[39] Chang-Sei Kim, Andrew M. Carek, Ramakrishna Mukkamala, Omer T. Inan, and Jin-Oh Hahn. Ballistocardiogram as proximal timing reference for pulse transit time measurement: Potential for cuffless blood pressure monitoring. *IEEE transactions on bio-medical engineering*, Nov 2015.

[40] Hg Klein and Dj Anstee. *Mollison's blood transfusion in clinical medicine*. 2008.

[41] P J Kling, R L Schmidt, R A Roberts, and J A Widness. Serum erythropoietin levels during infancy: associations with erythropoiesis. *The Journal of pediatrics*, 128(6):791–796, 6 1996.

[42] Jens Kraitl, Ulrich Timm, Elfed Lewis, and Hartmut Ewald. Optical Sensor Technology for a Noninvasive Continuous Monitoring of Blood Components. *Library*, 7572:757209–757209, 2010.

[43] F. Landreani, M. Morri, A. Martin-Yebra, C. Casellato, E. Pavan, C. Frigo, and E. G. Caiani. Ultra-short-term heart rate variability analysis on accelerometric signals from mobile phone. In *E-Health and Bioengineering Conference (EHB)*, pages 241–244, June 2017.

[44] J D Lane, L Greenstadt, D Shapiro, and E Rubinstein. Pulse transit time and blood pressure: an intensive analysis. *Psychophysiology.*, Jan 1983.

[45] Eric C Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N Patel. SpiroSmart : Using a Microphone to Measure Lung Function on a Mobile Phone. (Figure 1).

[46] Eric C Larson, Mayank Goel, Gaetano Boriello, Sonya Heltshe, Margaret Rosenfeld, and Shwetak N Patel. SpiroSmart : Using a Microphone to Measure Lung Function on a Mobile Phone. In *Ubicomp '12*, pages 280–289, 2012.

[47] Dongheng Li, David Winfield, and Derrick J Parkhurst. Starburst : A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops*, page 79, 2005.

[48] Marshall A Lichtman, Ernest Beutler, Thomas J Kipps, Josef T Prchal, and Marcel Levi, editors. *Williams' Hematology*. McGraw-Hill, New York, 7 edition, 2006.

[49] He Liu, Kamen Ivanov, Yadong Wang, and Lei Wang. Toward a smartphone application for estimation of pulse transit time. In *Sensors*, 2015.

[50] He Liu, Kamen Ivanov, Yadong Wang, and Lei Wang. Toward a Smartphone Application for Estimation of Pulse Transit Time. *Sensors*, 15(10):27303–27321, 2015.

[51] Mark R. MacKnet, Martin Allard, Richard L. Applegate, and James Rook. The Accuracy of Noninvasive and Continuous Total Hemoglobin Measurement by Pulse CO-oximetry in Human Subjects Undergoing Hemodilution. *Anesthesia and Analgesia*, 111(6):1424–1426, 2010.

[52] A Mariakakis, E Wang, S Patel, and J C Wen. A Smartphone - based System for Assessing Intraocular Pressure. In *IEEE EMBC '16*, 2016.

[53] Alex Mariakakis, Megan A Banks, Lauren Phillipi, Lei Yu, James Taylor, and Shwetak N Patel. BiliScreen: Smartphone-Based Scleral Jaundice Monitoring for Liver and Pancreatic Disorders. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(2):1–26, 2017.

[54] Alex Mariakakis, Sayna Parsi, Shwetak N Patel, and Jacob O Wobbrock. Drunk User Interfaces : Determining Blood Alcohol Level through Everyday Smartphone Tasks. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2018.

[55] Alex Mariakakis, Edward J Wang, Mayank Goel, and Shwetak N Patel. Challenges in realizing smartphone-based health sensing. *To appear in IEEE Pervasive Computing*, 2019.

[56] G V Marie, C R Lo, J Van, and D W Johnston. The relationship between arterial blood pressure and pulse transit time during dynamic and static exercise. *Psychophysiology.*, Sep 1984.

[57] Alissa Martin and Alexis A Thompson. Thalassemias. *Pediatric clinics of North America*, 60(6):1383–1391, 12 2013.

[58] Y Matoth, R Zaizov, and I Varsano. Postnatal changes in some red cell parameters. *Acta paediatrica Scandinavica*, 60(3):317–323, 5 1971.

[59] Reham Mohamed and Moustafa Youssef. HeartSense. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):1–18, 2017.

[60] Reham Mohamed and Moustafa Youssef. Heartsense: Ubiquitous accurate multi-modal fusion-based heart rate estimation using smartphones. *IMWUT*, 1:97:1–97:18, 2017.

[61] Uma P. Moravapalle, Amit Deshpande, Ashish Kapoor, Ramachandran Ramjee, and Priya Ravi. Blood count on a smartphone microscope: Challenges. In *Proceedings of the 18th International Workshop on Mobile Computing Systems and Applications*, HotMobile '17, pages 37–42, New York, NY, USA, 2017. ACM.

[62] R. Mukkamala, J. O. Hahn, O. T. Inan, L. K. Mestha, C. S. Kim, H. Töreyin, and S. Kyal. Toward ubiquitous blood pressure monitoring via pulse transit time: Theory and practice. *IEEE Transactions on Biomedical Engineering*, 62(8):1879–1901, Aug 2015.

[63] Rajalakshmi Nandakumar, Shyamnath Gollakota, and Nathaniel Watson. Contactless Sleep Apnea Detection on Smartphones. *The 13th Annual International Conference on Mobile Systems, Applications and Services*, pages 45–57, 2015.

[64] Andreas Patzak, Yuri Mendoza, Heiko Gesche, and Martin Konermann. Continuous blood pressure measurement using the pulse transit time: Comparison to intra-arterial measurement. *Blood pressure*, 24(4):217–221, April 2015.

[65] Rong-Chao Peng, Wen-Rong Yan, Ning-Ling Zhang, Wan-Hua Lin, Xiao-Lin Zhou, and Yuan-Ting Zhang. Cuffless and continuous blood pressure estimation from the heart sound signals. *Sensors*, 15(9):23653–23666, 2015.

[66] Harry Quigley and A.T Broman. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Opthalmol*, 90:262–267, 2006.

[67] David C Rees, Thomas N Williams, and Mark T Gladwin. Sickle-cell disease. *Lancet (London, England)*, 376(9757):2018–2031, 12 2010.

[68] David B Rein, Ping Zhang, Kathleen E Wirth, Paul P Lee, Thomas Hoerger, Nancy McCall, Ronald Klein, James M Tielsch, Sandeep Vijan, and Jinan Saaddine. The Economic Burden of Major Adult Visual Disorders in the United States. *Arch Ophthalmol*, 124, 2006.

[69] Mark J. Rice, Nikolaus Gravenstein, and Timothy E. Morey. Noninvasive Hemoglobin Monitoring: How Accurate is Enough? *Anesthesia and Analgesia*, 117(4):902–907, 2013.

[70] Js Ruckman. A Comparative Study of Total Hemoglobin Measurement Technology: Noninvasive Pulse Co-oximetry and Conventional Methods. 2011.

[71] Carla Sala, Erika Santin, Marta Rescaldani, Cesare Cuspidi, and Fabio Magrini. What is the accuracy of clinic blood pressure measurement? *American Journal of Hypertension*, 18(2):244–248, 2005.

[72] N. Shah, E. A. Osea, and G. J. Martinez. Accuracy of noninvasive hemoglobin and invasive point-of-care hemoglobin testing compared with a laboratory analyzer. *International Journal of Laboratory Hematology*, 36(1):56–61, 2014.

[73] Kim Smith-Whitley, Huaqing Zhao, Richard L Hodinka, Janet Kwiatkowski, Renee Cecil, Tamara Cecil, Avital Cnaan, and Kwaku Ohene-Frempong. Epidemiology of human parvovirus B19 in children with sickle cell disease. *Blood*, 103(2):422–427, 1 2004.

[74] Robert L Stamper. A History of Intraocular Pressure and Its Measurement. 88(1):16–28, 2011.

[75] Sue Stevens, Clare Gilbert, and Nick Astbury. How to measure intraocular pressure : applanation tonometry. 20(64):2007, 2007.

[76] J. Talts, R. Raamat, K. Jagomägi, and J. Kivastik. *An Influence of Multiple Affecting Factors on Characteristic Ratios of Oscillometric Blood Pressure Measurement*, pages 73–76. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[77] U. Timm, J. Kraitl, H. Gewiss, K. Stuepmann, M. Hinz, S. Koball, and H. Ewald. Novel Multi Wavelength Sensor Concept to Detect Total Hemoglobin Concentration, Methemoglobin and Oxygen Saturation. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 9332(Figure 1):1–9, 2015.

[78] H. Von Schenck, M. Falkensson, and B. Lundberg. Evaluation of 'HemoCue,' a new device for determining hemoglobin. *Clinical Chemistry*, 32(3):526–529, 1986.

[79] Edward Jay Wang, William Li, Doug Hawkins, Terry Gernsheimer, Colette Norby-slycord, and Shwetak N Patel. HemaApp: Noninvasive Blood Screening of Hemoglobin using Smartphone Cameras. In *UbiComp '16 Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 593–604, 2016.

[80] E.J. Wang, J. Zhu, W. Li, R. Rana, and S. Patel. HemaApp IR: Noninvasive hemoglobin measurement using unmodified smartphone cameras and built-in leds. In *UbiComp/ISWC 2017 - Adjunct Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 2017.

[81] WHO. Global health observatory (gho) data: Raised blood pressure, 2017.

[82] E. S. Winokur, D. D. He, and C. G. Sodini. A wearable vital signs monitor at the ear for continuous heart rate and pulse transit time measurements. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2724–2727, Aug 2012.

[83] Mico Yee-Man Wong, Carmen Chung-Yan Poon, and Yuan-Ting Zhang. An evaluation of the cuffless blood pressure estimation based on pulse transit time technique: a half year study on normotensive subjects. *Cardiovascular Engineering*, 9(1):32–38, Mar 2009.

[84] D F Woodhouse. Five g ME Applanation Tonometry. *Exp. Eye. Res.*, 15:509–512, 1973.

[85] World Health Organization. Nutritional anaemias. Report of a WHO group of experts., 1968.

[86] World Health Organization. Treatment of Severe Malaria. *Guidelines For The Treatment of Malaria*, pages 71–88, 2015.

[87] C. Yang and N. Tavassolian. Pulse transit time measurement using seismocardiogram, photoplethysmogram, and acoustic recordings: Evaluation and comparison. *IEEE Journal of Biomedical and Health Informatics*, PP(99):1–1, 2017.

[88] Guanqun Zhang, Mingwu Gao, Da Xu, N. Bari Olivier, and Ramakrishna Mukkamala. Pulse arrival time is not an adequate surrogate for pulse transit time as a marker of blood pressure. *Journal of Applied Physiology*, 111(6):1681–1686, 2011.

[89] Y L Zhang, Y Y Zheng, Z C Ma, and Y N Sun. Radial pulse transit time is an index of arterial stiffness. *Hypertension research : official journal of the Japanese Society of Hypertension.*, Jul 2011.