2020

Applications Of Wearable Sensors In Delivering Biologically Relevant Signals

Jordyn Scism
University of Vermont

APPLICATIONS OF WEARABLE SENSORS IN DELIVERING BIOLOGICALLY RELEVANT SIGNALS

A Thesis Presented

by

Jordyn E. Scism

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Science
Specializing in Biomedical Engineering

August, 2020

Defense Date: July 10th, 2020
Thesis Examination Committee:

Ryan McGinnis, M.D., Advisor
Eric Hernandez, Ph.D., Chairperson
Jason Bates, Ph.D.
Jeff Frolik, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

ABSTRACT

With continued advancements in wearable technologies, the applications for their use are growing. Wearable sensors can be found in smart watches, fitness trackers, and even our cellphones. The common applications in everyday life are usually step counting, activity tracking, and heart rate monitoring. However, researchers have developed ways to use these similar sensors for clinically relevant diagnostic measures, as well as, improved athletic training and performance. Two areas of interest for the use of wearable sensors are mental health diagnostics in children and heart rate monitoring during intense physical activity from new locations, which are discussed further in this thesis.

About 20% of children will experience an anxiety or depressive disorder. These disorders, if left untreated, can lead to comorbidity, substance abuse, and even suicide. Current methods for diagnosis are time consuming and only offered to those most at risk (i.e., reported or referred by a teacher, doctor, or parent). For the children that do get referred to a specialist, the process is often inaccurate. Researchers began using mood induction task to observe behavioral responses to specific stimuli in hopes to improve the accuracy of diagnostics. However, these methods involve long hours of training and watching videos of the activities. Recently, a few studies have focused on using wearable sensors during mood induction tasks in hopes to pick up on relevant movements to distinguish those with and without an internalizing disorder. The first study presented in this thesis focuses on using wearable inertial measurement units during the 'Bubbles' mood induction task. A decision tree was developed to identify children with internalizing disorders, accuracy of this model was 71% based on leave-one-subject-out cross validation.

The second study focuses on estimating heart rate using wearable photoplethysmography sensors at multiple body locations. Heart rate is an important vital sign used across a variety of contexts. For example, athletes use heart rate to determine whether they are hitting their desired heart rate zones during training and doctors can use heart rate as an early indicator of disease. With the advancements made in wearables, photoplethysmography can now be used to collect signals from devices anywhere on the body. However, estimating heart rate accurately during periods of intense physical activity remains a challenge due to signal corruption cause by motion artifacts. This study focuses on evaluating algorithms for accurately estimating heart rate from photoplethysmograms and determining the optimal body location for wear. A phase vocoder and Wiener filtering approach was used to estimate heart rate from the forearm, shank, and sacrum. The algorithm estimated heart rate to within 6.2 6.9, and 6.7 beats per minute average absolute error for the forearm, shank, and sacrum, respectively, across a wide variety of physical activities selected to induce varying levels of motion artifact. These results represent a 26.1%, 18.3%, 21.0% improvement from the estimates of an Apple Watch for the forearm, shank, and sacrum, respectively.

ACKNOWLEDGEMENTS

First, I would like to thank my family all of their love and support throughout both my undergraduate and graduate studies. They gave me the opportunity to be here today and have always encouraged me to pursue my dreams. They have also taught me that I can do anything I put my mind to, and for that I am always grateful. Secondly, I would like to thank my friends for the laughter and adventures that kept me sane throughout this experience. Without them, I would have likely driven myself crazy.

Next, I would like to thank my advisor, Dr. Ryan McGinnis, for his support over the past 4 years. His guidance and support have helped shape me into the engineer I am today. Through his courses and advising, I have learned so much about using my engineering skills to help others and share new ways of thinking. I would also like to thank the other members of my thesis committee, Dr. Jeff Frolik, Dr. Jason Bates, and Dr. Eric Hernandez, for all they have taught me throughout my college career and agreeing to be a part of my committee.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1. Background

Wearable sensors have the ability to collect an array of health-related data outside of research settings, and for long periods of time. Using sensor-based measures to understand the human body allows bias to be taken out of the data collected, and therefore, the diagnosis of diseases, disorders, and injuries. Wearable sensors also give the opportunity to remotely access data collected, analyze data in real-time, and observe pattern changes over time. This is extremely important during these times, where we are trying to have as little contact with other humans as possible. The remainder of this Chapter discusses different types of wearable sensors, their measurement modalities, and how they have been used in research and practice.

Inertial measurement units (IMUs), containing an accelerometer and gyroscope, are common wearables for obtaining information about the way people move. Many studies use these modalities for gait analysis [1]–[3]. These sensors have been incorporated into devices that can be worn almost anywhere on the body. This gives researchers more freedom when designing studies. They have also been paired with electrodes for electromyography (EMG), which captures the muscle activations underlying body movements [4].

Electrodes and optical sensors are also being incorporated into wearable sensors for capturing electrocardiographs (ECGs) and photoplethysmographs (PPGs), respectively [5]. PPG and ECGs are able to give important information about HR, heart rate variability (HRV), lung health, and heart health [6]–[10]. These measures are often important for doctors, as well as athletes.

More recently, new band aid or patch-like wearable sensors have been developed that are equipped with IMUs, ECGs, and/or PPGs. With the ability to have one or more of these tools in a small sensor that can be placed anywhere on the body, more data can be collected and used for answering research questions. These devices have become the heart of new approaches for point-of-care diagnostics [11]–[14], fitness tracking [6]–[8], gait analysis [1]–[3], continuous monitoring of patient activities [15], [16], would healing [17], and more feasible analysis of disease progression in Multiple Sclerosis (MS), Huntington's (HD), and Parkinson's disease (PD) [1], [2].

Wearable sensors have the ability to change the healthcare system, both in terms of treatment and rehabilitation methods, as well as diagnostic measures. With the ability to collect more data remotely, people living in rural areas may be able to have more access to specialist healthcare services and receive the quality of care currently experienced by people living in more urban environments [18], [19].

1.2. Wearables in Healthcare

Many researchers have become interested in using wearable sensors for diagnosing diseases/disorders, monitoring the progression of a disease/disorder, or for tracking responses to rehabilitation or medication [1], [3], [20]–[25]. The following two subsections discuss the use of wearable sensors in monitoring (1.2.1) and diagnosing (1.2.2) diseases and disorders.

1.2.1. Monitoring

Wearable sensors have become popular for monitoring patients with different diseases to assess disease progression, safety (i.e., fall risks), rehabilitation, and treatment efficacy [26]. People with MS, HD, and PD are at increased risk of falling [20]–[23], [27]. Studies using wearable IMUs are being conducted to determine when a person may be at a higher risk of falling and alerting them so they can make an informed decision to either stop what they are doing or be more cautious about their activities. This is important since falling can lead to other injuries, especially in older populations [26].

Other work has been done to monitor bipolar disorders [28]–[30]. Self- or parent-reports are often used for monitoring mental health disorders which often introduce biases [28], [29], [31]. Wearable sensors have become more popular in monitoring bipolar mood shifts since they can capture biologically relevant responses opposed to having the patient report their own symptoms. Studies have looked into collecting physiological signals like HR, body posture, activity recognition and respiration rates [28], [30]. Another study looked at using smart phone data to determine states and state changes in bipolar patients [29]. Understanding changes in patterns, like movement, social media usage, talking on the phone, and voice characteristics, throughout the day or week can collect important changes that may not be reported in a self-assessment or questionnaire [29]. In regards to bipolar disorders, wearables are able to reliably identify key changes in states that are not identifiable from self-assessments making them more reliable for monitoring the disorder so proper treatment can be made.

1.2.2. Diagnosing

The key to treating many diseases and disorders is to intervene early. Mental health disorders [32]–[34], arthritis [35], Alzheimer's [36], and congenital heart disease [37] are just a few examples where early diagnosis and intervention can have a huge effect in the way the disease progresses and is treated. With the ability to continuously monitor patient vitals and symptoms outside of a clinic or laboratory setting, data is able to capture important information that may not be seen at the doctor's office like an arrhythmia [38] or changes in blood pressure and blood oxygen saturation [39].

Mental health disorders, like anxiety and depression, can lead to suicide which is a potentially preventable issue if patients have early interventions [40]. Similarly to bipolar disorder, other mental health disorders are diagnosed and monitored through interviews and parent- and self-reports [31], [41], [42]. In attempts to minimize the bias involved in diagnosing children with these disorders, wearable sensors have been used to capture motion and voice responses during mood induction tasks (i.e., tasks developed to induce specific moods in the subject) [11]–[14]. These studies measure Negative valance in children (i.e., a term used to describe subdomains of anxiety and fear) [31] by startling the children with a fake snake [11], [13], [14] or making the children give a presentation in which they will be judged on how interesting it is [12]. While the tasks may seem intense, these studies show promising results for identifying children with anxiety and depression.

Overall, wearable sensors have the ability to capture relevant biological markers and responses outside of a laboratory setting. They have fast processing times and little to no human bias in them. This gives healthcare providers more access to important

information which allows for better point-of-care diagnostics and better treatment regimens.

1.3. Wearables in Athletic Performance

Recently, researchers have begun to evaluate the performance of wearable sensors during intense physical activities [43]–[46]. HR training has shown promising results in building endurance and overall aerobic fitness [6], [47]. Current methods for estimating HR during intense physical activities use PPG and accelerometry [48]–[51]. The studies also used clip- and belt-worn sensors which limit the location for wear. These studies are limited to collected PPG data from the wrist [49], [50], [52]–[55], finger [56], ear and forehead [57], and hip [51]. With these limited locations for wear, and having the sensor be attached by a clip or band, these sensor types may not be feasible for use in real-time practice or games where there is contact, but they could still be a useful tool for individual training.

With the ability to measure HR, HRV, motion, and forces from small wearable sensors, coaches, athletes, and athletic trainers could personalize workouts and trainings for specific needs and wants. Now, biofluidic skin sensors have the ability to analyze sweat composition, giving information about dehydration [58] which can benefit the athletes from cramping and losing energy during games or practices. Not only does this allow for optimal training methods, it can also decrease the risk of an injury for the athlete.

1.4. Gap Analysis

Wearables have been used to monitor and diagnose patients in lab and remotely, as well as track fitness and safety measures for athletes. These applications focus on common and pressing disease and disorders like MS, PD, HD [20]–[23], [27], COVID-19 [59]–[62], and heart disease [37], [62]. In athletics, most focus is on sensors for monitoring fitness activities in a gym [47], [63], [64] or using sensors to understand the impacts athletes encounter during game play and practice [65]–[68].

1.4.1. Childhood Mental Health

Some research has begun to use wearable sensors to monitor and diagnose mental health disorders like bipolar disorder [28], [29], [69], and anxiety and depression [11]–[14]. While the studies for identifying internalizing disorders in children have promising results, they focus on measuring reactions during task designed to induce fear and anxiety in the children [11], [11]–[14]. While these studies have used wearables to identify children with internalizing disorder, no one has looked at using mood induction tasks meant to induce positivity and feelings of joy in children for identification.

1.4.2. Athletic Performance

Wearables are also slowly starting to make their way into professional athletics, but accurate measures need to be made with these sensors. In attempts to collect and clean data during sports, researchers have looked at motion artifact cancellation algorithms to clean PPG signals [49]–[51], [55], [56], [70]. These studies use wearables to estimate HR during

intense physical activities through motion artifact cancellation methods, but new locations for sensor wear to estimate HR from has not been explored.

1.5. Objectives

This thesis will look into fulfilling the two gaps stated above through two different studies. First, the use of wearable sensors paired with machine learning and a positive mood induction task to aid in the diagnosis of internalizing disorders (i.e., anxiety and depression collectively) in children, and second, to determine new locations for accurate HR estimation using a wearable PPG and IMU sensor during intense physical activity. These objectives can be further specified by the following:

1. Children with and without internalizing disorders will wear IMUs secured by elastic bands during a positive valance mood induction task. Features will be extracted from the data collected to build a decision tree to classify the subjects into those with a disorder and those without. Statistical analyses will be done to determine the specificity, sensitivity, and accuracy of the model. The model performance will be compared to that of the Child Behavioral Checklist (CBCL), which is commonly used in practice [71].
2. New wearable PPG plus IMU sensors will be worn in multiple body location during a series of physical activities. Two ECG and one additional PPG sensor will also be worn to validate three HR estimation algorithms. The three algorithms will estimate HR and be compared to the Polar H10 chest strap, which is a commercial wearable ECG that is the current gold standard [72]. After validation, the optimal location for PPG collection will be determined.

These two studies will be discussed further in Chapters 2 and 3 and have the potential to advance wearable-enabled point-of-care diagnostics and remote health and performance monitoring. Both studies utilize new technologies for identifying subjective (internalizing disorders) and objective (heart rate) measures in humans.

1.6. Thesis Outline

The remainder of this thesis describes the two studies mentioned above. Chapter 2 discusses the use of wearable sensors paired with machine learning to identify children with internalizing disorders. The subsections of Chapter 2 discuss the background of childhood internalizing disorders, how they are diagnosed in practice and research, previous works, the methods of the conducted study, results and discussion, and conclusions and future works.

Chapter 3 follows a similar structure to Chapter 2 but discusses the use of wearable sensors and motion artifact cancellation algorithms to determine HR from PPG and accelerometer data at various new body locations. Again, a background on HR estimation is described, followed by how wearable sensors are used in practice and research for estimating HR for medical and athletic purposes. Next, the methods of the study are described, followed by the results and a discussion of them, and a conclusion with future work. Finally, the thesis is concluded in Chapter 4 with an overview of the works discussed and concluding remarks made about the research process.

CHAPTER 2: DIAGNOSIS OF INTERNALIZING DISORDERS IN CHILDREN USING WEARABLE SENSORS

2.1. Introduction

2.1.1. Background

Internalizing disorders (i.e., anxiety and depression) can develop early in life. Approximately 10.3% of children ages 3-17 have been diagnosed with anxiety and or depression in the U.S. [40]. However, approximately 20% of children will experience a depressive or anxiety disorder [73]. Of those diagnosed with depression, about three in four of them will also have anxiety, and one in two will have behavioral problems [40]. Children with anxiety are also likely to have behavioral problems and/or depression (approximately one in three for both) [40].

If left untreated, these disorders can lead to severe health problems [73] including chronic psychopathology, substance abuse [74], increased risk for suicide [75], and functional impairment [76]. Suicide is a potentially preventable issue, but it is the second leading cause of death for those between 10-34 years old [77]. It is imperative that these disorders are identified at an early age so that interventions can be applied when brain plasticity, and thus the likelihood of successes, is greatest.

2.1.2. Internalizing Diagnostics in Practice

Diagnosis for these disorders in children is usually conducted by clinicians who analyze multi-informant reports from the child them self, their parent or guardian, and any teachers the child may have [78]. There are incredibly long wait lists and periods for these services (approximately 96 days once referred [71]), and only the most affected

children (e.g., those with behavioral problems) are ever referred to see a psychiatrist/psychologist for assessment.

The multi-informant reports used have shown limitations. Since the reports are based on parental assessment of the child, they are often inaccurate [41], and when the child is at a young age (under 8 years old), their response can be unreliable [79]. With these limitations, it is hard to accurately diagnose children, and in turn, to give them proper treatment. To overcome this, research is being conducted using mood induction tasks to understand how children with a disorder may react differently to specific stimuli compared to children without a disorder [13], [73], [80].

2.1.3. Internalizing Diagnostics in Research

The National Institute of Mental Health has developed the Research Domain Criteria (RDoC) for examination of mental health. While the RDoC is a good framework, its use in psychological research has been limited because there was no way to assess the domains in children. In 2017, Waxler developed a behavioral battery assessment to measure Positive and Negative valance domains of the RDoC in preschool children. Her work focused on using 'Mood Induction Tasks' adapted from the LAB-TAB task and Trier-Social Stress Task for Children (TSST-C). These tasks were behaviorally coded by trained coders who had to achieve an intra-correlation coefficient of 0.7 or higher with a trainer before coding independently. This took a lot of time to achieve and there were multiple training steps to accomplish. Even with the trained coding, coders were not reliably able to diagnose children with internalizing disorders [31].

Other than the work from Waxler [31], studies have used questionnaires and fMRI scans in researching internalizing disorders and other mental health disorders [81]–[87]. All of these studies used machine learning to diagnose children with an array of disorders from internalizing and externalizing disorders diagnosable by the CBCL [83], [84], [87], post-traumatic stress disorder (PTSD) [88], Autism [86], and schizophrenia [85]. While these studies had promising results, most of them used fMRI scans as the new measure for diagnosing these disorders [81]–[84], which are not a feasible tool for this kind of diagnostics. MRI scans are time consuming and expensive, and not all insurance companies cover the costs. The other studies use questionnaires to determine the diagnosis and these have similar problems to the CBCL and behavior coding because they involve parent-reports and have natural human bias in the decision [86]–[88]. While these methods may be feasible for research, they are not feasible to be used in practice or as a screening tool.

2.1.4. Wearable Sensors in Internalizing Diagnostics Research

The use of wearables has just recently been introduced into research for diagnosing internalizing disorders. Within the past three years, wearable IMU sensors have been paired with machine learning to create accurate models for determining children with and without internalizing diagnosis [11], [13], [14], [73]. Wearables have also been used in attempts to distinguish bipolar disorder from ADHD [69]. All of these studies use supervised machine learning methods on data collected from wearables or audio recordings to determine the diagnosis of each subject. These studies showed promising results with accuracies around

80% suggesting that wearables and machine learning could be used to develop and effective tool for detecting internalizing disorders.

2.1.5. Gap Analysis

While previous works have used wearable sensors to help diagnose internalizing disorders in children, these studies have focused on tasks that measure Negative valance in the children [11]–[14]. Negative valance is the phenotypic expression of negativity and adverse motivation with five subdomains including anxiety and fear [31]. In these studies, the subjects perform the 'Snake Task' and 'Speech Task' during wearable sensor instrumented mood induction tasks [11]–[14], [73]. The Snake Task involves bringing the child into a dimly lit room, startling them with a fake snake, and then allowing them to touch the snake to ensure they know it's fake. Afterwards, the child is allowed free play time and a debriefing session [11], [13], [14]. The Speech Task involves having the child prepare a three-minute speech in which they will be judged on how interesting it is. They are given three-minutes to prepare and are interrupted with a buzzer twice during the speech to tell them the remaining time [12]. While these studies have had promising results, they focus on tasks that are intended to induce anxiety and fear into the children [11]–[14].

The study discussed in this chapter presents a method to identify internalizing disorders using a task to measure Positive valance, the 'Bubbles Task.' While this task does involve additional equipment to the IMUs, it is only an inexpensive bubble machine and bubble solution, not a terrarium, fake snake, or video camera like the previous studies [11]–[14]. Since the task measures Positive valence, it is intended to induce positive feelings like joy and reward, not negative feelings like fear and anxiety [31].

2.1.6. Objectives

The objectives of this study are first, to determine wearable sensor derived measures that best correlate with childhood internalizing symptoms, and second, to establish the accuracy of a machine-learning based method for classifying internalizing diagnosis in children.

2.2. Methods

2.2.1. Participants and Recruitment

Data were collected from a total of 63 (57% female) children and their primary caregivers (95.2% mothers) in a clinical research space at the University of Michigan. To be eligible for the study, participants had to speak fluent English and be between the ages of three and eight with a primary caregiver over the age of 18. Exclusion criteria included children suspected or diagnosed with a developmental disorder, having a serious medical condition, or taking medications that affect the central nervous system. Participants for this study were recruited from an ongoing study (Bonding Between Mothers and Children, PI: Maria Muzik) or from flyers posted around the community and at the University of Michigan Depression Center. A subset of this dataset, specifically the 'Bubbles Task', was used to complete the following study.

2.2.2. Data Collection

Subjects and their primary caregivers were brought into the university-based laboratory at the University of Michigan and provided written consent to complete an array of tasks. Studies were approved by the University of Michigan Institutional Review Board

(HUM00091788; HUM00033838). The children were equipped with two belt-worn IMUs (3-Space Sensor, YEI Technology, Portsmouth, OH, USA) – one secured around the head and the other around the sacrum (see Figure 1). These sensors were used to collect acceleration and angular velocity during the array of mood induction tasks.

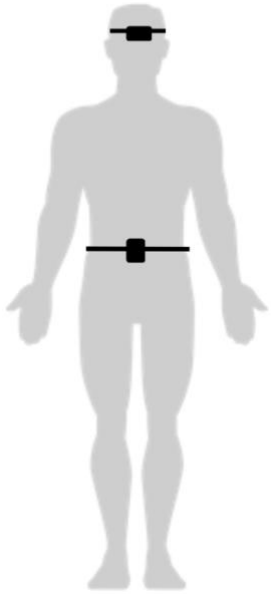

Figure 1. IMU locations on the subject for each task.

The task this study focuses on, the 'Bubbles Task', is a mood induction task intended to induce a positive feeling and emotions of joy. The task is designed to measure Positive valance, which is a term to describe phenotypic expressions of positivity and approach motivation [31]. The duration of this task was approximately 180 seconds. During the task, administers led the subjects into a room with a bubble machine on a table. The bubble machine was used to limit the negative emotions that could come from personally blowing bubbles from a wand, since the child may not be able to do it themselves. With the machine on, the administer gave scripted statements to encourage

positive emotions and behavior such as, “Try to pop the bubbles,” and, “Look at how fun this is.”

2.2.3. Clinical Measures

While at the lab, the children and caregivers also completed the Child Behavior Checklist (CBCL) questionnaire and a clinical interview. The CBCL is parent-completed and designed to assess child problem behavior in both clinical and research settings [31], [89], [90]. The CBCL consists of 120 items relating to child behavior across multiple domains and the frequency each item occurs. The CBCL has well established validity and reliability (see [90]). It takes about 15-minutes to complete and results in a global T-score relating to internalizing, externalizing, and total problems. The scale also results in scores for disorder-based subscales including Anxiety/Depression, Attention Deficit/Hyperactivity, and Oppositional Defiant Disorder. Only scales available in both versions (ages 1.5-5 and 6-18) were used in subsequent analyses. Subject demographic information including race, gender, and family income was also collected.

The diagnostic interview was conducted by trained clinical psychology doctoral students or postdoctoral fellows and lasted about 1-2 hours with each child’s caregiver. A version of The Schedule for Affective Disorders and Schizophrenia for School Aged Children and Lifetime Version (K-SADS-PL) which is validated for ages 6-18 and modified for preschoolers was used in this study [91]. This interview explored past and current symptoms of the child’s psychiatric disorders. Interviewers were supervised on at least a monthly basis by a licensed psychologist or psychiatrist and all cases were reviewed by all clinicians and their supervisor. The final diagnoses were a result of clinical consensus

using the best estimate procedures [92] based on the child and parent report, family history, and other self-report symptom checklists. The K-SADS is considered the gold-standard but is not commonly used in the 1-hour unstructured interviews conducted in practice.

2.2.4. Signal Processing and Feature Extraction

During the mood induction task, the children's motions were measured by acceleration (m/s^2) and angular velocity (radians/s) using two belt-worn IMUs. The data were sampled at approximately 300 Hz, down-sampled to 100 Hz, and low-pass filtered using a fourth-order Butterworth IIR filter (cutoff of 20 Hz) in software prior to use.

Of the N=63 subjects, N=48 (60.4% female) had useable headband data during the Bubbles Task. Data were unusable if the child removed or turned off the device during the activity, or if data were not labeled properly once downloaded. While the task was intended to be 180 seconds, only 150 seconds of the data were used to maximize the number of subjects with usable data.

IMU data from the head were separated into two phases during the task, the Initial (first 30 seconds) and Sustained (130-150 seconds) phase (see Figure 2), and signal features were extracted from the vector magnitude of acceleration and angular velocity.

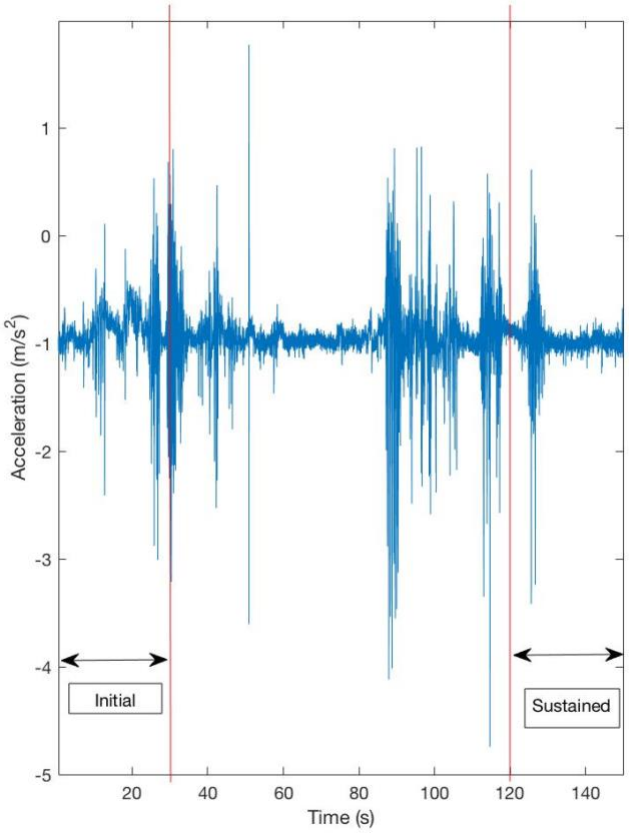

Figure 2. Raw acceleration from a subject's headband sensor. Orange lines show Initial and Sustained phase cutoffs.

Signal features included mean, root mean square (RMS), skewness, kurtosis, range, maximum, minimum, standard deviation, peak to RMS amplitude, signal power within frequency bands, and the location and height of peaks in the power spectrum and autocorrelation of the signal. This yielded 29 features for each signal (i.e., acceleration and angular velocity), resulting in a total of 58 features from each of the two phases. In addition to the two phases, the data were classified into segments of motion (when the child was actively playing with bubbles) and no motion (when the child stopped for a period of time). Child motion was determined by creating an envelope of the data through filters, full-wave

rectification and applying a threshold the amplitude of data needed to exceed to be considered motion.

Next, the start and stop timestamps of each motion segment were extracted for further feature extraction. Features extracted from these motion segments included total duration of motion, average duration of motion segments, standard deviation, skewness and kurtosis of motion segments, maximum and minimum of motion segments, and median of the motion segments. This yielded an additional eight features for each signal, resulting in a total of 16 features overall. With both sets of features (Initial/Sustained and overall motion), a total of 74 features were available to characterize each subject's motion. These features were used in a supervised learning environment to determine each subject's internalizing diagnosis.

2.2.5. Statistical Model for Identifying Internalizing Diagnosis

Supervised learning was used to train classification models relating IMU-derived features to the internalizing diagnosis as established via the K-SADS-PL with clinical consensus. Performance of the classifiers were established using a leave-one-subject-out (LOSO) cross-validation. This process partitions features from all but one subject (i.e., 47 of 48 subjects) into a training set before converting to z-scores and performing a Davies-Bouldin Index based feature selection to yield ten features with zero mean and unit variance that best discriminate between diagnostic groups. The score threshold used to determine diagnosis of each iteration was Youden's Index [93] based on the ROC curve of the training data.

The selected features (from the Initial/Sustained phases) were then used to train a decision tree classifier for predicting if the one remaining test subject has an internalizing diagnosis or not. This process was repeated until the diagnosis of each subject had been predicted. A visualization of this process can be seen in Figure 3.

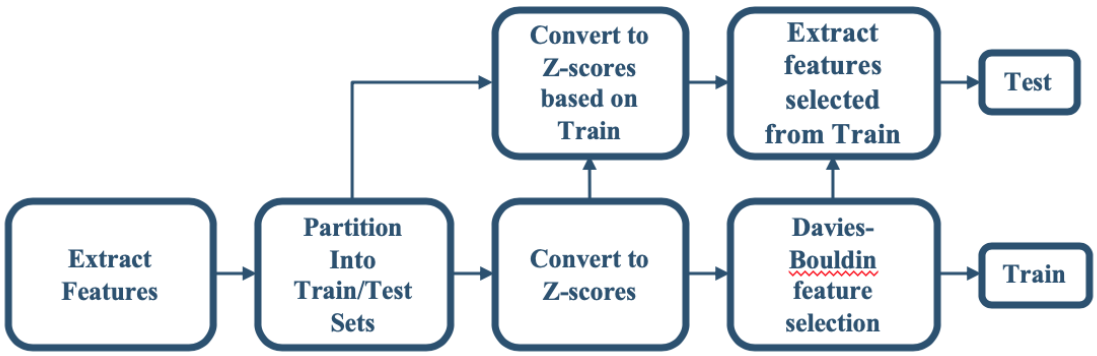

Figure 3. Flow chart of the process for training and testing the decision tree algorithm.

This was again repeated for the Initial/Sustained phases with the additional motion segment features added, but the Davies-Bouldin Index yielded 13 features.

The model performances were then assessed in several ways. The accuracy, sensitivity, and specificity were calculated for both the Initial and Sustained models and for the CBCL measures. The receiver operating characteristic (ROC) curve, which plots true positive rate against false positive rate, was plotted and the area under the ROC curve (AUC) was used to assess the discriminative ability of the classifiers [94]. Correlations between CBCL measures including Internalizing, Depression, and Anxiety Problems, as well as somatic symptoms (e.g., extreme focusing on physical pain [95]) because they are

common in children with internalizing disorders, and selected model features were also calculated.

A permutation test was also done to test the models' classification error rates (error rate = number of incorrect predictions/total number of predictions = 1 - classification accuracy) against random chance. To complete this test, the distribution of error rates for each model was approximated as a beta distribution parameterized by the number of incorrect predictions and total number of observations, as indicated in [96], [97], and randomly sampled 100 possible error rates from this distribution. Next, the model training procedure previously described was again repeated for 100 random permutations of the diagnostic labels and computed the classification error rate each time. Finally, a Mann-Whitney U-Test was used to determine if the error rates of the classification models were statistically different from random chance.

2.3. Results and Discussion

2.3.1. Results

Figure 4 shows the results from the permutation test for the decision tree models developed using features from the Initial+Motion and Sustained+Motion phases of the Bubbles Task. The error rates for the models developed are shown in blue and those due to random chance are in gray. The Initial+Motion model did not have statistically better error rates than random chance; however, the Sustained+Motion model did (p-value < 0.001).

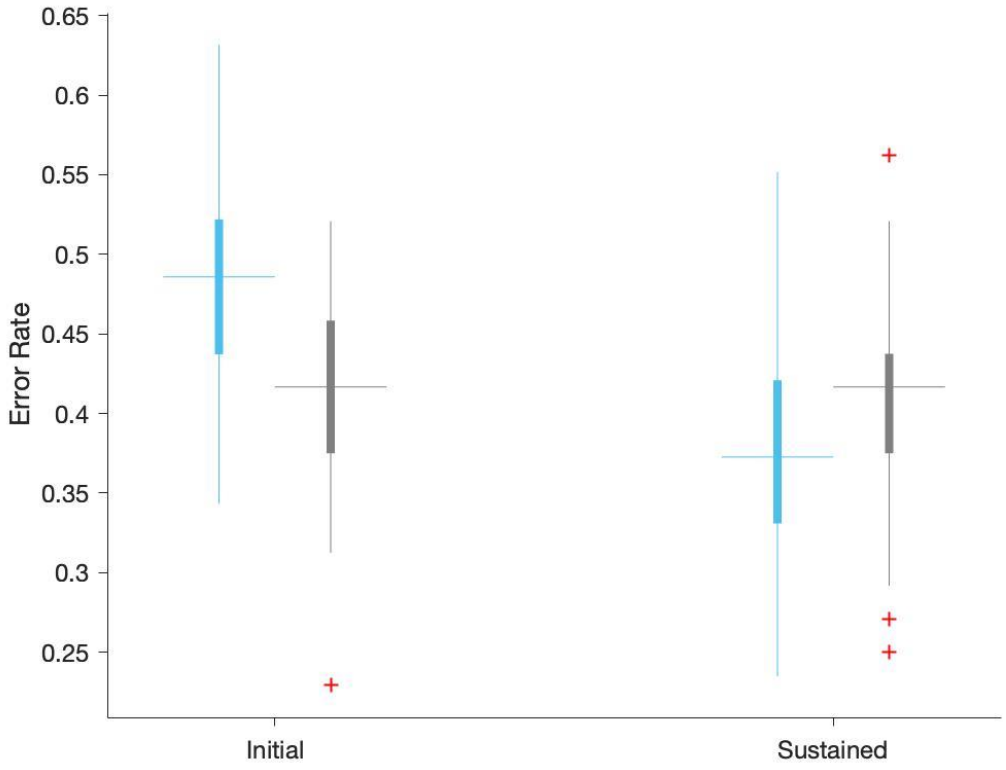

Figure 4. Error rates from the developed decision tree models (blue) compared to those due to random chance (gray).

Although the Initial+Motion model did not have better error rates than random chance, it was still compared to the Sustained+Motion model, which was the best performing model developed. The receiver operating curves (ROC) for the Initial+Motion and Sustained+Motion models are shown in Figure 5 reports the true and false positive rates. The AUC for the Initial+Motion and Sustained+Motion models were 0.39 and 0.81, respectively.

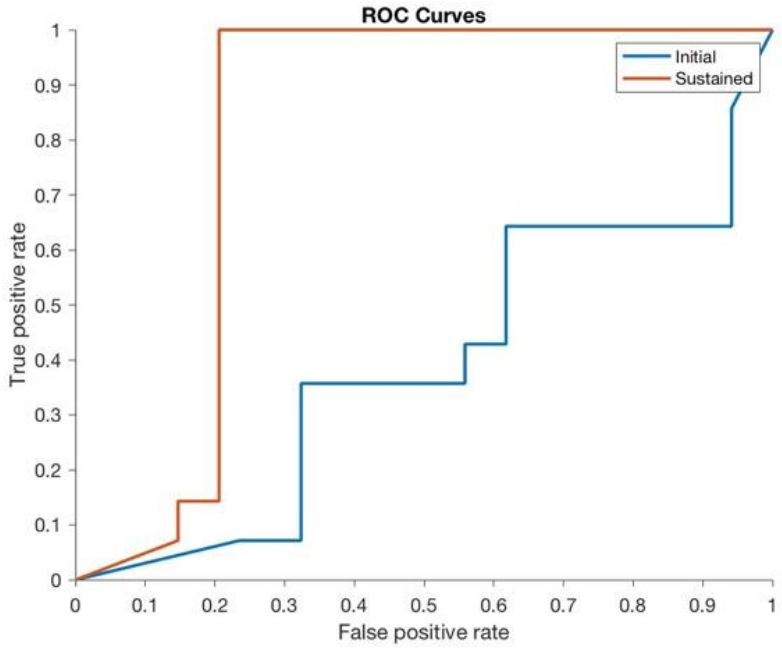

Figure 5. ROC curves for the decision tree models using Initial/Sustained features and motion segment features.

To better visualize the data values of the chosen features, violin plots were created for both the subjects with and without a diagnosis (see Figure 6). These plots compare the z-scores of each diagnosis for each selected feature of the Sustained+Motion model. The feature values for those without a diagnosis are represented in the dark blue violin plots and the feature values for those with a diagnosis are in the light blue violin plots.

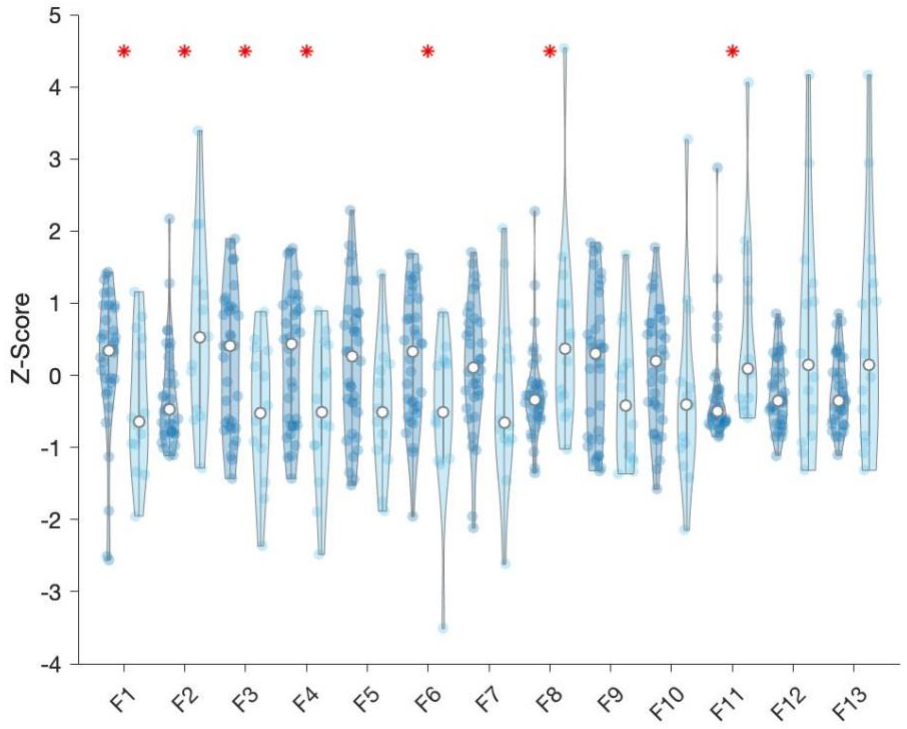

Figure 6. Violin plots of the selected features z-scores for both subjects without an internalizing disorder (dark blue) and with an internalizing disorder (light blue) from the Sustained+Motion model. The plots with a red asterisk above them were determined, from a Mann-Whitney U-Test, to be statistically significantly different between diagnostic groups.

The features selected for the Sustained+Motion model (see Table 1) can be formed into three categories: Power, Intensity, and Frequency.

Table 1. Sustained+Motion model selected features with respective category. Features with an asterisk note those statistically significantly different between diagnostic groups.

Feature Label	Feature	Category
F1*	Mean (Acc.)	Intensity
F2*	Skewness (Acc.)	Intensity
F3*	Covariance of height at 0 (Acc.)	Intensity
F4*	Covariance of height at 1st peak (Acc.)	Power
F5	Spectral frequency location of 3rd highest peak (Acc.)	Frequency
F6*	Total time in motion (Acc.)	Intensity
F7	Skewness of motion segments (Acc.)	Intensity
F8*	Peak-to-RMS (Ang. Vel.)	Intensity
F9	Spectral frequency location of 2nd highest peak (Acc.)	Frequency
F10	Spectral frequency location of 5th highest peak (Acc.)	Frequency
F11*	Kurtosis (Acc.)	Frequency
F12	Maximum (Ang. Vel.)	Intensity
F13	Range (Ang. Vel.)	Intensity

The accuracy, sensitivity, specificity and AUC for the CBCL measures at two thresholds can be seen in

Table 2. These values can be compared to the accuracy, sensitivity, specificity and AUC from the models developed (Table 3).

Table 2. CBCL measures for accuracy, sensitivity, specificity and AUC.

CBCL	Accuracy		Sensitivity		Specificity		AUC
Cutoff	70	55	70	55	70	55	70, 55
Internalizing	0.70	0.72	0.00	0.29	1.00	0.91	0.83
Anxiety Problems	0.77	0.70	0.21	0.43	1.00	0.82	0.80
Depressive Problems	0.77	0.73	0.21	0.43	1.00	0.85	0.78

Table 3. Developed models' accuracy, sensitivity, specificity and AUC.

	Accuracy	Sensitivity	Specificity	AUC
Initial	0.52	0.36	0.59	0.39
Sustained	0.75	0.64	0.79	0.81

The correlations computed between selected features and the CBCL measures (Internalizing Problems, Depression Problems, Anxiety Problems, and Somatic Symptoms) with significant correlations can be seen in Table 4.

Table 4. Sustained+Motion selected features with significant correlations to CBCL measures.

Feature	CBCL Measure	P-value
Acceleration skewness	Depression Problems	0.037
Acceleration kurtosis	Depression Problems	0.0097
Acceleration skewness	Somatic Symptoms	0.043
Acceleration covariance of height at 0	Somatic Symptoms	0.03
Acceleration kurtosis	Somatic Symptoms	0.037
Acceleration spectral frequency location of 2nd highest peak	Somatic Symptoms	0.013
Acceleration spectral frequency location of 3rd highest peak	Somatic Symptoms	0.019

2.3.2. Discussion

With the need for a more accessible and reliable screening tool for psychopathology in young children, the results presented above suggest that methods similar to these could fulfill that need. In comparison to the questionnaire-based parent-reported CBCL the wearable IMU derived Sustained+Motion model had similar accuracy (0.75 vs. 0.70-77), higher sensitivity (0.64 vs. 0.00-0.43), and slightly lower specificity (0.79 vs. 0.82-1.00). The AUC was relatively similar between the model and CBCL (0.81 vs. 0.78-0.83). The model developed has similar performance to the CBCL but uses a single wearable sensor, a bubble machine, and takes three minutes to collect the data. That is less time than it usually takes for a parent or caregiver to complete the CBCL and it is more sensitive.

Figure 6 shows clear differences in the mean values of the diagnostic groups from over half of the selected features, and the permutation test (Figure 4) showed the model was statistically better than what would be expected by random chance (p-value < 0.001 determined by Mann-Whitney U-Test). Also, some of the selected features are correlated with the subdomains of the CBCL including depression problems and somatic symptoms. In most of the features (8 out of 13), the subjects with an internalizing disorder had lower z-scores which may suggest for the correlations to the CBCL's depression problems since it has been suggested that people in a depressive state may move less than those who are not [29]. Differences in sustained response to positive stimuli between individuals with and without internalizing disorders has been demonstrated previously in older subjects, providing support for these results [80].

Other machine learning models developed for mental health diagnosis have also had similar results [12]–[14], [73]. These studies saw similar accuracies (75-81%) to that presented in this study. The results from studies suggest that wearables and machine learning provide similar results to the CBCL and could be used as an easy, accessible, and quick way to screen for mental health disorders in children.

While the results presented are promising, this study is not without limitations. One limitation to this study is the small sample size of subjects. This work should be further validated and reproduced with data collected from a larger sample size and more balanced diagnoses (this set only had 14 subjects with a diagnosis). A broader sample set would more accurately represent the population and those at risk for developing internalizing disorders. It could also allow for the examination of which disorder type, if any, has more of a response to positive valance activities compared to the other.

2.4. Conclusion and Future Work

2.4.1. Conclusion

The results presented suggest that wearable sensor data capturing a child's motion for 150 seconds while playing with bubbles can be used to identify young children with an internalizing disorder with high sensitivity. If implemented into an easy-to-use app, this model could be used as an inexpensive screening tool in pediatric offices, after the initial purchase of an IMU. Being inexpensive (after the initial cost of the IMUs) is important since almost 90% of children live in low-income and middle-income countries [42]. This could change the stigma around mental health and increase the accessibility of getting help early on and when appropriate.

2.4.2. Future Work

As mentioned above, future work should focus on correcting the limitations to this study, which includes a larger, more balance sample. This work would increase the validity of the results presented. It would also give a better representation to the population at risk and if one disorder is more identifiable by the Bubbles Task.

The next chapter of this thesis will discuss another application of wearable sensors in providing more relevant biological signals for health monitoring, including HR. With the additional measure of HR, even more information could be collected to help diagnose children with internalizing disorders, especially since they are disorders that typically reflect inward and have a sympathetic nervous system response [98], [99]. However, it has proven to be a challenge to estimate accurate HRs during intense physical activities due to motion artifacts [49].

The remainder of this thesis will discuss a method to remove motion artifacts from PPG signals. It also explores new body locations for measurement, which may be more comfortable for longer wear, and do not rely on a belt or clip to be adhered close enough to the skin's surface for good signal quality.

CHAPTER 3: HEART RATE ESTIMATION USING WEARABLE PPG SENSORS

3.1. Introduction

3.1.1. Background

While the movement-based measurements of Chapter 2 capture important information during the Bubbles Task, they do not reveal the child's physiological response. Changes in HR [100] and heart rate variability (HRV) have previously been linked to anxiety and depression during social-evaluation tasks [99]. For example, low resting high frequency HRV has been associated with symptoms of anxiety and depression [99].

Measures of HR and HRV are also important vital signs that can be used to evaluate fitness, lung health, and heart health more broadly [6]–[10]. These measures are often captured using PPG [101], which is typically worn at the wrist [49], [50], [53]–[55], and subject to substantial movement artifact. Movement artifacts are corruptions in the PPG signals due to activities like brushing teeth, washing dishes, intense physical activities and other activities people may participate in throughout the day.

To obtain an accurate measurement of HR and HRV, motion artifacts are required to be removed from the signals. Previous works conducted to minimize the effect of motion artifact on PPG signals leverage accelerometer and/or gyroscope data with the idea that those signals can be used to identify the motion artifacts in the PPG signal, so that it can be removed [7], [49]–[51], [53]–[55], [57], [102]. However, these studies considered data from sensors secured to the wrist [49], [50], [53]–[55], ear and forehead [57], hip [51], and finger [56].

3.1.2. Wearable Sensors for Measuring Athletic Performance

Currently, there are off-the-shelf wearable sensors for fitness tracking available in stores (i.e., Apple Watch, Fitbit, Polar, etc.). These technologies use PPG or ECG Sensors to obtain HR during physical activity. These are typically used for personal use, but some gyms like Orange Theory Fitness have adopted wearable sensors for their members to use in classes to optimize their workouts [63]. Other companies, like Myzone, are also using HR sensors to help members maximize their workouts [64]. These companies use research that shows HR intensity-based group fitness significantly alters aerobic fitness, body composition and overall health if followed [47]. The wearable sensors help their members hit target HR zones to maximize their workouts since they can see their HR in real-time.

Although these sensing modalities are commercially used, it is difficult to get accurate PPG measurements during intense physical activity due to soft tissue artifacts and discontinuous contact between the sensor and skin [52]. In attempts to improve the accuracy and reliability of PPG sensors, research is being done to reduce the effects of these artifacts [51], [52], [55], [56], [102]. A lot of research focus has also been on sensors to measure stress and strain on the body during training and performance and sweat or other biofluid composition during workouts [43], [45], [103].

3.1.4. Previous Work

Previous work in developing robust motion artifact cancellation has focused on clip or band worn sensors [51], [55], [56], [102], [102]. These sensors limit the locations of wear on the body and can have a poor adherence to the skin, making it harder to get a clear

PPG reading. Many of these studies also use acceleration or gyroscope data to remove corrupted frequencies from the PPG signals [49], [56], [102]. The ability to have all of the hardware in one small sensor would make the research more feasible and wear more comfortable. Motion artifact cancellation algorithms have been developed and work for band and clip type sensors, but to my knowledge, there has not been research done on validating them with patch-like sensors that have the ability to be worn anywhere on the body.

3.1.5. Hardware Overview

Epicore Biosystems, Inc. has recently developed a wearable band aid-like sensor equipped with an IMU and PPG sensor called the AIM System. The sensor is made out of a low durometer silicone and can be adhered anywhere on the skin using adhesives. Within the silicone device there are three sensors: an accelerometer, gyroscope, and optical with specification described in Table 5.

Table 5. Sensor specifications embedded in AIM System device.

Sensor	Range	Resolution	Sampling
Accelerometer	±2, ±4, ±8, ±16g	16-bit per axis	50-500 Hz
Gyroscope	±250, ±500, ±1000, ±2000°/s	16-bit per axis	50-500 Hz
Optical	528-nm (Green LED)	19-bits	100-1000 Hz

The AIM system has a 1 G-Bit (128 MB) memory capacity and Bluetooth 5 communication compatibility. It is powered by a lithium polymer, rechargeable battery (10 mAh) and can go from 10% to 90% battery level in 1-hour of charging. Epicore Biosystems designed all of this to fit into one small, flexible device which can be seen in Figure 7.

Figure 7. Epicore Biosystem, Inc.'s AIM System wearable sensor with dimensions.

3.1.6. Objectives

With the new hardware developed by Epicore Biosystems, Inc., the objectives of this study are first, to process and validate HR estimation algorithms using PPG during intense physical activity, and second, to determine new locations for extracting HR from PPG sensors on the body. These can be further specified by the following:

1. To design and execute a protocol that includes intense physical activated to induce motion artifacts into the PPG data. With this corrupted data, three HR estimation approaches will be applied to the data collected during each activity.

The estimated HRs will be validated using the Polar H10 chest strap, which is considered the gold standard hardware for HR tracking [72].

2. During each activity, PPG sensors will be worn at three new body locations: forearm, shank, and sacrum. The error from each location will be calculated and compared to determine new locations of wear on the body that still provide accurate results.

3.2. Methods

3.2.1. Subjects

Data were collected from a total of 8 healthy adults in a research laboratory setting at the University of Vermont. To be eligible for the study, participants had to have no known cardiac arrhythmias, musculoskeletal injuries, and be between the ages of 18 and 29. Subjects were recruited through the M-Sense Research Group at the University of Vermont.

3.2.2. Data Collection

Participants were brought into the university-based laboratory where they filled out a written consent form to participate in an array of tasks. Activities were approved by the local Institutional Review Board. Each participant was equipped with a 3-lead ECG (Biostamp, MC10, Lexington, MA, USA) in lead III configuration placed on the chest, an ECG chest strap (Polar H10, Polar Electro Inc., Bethpage, NY, USA), a PPG sensor (Apple Watch, Apple Inc., Cupertino, CA, USA) on the right wrist, and 3 PPG+IMU sensors (AIM System, Epicore Biosystems Inc., Cambridge, MA, USA) placed on the inner right

forearm, right tibial plateau, and sacrum (see Figure 8). These sensors were used to collect an array of ECG, PPG and accelerometer recordings during the tasks.

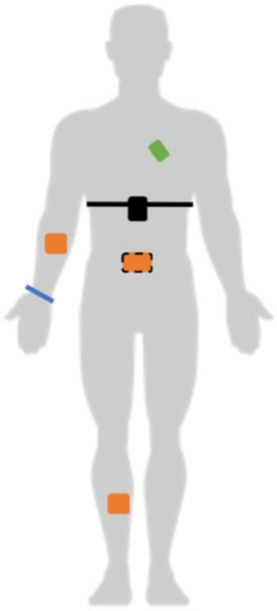

Figure 8. Sensors each subject was equipped with and their locations (orange-Aim System, blue-Apple Watch, green-Biostamp, black-Polar H10).

After being equipped with the sensors, the subjects were instructed to complete the following array of tasks: 1-minute standing calibration, five air squats, ten calf raises, ten push-ups (knee push-ups were acceptable), 30 seconds of jumping jacks, three 1-minute trials of walking on a treadmill (each trial at a different self-selected pace of slow, comfortable, and fast), three 1-minute trials of running on a treadmill (each trial at a different self-selected pace of slow, comfortable, and fast), and 1.5-minutes of sitting. Following the data collection, sensors were removed from the subject, cleaned, and data were downloaded for processing.

3.2.3. Signal Processing

Preprocessing

Of the eight subjects, only six (50% female) had useable data. Data were unusable if it had not successfully recorded and saved on its device (i.e., Polar H10, AIM System, or Biostamp). Data were sampled at approximately 250 Hz for the Biostamp, 100 Hz for the Epicore PPG and 250 Hz for the accelerometer. All data were processed using MATLAB 2018a.

Next, each subject's data from all sensors were divided into each activity. Start and stop timestamps were recorded using the MC10 Biostamp system. Since each sample from each measurement technology had an associated absolute timestamp, the start and end times from each activity recorded by the Biostamp system could be used to identify data corresponding to each trial. Once the data were segmented by activity, a Pan-Tompkins algorithm [104] was used to calculate the HR from the Biostamp 3-lead ECG sensor. The Polar H10 and Apple Watch output HRs, so no further processing was needed to be done.

The average HR was calculated for the Biostamp estimates, Polar H10 and Apple Watch for each subject and trial. The PPG signals from the AIM System were further processed using three HR estimate methods: Ensemble Empirical Mode Decomposition with Principal Component Analysis (EEMD-PCA) [105], Spectral Estimation and Median Filtering [48], and Weiner Filtering, Phase Vocoder and Viterbi Decoding (WFPH+VD) [50].

Method 1: EEMD-PCA Approach [105]

The EEMD-PCA approach was developed by Motin et al. in 2018. This method involves decomposing the raw PPG signal into intrinsic mode functions (IMFs) using EEMD. Next, the IMFs containing artifacts are automatically identified and rejected. The FFT of remaining IMFs are then used to calculate HR and respiratory rate (RR). Frequency ranges between 0.75 and 2.5 Hz are grouped for HR. Principle component analysis (PCA) was performed on the remained IMFs in the HR group. The principle components (PCs) were arranged so the first PC contained most of the variation present in the IMFs. The first PC was then used to determine HR.

Method 2: Spectral Estimation and Median Filter Approach [48]

Next, a motion artifact cancellation technique from the same authors of the EEMD-PCA approach was used. This approach involves a simple spectral analysis and median filter to improve the HR estimation. To further explain this technique, an 8 second window with 75% overlap is used across the PPG and three channels of acceleration; however, since the Polar H10 outputs a HR every second, an overlap over 87.5% was used instead. During each window, the signals are filtered using a fourth-order Butterworth filter and down-sampled to 25 Hz. Next, the FFT of each signal is taken. The dominant frequencies of the PPG are compared to the dominant frequencies of each channel of acceleration. Any dominant PPG frequencies within 1 Hz of the dominant acceleration frequencies was discarded. The remaining dominant frequency was used to calculate HR.

Each successive estimated HR was compared to that from the previous window. If the two HRs were too far apart (i.e., more than 30 BPM difference), the next dominant

uncorrupted peak from the PPG's FFT was used for the estimate. Also, the HR estimate could not be the same for more than two consecutive windows. If it was estimated for more than two, the next dominant uncorrupted peak was used for estimation. Once the HRs were estimated for each window, a median filter was used to smooth out the results.

Method 3: WFPV+VD Approach [50]

The final motion artifact method applied to the PPG data was developed by Temko in 2017. This author provided open-source code from their study [106]. This method involves using a Wiener filter, phase vocoder, and Viterbi decoding to estimate an accurate HR from a motion artifact corrupted signal. The method uses an 8 second window with a step size of 2 seconds; however, since the Polar H10 outputs a HR every second, a step size of 1 second was used.

During each window, the PPG and each channel of acceleration was filtered using a fourth-order Butterworth filter and down-sampled to 25 Hz. The Discrete Time Fourier Transform (DTFT) was also taken of each signal. Next, a Weiner filter was applied to the windows. According to Temko, the Weiner filter acts as a signal-to-noise attenuator since frequencies more likely to be corrupted by noise are given less importance when determining HR. The frequencies more likely to be corrupted by noise are found from the DTFT of the three channels of acceleration. Next, the estimated HR from the Weiner filter is refined using the Phase Vocoder technique which involves calculating the instantaneous frequency of the signal during each window. Finally, Viterbi decoding is used for final offline post-processing.

3.2.4. Validation

After estimating the HR for all subject, locations, and activities using the three methods mentioned above, the average absolute error was calculated between the estimated HRs and those obtained from the Polar H10 since it is considered the gold standard wearable ECG [72]. However, since the EEMD-PCA approach only provides a single HR estimate for each trial, the average HR was calculated for each trial from the other estimation methods and compared to the average Polar H10, Apple Watch, and Biostamp as well. The average absolute error was calculated between each sensor and trial and the Polar H10 data using the following equation:

$$\frac{1}{N} * \sum_{i=1}^{N} |truth(i) - estimation(i)|$$

Where N is the number of samples, truth is the Polar H10 data and estimation is the HR outputted from the Apple Watch, Biostamp or AIM system data after the artifact cancellation method.

3.3. Results and Discussion

3.3.1. Results

From the methods described above, the WFPD+VD [50] method outperformed both the EEMD-PCA [105] and Spectral Frequency with Median filtering approaches [48]. Figure 9, Figure 10, and Figure 11 show boxplots of the average absolute error at each location during each activity.

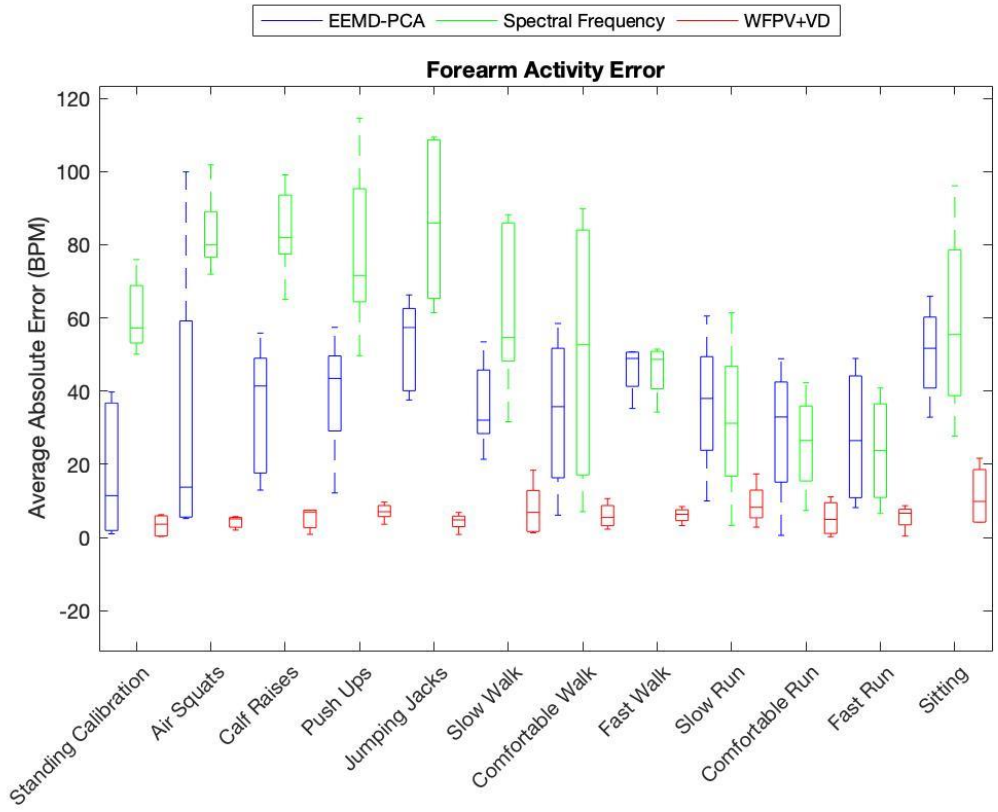

Figure 9. Average absolute forearm errors for the EEMD-PCA (blue), Spectral Frequency (green) and WFPD+VD (red) method during each trial.

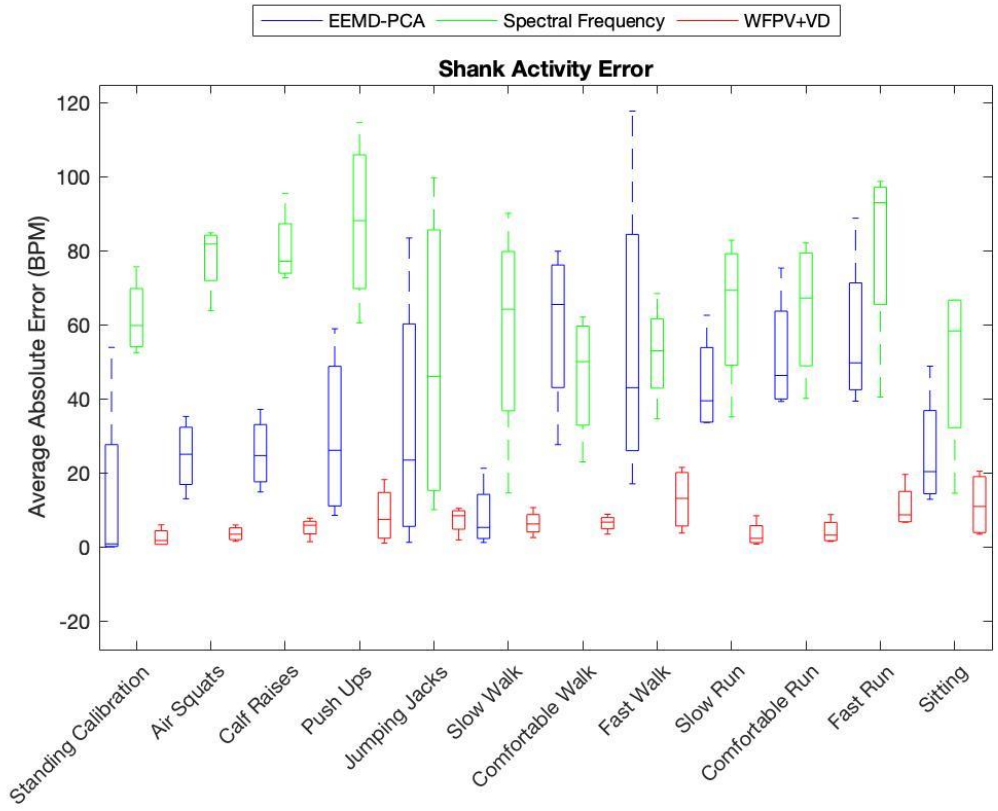

Figure 10. Average absolute shank errors for the EEMD-PCA (blue), Spectral Frequency (green) and WFPD+VD (red) method during each trial.

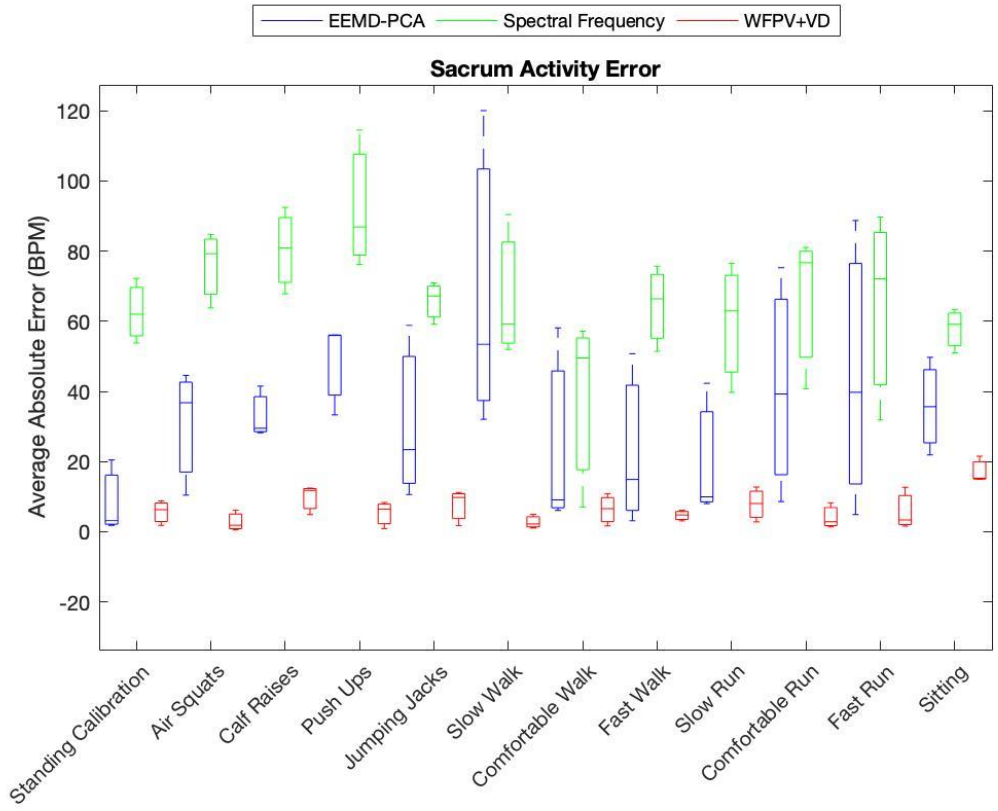

Figure 11. Average absolute sacrum errors for the EEMD-PCA (blue), Spectral Frequency (green) and WFPD+VD (red) method during each trial.

The average absolute error from the Pan-Tompkins algorithm used on the ECG data from the Biostamp sensor and the Apple Watch results can also be seen in boxplots below (Figure 12) with respect to the best performing location and method.

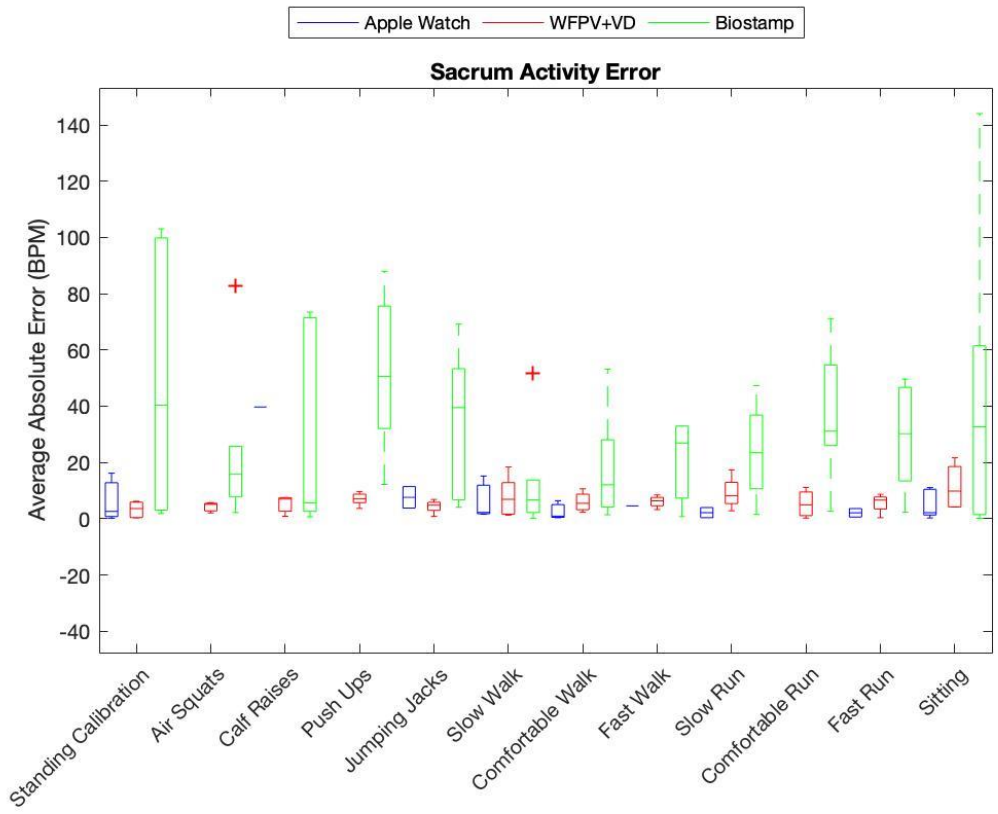

Figure 12. Average absolute errors for the Apple Watch (blue), WFPV+VD for the forearm (red), and Biostamp (green) sensors during each trial.

In total comparison, the average absolute error was calculated for each subject for each trial and location. It was then compared across all methods and forms of data collection modalities (i.e., the different types of sensors). These results can be seen in Table 6.

Table 6. Average Absolute Errors of each HR estimation method (1: EEMD-PCA, 2: Spectral Frequency, and 3: WFPV+VD) at each location for all sensors and trials.

Average Absolute Error Per Trial											
Activity	Forearm			Shank			Sacrum			Biostamp	Apple
	1	2	3	1	2	3	1	2	3		
Standing Calibration	9.6	60.8	3.3	14.3	61.9	2.5	7.6	62.7	5.6	48.1	6.3
Squats	5.5	83.4	4.2	5.9	78.0	3.6	3.3	76.0	2.8	25.0	NA
Calf raises	6.8	83.9	5.2	12.5	80.6	5.2	7.3	80.4	9.7	26.6	39.7
Push ups	6.2	78.8	7.0	8.9	87.8	8.5	4.7	92.6	5.2	51.5	NA
Jumping jacks	6.5	86.4	4.3	10.2	50.4	7.3	5.3	65.8	7.6	35.4	7.6
Slow walk	12.3	62.7	7.9	12.1	58.2	6.4	5.5	67.2	2.8	13.5	6.3
Comfortable walk	12.9	50.4	5.3	16.5	46.2	6.4	12.3	37.9	6.3	18.3	2.5
Fast walk	7.7	45.8	6.1	10.0	52.2	12.9	8.5	64.5	4.7	21.3	4.6
Slow run	13.9	31.8	9.2	8.5	64.1	3.5	10.4	59.7	7.9	23.9	2.2
Comfortable run	5.2	25.7	5.3	5.6	64.1	4.1	15.0	66.2	4.1	36.1	NA
Fast run	12.5	23.7	5.6	14.8	81.3	10.9	19.9	64.6	5.9	28.8	2.1
Sitting	28.0	58.7	11.4	18.6	49.4	11.4	11.9	57.8	17.2	45.4	4.5
Total	10.6	57.7	6.2	11.5	64.5	6.9	9.3	66.3	6.7	31.2	8.4

Figure 12 and Table 6 showed some trials, including Air Squats, Jumping Jacks, and Comfortable Running, where the Apple Watch did not collect data for any subjects. This was further investigated, and the number of missed trials by the Apple Watch can be seen in Table 7 below.

Table 7. Number of missed trials by the Apple Watch.

Activity	Number of missed trials
Standing Calibration	3
Squats	6
Calf raises	5
Push ups	6
Jumping jacks	4
Slow walk	3
Comfortable walk	3
Fast walk	5
Slow run	4
Comfortable run	6
Fast run	4
Sitting	0

The Polar H10 also missed 12 seconds worth of data during subject 7's Sitting Trial. The missed recordings corresponding estimates were discarded when calculating average HR during the trial.

3.3.2. Discussion

The results above suggest that the WFPD+VD approach [50] was the most robust for estimating HR during intense physical activity using the AIM System data. With the intense physical activities performed in this study, there were large motion artifacts. With the Spectral Frequency method [48], dominant PPG frequencies too close to dominant acceleration frequencies were discarded. It has been noted that during cyclical motions, like running and walking, that the frequency of steps can be similar to the frequency of HR, which may provide reasoning as to why it did not perform as expected [107]. These results

also suggest that the forearm is an ideal place for these new wearable PPG sensors to be worn, although the shank and sacrum also perform with similar errors.

This method also outperforms the Pan-Tompkins method [104] used on the Biostamp ECG data, and most of the Apple Watch HR estimates. The Pan-Tompkins method likely gave false reportings of HR due to the motion artifacts in the data since no cancellation algorithm was applied to this data. However, with the artifact cancellation, the PPG data is more reliable than the ECG data without a more advanced algorithm. Also, when importing the Apple Watch data, many trials had missing data (Table 7). No sampling rate for the Apple Watch PPG could be found other than "hundreds of samples a second," so interpolation could not be used to fill in missing data. With so many missing trials from the Apple Watch, especially in the activities that used more arm/wrist movements, it likely had trouble remaining in good contact with the subject's skin or the quality was not good enough for their algorithms to estimate a HR.

It is also surprising to see the Polar H10 have some missing data. This occurred during the sitting trial for subject 7. The sitting trial occurred last in the series of exercises and some subjects, including subject 1 who had to be discarded from the study since the Polar H10 did not collect any data for them, mentioned they noticed the Polar H10 device was sliding during some tasks. This may have been due to the use of the M-XXL band size used for this study, and Polar does sell a XS-S size as well.

The results of this study also show that the forearm, shank, and sacrum have relatively similar performance making all of them a new location to collect PPG data from. With this discovery, PPG and accelerometer data could be collected at the sacrum or shank to give not only HR information, but also information about gait biomechanics. This could

be beneficial for athletes and trainers who want to know more about technique, fitness, and forces experienced during real-time games or events and in training. These new locations also open up opportunities for many new research studies like fall risk detection and gait analysis.

While the two main objectives were fulfilled, this study still had some limitations. The first limitation was the small sample size. Only eight subjects were able to be recruited with only six having usable data for a majority of the trials and locations. The subjects were also all healthy adults with no known cardiovascular or musculoskeletal issues. Another limitation was the exercises being constrained to the laboratory setting, especially for the treadmill trials. It has been noted that walking and running kinematics have been substantially different when on a treadmill versus overground [108].

3.4. Conclusion and Future Work

3.4.1. Conclusion

This study provides evidence that PPG can be collected at locations other than the finger, wrist, ear, and forehead and provide relevant results. The results showed an average of 26.1%, 18.3%, and 21.0% improvement in estimating HR at the forearm, shank, and sacrum, respectively, compared to the Apple Watch data from all trials of this study. This work provides the foundation to further explore how well the AIM System hardware can track HR comfortably, during intense physical activities and other potential locations. The HR algorithm could be integrated into an app for use with the system since the hardware includes an accelerometer and PPG sensor all in one device. This would make future

research with the AIM System more feasible, and it could potentially become a tool used in the fitness and athletic training.

3.4.2. Future Work

Future work should focus on validating the algorithms with more subjects and data. A more diverse sample set in regard to subjects and exercises would also strengthen the results reported in this thesis. Another addition to this work would be to develop a novel technique for cancelling motion artifacts in PPG for the AIM System. Since the AIM System has an IMU and PPG all onboard, and the methods used in this paper did not, a novel algorithm to cancel motion artifacts for this new type of hardware would be beneficial.

After a more in-depth validation, this system could be applied to the work discussed in Chapter 2. With accurate HR estimations, this technology could be used to help understand diseases and disorders that affect the parasympathetic nervous system, along with musculoskeletal disorders as well. With the ability to measure full gait metrics and HR measures, many new studies could be conducted.

CHAPTER 4: CONCLUDING REMARKS

4.1. Internalizing Disorder Identification in Children

Wearable IMUs show promise in being used to screen children for internalizing disorders by having them play with a bubble machine for three minutes. The results from this study are also supported by previous research that also uses mood induction tasks and machine learning to identify internalizing disorders in children [11]–[14]. With an accuracy of 75%, and high sensitivity, this model performs well compared to the current questionnaire (CBCL) which could potentially be used as a screening tool. This provides a low-cost option, after the initial purchase of wearable IMUs and a bubble machine, to be used in pediatric offices as a screening tool. This could allow for more children with internalizing disorders to get help early on and dismantle the stigma around mental health disorders.

4.2. HR Estimation from New Body Locations

Additional physiological measure like HR could potentially strengthen the results found in the childhood internalizing disorder study. The second study, and new wearable sensor tool, of HR estimation during intense physical activities at new body locations has the potential to be used in studies like the one just mentioned. With average absolute error rates of 6.2, 6.9, and 6.7 BPM at the forearm, shank, and sacrum, respectively, compared to the gold standard Polar H10 chest strap, the results show promise for patch-like wearables with photoplethysmography and accelerometry to be useful in accurately estimating HR. With the ability for these patch-like sensors to be worn anywhere, there is a large set of applications for this sensor. This type of hardware also allows for good skin

contact through an array of exercises which maximized the signal quality and ability for HR estimation compared to devices like the Apple Watch (methods presented had a 26.1%, 18.3%, and 21.0% improvement on average at the forearm, shank, and sacrum, respectively, compared to the Apple Watch) and other clip- or band-worn sensors that can move freely about the skin.

4.3. Current and Future Work

With the current global pandemic caused by COVID-19, researchers have been putting efforts into finding ways to monitoring symptoms. Wearable devices now have many sensing abilities and can keep track of HR, sleep durations and disturbances, temperature, and more, making them a useful tool for tracking symptoms [59], [109]. The coronavirus has three main symptoms of respiratory distress, fever, and coughing, which can all be monitored with wearable sensors [109]. Monitoring of these symptoms, and tracking the data, can give a larger, clearer picture of what is going on. It also allowed healthcare providers to stay connected with patients who may already have underlying health problems, like cardiac arrhythmias, and continue treatment with minimal contact [60].

Researchers have also been looking for quick, effective, low cost ways to diagnose people with the novel coronavirus as well. Two companies, Cardea Bio [110] and Hememics Biotechnologies [111], have been putting efforts towards developing semiconductor biosensor chips that rapidly detect viral RNA, antibodies, or antigens from a nasal swap or blood test sample [112]. Cardea Bio has developed a hand-held biosensor chip that rapidly detects and identifies disease markers. This chip allows for testing to be

done outside of the lab, and it can test for multiple different molecular signals at once allowing to differentiate between someone who has COVID-19 or simply the influenza [110], [113]. Similarly, Hememics Biotechnologies created a hand-held testing device that can test for up to 17 different pathogens using a single drop of blood or nasal swab [61]. Hememics' technology also allows for the data to be uploaded to a cloud-based system which allows for mapping of the outbreaks geographically and can be used as a screening tool before travel or returning back to work [61], [111].

With current efforts focused on the global pandemic, future work for the studies presented in this thesis (Chapters 2 and 3) should focus on validation with larger sample sizes. With the framework for these new tools in place, larger, more diverse sample sizes could provide insight into where more improvement could be made in both studies. It would also make the methods and studies conducted more reliable.

4.4. Conclusion

Overall, the work in this thesis provides new applications for wearable sensors to be used in delivering reliable biologically and physiologically relevant signals. These new tools limit the amount of human bias injected into the measurement and results, which can be used to improve diagnostic measures, monitoring of diseases and disorders, and tracking athletic performance for safety and training measures.

BIBLIOGRAPHY

[1] P. B. Shull, W. Jirattigalachote, M. A. Hunt, M. R. Cutkosky, and S. L. Delp, "Quantified self and human movement: A review on the clinical impact of wearable sensing and feedback for gait analysis and intervention," *Gait Posture*, vol. 40, no. 1, pp. 11–19, May 2014, doi: 10.1016/j.gaitpost.2014.03.189.

[2] L. C. Benson, C. A. Clermont, E. Bošnjak, and R. Ferber, "The use of wearable devices for walking and running gait analysis outside of the lab: A systematic review," *Gait Posture*, vol. 63, pp. 124–138, Jun. 2018, doi: 10.1016/j.gaitpost.2018.04.047.

[3] R. D. Gurchiek *et al.*, "Remote Gait Analysis Using Wearable Sensors Detects Asymmetric Gait Patterns in Patients Recovering from ACL Reconstruction," in *2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, May 2019, pp. 1–4, doi: 10.1109/BSN.2019.8771038.

[4] M. Sartori, D. Farina, and D. G. Lloyd, "Hybrid neuromusculoskeletal modeling to best track joint moments using a balance between muscle excitations derived from electromyograms and optimization," *J. Biomech.*, vol. 47, no. 15, pp. 3613–3621, Nov. 2014, doi: 10.1016/j.jbiomech.2014.10.009.

[5] A. Sano *et al.*, "Recognizing Academic Performance, Sleep Quality, Stress Level, and Mental Health using Personality Traits, Wearable Sensors and Mobile Phones," *Int. Conf. Wearable Implant. Body Sens. Netw. Int. Conf. Wearable Implant. Body Sens. Netw.*, vol. 2015, Jun. 2015, Accessed: Jan. 29, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5431072/.

[6] R. E. De Meersman, "Heart rate variability and aerobic fitness," *Am. Heart J.*, vol. 125, no. 3, pp. 726–731, Mar. 1993, doi: 10.1016/0002-8703(93)90164-5.

[7] "Physical Fitness Levels vs Selected Coronary Risk Factors: A Cross-Sectional Study | JAMA | JAMA Network." https://jamanetwork.com/journals/jama/article-abstract/346661 (accessed Jan. 29, 2020).

[8] C. E. Garber *et al.*, "Quantity and Quality of Exercise for Developing and Maintaining Cardiorespiratory, Musculoskeletal, and Neuromotor Fitness in Apparently Healthy Adults: Guidance for Prescribing Exercise," *Med. Sci. Sports Exerc.*, vol. 43, no. 7, pp. 1334–1359, Jul. 2011, doi: 10.1249/MSS.0b013e318213fefb.

[9] P. Lehrer *et al.*, "Heart Rate Variability Biofeedback: Effects of Age on Heart Rate Variability, Baroreflex Gain, and Asthma," *Chest*, vol. 129, no. 2, pp. 278–284, Feb. 2006, doi: 10.1378/chest.129.2.278.

[10] Vrijkotte Tanja G. M., van Doornen Lorenz J. P., and de Geus Eco J. C., "Effects of Work Stress on Ambulatory Blood Pressure, Heart Rate, and Heart Rate Variability," *Hypertension*, vol. 35, no. 4, pp. 880–886, Apr. 2000, doi: 10.1161/01.HYP.35.4.880.

[11] E. W. McGinnis *et al.*, "Wearable sensors detect childhood internalizing disorders during mood induction task," *PLoS ONE*, vol. 13, no. 4, 2018, doi: 10.1371/journal.pone.0195598.

[12] E. W. McGinnis *et al.*, “Giving Voice to Vulnerable Children: Machine Learning Analysis of Speech Detects Anxiety and Depression in Early Childhood,” *IEEE J. Biomed. Health Inform.*, vol. 23, no. 6, pp. 2294–2301, Nov. 2019, doi: 10.1109/JBHI.2019.2913590.
[13] E. W. McGinnis *et al.*, “Movements Indicate Threat Response Phases in Children at Risk for Anxiety,” *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1460–1465, Sep. 2017, doi: 10.1109/JBHI.2016.2603159.
[14] R. S. McGinnis *et al.*, “Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning,” *PLOS ONE*, vol. 14, no. 1, p. e0210267, Jan. 2019, doi: 10.1371/journal.pone.0210267.
[15] Z. Zhu, T. Liu, G. Li, T. Li, and Y. Inoue, “Wearable Sensor Systems for Infants,” *Sens. Basel*, vol. 15, no. 2, pp. 3721–3749, 2015, doi: http://dx.doi.org/10.3390/s150203721.
[16] S. C. Mukhopadhyay, “Wearable Sensors for Human Activity Monitoring: A Review,” *IEEE Sens. J.*, vol. 15, no. 3, pp. 1321–1330, Mar. 2015, doi: 10.1109/JSEN.2014.2370945.
[17] T. Guinovart, G. Valdés-Ramírez, J. R. Windmiller, F. J. Andrade, and J. Wang, “Bandage-Based Wearable Potentiometric Sensor for Monitoring Wound pH,” *Electroanalysis*, vol. 26, no. 6, pp. 1345–1353, 2014, doi: 10.1002/elan.201300558.
[18] T. S. Caudill, R. Lofgren, C. D. Jennings, and M. Karpf, “Commentary: Health Care Reform and Primary Care: Training Physicians for Tomorrow’s Challenges,” *Acad. Med.*, vol. 86, no. 2, pp. 158–160, Feb. 2011, doi: 10.1097/ACM.0b013e3182045f13.
[19] L. Chan, L. G. Hart, and D. C. Goodman, “Geographic Access to Health Care for Rural Medicare Beneficiaries,” *J. Rural Health*, vol. 22, no. 2, pp. 140–146, 2006, doi: 10.1111/j.1748-0361.2006.00022.x.
[20] D. A. Kegelmeyer *et al.*, “Quantitative biomechanical assessment of trunk control in Huntington’s disease reveals more impairment in static than dynamic tasks,” *J. Neurol. Sci.*, vol. 376, pp. 29–34, May 2017, doi: 10.1016/j.jns.2017.02.054.
[21] R. Sun *et al.*, “Assessment of Postural Sway in Individuals with Multiple Sclerosis Using a Novel Wearable Inertial Sensor,” *Digit. Biomark.*, vol. 2, no. 1, pp. 1–10, Jan. 2018, doi: 10.1159/000485958.
[22] J. Barth *et al.*, “Biometric and mobile gait analysis for early diagnosis and therapy monitoring in Parkinson’s disease,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2011, pp. 868–871, doi: 10.1109/IEMBS.2011.6090226.
[23] Y. Nancy Jane, H. Khanna Nehemiah, and K. Arputharaj, “A Q-backpropagated time delay neural network for diagnosing severity of gait disturbances in Parkinson’s disease,” *J. Biomed. Inform.*, vol. 60, pp. 169–176, Apr. 2016, doi: 10.1016/j.jbi.2016.01.014.
[24] P. Bonato, “Advances in wearable technology and its medical applications,” in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug. 2010, pp. 2021–2024, doi: 10.1109/IEMBS.2010.5628037.

[25] M. M. Rodgers, V. M. Pai, and R. S. Conroy, “Recent Advances in Wearable Sensors for Health Monitoring,” *IEEE Sens. J.*, vol. 15, no. 6, pp. 3119–3126, Jun. 2015, doi: 10.1109/JSEN.2014.2357257.
[26] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, “A review of wearable sensors and systems with application in rehabilitation,” *J. NeuroEngineering Rehabil.*, vol. 9, no. 1, Art. no. 1, Dec. 2012, doi: 10.1186/1743-0003-9-21.
[27] M. J. Bradshaw, S. Farrow, R. W. Motl, and T. Chitnis, “Wearable biosensors to monitor disability in multiple sclerosis,” *Neurol. Clin. Pract.*, vol. 7, no. 4, pp. 354–362, Aug. 2017, doi: 10.1212/CPJ.0000000000000382.
[28] G. Valenza *et al.*, “Wearable Monitoring for Mood Recognition in Bipolar Disorder Based on History-Dependent Long-Term Heart Rate Variability Analysis,” *IEEE J. Biomed. Health Inform.*, vol. 18, no. 5, pp. 1625–1635, Sep. 2014, doi: 10.1109/JBHI.2013.2290382.
[29] A. Grünerbl *et al.*, “Smartphone-Based Recognition of States and State Changes in Bipolar Disorder Patients,” *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 140–148, Jan. 2015, doi: 10.1109/JBHI.2014.2343154.
[30] A. Lanata, G. Valenza, M. Nardelli, C. Gentili, and E. P. Scilingo, “Complexity Index From a Personalized Wearable Monitoring System for Assessing Remission in Mental Health,” *IEEE J. Biomed. Health Inform.*, vol. 19, no. 1, pp. 132–139, Jan. 2015, doi: 10.1109/JBHI.2014.2360711.
[31] E. Waxler, “Emotion Dysregulation Assessment in Young Children According to Research Domain Criteria,” University of Michigan, 2017.
[32] E. Holmes *et al.*, “Metabolic Profiling of CSF: Evidence That Early Intervention May Impact on Disease Progression and Outcome in Schizophrenia,” *PLoS Med.*, vol. 3, no. 8, p. e327, Aug. 2006, doi: 10.1371/journal.pmed.0030327.
[33] P. D. McGorry, “Early Intervention in Psychosis,” *J. Nerv. Ment. Dis.*, vol. 203, no. 5, pp. 310–318, May 2015, doi: 10.1097/NMD.0000000000000284.
[34] “Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions - McGorry - 2006 - Australian and New Zealand Journal of Psychiatry - Wiley Online Library.” https://onlinelibrary-wiley-com.ezproxy.uvm.edu/doi/full/10.1111/j.1440-1614.2006.01860.x (accessed Jun. 19, 2020).
[35] M. A. Quinn and P. Emery, “Window of opportunity in early rheumatoid arthritis: Possibility of altering the disease process with early intervention,” p. 5.
[36] S. DeKosky, “Early Intervention Is Key to Successful Management of Alzheimer Disease,” *Alzheimer Dis. Assoc. Disord.*, vol. 17, p. S99, Sep. 2003.
[37] S. W. Yun, “Congenital heart disease in the newborn requiring early intervention,” *Korean J. Pediatr.*, vol. 54, no. 5, pp. 183–191, May 2011, doi: 10.3345/kjp.2011.54.5.183.
[38] S. M. Abubakar, W. Saadeh, and M. A. B. Altaf, “A wearable long-term single-lead ECG processor for early detection of cardiac arrhythmia,” in *2018 Design, Automation Test in Europe Conference Exhibition (DATE)*, Mar. 2018, pp. 961–966, doi: 10.23919/DATE.2018.8342148.

[39] I. Korhonen, J. Parkka, and M. Van Gils, “Health monitoring in the home of the future,” *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 66–73, May 2003, doi: 10.1109/MEMB.2003.1213628.
[40] CDC, “Data and Statistics on Children’s Mental Health | CDC,” *Centers for Disease Control and Prevention*, Apr. 19, 2019. https://www.cdc.gov/childrensmentalhealth/data.html (accessed Apr. 30, 2019).
[41] D. J. Kolko and A. E. Kazdin, “Emotional/Behavioral Problems in Clinic and Nonclinic Children: Correspondence Among Child, Parent and Teacher Reports,” *J. Child Psychol. Psychiatry*, vol. 34, no. 6, pp. 991–1006, 1993, doi: 10.1111/j.1469-7610.1993.tb01103.x.
[42] C. Kieling *et al.*, “Child and adolescent mental health worldwide: evidence for action - ScienceDirect,” Oct. 16, 2011. https://www-sciencedirect-com.ezproxy.uvm.edu/science/article/pii/S0140673611608271 (accessed Jun. 19, 2020).
[43] S. Coyle, D. Morris, K.-T. Lau, D. Diamond, and N. Moyna, “Textile-Based Wearable Sensors for Assisting Sports Performance,” in *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, Jun. 2009, pp. 307–311, doi: 10.1109/BSN.2009.57.
[44] “Wearable sensors for monitoring the internal and external workload of the athlete | npj Digital Medicine.” https://www-nature-com.ezproxy.uvm.edu/articles/s41746-019-0149-2 (accessed Jun. 22, 2020).
[45] R. Ghaffari *et al.*, “Soft Wearable Systems for Colorimetric and Electrochemical Analysis of Biofluids,” *Adv. Funct. Mater.*, vol. n/a, no. n/a, p. 1907269, doi: 10.1002/adfm.201907269.
[46] T. Malkinson, “Current and emerging technologies in endurance athletic training and race monitoring,” in *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*, Sep. 2009, pp. 581–586, doi: 10.1109/TIC-STH.2009.5444434.
[47] J. Quindry, C. Williamson-Reisdorph, and J. French, “Health and fitness benefits using a heart rate intensity-based group fitness exercise regimen,” *J. Hum. Sport Exerc.*, vol. 15, no. 3, 2019, doi: 10.14198/jhse.2020.153.18.
[48] M. A. Motin, C. K. Karmakar, and M. Palaniswami, “PPG Derived Heart Rate Estimation During Intensive Physical Exercise,” *IEEE Access*, vol. 7, pp. 56062–56069, 2019, doi: 10.1109/ACCESS.2019.2913148.
[49] M. Boloursaz Mashhadi, E. Asadi, M. Eskandari, S. Kiani, and F. Marvasti, “Heart Rate Tracking using Wrist-Type Photoplethysmographic (PPG) Signals during Physical Exercise with Simultaneous Accelerometry,” *IEEE Signal Process. Lett.*, vol. 23, no. 2, pp. 227–231, Feb. 2016, doi: 10.1109/LSP.2015.2509868.
[50] A. Temko, “Accurate Heart Rate Monitoring During Physical Exercises Using PPG,” *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2016–2024, Sep. 2017, doi: 10.1109/TBME.2017.2676243.
[51] T. Shimazaki, S. Hara, H. Okuhata, H. Nakamura, and T. Kawabata, “Motion artifact cancellation and outlier rejection for clip-type ppg-based heart rate sensor,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine*

and Biology Society (EMBC), Milan, Aug. 2015, pp. 2026–2029, doi: 10.1109/EMBC.2015.7318784.
[52] P.-H. Lai and I. Kim, "Lightweight wrist photoplethysmography for heavy exercise: motion robust heart rate monitoring algorithm," *Healthc. Technol. Lett.*, vol. 2, no. 1, pp. 6–11, 2015, doi: 10.1049/htl.2014.0097.
[53] A. Cicone and H.-T. Wu, "How Nonlinear-Type Time-Frequency Analysis Can Help in Sensing Instantaneous Heart Rate and Instantaneous Respiratory Rate from Photoplethysmography in a Reliable Way," *Front. Physiol.*, vol. 8, 2017, doi: 10.3389/fphys.2017.00701.
[54] T. Schack, M. Muma, and A. M. Zoubir, "Computationally efficient heart rate estimation during physical exercise using photoplethysmographic signals," in *2017 25th European Signal Processing Conference (EUSIPCO)*, Kos, Greece, Aug. 2017, pp. 2478–2481, doi: 10.23919/EUSIPCO.2017.8081656.
[55] J. Xiong, L. Cai, D. Jiang, H. Song, and X. He, "Spectral Matrix Decomposition-Based Motion Artifacts Removal in Multi-Channel PPG Sensor Signals," *IEEE Access*, vol. 4, pp. 3076–3086, 2016, doi: 10.1109/ACCESS.2016.2580594.
[56] H. Han and J. Kim, "Artifacts in wearable photoplethysmographs during daily life motions and their reduction with least mean square based active noise cancellation method," *Comput. Biol. Med.*, vol. 42, no. 4, pp. 387–393, Apr. 2012, doi: 10.1016/j.compbiomed.2011.12.005.
[57] J. A. C. Patterson, D. C. McIlwraith, and G.-Z. Yang, "A Flexible, Low Noise Reflective PPG Sensor Platform for Ear-Worn Heart Rate Monitoring," in *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, Jun. 2009, pp. 286–291, doi: 10.1109/BSN.2009.16.
[58] D. Morris *et al.*, "Wearable technology for bio-chemical analysis of body fluids during exercise," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Aug. 2008, pp. 5741–5744, doi: 10.1109/IEMBS.2008.4650518.
[59] A. Kapoor, S. Guha, M. Kanti Das, K. C. Goswami, and R. Yadav, "Digital healthcare: The only solution for better healthcare during COVID-19 pandemic?," *Indian Heart J.*, vol. 72, no. 2, pp. 61–64, 2020, doi: 10.1016/j.ihj.2020.04.001.
[60] Lakkireddy Dhanunjaya R. *et al.*, "Guidance for Cardiac Electrophysiology During the COVID-19 Pandemic from the Heart Rhythm Society COVID-19 Task Force; Electrophysiology Section of the American College of Cardiology; and the Electrocardiography and Arrhythmias Committee of the Council on Clinical Cardiology, American Heart Association," *Circulation*, vol. 141, no. 21, pp. e823–e831, May 2020, doi: 10.1161/CIRCULATIONAHA.120.047063.
[61] H. B. Inc, "HEMEMICS Biotechnologies, Inc. Receives HHS Support to Develop Rapid Antigen, Antibody Diagnostic to Identify COVID-19 Infected Americans." https://www.prnewswire.com/news-releases/hememics-biotechnologies-inc-receives-hhs-support-to-develop-rapid-antigen-antibody-diagnostic-to-identify-covid-19-infected-americans-301041698.html (accessed Jun. 25, 2020).
[62] E. Z. Gorodeski *et al.*, "Virtual Visits for Care of Patients with Heart Failure in the Era of COVID-19: A Statement from the Heart Failure Society of America," *J.*

Card. Fail., vol. 26, no. 6, pp. 448–456, Jun. 2020, doi: 10.1016/j.cardfail.2020.04.008.
[63] "Heart-Rate Based HIIT Workout | Orangetheory Fitness US." https://www.orangetheory.com/en-us/workout/ (accessed Jun. 22, 2020).
[64] M. Ltd *, "Group Fitness Tracking Software | Wearable Fitness Trackers." https://www.myzone.org (accessed Jun. 22, 2020).
[65] L. C. Wu *et al.*, "Detection of American Football Head Impacts Using Biomechanical Features and Support Vector Machine Classification," *Sci. Rep.*, vol. 8, no. 1, Art. no. 1, Dec. 2017, doi: 10.1038/s41598-017-17864-3.
[66] J. D. Ralston, W. Meiring, M. Cieslak, A. Asturias, and S. T. Grafton, "Wearable sensors for head impact dosimetry," in *2017 IEEE SENSORS*, Oct. 2017, pp. 1–3, doi: 10.1109/ICSENS.2017.8234233.
[67] A. Alderson, "Sports Tech - wearable sensors [Technology Rugby]," *Eng. Technol.*, vol. 11, no. 6, pp. 76–77, Jul. 2016, doi: 10.1049/et.2016.0607.
[68] G. P. Siegmund, K. M. Guskiewicz, S. W. Marshall, A. L. DeMarco, and S. J. Bonin, "Laboratory Validation of Two Wearable Sensor Systems for Measuring Head Impact Severity in Football Players," *Ann. Biomed. Eng.*, vol. 44, no. 4, pp. 1257–1274, Apr. 2016, doi: 10.1007/s10439-015-1420-6.
[69] G. L. Faedda *et al.*, "Actigraph measures discriminate pediatric bipolar disorder from attention-deficit/hyperactivity disorder and typically developing controls," *J. Child Psychol. Psychiatry*, vol. 57, no. 6, pp. 706–716, Jun. 2016, doi: 10.1111/jcpp.12520.
[70] B. Lee, J. Han, H. J. Baek, J. H. Shin, K. S. Park, and W. J. Yi, "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry," *Physiol. Meas.*, vol. 31, no. 12, pp. 1585–1603, Dec. 2010, doi: 10.1088/0967-3334/31/12/003.
[71] M. Tandon, E. Cardeli, and J. Luby, "Internalizing Disorders in Early Childhood: A Review of Depressive and Anxiety Disorders," *Child Adolesc. Psychiatr. Clin. N. Am.*, vol. 18, no. 3, pp. 593–610, Jul. 2009, doi: 10.1016/j.chc.2009.03.004.
[72] R. Gilgen-Ammann, T. Schweizer, and T. Wyss, "RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise," *Eur. J. Appl. Physiol.*, vol. 119, no. 7, pp. 1525–1532, Jul. 2019, doi: 10.1007/s00421-019-04142-5.
[73] R. S. McGinnis *et al.*, "Wearable sensors and machine learning diagnose anxiety and depression in young children," in *2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Las Vegas, NV, USA, Mar. 2018, pp. 410–413, doi: 10.1109/BHI.2018.8333455.
[74] H.-U. Wittchen and H. Sonntag, "Nicotine consumption in mental disorders: a clinical epidemiological perspective," *Eur. Neuropsychopharmacol.*, vol. 10, p. 119, Sep. 2000, doi: 10.1016/S0924-977X(00)80014-0.
[75] M. S. Gould *et al.*, "Psychopathology Associated With Suicidal Ideation and Attempts Among Children and Adolescents," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 37, no. 9, pp. 915–923, Sep. 1998, doi: 10.1097/00004583-199809000-00011.

[76] M. G. Craske and M. B. Stein, “Anxiety,” *The Lancet*, vol. 388, no. 10063, pp. 3048–3059, Dec. 2016, doi: 10.1016/S0140-6736(16)30381-6.
[77] “NIMH » Suicide.” https://www.nimh.nih.gov/health/statistics/suicide.shtml (accessed May 05, 2019).
[78] J. Garber and K. M. Kaminski, “Laboratory and Performance-Based Measures of Depression in Children and Adolescents,” *J. Clin. Child Psychol.*, vol. 29, no. 4, pp. 509–525, Nov. 2000, doi: 10.1207/S15374424JCCP2904_5.
[79] A. G. Renouf and M. Kovacs, “Concordance between Mothers’ Reports and Children’s Self-Reports of Depressive Symptoms: A Longitudinal Study,” *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 33, no. 2, pp. 208–216, Feb. 1994, doi: 10.1097/00004583-199402000-00008.
[80] C. Lord *et al.*, “The Autism Diagnostic Observation Schedule–Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism,” p. 19.
[81] H. Hart *et al.*, “Pattern classification of response inhibition in ADHD: Toward the development of neurobiological markers for ADHD,” *Hum. Brain Mapp.*, vol. 35, no. 7, pp. 3083–3094, 2014, doi: 10.1002/hbm.22386.
[82] R. Iannaccone, T. U. Hauser, J. Ball, D. Brandeis, S. Walitza, and S. Brem, “Classifying adolescent attention-deficit/hyperactivity disorder (ADHD) based on functional and structural imaging,” *Eur. Child Adolesc. Psychiatry*, vol. 24, no. 10, pp. 1279–1289, Oct. 2015, doi: 10.1007/s00787-015-0678-4.
[83] J. R. Sato *et al.*, “Default mode network maturation and psychopathology in children and adolescents,” *J. Child Psychol. Psychiatry*, vol. 57, no. 1, pp. 55–64, Jan. 2016, doi: 10.1111/jcpp.12444.
[84] J. R. Sato *et al.*, “Association between abnormal brain functional connectivity in children and psychopathology: A study based on graph theory and machine learning,” *World J. Biol. Psychiatry*, vol. 19, no. 2, pp. 119–129, Feb. 2018, doi: 10.1080/15622975.2016.1274050.
[85] D. Greenstein, B. Weisinger, J. D. Malley, L. Clasen, and N. Gogtay, “Using Multivariate Machine Learning Methods and Structural MRI to Classify Childhood Onset Schizophrenia and Healthy Controls,” *Front. Psychiatry*, vol. 3, 2012, doi: 10.3389/fpsyt.2012.00053.
[86] D. Bone, S. L. Bishop, M. P. Black, M. S. Goodwin, C. Lord, and S. S. Narayanan, “Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion,” *J. Child Psychol. Psychiatry*, vol. 57, no. 8, pp. 927–937, Aug. 2016, doi: 10.1111/jcpp.12559.
[87] K. Carpenter, P. Sprechmann, M. Fiori, R. Calderbank, H. Egger, and G. Sapiro, “Questionnaire simplification for fast risk analysis of children’s mental health,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014, pp. 6009–6013, doi: 10.1109/ICASSP.2014.6854757.
[88] G. N. Saxe, S. Ma, J. Ren, and C. Aliferis, “Machine learning methods to predict child posttraumatic stress: a proof of concept study,” *BMC Psychiatry*, vol. 17, no. 1, p. 223, Jul. 2017, doi: 10.1186/s12888-017-1384-1.
[89] T. M. Achenbach, C. T. Howell, H. C. Quay, C. K. Conners, and J. E. Bates, “National Survey of Problems and Competencies among Four- to Sixteen-Year-

Olds: Parents' Reports for Normative and Clinical Samples," *Monogr. Soc. Res. Child Dev.*, vol. 56, no. 3, pp. i–130, 1991, doi: 10.2307/1166156.
[90] "Preschool (CBCL-LDS, C-TRF)," *ASEBA*. https://aseba.org/preschool/ (accessed Jun. 25, 2020).
[91] M. Gaffrey and J. Luby, "Kiddie schedule for affective disorders and schizophrenia-early childhood version (K-SADS-EC)," St. Louis, MO: Washington University School of Medicine, 2012.
[92] M. Maziade *et al.*, "Reliability of best-estimate diagnosis in genetic linkage studies of major psychoses: Results from the Quebec pedigree studies," *Am. J. Psychiatry*, vol. 149, no. 12, pp. 1674–1686, 1992, doi: 10.1176/ajp.149.12.1674.
[93] F. Habibzadeh and M. Yadollahie, "Number Needed to Misdiagnose: A Measure of Diagnostic Test Effectiveness," *Epidemiology*, vol. 24, no. 1, pp. 170–170, 2013.
[94] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.
[95] "Somatic symptom disorder - Symptoms and causes - Mayo Clinic." https://www.mayoclinic.org/diseases-conditions/somatic-symptom-disorder/symptoms-causes/syc-20377776 (accessed May 05, 2019).
[96] A. Isaksson, M. Wallman, H. Göransson, and M. G. Gustafsson, "Cross-validation and bootstrapping are unreliable in small sample classification," *Pattern Recognit. Lett.*, vol. 29, no. 14, pp. 1960–1965, Oct. 2008, doi: 10.1016/j.patrec.2008.06.018.
[97] G. Santafe, I. Inza, and J. A. Lozano, "Dealing with the evaluation of supervised classification algorithms," *Artif. Intell. Rev.*, vol. 44, no. 4, pp. 467–508, Dec. 2015, doi: 10.1007/s10462-015-9433-y.
[98] A. Fiol-Veny, A. De La Torre-Luque, M. Balle, and X. Bornas, "Altered Heart Rate Regulation in Adolescent Girls and the Vulnerability for Internalizing Disorders," *Front. Physiol.*, vol. 9, 2018, doi: 10.3389/fphys.2018.00852.
[99] T. P. Beauchaine and J. F. Thayer, "Heart rate variability as a transdiagnostic biomarker of psychopathology," *Int. J. Psychophysiol.*, vol. 98, no. 2, pp. 338–350, Nov. 2015, doi: 10.1016/j.ijpsycho.2015.08.004.
[100] S. A. Callahan and S. M. Panichelli-Mindel, "DSM-IV and internalizing disorders: Modifications, limitations, and utility," *Sch. Psychol. Rev.*, vol. 25, no. 3, p. 297, Sep. 1996.
[101] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, Feb. 2007, doi: 10.1088/0967-3334/28/3/R01.
[102] H. Han, M.-J. Kim, and J. Kim, "Development of real-time motion artifact reduction algorithm for a wearable photoplethysmography," in *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, France, Aug. 2007, pp. 1538–1541, doi: 10.1109/IEMBS.2007.4352596.
[103] L. Wang and K. J. Loh, "Wearable carbon nanotube-based fabric sensors for monitoring human physiological performance," *Smart Mater. Struct.*, vol. 26, no. 5, p. 055018, May 2017, doi: 10.1088/1361-665X/aa6849.
[104] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985, doi: 10.1109/TBME.1985.325532.

[105] M. A. Motin, C. K. Karmakar, and M. Palaniswami, “Ensemble Empirical Mode Decomposition With Principal Component Analysis: A Novel Approach for Extracting Respiratory Rate and Heart Rate From Photoplethysmographic Signal,” *IEEE J. Biomed. Health Inform.*, vol. 22, no. 3, pp. 766–774, 2018, doi: 10.1109/JBHI.2017.2679108.
[106] andtem2000, *andtem2000/PPG*. 2020.
[107] G. Blain, O. Meste, A. Blain, and S. Bermon, “Time-frequency analysis of heart rate variability reveals cardiolocomotor coupling during dynamic cycling exercise in humans,” *Am. J. Physiol.-Heart Circ. Physiol.*, vol. 296, no. 5, pp. H1651–H1659, May 2009, doi: 10.1152/ajpheart.00881.2008.
[108] B. M. Nigg, R. W. De Boer, and V. Fisher, “A kinematic comparison ofverground and treadmill running,” *Med. Sci. Sports Exerc.*, vol. 27, no. 1, pp. 98–105, 1995.
[109] X.-R. Ding *et al.*, “Wearable Sensing and Telehealth Technology with Potential Applications in the Coronavirus Pandemic,” *IEEE Rev. Biomed. Eng.*, pp. 1–1, 2020, doi: 10.1109/RBME.2020.2992838.
[110] “Cardean Transistors™ made available to companies and government agencies willing to build handheld Coronavirus detection devices,” *PRWeb*. https://www.prweb.com/releases/cardean_transistors_made_available_to_companies_and_government_agencies_willing_to_build_handheld_coronavirus_detection_devices/prweb17004954.htm (accessed Jun. 25, 2020).
[111] “Hememics – Rapid Point-of-Care Diagnostics.” https://www.hememics.com/ (accessed Jun. 25, 2020).
[112] “Biosensors May Hold the Key to Mass Coronavirus Testing - IEEE Spectrum,” *IEEE Spectrum: Technology, Engineering, and Science News*. https://spectrum.ieee.org/view-from-the-valley/biomedical/diagnostics/biosensors-key-mass-coronavirus-testing?utm_source=circuitsandsensors&utm_medium=email&utm_campaign=circuitsandsensors,06-02-20&mkt_tok=eyJpIjoiWXpRNFltTmlOalZpWkdFdyIsInQiOiJXenBDeitIK3p4a29ydXlHR3dCbitsR29TQ3huNnZcL0JEUTV4WlZtTW84elRiem40QXZSZE1xWVdhVlhBWVNEV0JMMjdMMDR0Ym1VRUg2dlo0cG1GYmxZN1JzTVNpTkJLcnprb21JWXlTMU1Ka2wwc2k3cGpuWXN2R1Vnc3YzV2MifQ%3D%3D (accessed Jun. 25, 2020).
[113] B. R. Goldsmith *et al.*, “Digital Biosensing by Foundry-Fabricated Graphene Sensors,” *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Jan. 2019, doi: 10.1038/s41598-019-38700-w.