DEVELOPMENT OF AN INERTIAL MEASUREMENT-BASED ASSESSMENT OF DISEASE SEVERITY IN CHRONIC FATIGUE SYNDROME

by

Kenneth Turner McGill Palombo

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

IN

MECHANICAL ENGINEERING

Mechanical Engineering

The University of Utah

August 2020

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Kenneth Turner McGill Palombo**

has been approved by the following supervisory committee members:

Shad Roundy	, Chair	**April 20, 2020** Date Approved
Andrew Merryweather	, Member	**April 20, 2020** Date Approved
Suzanne Vernon	, Member	**April 20, 2020** Date Approved

and by **Bruce K. Gale** , Chair/Dean of

the Department/College/School of **Mechanical Engineering**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

More than two million Americans suffer from myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS). While ME/CFS is still poorly understood, a recent upsurge in related research has identified the disease's core symptoms, including post-exertional malaise (PEM) and unrefreshing sleep. The FDA has yet to approve any treatments for ME/CFS, partially due to a lack of validated efficacy endpoints.

The central focus of this research is to develop ME/CFS efficacy endpoints using a non-invasive, inertial measurement-based approach. Accessible endpoints will provide a way to properly evaluate potential treatments for ME/CFS. In this research, an inertial measurement-based assessment of upright activity—referred to as UpTime—is validated. Furthermore, potential efficacy endpoints corresponding to unrefreshing sleep are evaluated using data from an Oura Ring.

Using a Kalman filter, inertial measurement unit (IMU) data can be converted to optimized leg angle estimates. These angle estimates can then be converted to personalized daily UpTime scores.

In a six-day, case-control study conducted at the Bateman Horne Center, UpTime and Oura Ring sleep data were collected from 15 subjects (five controls, five moderate-level ME/CFS, and five severe-level ME/CFS). Analysis of UpTime data collected during this study indicates that each of these groups spends different proportions of their days upright and active. This result indicates that UpTime can accurately detect upright

posture that can be a surrogate for disease severity. UpTime is, therefore, a reliable endpoint for evaluating treatment efficacy. While the sleep data collected by the Oura Rings was interesting, our analysis did not yield any significant insights; this is most likely due to short data collection periods and insufficient sampling.

Future studies with extended data collection periods are required to fully evaluate the value of the Oura Ring as a measurement tool for unrefreshing sleep. Together, Oura Rings and IMUs could provide a platform to more easily measure efficacy endpoints related to the core symptoms of ME/CFS. With the added perspective of large-scale studies, this sensor-based platform could provide a recovery path for individuals struggling with ME/CFS.

TABLE OF CONTENTS

LIST OF TABLES

Tables

LIST OF FIGURES

Figures

LIST OF SYMBOLS

ϕ	Roll, angle relative to a global x-axis
θ	Pitch, angle relative to a global z-axis
ψ	Yaw, angle relative to a global y-axis
a_x	Acceleration along the x-axis
a_y	Acceleration along the y-axis
a_z	Acceleration along the z-axis
p	Body fixed rotation rates about the x-axis, recorded by the gyroscope
q	Body fixed rotation rates about the y-axis, recorded by the gyroscope
r	Body fixed rotation rates about the z-axis, recorded by the gyroscope
ϕ_{Acc}	Accelerometer-based measurement of roll
θ_{Acc}	Accelerometer-based measurement of pitch
a_{linear}	Linear acceleration affecting accelerometer measurements
a_{total}	Total acceleration, measured by the accelerometer
g	Gravitational acceleration constant $\left(9.81\,\frac{m}{s^2}\right)$
$\dot{\phi}_G$	Euler angle rate about the global x-axis, calculated from the gyroscope data
$\dot{\theta}_G$	Euler angle rate about the global z-axis, calculated from the gyroscope data
$\dot{\psi}_G$	Euler angle rate about the global y-axis, calculated from the gyroscope data
ϕ_G	Angle relative to global x-axis, specifically, integrated transformed gyroscope data
θ_G	Angle relative to global z-axis, specifically, integrated transformed gyroscope data

ψ_G Angle relative to global y-axis, specifically, integrated transformed gyroscope data

$\hat{\phi}_k$ Estimation of roll for the current time-step

$\hat{\theta}_k$ Estimation of pitch for the current time-step

$b_{\hat{\phi}_k}$ Bias error associated with the current estimate of roll

$b_{\hat{\theta}_k}$ Bias error associated with the current estimate of pitch

$\hat{\boldsymbol{x}}_k$ Estimated state vector for the current time-step

A State matrix

$\hat{\boldsymbol{x}}_{k-1}$ Estimated state vector for the previous time-step

B Input matrix

$\boldsymbol{u}_k$ Input vector for the current time-step

$\hat{\boldsymbol{y}}_k$ Estimated output vector for the current time-step

C Output matrix

$\dot{\phi}_{G_k}$ Euler angle rate about the global x-axis for the current time-step, calculated from the gyroscope data

$\dot{\theta}_{G_k}$ Euler angle rate about the global x-axis for the current time-step, calculated from the gyroscope data

$\boldsymbol{y}_k$ Measurement vector

ϕ_{Acc_k} Accelerometer-based measurement of roll for the current time-step

θ_{Acc_k} Accelerometer-based measurement of pitch for the current time-step

dt Time-step between discrete measurements

K_k Kalman Gain

$\hat{\boldsymbol{x}}_k^-$ A priori predicted state estimates

P_k^- A priori estimate of the error covariance matrix for the current time-step

P	Error covariance matrix
Q	Process noise covariance matrix
R	Measurement noise covariance matrix
I	Identity matrix
q_r	Quaternion real component
q_i	Quaternion unit vector pointing along the *i*-axis
q_j	Quaternion unit vector pointing along the *j*-axis
q_k	Quaternion unit vector pointing along the *k*-axis
$\angle Leg$	Absolute leg angle, measured from vertical

CHAPTER 1

INTRODUCTION

More than two million Americans suffer from myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) [1], with an annual cost of $24 billion [2]. While our understanding of the etiology of ME/CFS is currently incomplete, studies have shown that the disease commonly occurs following viral infection and other acutely stressful events, impacting women more frequently than men at a rate of 6:1 [3]. A recent upsurge in ME/CFS research has led to an understanding of the disease's core symptoms: (1) fatigue as a response to physical exertion, (2) post-exertional malaise (PEM), (3) unrefreshing sleep, (4) cognitive impairment, and (5) orthostatic intolerance (OI) [2]. Individuals suffering from ME/CFS experience drastic fluctuations of symptoms in terms of both intensity and variety, including numerous manifestations beyond core symptoms [4]. This variation makes assessing treatment efficacy particularly challenging, hence the disease's poor test-retest reliability [5]. While the scientific community's understanding of ME/CFS is continuously improving, no cure has been discovered. Patients often suffer from ME/CFS for years and sometimes even until death [6].

PEM causes individuals with ME/CFS to become disproportionately fatigued following mental or physical exertion. It is regarded as the distinctive symptom of ME/CFS [7]. The onset of PEM often lags behind the related exertion by multiple days

[8]. As a result of this symptom, individuals with ME/CFS generally suffer from a diminished exercise tolerance and can have difficulty performing mundane tasks such as routine cleaning, grocery shopping, and even showering.

OI refers to the onset of symptoms which occur when standing upright; these symptoms can be alleviated by reclining. While the exact cause of OI remains unknown, Dr. van Campen's research suggests that significantly lower blood volume is common among adults with ME/CFS who experience OI [9]. Sub-normal blood volume is likely the cause of the circulation-related issues many ME/CFS patients endure such as dizziness, headaches, weakness, and nausea. These are the most common symptoms of OI, all of which occur as a result of prolonged upright posture.

As is typical for many who suffer from ME/CFS, PEM and OI are often coupled with unrefreshing sleep, leading to a housebound or even bedridden existence. ME/CFS patients commonly report a constant feeling of sleepiness, yet they often wake feeling unrefreshed despite the lack of surface-level sleep-related symptoms such as reduced overall sleep times, interrupted sleep, or increased sleep onset latency [10]. Dr. Jackson notes that the sleepiness commonly present among ME/CFS patients could simply be a result of an individual tendency to monitor sleep cycles closely. She also proposes that affected individuals may experience arousals during sleep that go undetected by current sleep scoring methods. Despite an insufficient understanding about why unrefreshing sleep symptom occurs, those with ME/CFS go on with their lives without the restorative benefits of sleep.

The FDA has yet to approve any treatments—physical or pharmaceutical—for ME/CFS. To some extent, this lack of FDA-approved treatments is due to a lack of

validated efficacy endpoints [2]. In clinical trials, efficacy endpoints are used to reliably monitor the improvement of subjects as a result of a prescribed treatment. In the case of ME/CFS, some potential efficacy endpoints are (1) the maximum amount of time an individual can spend upright and active without triggering PEM and (2) the extent to which a patient feels refreshed following eight hours of uninterrupted sleep. In recent years, researchers have developed some ME/CFS efficacy endpoints using blood tests [11] and other invasive methods [12]. The central focus of this research is to develop efficacy endpoints using a completely non-invasive, inertial measurement-based approach. More accessible efficacy endpoints will provide a way to properly evaluate potential treatments for ME/CFS, especially if these endpoints correspond to the disease's core symptoms.

Researchers at the Bateman Horne Center (BHC) in Salt Lake City, Utah recently discovered an endpoint that shows promise as a reliable assessment of functional impairment among patients with ME/CFS. In studies conducted by the BHC, subjects were asked to fill out questionnaires, identifying how much time they spent upright during the previous 24 hours. Responses to these questionnaires indicate that patients diagnosed with severe ME/CFS spend less than five hours in an upright position. In comparison, those diagnosed with moderate ME/CFS spend between five and eight hours in an upright position [transcript in progress]. The BHC refers to this measurement of uprightness as hours of upright activity (HUA).

Due to its strong correlation with PEM, HUA is a simple and clear way to gauge disease severity among individuals with ME/CFS [transcript in progress]. While HUA is a valuable efficacy endpoint, its deficiencies are significant. The primary weakness of

HUA is the inaccuracy of its current data collection method—questionnaire [13]; it is unreasonable to expect patients to accurately recall the amount of time they spent in an upright position the previous day. Another weakness of HUA is the low resolution offered by the measurement. “Hours” of upright activity is just that, a measurement recorded as whole integers in units of hours. Due to the inaccuracy and low resolution of data collected from HUA questionnaires, the only way to obtain a measurement of upright activity with a higher level of precision involves significant alterations to the current measurement process.

In order to validate treatments for ME/CFS, an objective and accurate measurement is required. Currently, pharmaceutical companies and independent researchers struggle to receive FDA-approval for their treatments due to inaccurate measurements, such as HUA. Improving the resolution of BHC’s upright activity assessment will allow them (and others who use the improved process) to objectively and accurately measure the efficacy of their ME/CFS treatments.

To address HUA’s shortcomings, an improved method for evaluating upright activity is proposed. Inertial measurement units (IMU) are commonly used to gauge activity level, e.g., Apple watches, Fitbits, pedometers. Many of these and other wearables use IMUs to detect workouts automatically [14]. However, these commercially available devices are calibrated to help healthy people achieve their fitness goals. Using these devices to monitor disease states is challenging because we are unable to tailor the higher-level algorithms used by these devices to more appropriately fit our unique application. No research has used IMUs to directly measure an efficacy endpoint for ME/CFS. Using an IMU, it is possible to continuously and accurately measure upright

activity, thus providing an effective method to assess disease severity among individuals with ME/CFS. By continuously measuring the uprightness of the lower legs, we can obtain a measurement referred to as UpTime. The advantages of this approach are two-fold. The first advantage is that healthcare providers will no longer need to rely upon the accuracy of a patient's memory to approximate upright activity. The second advantage comes from increasing the resolution of the measurement from hours—HUA—to seconds— UpTime.

The primary goal of our research is to validate an improved method to assess upright activity. As a secondary goal, we aim to provide an efficacy endpoint corresponding to unrefreshing sleep. Experience has led researchers to believe that individuals with ME/CFS spend more time "awake" during sleeping hours and maintain consistently lower heart rate variability (HRV) than their healthy counterparts [10,15]. Furthermore, the most vital factors related to refreshing sleep are sleep continuity [16] and time spent in restorative—REM and deep—sleep cycles. To formally evaluate UpTime and these sleep-related measures as efficacy endpoints for ME/CFS disease severity, this thesis evaluates the results of a study wherein a healthy control group and an experimental group of ME/CFS patients were outfitted with Oura Rings and Shimmers—the IMU used for this study—for six consecutive days. During this six-day period, the Oura Rings tracked sleep patterns and recorded nightly HRV scores, while the Shimmers continuously measured UpTime. Statistical tests and other comparisons were used to evaluate the correlation between data collected by the Oura Ring and unrefreshing sleep. We review the results of this study later in this document.

Our research simplifies symptom severity evaluation among patients with

ME/CFS. We expect that the Oura Ring and the Shimmers together will provide a platform to more easily measure efficacy endpoints related to the core symptoms of ME/CFS. As a result, assessing the long-term efficacy of treatments for patients with ME/CFS will significantly improve the evaluation of disease severity in terms of both ease and accuracy. These changes will enable the development of effective treatments, thus providing a path to recovery for individuals struggling with ME/CFS.

In this thesis, we describe the development and testing of an IMU-based measurement of UpTime. Chapter 2 details how a Kalman filter is used to convert raw IMU data (recorded by the Shimmer) to angle estimates. In Chapter 3, the accuracy of these angle estimates is confirmed by comparison to a motion capture system. Chapter 4 details the methods by which the Shimmers were prepared for clinical use. The BHC used both the prepared Shimmers and the Oura Rings to collect data in a small case-control study. Chapters 5 and 6 detail the data processing and evaluation methods used to compare the study results—including both UpTime scores and Oura Ring measurements. We conclude this thesis with final comments about the value of these measurements as efficacy endpoints and recommend pathways for directing future work.

CHAPTER 2

IMPLEMENTING A SENSOR FUSION ALGORITHM FOR THE IMU

2.1 Introduction

As expressed in Chapter 1, the chief objective of this research is to determine UpTime by continuously measuring lower leg orientation. Calculating UpTime is a two-step process. First, we measure lower leg angle by filtering IMU data, as described in this chapter. Second, we evaluate this measured angle to determine if the leg is upright. Distinguishing leg uprightness is crucial because it relates to the HUA questionnaire, which quantifies daily time spent in upright postures (see Figure 2.1).

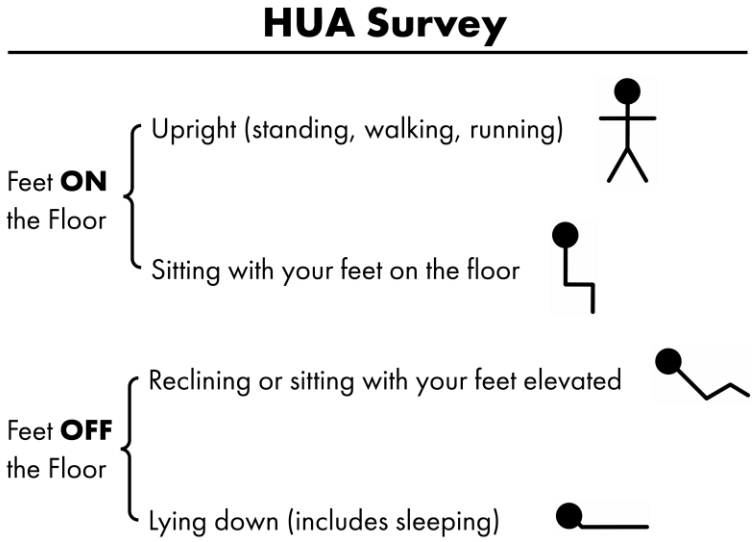

Figure 2.1. HUA survey. Postures are separated into two groups: (1) upright with feet on the floor and (2) not upright with feet off the floor. HUA is calculated by summing the time spent with feet on the floor—based on patient estimates.

Specifically, the HUA questionnaire seeks to identify how much time ME/CFS patients spend with their feet on the floor. Our goal is to replace the HUA questionnaire with an IMU-based measurement of upright activity called UpTime. We chose to place an IMU on each lower leg rather than the foot. Lower leg angles allow us to accurately assess whether the feet are on the floor (lower legs vertical) or off the floor (lower legs reclined/horizontal) while maximizing user comfort. By minimizing required sensors and maximizing user comfort, we hope that IMU-based UpTime will become a widely adopted efficacy endpoint for various treatments. User comfort during passive data collection is especially important over long periods of time; to fully evaluate the efficacy of ME/CFS interventions, clinics will need to track UpTime for several months.

The Shimmer, a commercially available IMU, was selected for use in this research due to its small and lightweight design, data logging capacity, ample battery life, and previous use in related work [17,18]. Using an internal SD card, the Shimmer can simultaneously record accelerometer, gyroscope, and magnetometer data at a range of specified frequencies for extended periods. Accurate orientation estimations can be obtained using only the accelerometer and gyroscope (refer to Chapter 3 for verification of this claim). Excluding the Shimmer's magnetometer from our research simplified the data filtering process and extended battery life, thereby lengthening the duration of possible data collection periods.

Combining data from multiple sensors, otherwise known as sensor fusion, has been extensively reviewed in the literature [19]. Sensor fusion reduces measurement uncertainty by merging data from multiple sensors. This chapter describes the sensor

fusion algorithm we used to merge the Shimmer's raw accelerometer and gyroscope data to determine lower leg angle, measured from vertical.

2.2 Pre-Conditioning Data

For this research, measuring lower leg angle requires measurements of roll (ϕ, angle relative to a global x-axis) and pitch (θ, angle relative to a global z-axis). Rotation about the y-axis, yaw (ψ), offers no useful information about the lower leg's angle relative to vertical; this can be seen in Figure 2.2.

Raw acceleration data collected by the Shimmer has three components—acceleration along the x-axis (a_x), acceleration along the y-axis (a_y), and acceleration along the z-axis (a_z).

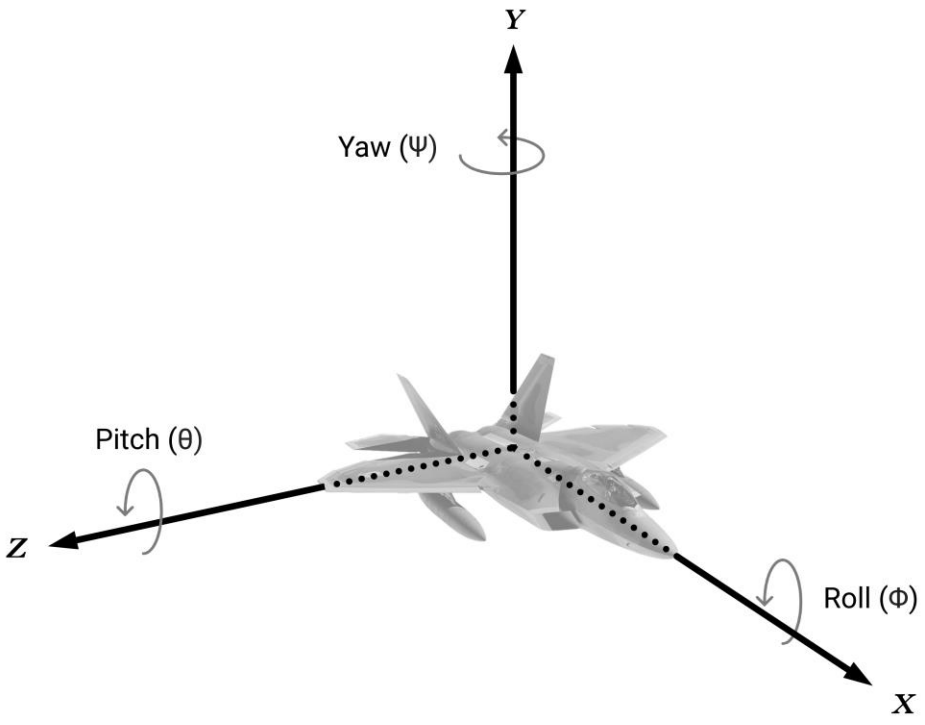

Figure 2.2. Illustration of a 3-dimensional coordinate system with roll, pitch, and yaw.

These measurements are recorded in units of $\frac{m}{s^2}$. Raw gyroscope data is recorded with each component measuring the rate at which the Shimmer rotates around its body frame of reference, yielding one gyroscopic measurement about the x-axis (p), one about the y-axis (q), and one about the z-axis (r). Measurements taken by the Shimmer's gyroscope are recorded in units of $\frac{deg}{s}$.

Based on the accelerometer data alone, measurements of roll (ϕ_{Acc}) and pitch (θ_{Acc}) can be derived using trigonometry [20]. Note that the subscript "Acc" indicates that these measurements are calculated using accelerometer data. These accelerometer-based measurements are given in equations 2.1 and 2.2.

$$\phi_{Acc} = \tan^{-1}\left(\frac{a_y}{a_x^2 + a_z^2}\right) \tag{2.1}$$

$$\theta_{Acc} = \tan^{-1}\left(\frac{-a_x}{a_y^2 + a_z^2}\right) \tag{2.2}$$

Equations 2.1 and 2.2 allow us to accurately find the direction of gravity while the IMU rotates in place. However, these angle measurements become inaccurate as soon as the device begins to move in space. Applied external linear accelerations cause movements (a_{linear}), which change the total acceleration (a_{total}) read by the accelerometer, as shown in equation 2.3.

$$a_{total} = \sqrt{a_x^2 + a_y^2 + a_z^2} = g + a_{linear} \tag{2.3}$$

Linear accelerations likely to occur during our study included walking, running, and even driving a car. Passing accelerometer measurements through a low-pass filter will remove linear acceleration inputs, thus isolating gravity. Low-pass filtering also helps to clean up raw accelerometer measurements, as this sensor type tends to record noisy data. Unfortunately, low-pass filtering accelerometer data causes slight signal attenuation, especially during high-frequency motion such as walking—an effect that can be mitigated by fusing accelerometer data and gyroscope data. A critical strength of accelerometers is that their readings do not drift over time. As a result, they can be used to reliably evaluate orientation over long periods.

The Shimmer's gyroscope measures angular velocity with respect to a frame of reference which is affixed to the device itself. Because we are interested in the device's orientation relative to the earth, we convert the gyroscope's body frame angle rates (p, q, and r) to Euler angle rates ($\dot{\phi}_G$, $\dot{\theta}_G$, and $\dot{\psi}_G$), which are measured with respect to a global or inertial frame of reference. This process is done using the following transformation:

$$\begin{bmatrix} \dot{\phi}_G \\ \dot{\theta}_G \\ \dot{\psi}_G \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin(\theta) \\ 0 & \cos(\phi) & \cos(\theta)\sin(\phi) \\ 0 & -\sin(\phi) & \cos(\theta)\cos(\phi) \end{bmatrix}^{-1} \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{2.4}$$

$$= \begin{bmatrix} p + q\sin(\phi)\tan(\theta) + r\cos(\phi)\tan(\theta) \\ q\cos(\phi) - r\sin(\phi) \\ q\sin(\phi)/\cos(\theta) + r\cos(\phi)/\cos(\theta) \end{bmatrix}$$

By integrating the transformed gyroscope data, we attain accurate angle measurements defining the device's orientation in space ($\phi_G, \theta_G,$ and ψ_G). Unfortunately, integrating gyroscope data amplifies the effect of bias errors, causing drift over time [21].

Due to this weakness, gyroscopes can be used to accurately measure orientation for short periods but are not suited for extended use. Unlike accelerometers, gyroscopes are not affected by linear acceleration and are therefore equally suited for quasi-static and dynamic applications.

On their own, both the accelerometer and gyroscope measurements have weaknesses. Accelerometer data is consistent over extended periods but tends to be noisy and can confound linear acceleration and the effect of gravity. Gyroscope data has inherent bias errors that, when integrated, lead to drifting angle measurements. Despite these weaknesses, it is possible to produce a reliable estimation of orientation with reduced noise and without drift by combining data from both sensors—a process called sensor fusion. The goal of this process is to retain the strengths of each sensor while minimizing their flaws.

2.3 Kalman Filtering

Sensor fusion can be accomplished using several acceptable algorithms. Simple complimentary filters [22], nonlinear complementary filters [23], multiple model adaptive filters [24], and many other techniques [25] have been researched extensively. Of the numerous available sensor fusion techniques, we chose to use a Kalman filter due to the stochastic nature of our collected IMU data. Furthermore, the Kalman filter is computationally cheaper than other nonlinear filtering methods, and we anticipated the collected data to be on the order of multiple gigabytes.

The Kalman filter, invented by Rudolf E. Kalman, is a sensor fusion method designed to minimize estimation error [26]. In our application, a Kalman filter

accomplished this error optimization by iteratively combining roll and pitch measurements provided by the Shimmer's accelerometer and gyroscope (see equations 2.1, 2.2, and 2.4). Based on parameters set during its construction, the Kalman filter provides angle estimates which are more accurate than measurements provided by either sensor.

2.3.1 Defining the State-Space Form of the System

A Kalman filter is a type of optimal state estimator which minimizes the quadratic cost function [27]. The Kalman filter used in this research follows the structure provided in Philip Salmony's "Course on IMU Attitude Estimation" [28]. Our variables of interest—estimated roll $(\hat{\phi}_k)$, estimated pitch $(\hat{\theta}_k)$, and their respective biases $(b_{\hat{\phi}_k}\ and\ b_{\hat{\theta}_k})$—make up the estimated state vector $(\hat{\boldsymbol{x}}_k)$, as shown in equation 2.5.

$$\hat{\boldsymbol{x}}_k = \begin{bmatrix} \hat{\phi}_k \\ b_{\hat{\phi}_k} \\ \hat{\theta}_k \\ b_{\hat{\theta}_k} \end{bmatrix} \tag{2.5}$$

Note that the subscript k is used because all measurements and estimates are discrete, rather than continuous. Furthermore, we use a "hat" over a variable to denote an estimate.

Equations 2.6 and 2.7 show the estimated state vector $(\hat{\boldsymbol{x}}_k)$ and estimated output vector $(\hat{\boldsymbol{y}}_k)$, which form the discrete state-space model of our system:

$$\hat{\boldsymbol{x}}_k = A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_k \tag{2.6}$$

$$\hat{\boldsymbol{y}}_k = C\hat{\boldsymbol{x}}_k \tag{2.7}$$

In equation 2.6, the subscript $k-1$ indicates values corresponding to the previous time-step. This equation also marks the first time we see the input vector, $\boldsymbol{u}_k$, which consists of our gyroscope's transformed angle rates, as shown in equation 2.8.

$$\boldsymbol{u}_k = \begin{bmatrix} \dot{\phi}_{G_k} \\ \dot{\theta}_{G_k} \end{bmatrix} \tag{2.8}$$

Similarly, we built our measurement vector $(\boldsymbol{y}_k)$—which is different from the estimated output vector $(\hat{\boldsymbol{y}}_k)$—using our accelerometer's measurements of roll $\left(\phi_{Acc_k}\right)$ and pitch $\left(\theta_{Acc_k}\right)$ (see equation 2.9).

$$\boldsymbol{y}_k = \begin{bmatrix} \phi_{Acc_k} \\ \theta_{Acc_k} \end{bmatrix} \tag{2.9}$$

The final pieces of the system's discrete state-space model are the A, B, and C matrices. These matrices—which define the relationships between estimated state vector, estimated input vector, and estimated output vector—are shown in equations 2.10, 2.11, and 2.12.

$$A = \begin{bmatrix} 1 & -dt & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -dt \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.10}$$

$$B = \begin{bmatrix} dt & 0 & 0 & 0 \\ 0 & 0 & dt & 0 \end{bmatrix} \tag{2.11}$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{2.12}$$

2.3.2 The Kalman Filter Equations

It is necessary to model the system in state-space form, as these relationships are used to implement a Kalman filter. The discrete Kalman filter equation combines equations 2.6 and 2.7 to form equation 2.10.

$$\hat{\boldsymbol{x}}_k = A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_k + K_k\left(\boldsymbol{y}_k - C(A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_k)\right) \tag{2.10}$$

In this equation, the Kalman Gain, K, adjusts to compensate for differences between the inputs (angles measured by the gyroscope) and the previous predicted states. Estimates of the current states, $\hat{\boldsymbol{x}}_k$, are predicted using the state estimates from the previous timestep, $\hat{\boldsymbol{x}}_{k-1}$, and the current inputs, $\boldsymbol{u}_k$ (see equation 2.11). This equation is the first of five specific Kalman filter equations (KFE) needed to implement a Kalman filter algorithm.

KFE #1: $$\hat{\boldsymbol{x}}_k^- = A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_k \tag{2.11}$$

Each of these five KFEs falls into one of two categories: (1) prediction and (2) update. As shown in equation 2.11, KFE #1 falls into the prediction category because it uses estimates from the previous timestep to predict estimates for the current timestep. Predicted state estimates are calculated without using the current measurements and are therefore referred to as *a priori*, literally “something that can be known without prior

experience." A priori estimates are denoted using a superscript minus—e.g., $\widehat{\boldsymbol{x}}_k^-$.

During the prediction step, we also define the a priori estimate of the error covariance matrix (P_k^-). The formula for this estimate is KFE #2 and is shown in equation 2.12.

KFE #2: $$P_k^- = AP_{k-1}A^T + Q \tag{2.12}$$

The error covariance matrix, P, is initially set by us (as shown in equation 2.13) but is modified during each iteration through the Kalman filter.

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.13}$$

The diagonal values of this matrix indicate how confident we are that the previous state values are correct. When setting the initial values, lower numbers indicate greater certainty in the initial guess. As the Kalman filter iterates through each discrete set of measurements, it determines the confidence level of the previous timestep's state estimates and adjusts the covariance matrix accordingly.

Covariance matrices are used as tuning parameters for the Kalman filter, which allow us to estimate lower leg angles more accurately. Our Kalman filter has three covariance matrices: (1) an error covariance, P, (2) a process noise covariance, Q, and (3) a measurement noise covariance, R. In all three cases, we cannot know specifics about the error or noise involved in the Kalman filtering process. Fortunately, we can assume

that these sources of measurement error are zero-mean Gaussians, i.e., they are normally distributed with their means centered at zero. This assumption allows us to account for each source of error simply by defining a corresponding covariance matrix.

The process noise covariance, Q, (seen previously in equation 2.13) is associated with the noise in the states. Corresponding to the number of variables in our state vector, Q is a 4×4 diagonal matrix. The diagonal values of this matrix are set depending on how sure we are about the model dynamics, where smaller values indicate a more accurate model. Because we are using the Kalman filter generally, rather than modeling a specific system, we set the diagonal values of Q to a low value of 0.01 (see equation 2.14).

$$Q = \begin{bmatrix} .01 & 0 & 0 & 0 \\ 0 & .01 & 0 & 0 \\ 0 & 0 & .01 & 0 \\ 0 & 0 & 0 & .01 \end{bmatrix} \tag{2.14}$$

The measurement noise covariance, R, is a 2×2 diagonal matrix corresponding to the inherent noise levels of sensors used. In this matrix, the diagonal values are set depending on the noise of the sensors used in the system, with larger values indicating greater sensor noise. The Shimmer's low-noise accelerometer is a KXRB5-2042 device from Kionix. The Shimmer's gyroscope is the MPU-9150 chip from Invensense. Based on the datasheets for both sensors, we chose to set the diagonal values of R to 10—a relatively high value (see equation 2.15).

$$R = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 10 \end{bmatrix} \tag{2.15}$$

The second step of Kalman filtering is the update step. The first thing to update is the Kalman Gain for the current timestep (K_k) using the a priori error covariance (P_k^-), the output matrix (C), and the measurement noise covariance (R). This is KFE #3 (shown in equation 2.16).

KFE #3: $$K_k = \frac{P_k^- C^T}{C P_k^- C^T + R} \tag{2.16}$$

Now that the requisite predictions have been made, we can compare the a priori estimated states against the current sensor measurements. KFE #4 does exactly this by simplifying equation 2.10 using the a priori state estimate. This simplification is shown in equation 2.17.

KFE #4: $$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + K_k(\boldsymbol{y}_k - C\hat{\boldsymbol{x}}_k^-) \tag{2.17}$$

KFE #5 updates the error covariance, P_k, (as shown in equation 2.18) for use in the next iteration of the Kalman filter; this is the final KFE.

KFE #5: $$P_k = (I - K_k C)P_k^- \tag{2.18}$$

By cycling through the prediction and update KFEs, the Kalman filter will

compute optimal, unbiased estimates of lower leg angles with minimal variance. This process is repeated until an accurate, Kalman-filtered angle estimate has been calculated for each discrete set of IMU measurements. For clarity, a summary of all five KFEs is presented in Table 2.1.

The Kalman filter computes optimized roll and pitch estimates for each set of discrete sensor measurements. Each pair of roll and pitch estimates is then combined into an estimate of lower leg angle measured from vertical ($\angle Leg$). This process is done using quaternions.

Table 2.1. Kalman filter equations, separated into two categories: (1) predict and (2) update.

Predict	
KFE #1:	$\hat{\boldsymbol{x}}_k^- = A\hat{\boldsymbol{x}}_{k-1} + B\boldsymbol{u}_k$
KFE #2:	$P_k^- = AP_{k-1}A^T + Q$
Update	
KFE #3:	$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_k^- + K_k(\boldsymbol{y}_k - C\hat{\boldsymbol{x}}_k^-)$
KFE #4:	$K_k = \frac{P_k^- C^T}{CP_k^- C^T + R}$
KFE #5:	$P_k = (I - K_k C)P_k^-$

A quaternion is a vector of four elements (q_r, q_i, q_j, and q_k) used to define a rotation within a three-dimensional coordinate system. Together, elements q_i, q_j, and q_k

form a vector, and q_r defines how much to rotate about that vector. The equations for this rotation are included below (equations 2.19-2.23).

$$q_r = \cos\left(\frac{\hat{\phi}}{2}\right)\cos\left(\frac{\hat{\theta}}{2}\right) \tag{2.19}$$

$$q_i = \sin\left(\frac{\hat{\phi}}{2}\right)\cos\left(\frac{\hat{\theta}}{2}\right) \tag{2.20}$$

$$q_j = \cos\left(\frac{\hat{\phi}}{2}\right)\sin\left(\frac{\hat{\theta}}{2}\right) \tag{2.21}$$

$$q_k = -\sin\left(\frac{\hat{\phi}}{2}\right)\sin\left(\frac{\hat{\theta}}{2}\right) \tag{2.22}$$

$$\angle Leg = \tan^{-1}\left(\frac{\sqrt{q_i^2+q_j^2+q_k^2}}{q_r}\right) \tag{2.23}$$

2.4 Conclusion

In summary, the Kalman filter is an iterative process that combines a given state-space model with recorded data to optimally estimate a value that cannot be directly measured. In our case, the value that cannot be directly measured is lower leg angle.

This Kalman filter was implemented using MATLAB 2019a (Appendix F). With the Kalman filter fully defined, the next step was to collect sample Shimmer data and evaluate the accuracy of the filtered estimates. This evaluation is detailed in the next chapter.

CHAPTER 3

SHIMMER ACCURACY CONFIRMATION STUDY

3.1 Introduction

This research validates a surrogate measure of upright activity for clinical use. According to recent research conducted by the Bateman Horne Center (BHC) in Salt Lake City, Utah, daily upright activity corresponds to ME/CFS disease severity [transcript in progress]. In their research, the BHC gauged upright activity using lower leg orientation. However, the BHC collected data using questionnaires, which is a crude method to approximate actual leg orientation. Replacing the questionnaires with Shimmer devices will greatly improve the accuracy of leg orientation data. For this reason, we chose to outfit individuals with a pair of Shimmers, one on each lower leg, to obtain a continuous measurement of lower limb orientation. This IMU-based approach allows for a much higher resolution measurement of upright activity because the Shimmer can record data at frequencies over 2 kHz.

Human motion is limited to ultra-low frequencies—less than 10 Hz [29]. To protect our signal from aliasing, we needed a sample rate at least twice this limit [30]. Therefore, a sample rate of at least 20 Hz was required; however, collecting data above this minimum sample rate would improve filter performance. For this application, we selected a frequency of 30 Hz to maximize battery life while maintaining the detail of the

data collected by the Shimmers.

In Chapter 2, the function of a Kalman filter was explained in detail. Because we cannot measure limb orientation directly, we used a Kalman filter to indirectly obtain an optimal estimation instead. For this reason, it was necessary to confirm the accuracy of our data collection and filtering methods. This accuracy confirmation study used a VICON nine-camera motion capture system as the baseline to evaluate the accuracy of the Kalman filtered IMU data.

3.2 Methods

3.2.1 Operating Specifications—Shimmer Device

As discussed in Chapter 2, the Kalman filter employed in this research estimates orientation using three-dimensional accelerometer and magnetometer data as inputs. For the accuracy confirmation study, researchers configured the Shimmers using Shimmer Sensing's ConsensysBasic software package—build v1.6.0. All Shimmers were programmed to record data at a sample rate of 30 Hz. The Shimmer's low-noise accelerometer was set to an output range of ± 2 g, and the gyroscope was set to an output range of ± 500 deg/sec.

3.2.2 Operating Specifications—VICON Motion Capture System

A nine-camera VICON motion capture system, consisting of seven VICON MXF40 cameras and two VICON MXF20 cameras, was used to collect the baseline data in this accuracy confirmation study (see Figure 3.1). This VICON system is located in the Large Robotics Lab at the University of Utah (Building MEK, Room 0112). Optical

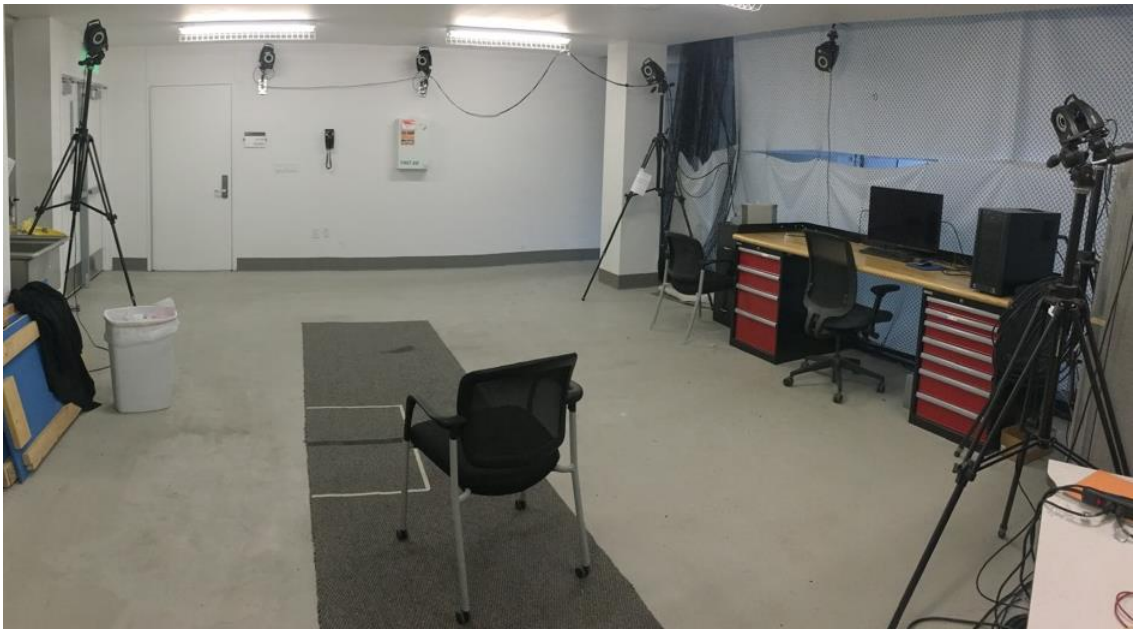

Figure 3.1. Image of VICON motion capture system used in this accuracy confirmation study. The VICON system included seven VICON MXF40 cameras and two VICON MXF20 cameras, each running at 60 Hz. The capture volume was approximately 8’ x 10’ x 2’.

motion capture uses an array of infrared-emitting cameras to locate retroreflective markers in three-dimensional space. For this study, we used VICON’s software package, NEXUS 2.9.2, to operate the VICON motion capture system. Each of the nine cameras recorded motion capture data at a sample rate of 60 Hz.

3.2.3 Markerset—Truncated Rizzoli Lower Body Protocol

When tracking optical motion data, the precise placement of retroreflective markers is critical. Each body segment requires at least three markers to properly collect all six degrees of freedom—three rotational and three translational. While the underlying principles of marker placement are relatively simple, using standard markersets is common in the field of motion capture. Over time, these standard markersets have

amassed a proven history of use in established motion capture labs and are preprogrammed into many motion capture software packages.

The Movement Analysis Laboratory in the Rizzoli Orthopedics Institute has developed multiple standard markerset protocols for tracking various segments of the body. Because our accuracy confirmation study was based solely around the biomechanics of the lower leg, we chose to use a truncated form of the Rizzoli Lower Body Protocol (26) [31]. This markerset was designed to track the motion of the lower body from the pelvis down. In order to fit our needs more closely, we removed all markers unrelated to the tracking of the lower leg and foot segments. This simplification produced the markerset shown in Figure 3.2.

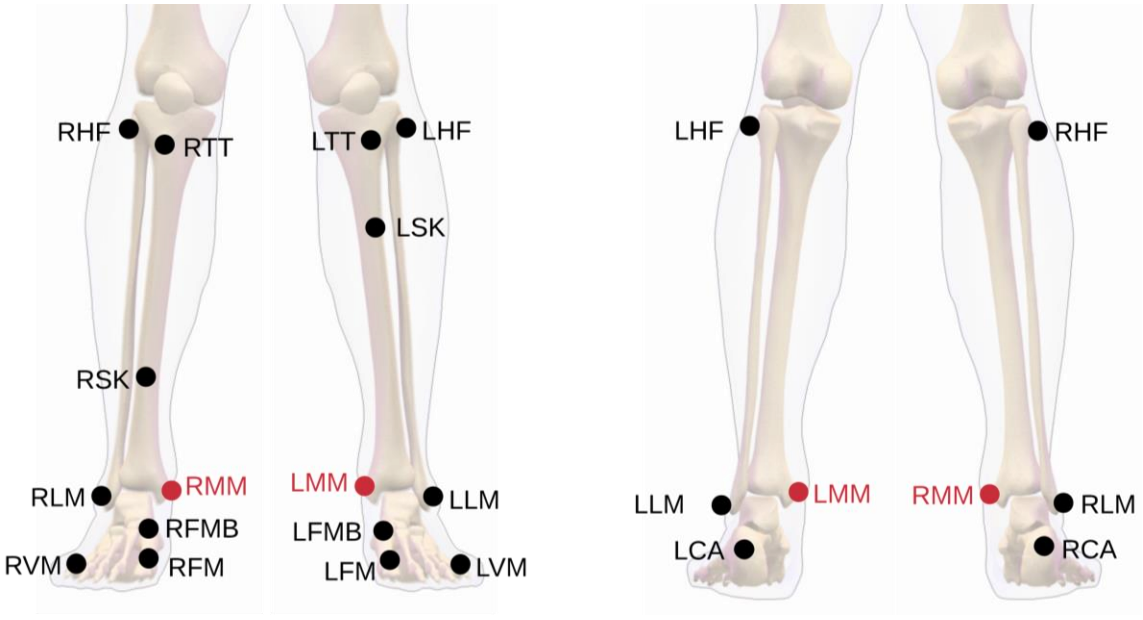

Figure 3.2. Diagram showing the truncated form of the Rizzoli Lower Body Protocol Markers. See OptiTrack's website for marker descriptions [31].

3.2.4 Accuracy Confirmation Study Description

3.2.4.1 Prescribed Positions and Motions

With the VICON Motion Capture system and Shimmers prepared, we developed a simple choreography, or sequence of postures, for the subjects to follow while collecting data. The IMU-based system validated in this research was and will continue to be used for diverse groups of people. With this purpose in mind, we chose to gather data from a small group of subjects with varying body shapes and gait cycles. Three subjects participated in the study; their characteristics are listed in Table 3.1.

One at a time, each subject was instructed to take a seated position with both lower legs vertical. Next, we began collecting data on both systems—VICON and Shimmer. The subject then stomped both feet against the ground to provide a reference point, which was used during post-processing to align the IMU and motion capture data. Finally, the subject followed a series of postures (shown in Figure 3.3), holding each for five seconds. This sequence of postures was explicitly developed to push the limits of the Shimmer's motion capture abilities and encompass the full range of lower leg angles that would be seen in a week-long study, from vertical to horizontal.

Table 3.1 Subject characteristics for accuracy confirmation study—all are healthy.

Subject	Age (years)	Gender	Height (meters)	Weight (kg)
1	51	Female	1.72	65.8
2	27	Male	1.80	72.6
3	26	Female	1.62	49.0

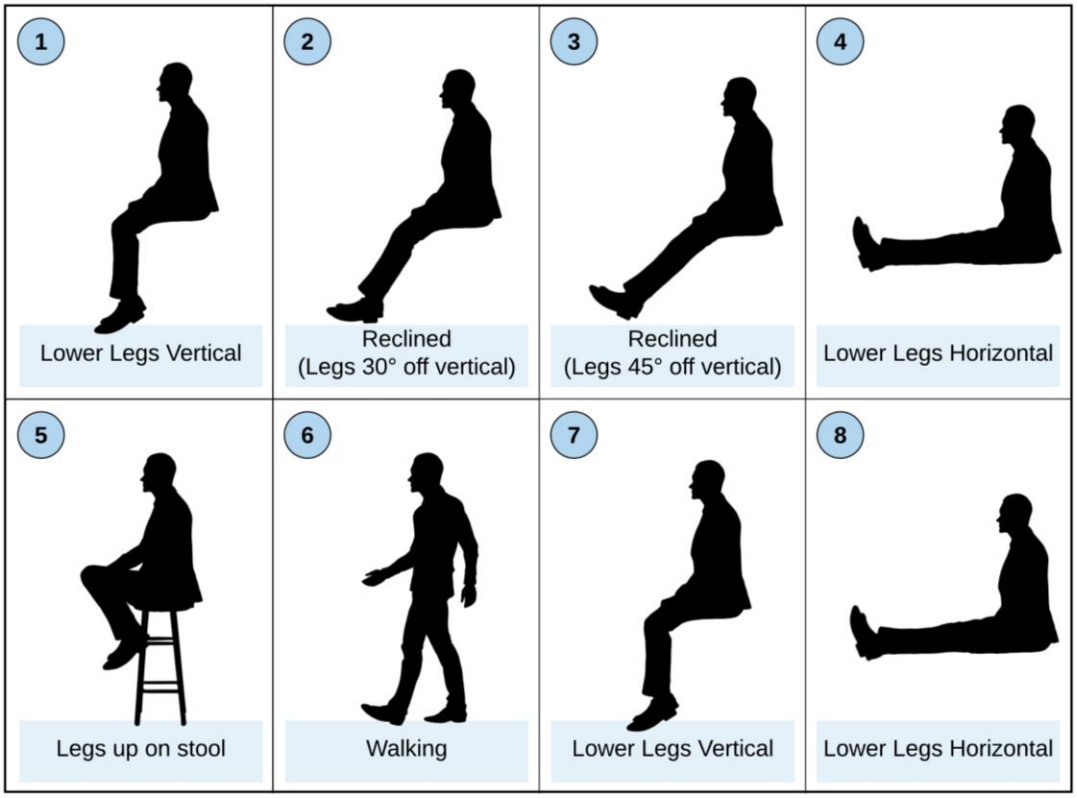

Figure 3.3. Sequence of postures imitated by subjects during accuracy confirmation study.

3.2.5 Calculation of UpTime

The goal of this accuracy confirmation study was to compare measurements obtained from the Shimmer and VICON systems to assess the accuracy of the former. To assess UpTime using the data collected, it was necessary to first calculate lower limb angle. As described in Chapter 2, the Kalman filter was used to calculate lower limb angle from the Shimmer data. To calculate lower limb angle from the VICON data, each choreographed trial was exported as a comma-separated values (CSV) file. Two markers per leg, LHF/RHF and LLM/RLM (see Figure 3.2), were used as the endpoints of two virtual lines. The angles of these virtual lines were then determined using basic principles of trigonometry. When comparing VICON angles to Shimmer angles, root mean squared

error (RMSE) calculations showed that the two measurements differed by an average of 0.53 degrees for all three subjects (see Figure 3.4). RMSE was 0.80 degrees for subject 1, 0.13 for subject 2, 0.66 for subject 3. Most of this error occurred during the walking sequence from 30 to 40 seconds. Having collected angle data from both systems—VICON and Shimmer—we proceeded to calculate UpTime using MATLAB 2019a software.

A custom MATLAB function was created to calculate UpTime. This UpTime function uses the angle data from both legs to calculate the amount of time spent in an upright position. This function requires three inputs: (1) angle vector for the left leg in

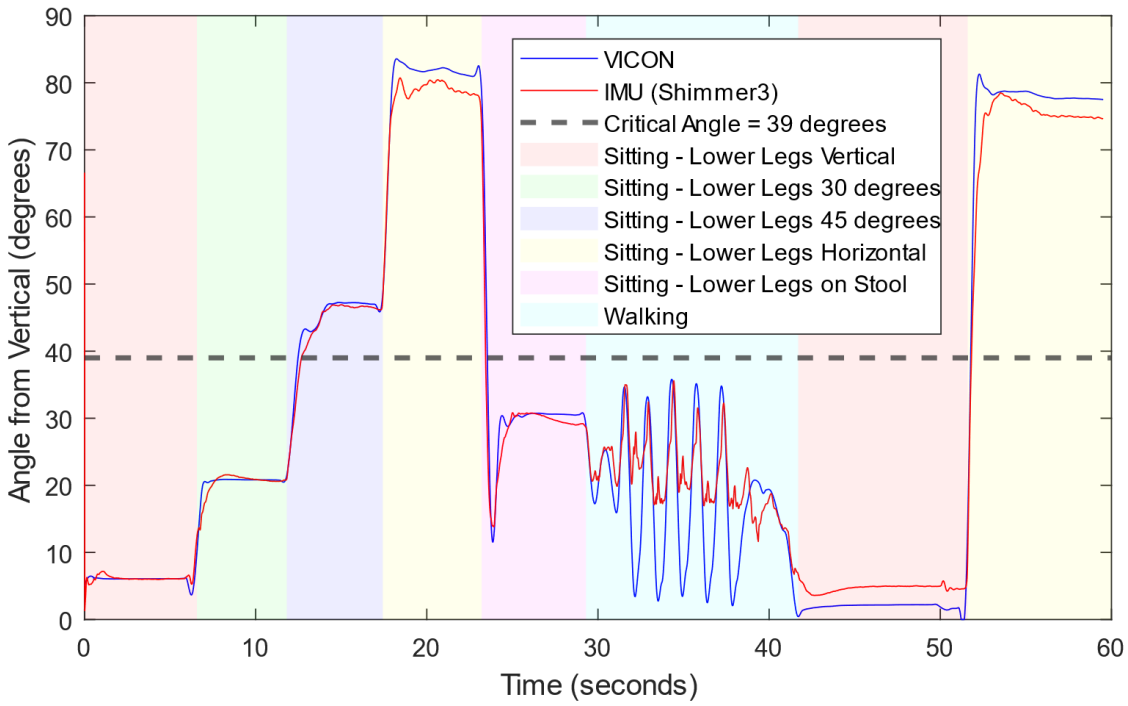

Figure 3.4. Comparison of angular data from VICON and Shimmers for one subject. Most of the error is due to dynamic activity, between 30 and 40 seconds, when the subject was walking. A lower leg angle below the critical angle (indicated by a horizontal line) means the subject was upright. Likewise, a lower leg angle above the critical angle means the subject was not upright.

units of degrees, (2) angle vector for the right leg in units of degrees, and the (3) sample rate of both angle vectors—sample rate must be the same for both legs—in units of Hz. Within the body of the function, the critical angle is defined in units of degrees. The critical angle is used to determine if the lower leg is upright or not upright (see Figure 3.5). If the measured angle of the lower leg is less than the critical angle, then the leg is classified as upright. If the measured angle of the lower leg is greater than the critical angle, then the leg is classified as not upright.

Due to a lack of pre-existing data, it was difficult to choose a specific critical angle that accurately differentiates all angles as either upright or not upright. For this reason, our UpTime function permits the user to define the critical angle as any angle between zero and 90 degrees from vertical. We chose a critical angle of 39 degrees for

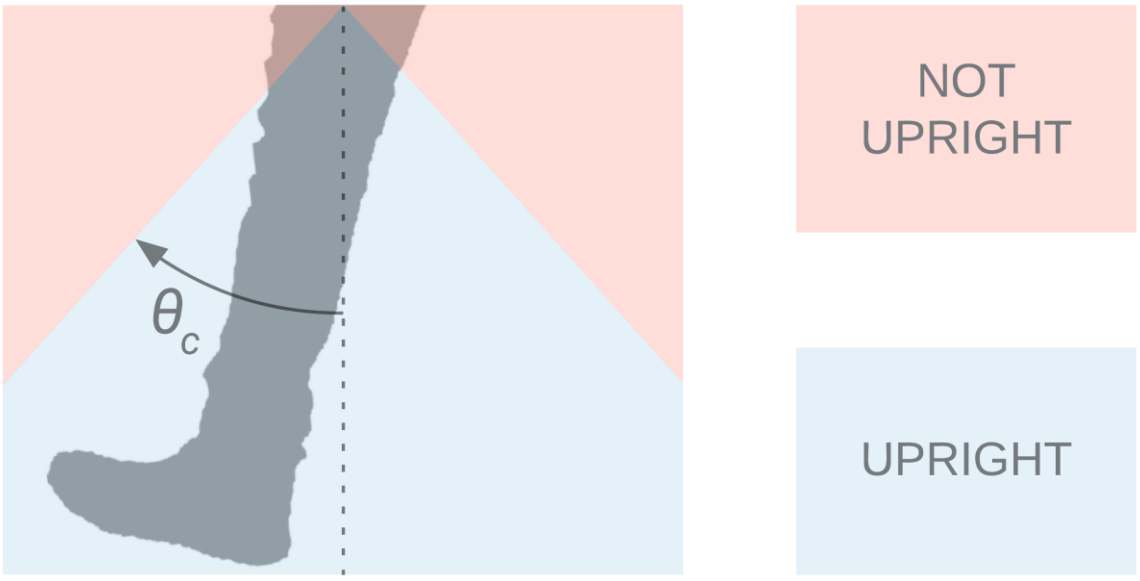

Figure 3.5. The angle of each lower leg is compared to the critical angle, θ_c, to determine uprightness. In this figure, the lower leg would be classified as upright because the leg angle is less than the critical angle. If the lower leg was horizontal, then the angle of the lower leg would exceed the critical angle and would, therefore, be classified as not upright.

this study. This decision was made based on data collected by researchers at the BHC. Their data indicates that individuals with ME/CFS find relief from symptoms by elevating the lower legs. Based on this knowledge, we knew that a critical angle close to 45 degrees would allow us to identify times when diseased individuals were seeking relief. Our rationale for selecting a critical angle of 39 degrees is provided in Chapter 6 of this thesis.

At each point in time, both Shimmers—one per leg—record data from their accelerometers and gyroscopes. For each Shimmer, the recorded data for each trial period is an N x 8 matrix, with each row corresponding to a specific point in time when the row of data was recorded. The eight columns correspond to (1-2) the date and time the data was recorded, (3-5) three axes of accelerometer data, and (6-8) three axes of gyroscope data. Each matrix is then converted into a single vector of lower leg angle data, with each value corresponding to a row in the Shimmer matrix. Finally, the UpTime function takes these vectors of leg angle data and classifies each leg as upright or not upright at each recorded point in time. Based on this classification, an integer is assigned to each value in the UpTime vector. If neither leg is upright, an integer of zero is assigned. If only one leg is upright, an integer of one is assigned. If both legs are upright, an integer of two is assigned. After both leg angle vectors are evaluated, all values in the UpTime vector are combined to form one total measurement of UpTime—which is expressed as a percentage of the day spent with the legs in an upright orientation. Both the total UpTime measurement and the UpTime vector are outputs of the UpTime function (see Figure 3.6).

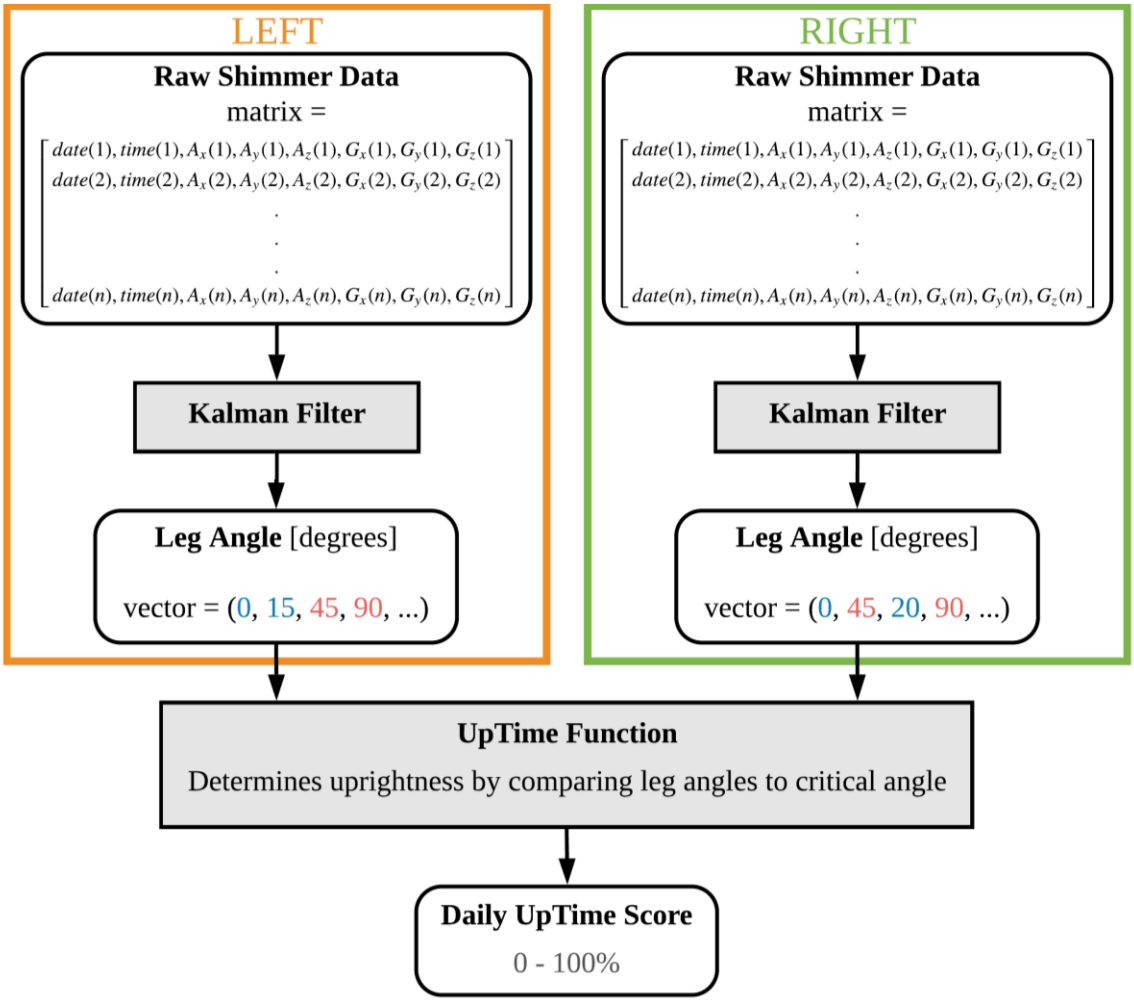

Figure 3.6. Flow chart illustration of the UpTime function. This example shows how two leg angle vectors are used to create the UpTime vector and total measurement of UpTime. Four leg angles have been arbitrarily created to illustrate all possible combinations of leg angles—both legs upright, one leg upright (left or right), and neither leg upright.

3.3 Results & Discussion

The UpTime function was used to calculate a value of UpTime corresponding to both the VICON system and the Shimmer. When reviewing the results for all three subjects (Table 3.2), we found that the Shimmer has an average error of 1.88% when compared to the VICON system, which represented 100% accurate measurements.

Table 3.2 UpTime Data for both the VICON system and the Shimmer. UpTime has been evaluated for one-minute of data in this table. Note that the final study will evaluate UpTime for multiple 24-hour periods.

System	UpTime (%)		
	Subject 1	Subject 2	Subject 3
VICON	29.61	31.47	24.79
Shimmer	29.74	30.67	25.45
Error	*2.54%*	*0.42%*	*2.67%*

This error is likely due to the method used to construct the Kalman filter. Rather than modeling our IMU-based system specifically, we built the Kalman filter using a general approach and then used this filter to estimate orientation angles. However, this small amount of error was deemed negligible for our application. Subject-to-subject differences in measurement accuracy were also acceptably low.

Most of the error between the Shimmer and VICON angles occurs during the walking phase (see Figure 3.4). During this period of dynamic motion, the IMU accurately estimates angles above 20 degrees. Due to the way UpTime is calculated, we expect the IMU to accurately measure UpTime even when the collected data involves a higher proportion of active movements. Switching from one posture to another will be much less frequent in unchoreographed data collection, especially for individuals with ME/CFS. For this reason, we expect that the Shimmer will be even more accurate in our study than in this accuracy confirmation study.

3.4 Conclusion

In this chapter, the MATLAB function used to calculate UpTime was described in detail. For the same choreographed trials, calculations of UpTime for both VICON and Shimmer were compared, with VICON representing 100% accurate measurements. We found a 1.88% error between the Shimmer and that VICON baseline, indicating that the Shimmer is sufficiently accurate to measure ME/CFS patient uprightness. In the end, switching from a questionnaire-based assessment of HUA to an IMU-based measurement of UpTime will offer a dramatic improvement in terms of both accuracy and resolution.

CHAPTER 4

PREPARING THE SHIMMER & OURA RING FOR CLINICAL USE

4.1 Introduction

Previous chapters have discussed and verified a method that converts IMU data to angles. Later, we will describe how this angle estimation was used in a case-control study to measure upright activity (or UpTime). Before this study could occur, we needed to prepare the Shimmer for prolonged use in diverse settings.

In this chapter, we describe the methods and procedures followed to outfit study participants with the Shimmers. The main concerns we address in this chapter are waterproofing the Shimmer and designing the user's experience with the Shimmer. At the end of this chapter, we explain the Shimmer attachment procedure in detail.

4.2 Shimmer – User Experience

In our study, we asked participants to wear two Shimmers—one per leg—for a full week. With this extended period of continuous use in mind, it was essential to sufficiently prepare the devices so that the user had as little discomfort as possible. We wanted to outfit each subject with the devices and then let them live their lives almost entirely uninhibited.

While the compact form factor of the Shimmer met our need for a small and

lightweight design, we anticipated that sleeping with the Shimmer attached to their lower leg could cause mild discomfort; this concern was especially important due to the increased sensitivity of our experimental group. To minimize the discomfort of sleeping with the Shimmer, we chose to place the device on the lateral side of the lower leg, approximately two inches above the malleolus. Nesting the Shimmer in the natural curve of the lower leg minimized pressure on the device during sleep, thus maximizing comfort. Apart from minimal discomfort during sleep, the only other issue we identified in terms of user experience was accidental switching or pressing of the ON/OFF Switch and/or User Button (see Figure 4.1).

Pressing the User Button starts and stops data collection. This button is pressed accordingly to initiate and terminate data collection for each trial. We 3D printed covers

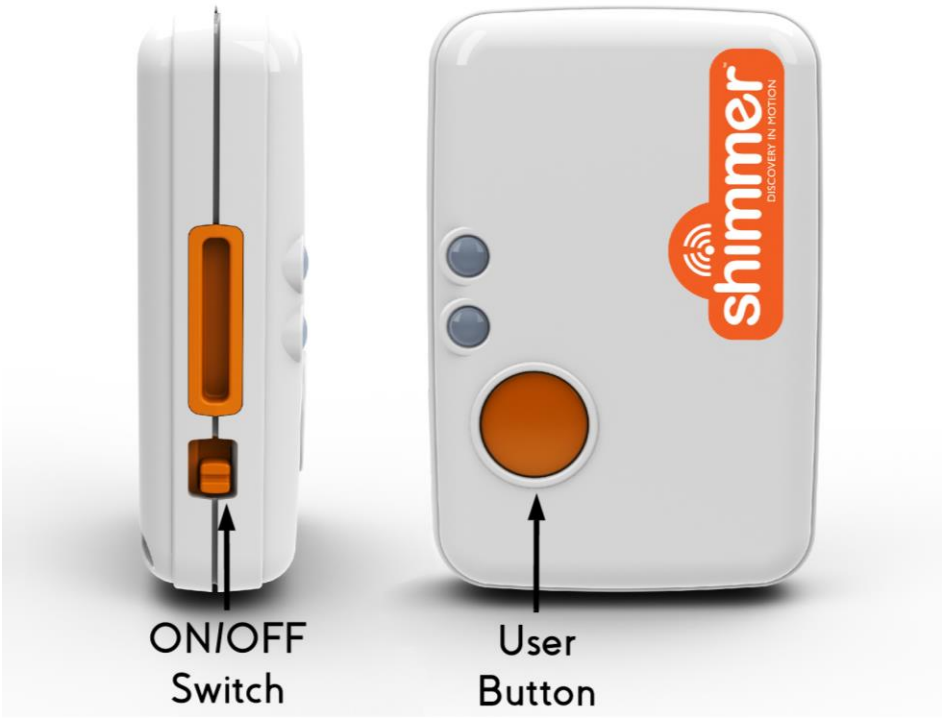

Figure 4.1. Shimmer unit with ON/OFF Switch and User Button labeled. Side view (left) and top view (right).

to be placed over the top of the Shimmers before they were attached to a subject's lower leg. These covers prevent study participants from accidentally pressing the User Button during data collection. The design of this cover minimizes added volume and weight; the cover adds an extra 3 grams of weight and 2 millimeters of height to the Shimmer (see Figure 4.2).

4.3 Shimmer – Waterproofing

We anticipated that many participants would appreciate being able to shower without having to remove or worry about the device. This concern was especially relevant due to the extended and continuous nature of each trial. To improve user comfort, we developed two-layers of waterproofing for the Shimmer that would allow users to shower during their trial. Figures in section 4.4 may help the reader to visualize the solutions described below.

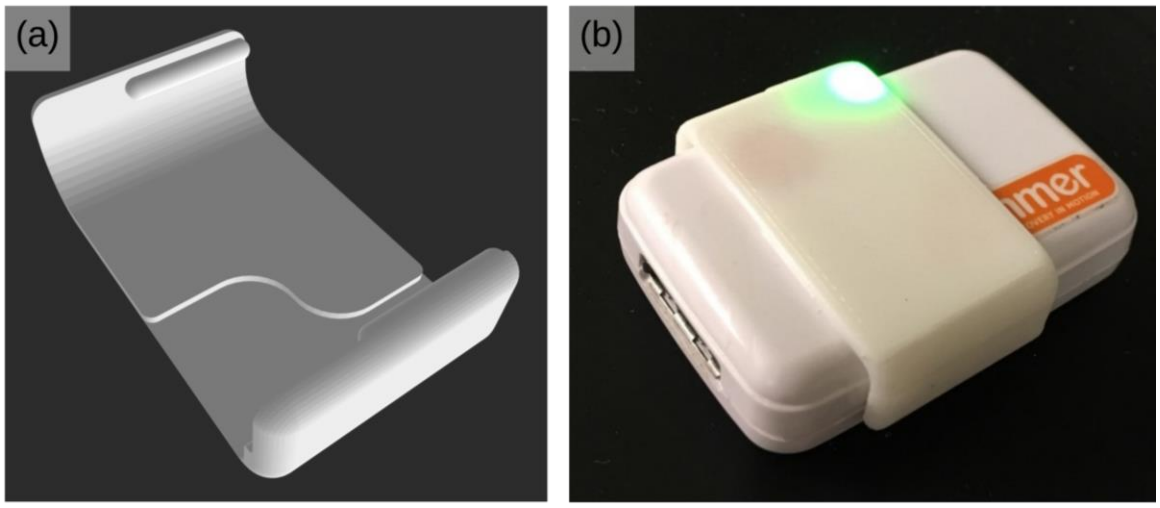

Figure 4.2. (a) Rendering of the cover placed over the Shimmer during data collection. (b) Shimmer with the 3D printed cover attached.

A detached finger of a nitrile glove, referred to hereafter as a nitrile sleeve, provided the first layer of waterproofing. We chose to use nitrile gloves due to their waterproof and puncture-resistant nature. Furthermore, we wanted to avoid potential allergic reactions to latex gloves, the most common alternative to nitrile gloves. We placed each Shimmer inside of a nitrile sleeve along with a desiccant packet. The desiccant packet absorbs any moisture existing within the Shimmer before placement inside of the nitrile sleeve. With the Shimmer and desiccant packet nested inside, we then sealed the nitrile sleeve shut with electrical tape, which was chosen for its elastic nature. This containment method protects the Shimmer from all possible sources of moisture, including moisture within the device, sweat from the user's skin, and outside sources of moisture.

Tegaderm film dressings provided the second layer of waterproofing. Tegaderm is a transparent, waterproof, and flexible adhesive which is most commonly used to protect IV insertion sites. For our purposes, Tegaderm provided a way to comfortably and securely attach the Shimmer to each participant's leg. This approach provided a natural way to deter participants from removing or otherwise adjusting the Shimmer. At the same time, the Tegaderm also provided a redundant layer of water protection. This second layer of waterproofing protects against external sources of water but does nothing to protect the Shimmer from the wearer's sweat, unlike the nitrile sleeve.

4.4 Shimmer – Attachment Procedure

Before the start of each trial in the study, the Shimmers were properly configured and given a trial name using Shimmer Sensing's ConsensysBasic software package—

build v1.6.0.

After the devices were configured, we placed each one back side down on a level surface (see Figure 4.3(a)). Next, we initiated data collection by pressing the user button once. To conceal the flashing lights, a potential source of discomfort for the experimental group, we covered the LEDs with tape (see Figure 4.3(b)). Next, we attached the 3D printed cover to the Shimmer to protect the user button (see Figure 4.3(c)). Then we placed the Shimmer in a nitrile sleeve along with a desiccant packet (see Figures 4.3(d) and 4.3(e) for orientation). We used electrical tape to seal the Shimmer within the nitrile sleeve and wrapped the entire package with gauze to allow air to circulate between the device and the lower leg (see Figures 4.3(f) and 4.3(g)).

Using an alcohol wipe, we disinfected the surface of the lower leg (see Figure 4.3(i)). Next, we attached the nitrile sleeve package to the lower leg using a Tegaderm film dressing. To prevent the Tegaderm from wrinkling and sticking to itself, we placed the Tegaderm (sticky side up) on a flat surface. We then carefully pressed the gauze-wrapped nitrile sleeve package onto the center of the sticky side of the Tegaderm dressing (see Figure 4.3(j)). Next, we lifted the Tegaderm, with nitrile sleeve package attached, from the table and attached everything to the lower leg (see Figures 4.3(j) and 4.3(k)). Finally, the Tegaderm newly attached nitrile sleeve package was loosely wrapped in Coban for an additional layer of support (see Figure 4.3(l)).

4.5 Conclusion

In conclusion, it was important to develop a way to permit wearers of the Shimmer to shower during data collection. This waterproofing procedure relies on two

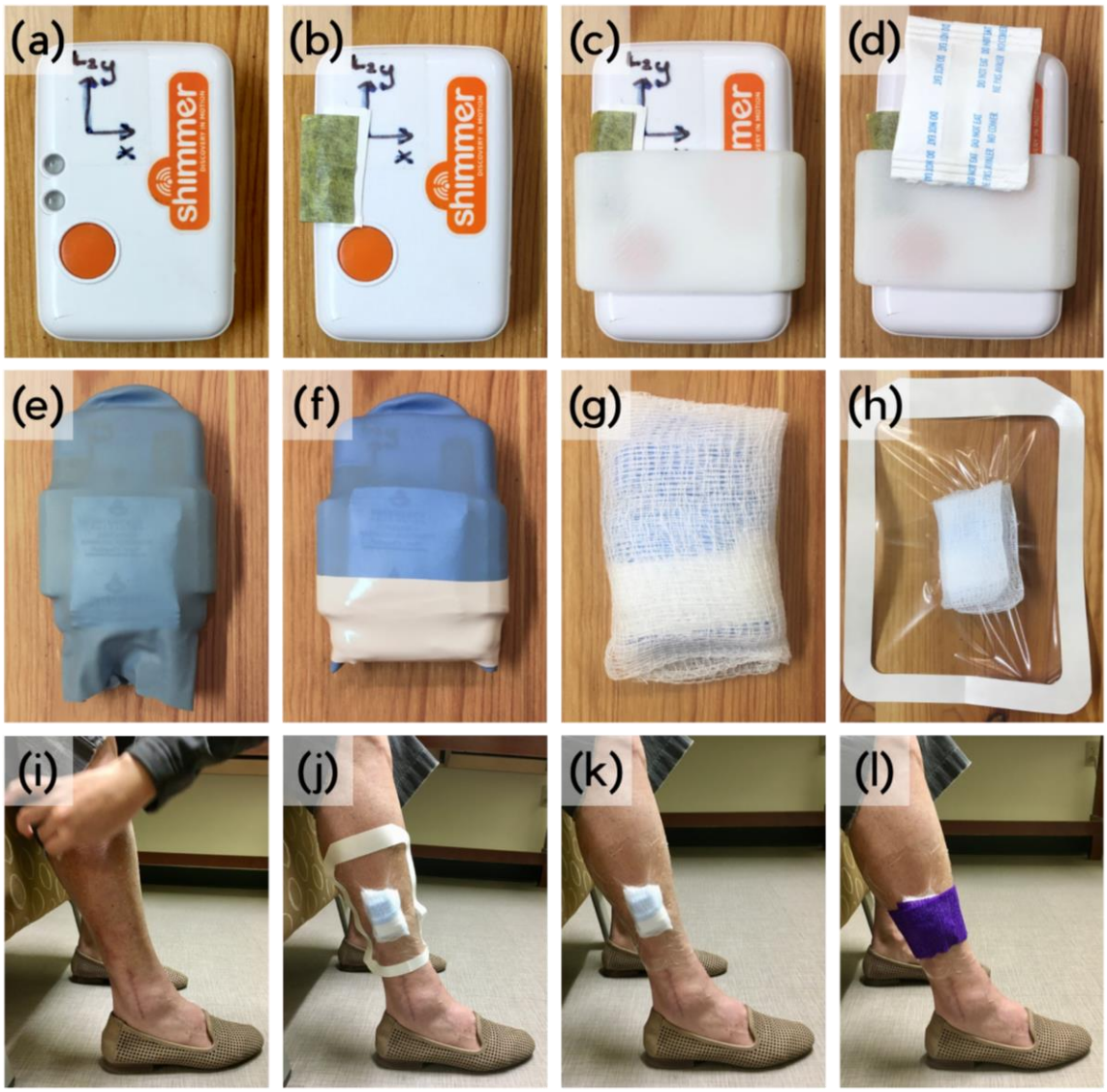

Figure 4.3. Shimmer attachment protocol. (a) Shimmer unit; (b) LEDs covered with tape to block light leakage; (c) 3D printed cover attached; (d) desiccant packet in place; (e) Shimmer, cover, and desiccant packet secured inside of the nitrile sleeve; (f) nitrile sleeve sealed shut with electrical tape; (g) entire package wrapped in gauze; (h) gauze wrapped IMU package attached to Tegaderm; (i) preparing the surface of the leg for device attachment using an alcohol wipe; (j) IMU package attached to lower leg with Tegaderm film; (k) outer layer of Tegaderm removed; and (h) Coban wrap applied to lower leg to provide extra support to IMU package.

layers of waterproofing: (1) the nitrile sleeve and (2) the Tegaderm film dressing. A procedure to outfit individuals with the Shimmer was then described. This attachment protocol was used for each participant in the study discussed in Chapter 5.

CHAPTER 5

RAW IMU DATA ANALYSIS METHODS

5.1 Introduction

This chapter explains the process of converting measurements collected by the IMU devices (Shimmers, in our case) to daily UpTime scores. Each daily UpTime score is reported as a percentage and represents the portion of the day spent upright and active; this measurement is called "UpTime." We use the orientation of the lower legs—calculated using a Kalman filter—to determine UpTime. Prior research conducted by the BHC has identified that UpTime varies depending on ME/CFS disease severity. To further investigate this claim, we planned and executed a clinical study evaluating UpTimes for diseased and healthy groups. Once validated, UpTime may be used as an endpoint to evaluate ME/CFS treatment efficacy.

5.2 Case-Control Study Design

For the planned Case-Control study, a total of fifteen subjects were outfitted with a Shimmer on each ankle and an Oura Ring on a finger of their choice. Each subject wore these devices for six days—starting on a Monday and ending on a Saturday. The fifteen subjects were divided into groups based on disease level—five subjects without ME/CFS (the controls), five subjects with moderate-level ME/CFS, and five subjects with severe-

level ME/CFS. Due to limited Shimmer availability, all subjects were staggered—with one or two subjects participating each week—from September 30, 2019 to December 19, 2019.

Each subject's six-day data collection period was broken into two "phases." For all subjects, phase one began on a Monday—when the subject traveled to the BHC to be outfitted with the devices—and ended 72 hours later on the following Thursday. Phase two began on Thursday (where phase one ended), lasting another 72 hours before ending on Sunday. The study was broken into two phases for multiple reasons. Primarily, the Shimmer batteries lasted approximately 72 hours using the specifications chosen for this study. Secondly, the data collected during phase one was meant to be a baseline against which the data from phase two would be compared; at the beginning of phase two, each subject performed the NASA Lean Test—meant to cause the onset of PEM for subjects with ME/CFS, but have no effect on the control group (although we did expect the weekend to have an effect on UpTime for the control group). Subjects were instructed to go about their lives in a normal manner during the study.

During each trial period, the subject filled out multiple nightly questionnaires, including the HUA Survey—most of these were used for research conducted separately by the BHC. Once all fifteen subjects completed their six-day trial period, the task of data processing began. This chapter details the method used to process the raw data collected by the Shimmers into daily UpTime scores. Chapter 6 describes the findings that resulted from statistical tests on these UpTime scores.

5.3 Data Preprocessing & Processing

This section covers the methods used to properly format the Shimmer data files (preprocessing) and compute UpTime measurements using the formatted Shimmer data (processing). Preprocessing involves restructuring the Shimmer data files into a format that can be efficiently read into MATLAB. Processing involves (1) using MATLAB to convert the formatted Shimmer data to angle data and then (2) calculating UpTime from the angle data.

In our study, we used ConsensysBasic to extract raw sensor data from the Shimmers as comma-separated values (CSV) files. Each Shimmer was programmed so that these CSV files contained eight columns of data:

1. Date of sensor measurements
2. Time of sensor measurements
3. Accelerometer (x-axis)
4. Accelerometer (y-axis)
5. Accelerometer (z-axis)
6. Gyroscope (x-axis)
7. Gyroscope (y-axis)
8. Gyroscope (z-axis)

Data preprocessing is based on this structuring of Shimmer data. We share this information so that other researchers can modify the following preprocessing method to fit the structure of other IMU hardware systems.

5.3.1 Data Preprocessing

Extracting Shimmer data for all 15 subjects yielded a total of 60 CSV files for the

entire study—four per subject. For each subject, these four CSV files correspond to data collected by the Left and Right Shimmers for each of the study's two phases: (1) Phase one – Left, (2) Phase one – Right, (3) Phase two – Left, and (4) Phase two – Right. For future reference, "paired" files are a Left and Right file from the same phase or day. All CSV files were preprocessed using a four-step approach: (1) trim the file, (2) divide the file into discrete days, (3) check the file for gaps, and (4) align paired files. These four steps were either performed manually, using the program TextEdit (Version 1.15) or MATLAB, depending on the step.

The purpose of the first step, file trimming, was to match start and stop times of paired files in a single phase. For all data, paired CSV files were opened side-by-side using TextEdit. Both datasets were then trimmed according to their date/time columns so that they began and ended at the same time. In cases where Shimmers recorded data surpassed 72 hours, both datasets were trimmed to end 72 hours after the start time. Using this process, paired files were synchronized for all subjects.

After trimming was complete, we divided each CSV file into discrete days. This task was completed manually using TextEdit. Recall that the study was divided into two phases—each lasting 72 hours—to minimize the need for subjects to return to the BHC. Our goal during this preprocessing step was to break each *phase* file into three *day* files, with each *day* file covering 24 hours. To accomplish this task, we made three duplicates of each CSV file and then manually cut each duplicate to 24 unique hours of data. Consequently, all phase one files were split into three files—Day 1, Day 2, and Day 3; and all phase two files were split into three files—Day 4, Day 5, and Day 6. Repeating this process for all data yielded a total of 90 single day files, six per subject.

Due to device instability, some Shimmers unexpectedly stopped and started recording data partway through a phase, resulting in multiple-hour gaps in the recorded data. The third step in preprocessing was identifying these randomly occurring gaps using a custom MATLAB script (Appendix A). This gap-checking script imported a CSV file and then isolated the data's time column. Each entry in this column was converted from a string (of the format 'hh:mm:ss') to a "duration" value—a built-in MATLAB data type. The gap-checking script compared successive duration values, checking for gaps longer than one second. If a gap was identified, the script delivered a message indicating its location within the file. Otherwise, the file was advanced to the last stage of preprocessing.

The fourth and final preprocessing step was aligning paired day files. While the first and last entries of paired day files had matching timestamps, the number of datapoints within these paired files varied slightly. Operating at our defined specifications (see Chapter 3.2.1), all Shimmer devices recorded data as N x 8 matrices—where N is the number of rows. At a sample rate of 30 Hz, each day's CSV files are expected to have 2,592,000 rows; however, slight differences in *actual* sample rates led to differences in the number of rows contained in paired CSV files. These differences—generally only a few dozen—were insignificant given the total rows in each file. Regardless, paired files needed to be equal in length to calculate UpTime—a problem that needed to be remedied.

Our solution was to create a MATLAB script which ensured timestamp alignment throughout the entirety of paired files (Appendix B). This script first imported a day's paired CSV files, saving each file in matrix form. Using the following method, the longer matrix was trimmed to match the shorter matrix. To keep the timestamps of the paired

files matched as much as possible, we opted not to remove rows from the beginning or end of the longer matrix. Instead, this script determined the number of rows to be removed from the longer matrix and then cuts rows at constant intervals throughout its length. At the end of the alignment process, both matrices had equal lengths with matching first and last timestamps.

5.3.2 Data Processing

Following preprocessing, we began calculating UpTime. Our UpTime calculator—another custom MATLAB script—was used to convert paired Shimmer files to daily UpTime scores (Appendix C). Before running this MATLAB script, the operator first specified the subject ID (01 to 15) and days of interest (any combination of integers 1 through 6, corresponding to Monday through Saturday). When executed, the UpTime calculator followed a three-step process: (1) reading the CSV file and formatting its data, (2) converting the formatted sensor data to angle data, and (3) using angle data to calculate an UpTime score.

The first processing step—reading and formatting data—imported the specified CSV files (Appendix D). Once imported, date and time columns were deleted to minimize RAM usage and speed up processing time. For each day, this process yielded a set of paired N x 6 matrices of double-precision floating-point values. The six columns of each matrix were comprised of 3-axis accelerometer and 3-axis gyroscope measurements recorded by the Shimmers.

In the second processing step, each matrix was fed into a filtering function based on Phillip Salmony's Kalman filter, which is described in Chapter 2 [28]. This Kalman

filter converted sensor measurements into an N-length vector of angle estimates for each leg, taking into account the sensitivity of each Shimmer (Appendix E and F). The resulting angle vector contained a full day of continuous angle measurements, with angles measured thirty times each second.

In the third and final processing step, paired angle vectors were used to calculate a daily UpTime score. Our UpTime function took paired angle vectors and compared each measured angle to the critical angle (39 degrees). At each point throughout the day, the lower leg was determined to be either upright (measured angle < critical angle) or not upright (measured angle > critical angle) (see Figures 3.5 & 3.6). The final output of the UpTime function was an "UpTime score," which indicated the portion of the day spent with the lower legs in an upright orientation.

All paired day files were processed using this method to obtain daily UpTime scores. Data for all 15 study participants were processed using this method, resulting in six daily UpTime scores each (see Table 6.1).

5.4 Choosing a Critical Angle

Before moving to a discussion of UpTime scores, we must first explain two points: (1) the rationale behind selecting a critical angle of 39 degrees, and (2) the use of UpTime measurements where Shimmer failures caused data to be incomplete—found in Section 5.5.

This thesis is primarily focused on estimating a subject's daily upright activity as an estimate of ME/CFS disease severity. To determine uprightness, we evaluate leg angles to be either upright or not upright. This distinction is defined by the critical

angle—measured as degrees from vertical—which is key to calculating UpTime (see Figure 3.5). It is, therefore, essential that we select a proper critical angle to distinguish between lower leg postures.

The critical angle must correctly differentiate between upright and reclined/supine postures. This differentiation is needed to assess a person's activity level. Patients with ME/CFS can elevate their legs to find relief from symptoms related to OI and PEM. Generally, this elevation translates to patients lying in bed or sitting in a chair with the feet propped at least six inches off the ground. For this reason, we expected the critical angle to be somewhere around 45 degrees from vertical to capture the full gamut of ameliorative postures.

It was also important to select a critical angle that could be increased or decreased by a few degrees with a minimal change in UpTime scores. To select a critical angle that is insensitive to change in this way, we created plots of day 1 data for three subjects—one subject per disease group. These plots show how UpTime scores change depending on the selected critical angle (see Figure 5.1).

These plots are shown for illustration purposes. An analysis of all ninety days is included below. All three plots show a flat section—highlighted in red—which indicates the range where UpTime is less sensitive to the critical angle. Visually, we can see that the critical angle can increase or decrease within this highlighted range without affecting UpTime significantly. While these highlighted ranges vary from day to day, they overlap for critical angles between 45 and 60 degrees. Due to this commonality, we expected the appropriate critical angle to be within this range.

Given that we were interested in where the sensitivity plots flatten out, we chose

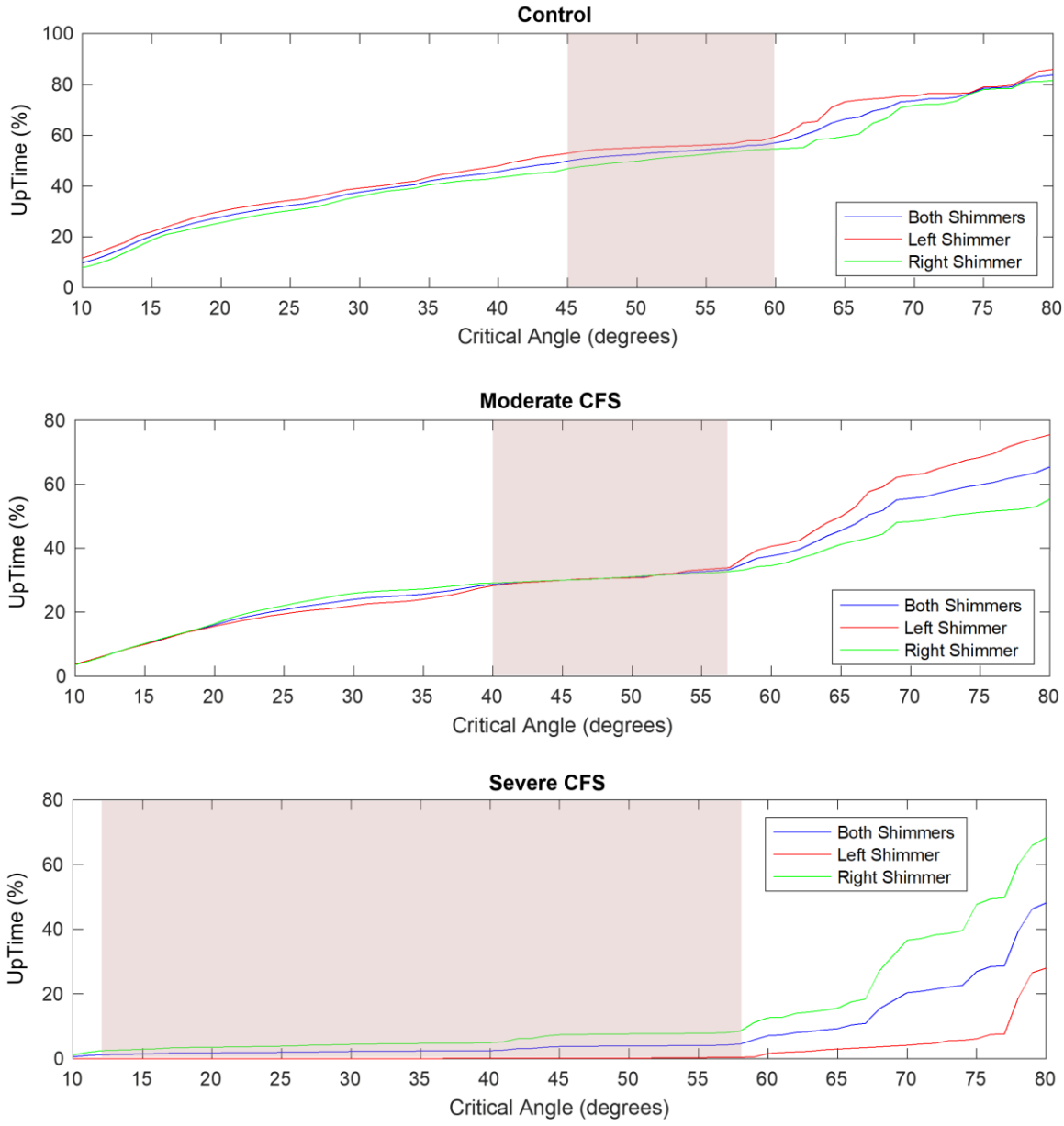

Figure 5.1. Sensitivity plots for UpTime and critical angle. The top plot corresponds to subject 01, day 1—a healthy control. The middle plot corresponds to subject 07, day 1—a subject with moderate ME/CFS. The bottom plot corresponds to subject 11, day 1—a subject with severe ME/CFS. The areas highlighted in red indicate where UpTime appears to be insensitive to changes in critical angle. These areas overlap between 45 and 57 degrees.

to pursue an analysis of the slope $\left(\frac{dU}{dA}\right)$ of these curves. By definition, UpTime becomes less sensitive to changes in critical angle (i.e., the sensitivity plot flattens out) as the local slope of this sensitivity plot approaches zero. For all ninety days of data, we calculated $\frac{dU}{dA}$ corresponding to critical angles from 10 to 80 degrees. Average $\frac{dU}{dA}$ values were then plotted along with standard deviations to identify the critical angle with the least variation and the slope closest to zero (see Figure 5.2).

We were also interested in variations in UpTime scores arising from calculation methods. It is preferred to calculate UpTime using angles from both the left and the right legs simultaneously; UpTime can also be calculated using data from only one leg when data from the other leg is unavailable. Figure 5.2 includes $\frac{dU}{dA}$ means and standard

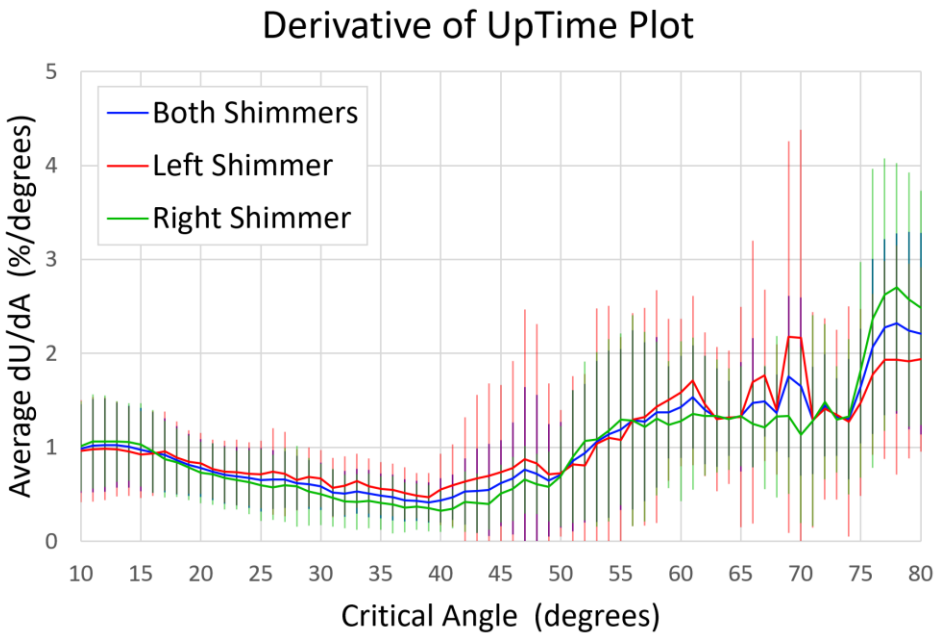

Figure 5.2. Average $\frac{dU}{dA}$ values. Vertical lines are +/- one standard deviation. This was done three times—calculating UpTime using only the left Shimmer, only the right Shimmer, and using both Shimmers. The critical angle with the least variation and smallest slope is 39 degrees.

deviations from all three of these UpTime calculation methods (or “Shimmer types”).

By combining all ninety days of data in Figure 5.2, we can see that the best choice for critical angle is somewhere between 37 and 41 degrees; these critical angles are optimal because their corresponding average $\frac{dU}{dA}$ values are closest to zero—indicating that corresponding sensitivity plots are flattest within this range. Furthermore, the standard of deviations for these angles—indicated by vertical lines—are relatively minimal. We ultimately selected 39 degrees because it meets both conditions—it is close to zero on the $\frac{dU}{dA}$ curve with low variation between trials.

The original assessment of upright activity, HUA, separates postures into two groups: (1) feet on the floor and (2) feet off the floor (see Figure 2.1). This distinction is meant to identify upright (exhaustive) and not upright (ameliorative) postures. HUA-related research performed at the BHC identified the act of elevating the feet six inches off the floor as the threshold between these two posture types. Depending on the length of the legs and the height of the chair, elevating the legs six inches could yield a range of lower leg angles—we estimated this range to be between 30-60 degrees from vertical. Our identified critical angle of 39 degrees falls within this expected range and maintains a clinically significant separation of posture types.

5.5 The Issue of Single-Shimmer Data

In this section, we will present our argument validating UpTime scores calculated using data from a single leg. First, we will review the causes of single-leg data—Shimmer failures.

5.5.1 Summary of Shimmer Failures

Before beginning our study, we performed a handful of "battery drain" tests to verify that the Shimmers would last three full days at our operating specifications. These preliminary tests simply involved running the Shimmers from full to empty batteries while collecting data and then checking the duration of the data collected. In all test cases, the Shimmers recorded data for at least the expected 72 hours. Despite these preparatory measures, several Shimmer failures occurred during the actual study; in all cases, these failures led to incomplete data. Table 5.1 is a summary of all data collected during our study, color-coded by number of hours collected for each 24-hour day. Complete days, i.e., days with 24 hours of data, are shown in green. As the duration of data collected for a given day drops below 24 hours, the cell color changes to yellow and then to red. Days with 20+ hours of data were considered to be complete.

Shimmer failures were separated into three categories: (1) battery failures, (2) data gaps, and (3) corrupted data. Any trial where the Shimmer battery died more than four hours before the end of the 72-hour phase length was counted as truncated due to battery failure. Of the 60 phase files shown in Table 5.1, nine have truncated data due to this failure type. The likely cause of battery failures is insufficient charging. We planned a minimum of six hours for each Shimmer to charge fully, but the actual charge time during the study may have varied slightly due to scheduling difficulties. Additionally, we saw multiple battery failures occur where the Shimmer died less than eight hours into the phase; the cause of these failures is unknown. In the format SUBJECT_PHASE_LEG, files with failures of this type are (1) 03_P1_R, (2) 04_P1_L, (3) 04_P2_R, (4) 06_P2_L, (5) 06_P2_R, (6) 12_P2_R, (7) 13_P2_L, (8) 13_P2_R, and (9) 15_P2_L.

Table 5.1. Summary of collected data color-coded by duration. Green cells are 24 hours long with cell color shifting to yellow and then red as duration decreases.

Group	Subject	Leg	Duration of Collected Data (hours)					
			Phase 1—*Before NASA Lean Test*			Phase 2—*After NASA Lean Test*		
			Day 1 *Monday*	Day 2 *Tuesday*	Day 3 *Wednesday*	Day 4 *Thursday*	Day 5 *Friday*	Day 6 *Sunday*
CONTROL	1	L	24	24	24	24	24	24
		R	24	24	24	24	24	24
	2	L	24	24	22.5	24	24	24
		R	24	24	22.5	24	24	24
	3	L	24	24	23	24	24	24
		R	7.2	0	0	24	24	24
	4	L	6.5	0	0	5.75	0	17.5
		R	24	24	24	24	24	7.5
	5	L	24	24	21	24	24	24
		R	24	19	6.5	24	24	24
MODERATE ME/CFS	6	L	24	24	23.5	24	24	5.5
		R	24	24	23.5	24	24	5.5
	7	L	24	24	24	24	24	24
		R	24	24	24	24	24	24
	8	L	24	24	22	24	24	24
		R	24	24	22	24	24	24
	9	L	24	24	24	24	24	24
		R	24	24	24	24	24	24
	10	L	24	24	23.5	24	24	24
		R	24	24	23.5	24	24	24
SEVERE ME/CFS	11	L	24	24	21	24	24	24
		R	24	24	21	24	24	24
	12	L	24	24	23	4.5	21	24
		R	24	24	23	24	24	3.5
	13	L	24	24	22	24	13	0
		R	24	24	22	24	4.5	0
	14	L	24	24	20	11	3.5	15.5
		R	24	24	20	24	24	24
	15	L	0	0	0	24	22.5	0
		R	0	0	0	24	24	24

The second Shimmer failure type is data gaps. In these cases, the Shimmer randomly stopped recording data partway through the trial and then randomly began collecting data again any number of hours later. These failures were identified via the gap checking MATLAB script discussed in Chapter 5.2.1. Four of the 60 total phase files shown in Table 5.1 are classified as failures due to gaps in the recorded data. Files with failures of this type are (1) 04_P2_L, (2) 05_P1_R, (3) 12_P2_L, and (4) 14_P2_L.

It is also worth noting that another two Shimmer files were marked by ConsensysBasic as corrupted and were therefore unrecoverable. The cause of this failure type is unknown; however, it is recommended that future studies avoid using Shimmers as a result of these unpredictable and unexplainable failures. Files with failures of this type are (1) 15_P1_L and (2) 15_P1_R.

In total, at least one type of Shimmer failure occurred in 15 of the 60 total phase files—a 25% failure rate. Due to these failures, 18 of the 90 total days have data from only one leg and another seven lack data from both legs. While the seven days without data from either leg are completely lost, there is still hope for the 18 days with data from only one leg. If we can prove the validity of UpTime measurements calculated from single-leg data, these 18 days (20% of all data) can still be of use. To justify including these trials, we must first confirm that UpTime is the same, regardless of whether it is calculated using one or two Shimmers.

Moreover, determining whether two IMUs are needed to determine UpTime is a clinically meaningful exercise. The burden on the user is lower if they only have to wear one device rather than two. In the long term, this potential change could be meaningful if UpTime clinics choose to use UpTime as a biomarker for treatment efficacy.

5.5.2 Validation of Single-Shimmer UpTime Calculations

The mean plot shown in Figure 5.3 compares UpTime to Shimmer type; levels for Shimmer type are (1) UpTime calculated using data from both Shimmers, (2) UpTime calculated using data from the left Shimmer only, and (3) UpTime calculated using data from the right Shimmer only. Figure 5.3 shows substantial overlap between all confidence intervals. This overlap indicates that there may not be a significant difference between UpTime calculated using different Shimmer types. Looking at a mean plot is a somewhat subjective analysis. To confirm this conclusion, we need to perform an objective statistical test.

A pairwise t-test allows us to perform and compare the results of multiple paired t-tests, allowing us to compare all three Shimmer types at once. We can use this statistical

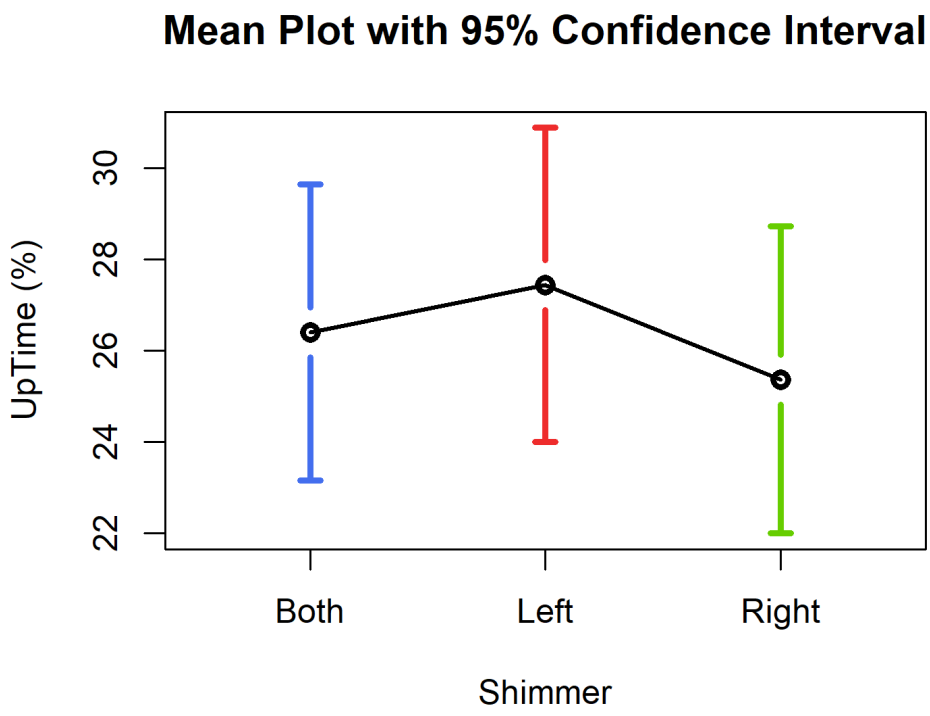

Figure 5.3. Mean plot of UpTime data grouped by Shimmer.
Shimmer type confidence intervals overlap heavily.

test to evaluate our UpTime data because all of the necessary assumptions are met: (1) UpTime is a continuous variable, (2) the observations are independent of one another, (3) UpTime scores are normally distributed, and (4) the data contains no outliers. Figure 5.4 shows that UpTime measurements follow a nearly normal distribution. As further confirmation of normality, the Shapiro-Wilk test produced a p-value of 0.069, indicating that the distribution of UpTime is not significantly different from a normal distribution (at the $\alpha = 0.05$ significance level).

As for the paired t-tests, the null hypothesis is that the mean difference between the paired samples is zero. The pairwise comparison was performed in R Studio (Version 1.1.463) using the Bonferroni multiple-comparison correction factor. The results of this test confirm that UpTime does not differ by Shimmer type (see Table 5.2)

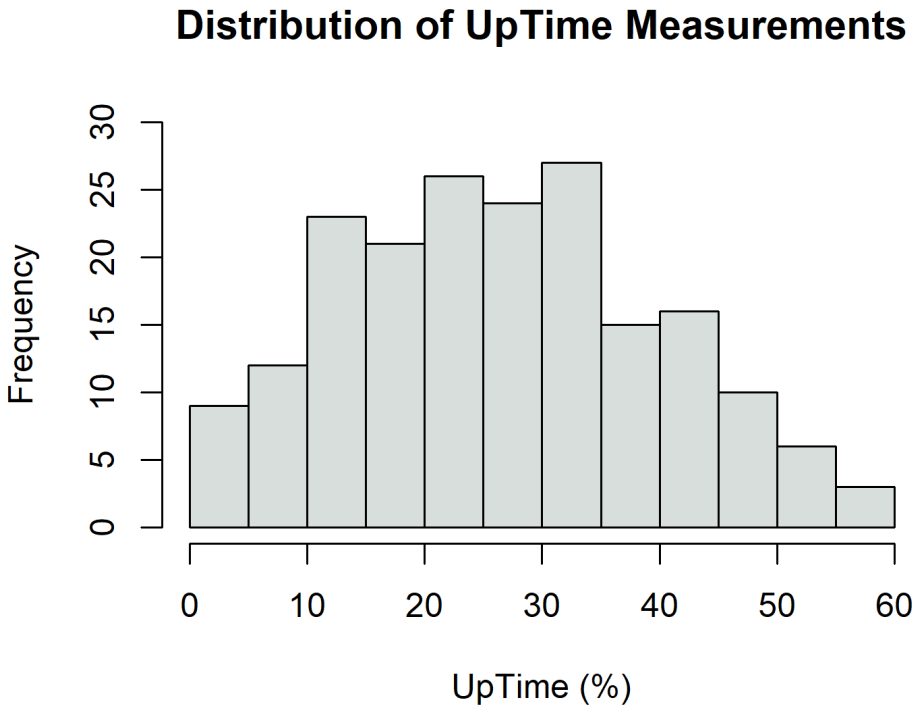

Figure 5.4. This bar plot shows that UpTime measurements are normally distributed.

Table 5.2. Results of the pairwise t-test for UpTime and Shimmer type. At the $\alpha = 0.05$ significance level, we can conclude that UpTime is the same for all Shimmer types.

	Both	**Left**
Left	0.14	-
Right	0.14	0.14

**P-value adjustment method: Bonferroni*

The mean plot (Figure 5.3) and the pairwise t-test both prove that the means for UpTimes calculated using each method are not statistically different ($p > 0.05$). Note that the power for this statistical test is 0.19, indicating that there is a fair chance we have performed a Type II error; a larger sample size is needed to confirm this conclusion. While no statistically significant difference was found, we need to take this conclusion a step further to prove that the UpTime scores for each Shimmer type are close to identical. In Figure 5.5, each line represents one day with the points on the left and right corresponding to UpTime scores calculated using only the left or right Shimmer. Left and right data are available for 64 of the 90 total days. In most cases, the lines connecting these paired UpTime scores are nearly horizontal, indicating that UpTime is very similar regardless of Shimmer type. Furthermore, even though some lines are further from horizontal, the average left and right UpTimes are within 2% of each other.

Lastly, an UpTime score calculated using both Shimmers is the average of the left and right legs' UpTimes. Figure 5.5 compares left to right UpTimes. A plot comparing left or right UpTime to an UpTime calculated from both devices would have slopes cut in half from what is seen in Figure 5.5.

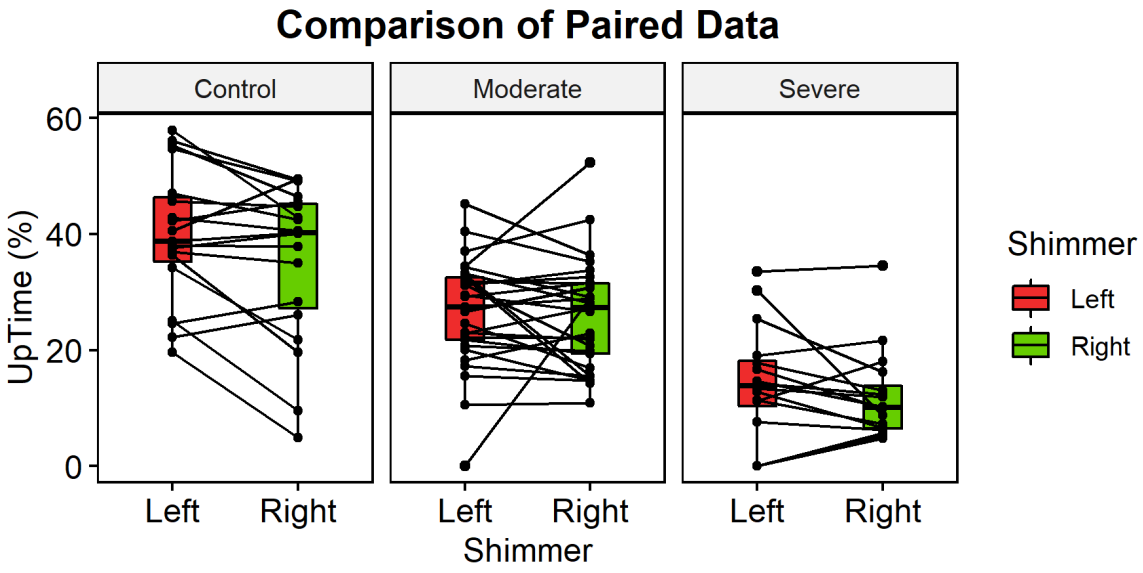

Figure 5.5. Plots of paired UpTimes separated by group. Paired left and right UpTimes are connected by lines. Note that most lines are nearly horizontal. Left and right means (indicated by underlying boxplots) are similar for all groups.

5.6 Conclusion

In conclusion, 25 of the 90 days in our study lack data from one or both Shimmers. Eighteen of these days have data from only one Shimmer. We have shown that UpTime calculated using one Shimmer—either left or right—is not statistically different from UpTime calculated using both Shimmers. Furthermore, left-only and right-only calculation methods give UpTime measurements that vary on average by only 2% (Figure 5.5). We conclude that this evidence is sufficient to justify the inclusion of daily UpTime scores calculated using data from a single Shimmer in our study results. By including days with single-Shimmer data, we reduce the number of days lost due to Shimmer failures from 25 to 7, thereby salvaging 72% of all lost data.

In future studies, we do not recommend using Shimmers due to the 15% failure rate. Alternative IMUs exist in the market. However, in-house development of an IMU

device is a viable option. A custom IMU device would allow for the possibility of a larger battery and more ergonomic form-factor.

The validation of single-Shimmer UpTime scores means that, for added comfort, subjects in future related studies could be required to wear a single IMU. Even so, the use of redundant IMUs may still be wise, depending on expected failure rates of a custom IMU device. Furthermore, redundant IMUs would mitigate the effect of asymmetrical leg postures on UpTime calculations.

CHAPTER 6

STUDY RESULTS & DISCUSSION

6.1 Introduction

Previously, we have discussed the value of using IMUs to measure upright activity. After selecting an IMU—the Shimmer—and proving its accuracy, we planned a case-control study that was carried out by the BHC. With subject data properly processed, we now move into a comparison of measurements.

The bulk of this analysis focuses on UpTime comparisons, beginning with a comparison by group. In our study, subject groups corresponded to three distinct disease levels: (1) controls, (2) moderate ME/CFS, and (3) severe ME/CFS. After comparing UpTimes by group, we will compare all UpTime measurements before and after the NASA 10-minute Lean Test—also referred to as the "NASA Lean Test" or just "Lean Test." Finally, we will compare IMU-based UpTime measurements to results from the original HUA questionnaire. Following these UpTime-related analyses, we will review a variety of measurements collected by the Oura Ring.

The broader effect of any conclusions made in this chapter is limited due to the low number of samples collected; only five subjects were included in each group. This study was designed to be a sort of pilot study. In that sense, it was meant to flesh out the process of using IMUs to accurately measure UpTime and point researchers at the BHC

towards areas of interest. Future studies with sufficiently large, randomized samples are required to improve the generalizability of these results.

6.2 Differences Between Disease Groups

All UpTime scores are provided in Table 6.1, along with daily and weekly averages. Due to differences in activity levels brought on by the presence and severity of ME/CFS, we expected the control group to have the highest UpTime and the severe ME/CFS group to have the lowest UpTime, with the moderate ME/CFS group's UpTime somewhere in the middle. The weekly mean column in Table 6.1 supports this expectation.

Controls generally had average weekly UpTimes above 30%. Subjects with moderate ME/CFS generally had UpTimes between 20 – 30%. Subjects with severe ME/CFS averaged daily UpTime scores below 20%. The non-overlapping group confidence intervals (shown by the vertical colored lines in Figure 6.1) are evidence indicating that UpTime differs by disease level.

The results of a linear mixed-effects model further substantiate these group UpTime differences. Qualifying assumptions for the linear mixed model—linearity, normally distributed residuals, and homoscedasticity—were met by the UpTime data (Figure 6.2). Because this research was developed as a pilot study—meant to identify areas of interest to be pursued in future (externally funded) studies—the sample sizes are smaller than we typically expect for statistical tests. Therefore, the results of this analysis should be interpreted as preliminary evidence for future full-scale studies.

Table 6.1. UpTime scores for all subjects. Cells are color-coded (high UpTime = green, medium UpTime = white, low UpTime = red). A dash indicates missing data.

Group	Subject	UpTime (%)						
		Day 1 Monday	Day 2 Tuesday	Day 3 Wednesday	Day 4 Thursday	Day 5 Friday	Day 6 Saturday	Weekly Mean
Control	1	44.79	50.86	51.90	52.77	50.36	27.98	**46.44**
	2	43.92	37.96	45.22	41.72	39.48	38.88	**41.20**
	3	50.40	51.24	44.04	27.97	12.25	17.31	**33.87**
	4	40.21	46.06	47.31	50.28	32.32	-	**43.24**
	5	45.05	16.13	24.76	26.44	35.97	24.11	**28.74**
	Daily Mean	**44.87**	**40.45**	**42.65**	**39.84**	**34.07**	**27.07**	**38.16**
Moderate	6	25.10	43.41	39.73	30.45	40.79	-	**35.90**
	7	28.19	20.55	15.12	28.74	28.67	20.73	**23.66**
	8	31.65	37.84	22.15	32.04	20.26	32.52	**29.41**
	9	26.06	16.39	14.50	24.27	23.45	31.73	**22.73**
	10	17.43	10.71	22.47	30.70	28.08	20.59	**21.66**
	Daily Mean	**25.69**	**25.78**	**22.80**	**29.24**	**28.25**	**26.39**	**26.36**
Severe	11	2.41	2.64	2.86	9.70	19.55	9.27	**7.74**
	12	20.35	6.88	13.35	13.81	11.84	0.04	**11.05**
	13	13.34	14.66	12.63	15.43	-	-	**14.01**
	14	20.81	12.43	-	18.70	18.04	21.52	**18.30**
	15	-	-	-	34.08	6.76	11.04	**17.29**
	Daily Mean	**14.22**	**9.15**	**9.62**	**18.34**	**14.05**	**10.47**	**12.64**

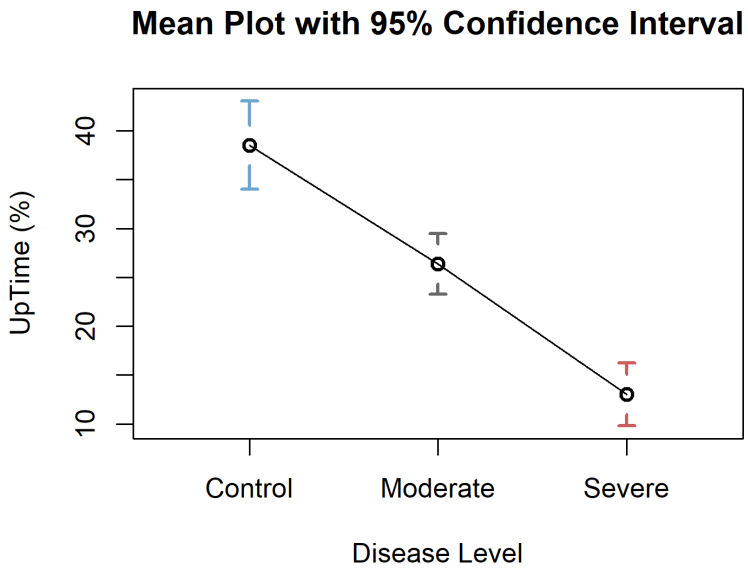

Figure 6.1. Mean plot of UpTime separated by disease level. Vertical lines represent 95% confidence intervals. These intervals do not overlap for groups, indicating significant differences in group UpTimes.

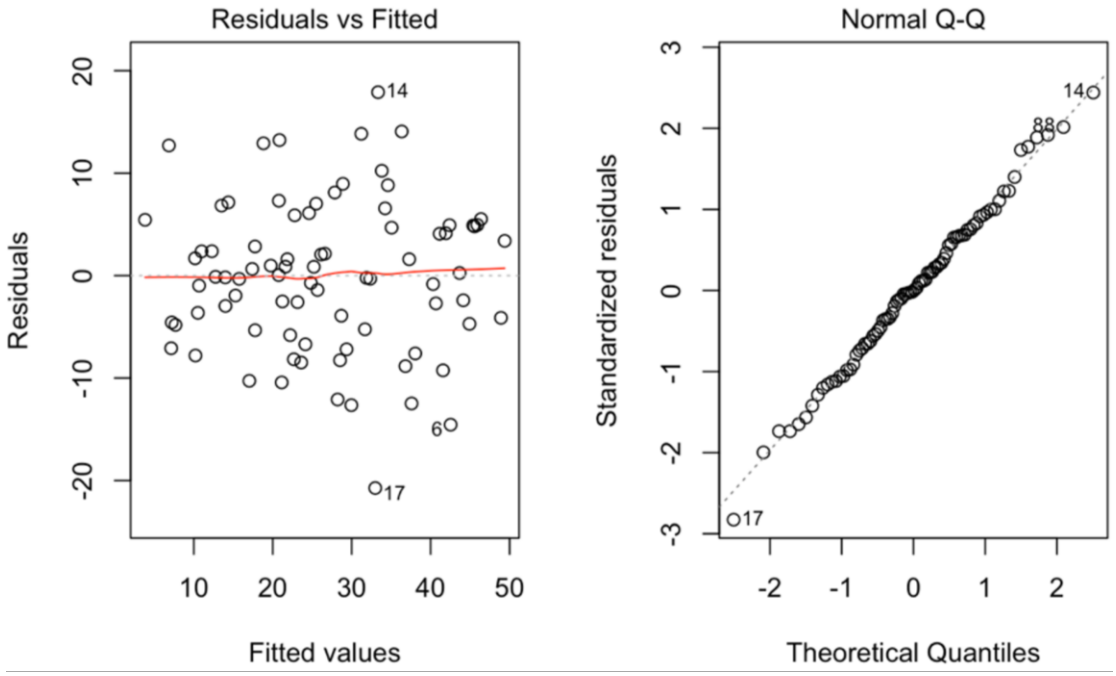

Figure 6.2. Residuals plot (left) and Q-Q plot (right) for UpTime data. The residuals plot shows that our collected UpTime data has constant variance, i.e., it is homoscedastic. The Q-Q plot shows all points close to the regression line; thus, the data is normally distributed.

The results of a linear mixed-effects model further substantiate these group UpTime differences. Qualifying assumptions for the linear mixed model—linearity, normally distributed residuals, and homoscedasticity—were met by the UpTime data (Figure 6.2). Because this research was developed as a pilot study—meant to identify areas of interest to be pursued in future (externally funded) studies—the sample sizes are smaller than we typically expect for statistical tests. Therefore, the results of this analysis should be interpreted as preliminary evidence for future full-scale studies.

With the proper assumptions met, we fit all UpTime scores to a linear mixed model, including a fixed effect (Disease Level), a blocking factor (Day), and a random effect (Subject). In R, this was done using the "lmer" function in the "lme4" library; this was done using the following formula: *Uptime ~ DiseaseLevel + Day + (1|Subject)*. The results of this model are shown in Table 6.2. The results of Table 6.2 confirm that mean Disease Level UpTimes are not equal ($p < 0.0001$). The power associated with this one-way ANOVA is 1, indicating that these results will likely be valid in full-scale studies.

To further expand upon these conclusions, we used Tukey's HSD (honestly significant difference) test. Up to this point, we have only shown that the group UpTime means are not the same. Tukey's HSD test identified which specific group differences exist. For all pairs of means, the calculated p-values are far less than 0.05, meaning that each group's mean UpTime is different from all other groups (Table 6.3).

With the combined results of the mean plot, the linear mixed model, and Tukey's HSD test, we can confidently state that UpTime differs for all disease levels. Using the UpTime scores in Table 6.1, we can define the UpTime expected for each group. Healthy individuals are expected to have weekly UpTime scores between 30% and 50%. Patients

Table 6.2. Results of disease level UpTime comparison. Used a linear mixed-effects model in the form of a Type III Analysis of Variance Table with Satterthwaite's method. P-values indicate a significant difference in UpTime by disease group.

Factor	Sum of Squares	Mean Sum of Squares	F value	P-value
Disease Level	3572.5	1786.3	27.79	8.22e-06***
Day	440.2	88.0	1.37	0.247

Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

with moderate ME/CFS are expected to have weekly UpTime scores between 20% and 30%. Patients with severe ME/CFS are expected to have weekly UpTime scores below 20% (Figure 6.3). These conclusions align with the observations of the BHC and their understanding of ME/CFS. Symptoms of this disease—such as post-exertional malaise (PEM) and orthostatic intolerance (OI)—limit a patient's ability to remain upright. As disease severity increases, so do these limitations. We now objectively see that UpTime corresponds to the presence and severity of ME/CFS.

6.3 Comparing UpTime Before vs. After NASA Lean Test

Next, we looked for differences before and after the NASA Lean Test, i.e., phase one versus phase two. The NASA 10-minute Lean Test requires subjects to stand straight upright and lean against a wall, with only the shoulder blades contacting the wall, and heels six inches from the wall [32]. In our study, the purpose of this test is to induce Post-Exertional Malaise (PEM) in subjects with ME/CFS. For this reason, we expected to see ME/CFS group UpTimes decrease after the test.

Table 6.3. Tukey's HSD test—used to clarify differences between disease groups. P-values indicate that all disease groups have unique UpTime averages.

Disease Level	Difference	Lower	Upper	P-value (adjusted)
Moderate-Control	-12.19	-18.17	-6.20	2.13e-05***
Severe-Control	-25.54	-31.82	-19.25	0.00e+00***
Severe-Moderate	-13.35	-19.64	-7.06	9.80e-06***

Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

In this comparison, a baseline UpTime score was calculated by averaging the three days before the Lean Test: Monday, Tuesday, and Wednesday. This baseline was used for comparison when reviewing the proceeding days: Thursday, Friday, and Saturday. Therefore, the variable "Number of Days after Lean Test" has the following levels:

- Baseline *(average UpTime for Monday, Tuesday, and Wednesday)*
- 1 Day after Lean Test *(Thursday's UpTime)*
- 2 Days after Lean Test *(Friday's UpTime)*
- 3 Days after Lean Test *(Saturday's UpTime)*

UpTime averages for each group, shown in Figure 6.4(a-c), do not decrease following the Lean Test as expected. Rather, mean UpTimes for ME/CFS groups spike one day after the test.

Disease Severity	Average UpTime (%)
Control	46.4
	41.2
	33.9
	43.2
	28.7
	38.2
Moderate	35.9
	23.7
	29.4
	22.7
	21.7
	26.4
Severe	7.7
	11.0
	14.0
	18.3
	17.3
	12.6

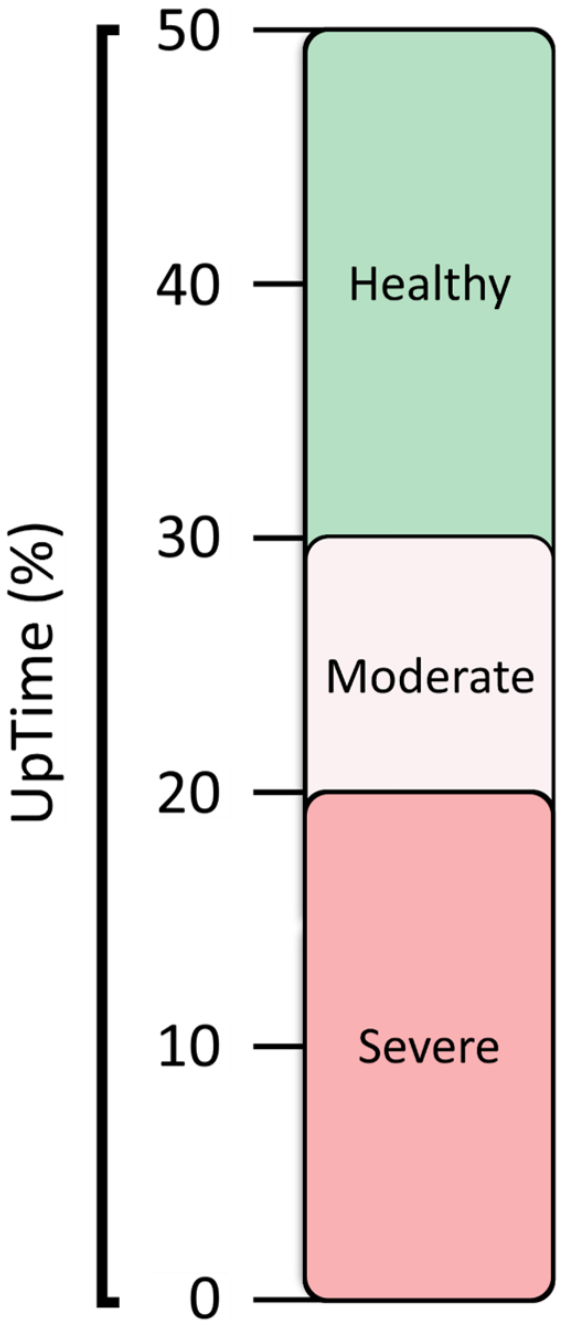

Figure 6.3. UpTime predictions for each group. Subject weekly mean UpTime scores are shown on the left, with bolded group averages. The scale on the right indicates the ranges of expected UpTime scores expected for each group.

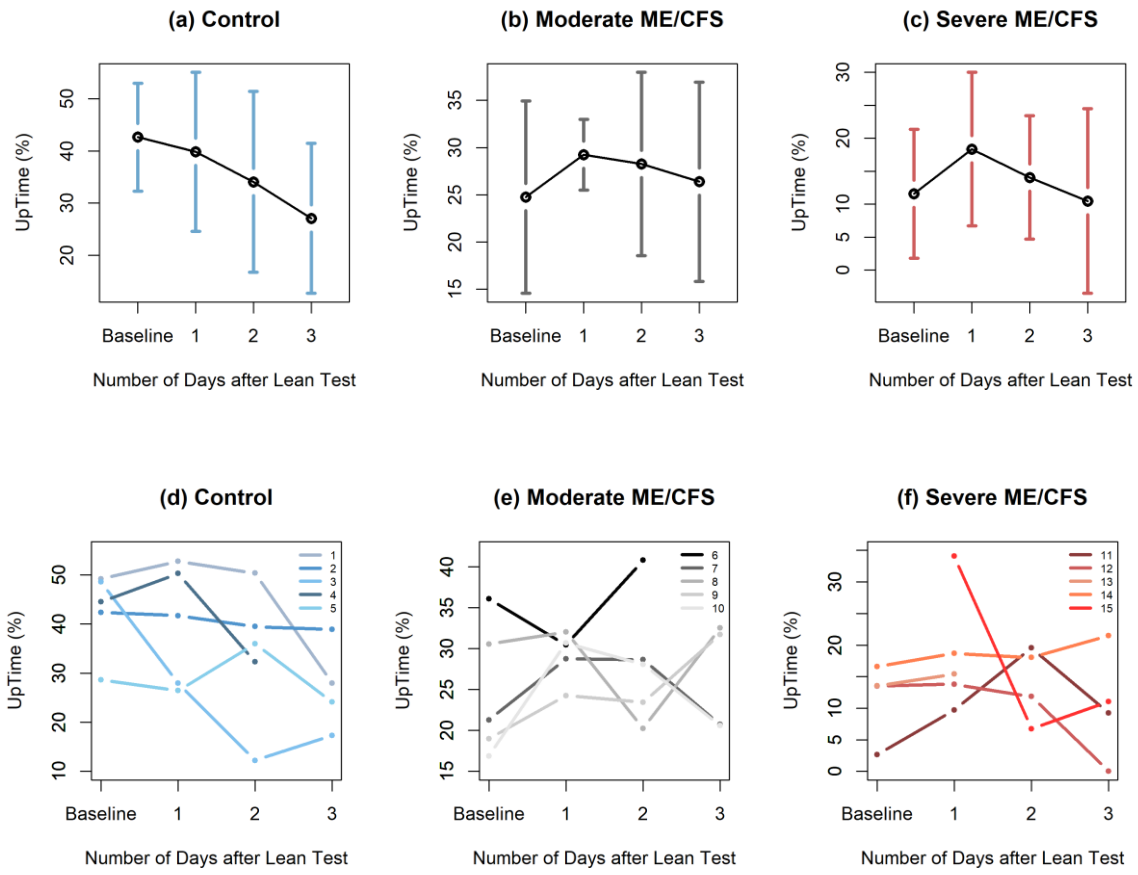

Figure 6.4. Group and individual UpTime plots. Group mean plots for UpTime (a-c); large confidence intervals are the result of UpTime variation within groups and data lost due to Shimmer failures. Individual UpTime trends (d-f).

Interestingly, the control group alone decreases after the Lean Test; however, this change is due to weekend relaxation rather than PEM. Furthermore, the ME/CFS groups' UpTime spikes could have been a direct result of participating in the NASA Lean Test. A 5-10% increase in UpTime equals roughly 1-2 hours of upright activity, which could easily be the amount of time required to drive to the BHC, take the Lean Test, and drive home.

Because the confidence intervals in Figure 6.4(a-c) overlap so heavily, we do not expect to find any significant difference in mean UpTimes before and after the NASA

Lean Test. This suspicion is confirmed by the high p-values shown in Table 6.4, which indicate that there are no significant differences in UpTime by day.

Despite the trends seen in Figure 6.4, we find ourselves forced to reject the consensus that activity decreases after PEM is induced. This finding can be explained in a few different ways. For one thing, we knew that sampling was insufficient for this study. It is possible that the ANOVA was unable to detect differences in means due to the scarcity of data available for this test. The power levels associated with these conclusions are 0.43 for controls, 0.14 for moderate ME/CFS, and 0.25 for severe ME/CFS. Because these power levels are so low, it is possible that we may have committed a Type II error—failing to reject a false null hypothesis. For this reason, we recommend that this test be performed again during a full-scale study.

Table 6.4. Single-Factor ANOVA tables comparing UpTime before and after Lean Test. For each disease level, UpTime means do not differ significantly.

	Factor	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F value	P-value
CONTROL	Days After Lean Test	3	630	210.0	1.658	0.218
	Residuals	15	1900	126.6	-	-
MODERATE ME/CFS	Days After Lean Test	3	58.8	19.59	0.431	0.734
	Residuals	15	682.0	45.47	-	-
SEVERE ME/CFS	Days After Lean Test	3	167.4	55.80	0.905	0.465
	Residuals	15	801.5	61.65	-	-

Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Experimental design flaws could be another potential cause for this result. On the first day of each trial, the subject traveled to and from the BHC to be equipped with the Shimmers and Oura Ring. Due to the extreme sensitivity of ME/CFS patients, this travel alone could have unintentionally induced PEM. With patients experiencing PEM throughout both phases (rather than just during days 4 through 6), we would expect to see constant UpTime scores. Requiring subjects to travel to and from the BHC could easily have affected UpTime scores for both ME/CFS groups. Future studies should consider home-visits to reduce this effect.

The floor effect could be an alternative explanation for these unexpected results; UpTime can only go so low. Baseline UpTimes for the ME/CFS groups could already be at minimum allowable levels. Further UpTime reductions could mean a significant decrease in lifestyle. (The quality of life for an individual with ME/CFS is already very low). Some subjects in the moderate ME/CFS group have part-time jobs; taking a few days off to recover from PEM may not be an option. For the severe ME/CFS group, it simply may not be possible to lower UpTime from their average four hours per day without going stir-crazy.

Lastly, constant ME/CFS UpTime scores could be a result of self-medication. Except for the morning of the Lean Test, ME/CFS subjects were permitted to take their prescribed medication(s) throughout the study. Subjects may have medicated more heavily following the NASA Lean Test to mitigate the effects of PEM, thus unintentionally flattening UpTime.

Whatever the reason, it is indisputable that the NASA Lean Test had no statistically significant effect on UpTime. A better experiment design would track each

subject for a longer period before and especially after the NASA Lean Test, thus establishing more accurate baseline UpTime scores for each subject. However, limitations in funding and time prohibited these design improvements. Further investigation is outside of the scope of this thesis but may provide deeper insight into the causes and effects of PEM.

6.4 HUA vs. UpTime

Moving on from the NASA Lean Test, we turn to an evaluation of HUA as a proxy for IMU-based UpTime scores. Until this study, the only tool researchers at the BHC had to evaluate daily upright activity was HUA—a questionnaire that crudely captures the amount of time an individual spends with the feet on the floor each day (see Figure 5.3). One purpose of this study was to develop an objective evaluation of upright activity with a high level of resolution; our IMU-based solution, using Shimmers to measure UpTime, has been described fully in the preceding chapters of this thesis. In this section, we look to compare UpTime (measured by the Shimmer devices) to HUA (calculated based on responses to the questionnaire).

Historically, HUA was reported in units of hours; however, we have converted HUA to a percentage of the day to accommodate its comparison to UpTime. Subjects were asked to fill out a HUA survey each day during the study; the results of these surveys are plotted with corresponding UpTime measurements in Figure 6.5. This figure shows that subjects generally tend to overestimate their UpTime. Indeed, a correlation plot—broken up by disease level—shows that HUA and UpTime are not correlated for both ME/CFS groups (Figure 6.6).

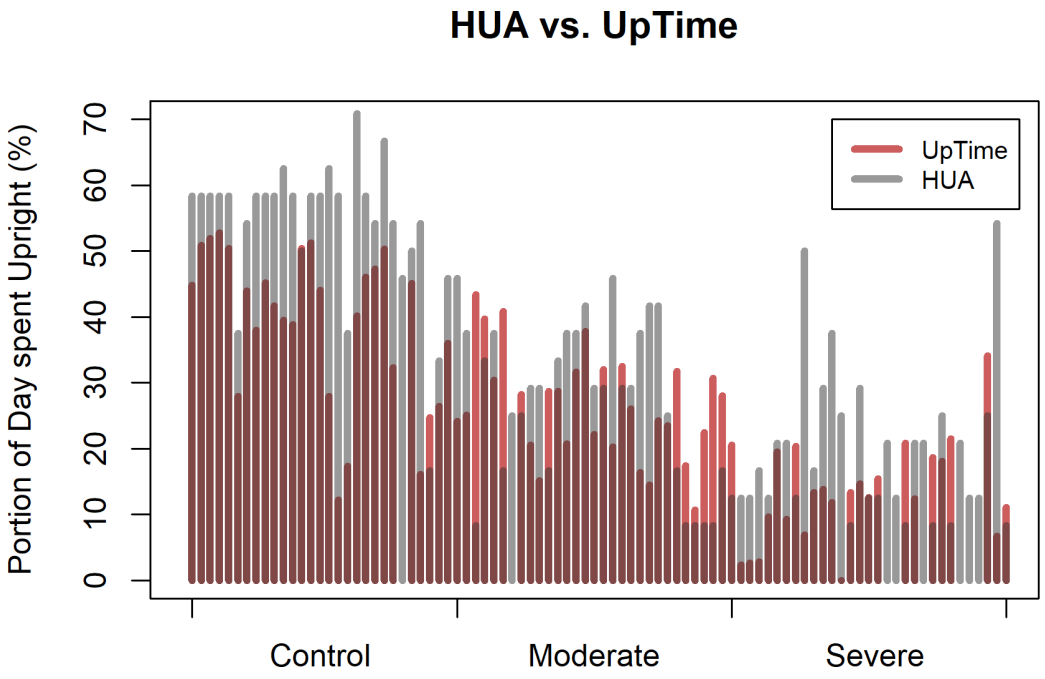

Figure 6.5. HUA and UpTime compared. The study's combined ninety days of data are displayed side-by-side. An entry where UpTime is not shown indicates a lack of Shimmer data.

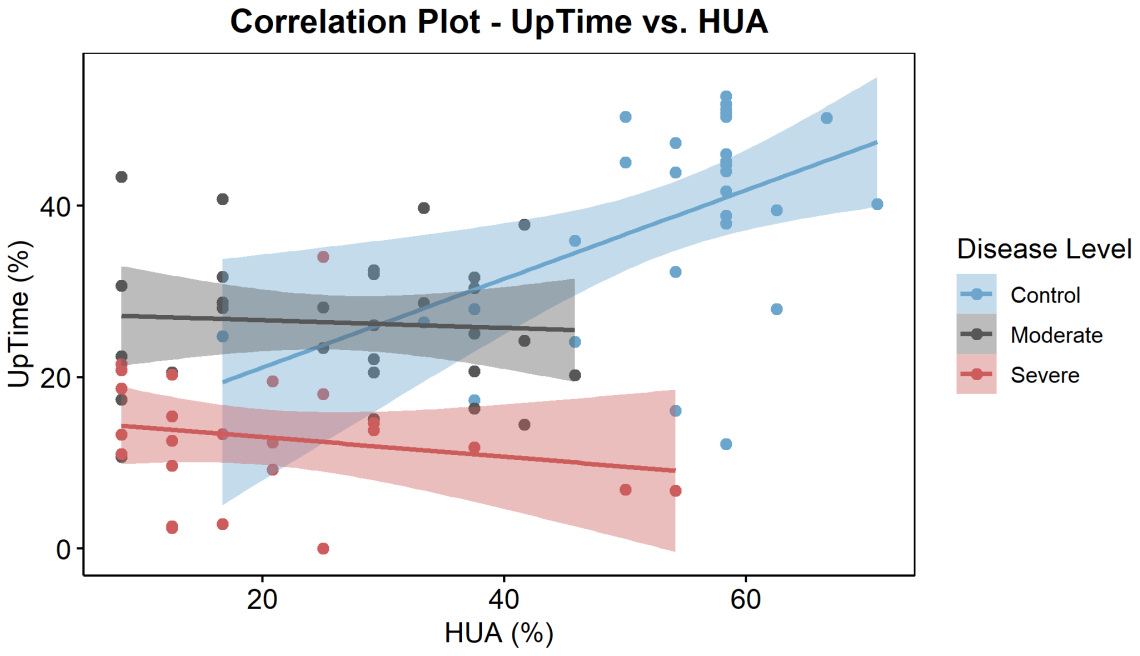

Figure 6.6. Correlation plots between UpTime and HUA (plotted separately according to disease level).

Both ME/CFS groups reported a wide range of HUA scores, while UpTime remained relatively invariant. This non-correlation is illustrated by the horizontal grey and red lines in Figure 6.6. Conversely, the control group estimated UpTime with some level of accuracy. We see a positive, linear correlation between UpTime and HUA for this group shown by the blue line in Figure 6.6. However, a multitude of blue outliers suggests the weakness of this correlation. The non-correlation between HUA and UpTime can be seen in Figure 6.7. HUA and UpTime follow similar trends for controls, but ME/CFS group comparisons differ widely. For all disease levels, HUA generally appears to overestimate UpTime by 5-10%.

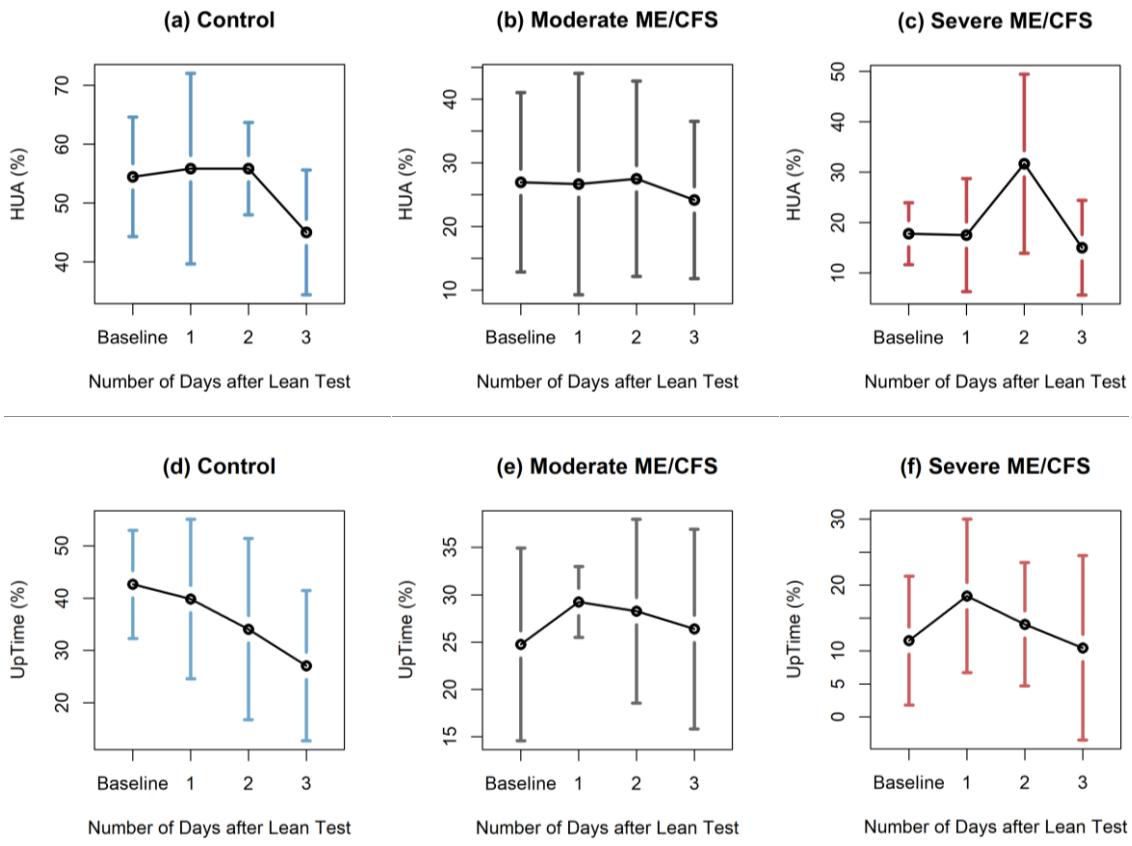

Figure 6.7. Mean plots for HUA (a-c) and UpTime (d-f), separated by disease level. This figure allows for trend comparison between the two measurements by disease level.

A paired t-test comparing all HUA and UpTime scores yielded a p-value of 2.72e-05, confirming that the two measurements are significantly different. The corresponding 95% confidence interval for the true mean difference is (4.17, 10.91). This interval means that we are 95% confident that the average difference between HUA and UpTime is between 4.17 and 10.91%. Given the range of expected UpTime for the Moderate ME/CFS group appears to be between 20-30%, any measurement system needs to have less than 10% UpTime error to accurately identify ME/CFS disease severity. For this reason, the increased resolution and accuracy of IMU-based UpTime will provide considerable benefit to clinicians monitoring patient improvement.

6.5 Oura Ring Data

A secondary interest of this research was to explore contributors to unrefreshing sleep. The Oura Ring collects an incredible amount of sleep-related data; for this reason, we hoped its use would further our understanding of the causes and effects of unrefreshing sleep. In this section, we review and discuss the Oura Ring measurement types that were expected to have some link to sleep efficacy and overall wellness: (1) Sleep Score, (2) Hours of Sleep, (3) Awake Time, (4) Readiness Score, (5) Lowest Resting heart rate (HR), (6) Activity Score, and (7) HRV. Much of the data processing done by the Oura Ring is proprietary, but each measurement type will be explained with as much detail as Oura has provided on their website [33]. Because the Oura Ring was not the primary interest of this thesis, analysis of each measurement type was only performed at a surface level. If the findings of an initial evaluation appeared significant, further analysis was performed.

6.5.1 Sleep Score

The Oura Ring estimates a daily Sleep Score, measured between 0 and 100%, which is a general evaluation of sleep quality. Sleep Score is a weighted sum of time spent in each stage of sleep stage (REM, light, and deep), sleep efficiency, sleep latency, restfulness, and sleep timing. Oura has not detailed the specific combination of these variables, but they do identify a sleep score of approximately 80% as normal for healthy adults. Because individuals with ME/CFS experience unrefreshing sleep, we expected their Sleep Scores to fall below this 80% threshold. Instead, we found that all disease level mean Sleep Scores were comparable—none of which were too far from "normal" (Figure 6.8(a)).

6.5.2 Hours of Sleep

The Oura Ring records total sleep time as "Hours of Sleep" (Figure 6.8(b)). This measurement type shows that all groups are within the normal range for healthy adults—

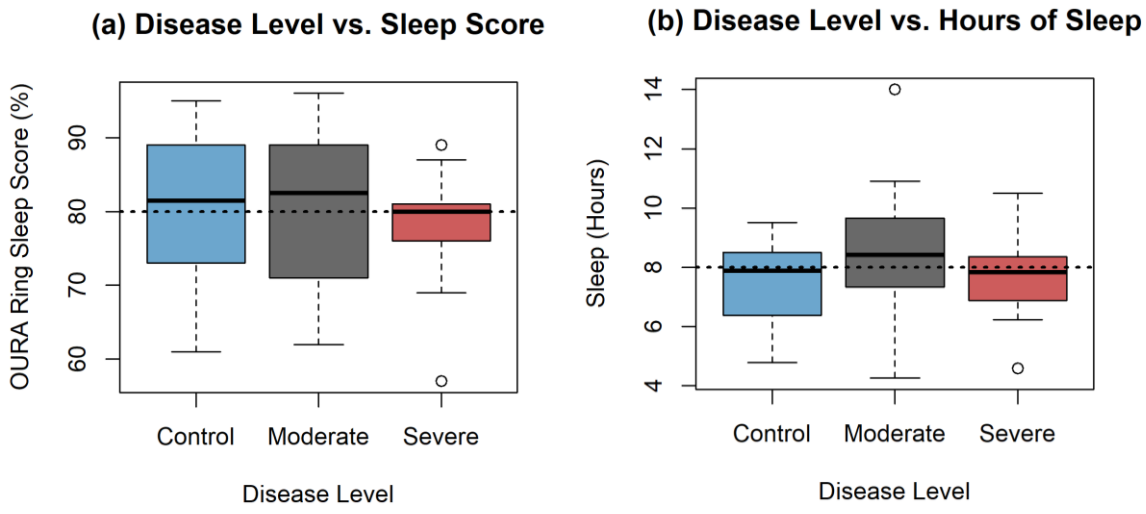

Figure 6.8. Oura Ring Sleep Scores (a) and Hours of Sleep (b), separated by disease level. Horizontal dotted lines identify normal scores for each measurement type.

7 to 8 hours. However, it is important to remember that ME/CFS patients are not healthy, and what is considered “normal” sleep quality/quantity is likely different for them. We conclude that the unrefreshing nature of sleep for individuals with ME/CFS is unrelated to sleep duration.

6.5.3 Awake Time

Another way to evaluate sleep quality is by measuring Awake Time. The Oura Ring calculates this value by summing the time spent awake in bed before falling asleep (sleep latency) and after falling asleep. A night of restless sleep will yield a higher Awake Time. Generally, a higher Awake Time will lead to a feeling of tiredness the next day. Because ME/CFS groups experience constant fatigue, we expected that they might have higher Awake Times; however, this measurement’s only clear trend is Awake Time increasing over the weekend for controls (Figure 6.9(a)).

Personal scores for subjects with ME/CFS tend to be a bit erratic—possibly due to measurement reliability. We cannot draw any useful conclusions about the connection between Awake Time and unrefreshing sleep from this analysis. Furthermore, Figure 6.9 shows no indication that sleep quality increases or decreases after the NASA Lean Test—performed on day 4.

6.5.4 Readiness Score

The Oura Ring’s Readiness Score—measured from 0 to 100%—is a combination of resting heart rate, body temperature, previous night sleep, and previous day activity. Oura has not explained how these contributors are used to calculate Readiness Score. We

do know that a Readiness Score below 70% indicates that the user should take the day to recover and avoid overexertion. All groups averaged Readiness Scores above 70% every day of the study (Figure 6.10).

Figure 6.10 shows that ME/CFS groups tended to have higher Readiness Scores than the control group, indicating that they were better prepared for days of intense physical activity. Oura's algorithm calibration process could explain this unexpected

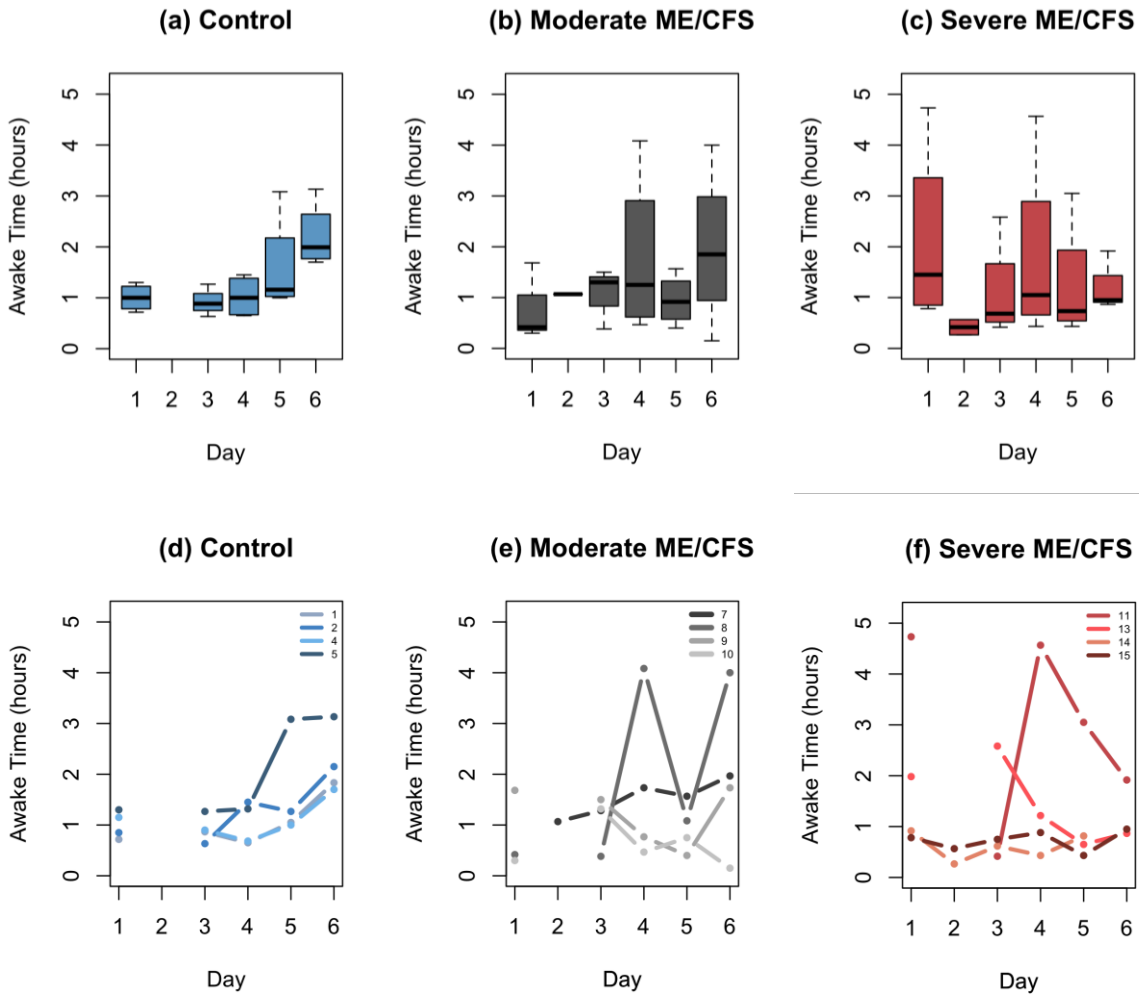

Figure 6.9. Mean plots of Awake Time. For each group (a-c). Individual daily Awake Time measurements for each group (d-f). Some data were lost due to Oura Ring failures, including all of day 2 for the controls.

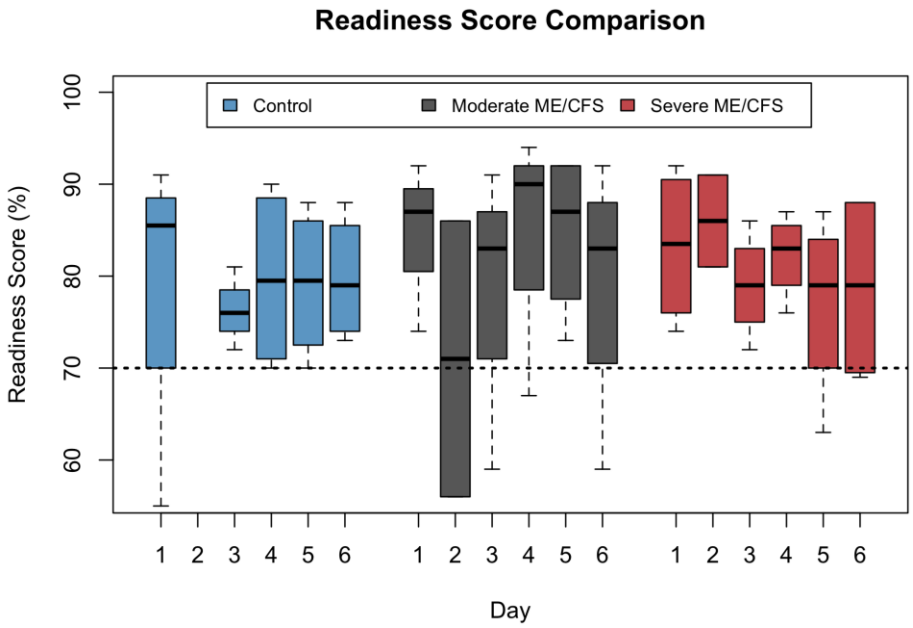

Figure 6.10. Oura Ring Readiness Score separated by group and by day. The threshold of general readiness, 70%, is identified by a horizontal dotted line. Some data were lost due to Oura Ring failures, including all of day 2 for the controls.

result. Readiness Score was calibrated for healthy subjects rather than for people with ME/CFS; in other words, the low activity levels of the ME/CFS groups may have tricked the algorithm into thinking they were more prepared for higher activity when, in fact, they were not. Like the Sleep Score, this unexpected result could be the effect of evaluating a sick population using normal baselines calibrated for healthy individuals.

6.5.5 Lowest Resting HR

During sleep, the Oura Ring measures the user's heart rate at 10-minute intervals. The lowest of these measurements is called Lowest Resting HR; a lower resting HR indicates better sleep quality and better overall health. We expected both ME/CFS groups to have resting HR scores above their healthy counterparts. Additionally, we expected the

onset of PEM post-Lean Test to cause Lowest Resting HR measurements to rise for days 4, 5, and 6. However, the data shown in Figure 6.11 refutes our expectations on both counts.

While group means for Lowest Resting HR did not meet our expectations, we do see an interesting trend in Figure 6.11. For some reason, of all three disease levels, the moderate ME/CFS group has the lowest scores for this measurement type. An ANOVA test confirms that Lowest Resting HR scores do indeed differ significantly by disease level (Table 6.5). Further clarification from Tukey's HSD test shows that Lowest Resting HR averages differ for every disease level (Table 6.6).

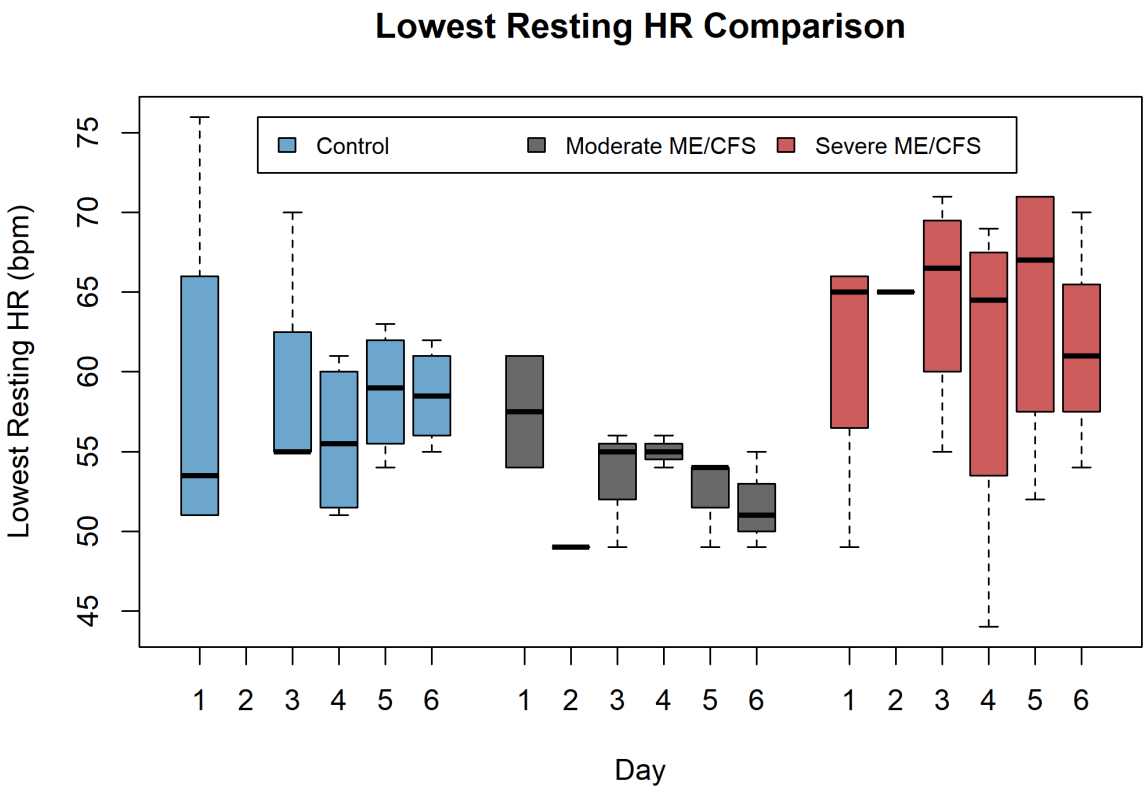

Figure 6.11. Lowest Resting Heart Rate means, as measured by the Oura Ring. Data is separated by group and by day. Some data were lost due to Oura Ring failures, including all of day 2 for the controls.

Table 6.5. Multiple-factor ANOVA table for Lowest Resting HR data. Data collected by Oura Rings. Disease levels are shown to have statistically significant means.

Factor	Degrees of Freedom	Sum of Squares	Mean Sum of Squares	F value	P-value
Disease Level	2	775.4	387.7	30.337	5.03e-08***
Day	5	35.4	7.1	0.555	0.734
Subject	8	1585.5	198.2	15.509	6.95e-09***
DiseaseLevel:Day	9	87.3	9.7	0.759	0.654
Residuals	31	396.1	12.8	-	-

Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Table 6.6. Tukey's HSD test comparing Lowest Resting HR by group. Data recorded by Oura Rings. All groups are shown to have statistically different means.

Disease Level	Difference	Lower	Upper	P-value (adjusted)
Moderate-Control	-4.65	-7.66	-1.64	0.0017**
Severe-Control	4.71	1.96	7.46	0.0006***
Severe-Moderate	9.36	6.39	12.34	0.0000***

Significance codes: 0 '*' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

While we are not surprised to see that each group differs in terms of Lowest Resting HR scores, the actual order defied our expectations:

1. Lowest Resting HR: Moderate ME/CFS
2. Middle Resting HR: Controls
3. Highest Resting HR: Severe ME/CFS

One possible explanation is that Lowest Resting HR varies widely from person to person, even for healthy adults. The unexpected results here could easily be explained by the presence of normal measurement variation, compounded by sub-optimal sample size. Each subjects' Lowest Resting HR scores (Figure 6.12) reflect this strange ordering. (It is also worth noting that the Oura rings failed to record Lowest Resting HR on several occasions, as evidenced by missing data in Figure 6.12.)

Even though the control group is made up of non-ME/CFS subjects, this group is only healthy insomuch as its subjects do not have ME/CFS. Some controls have other ailments, including heart conditions, which could also explain the unexpectedly high resting HR of the control group.

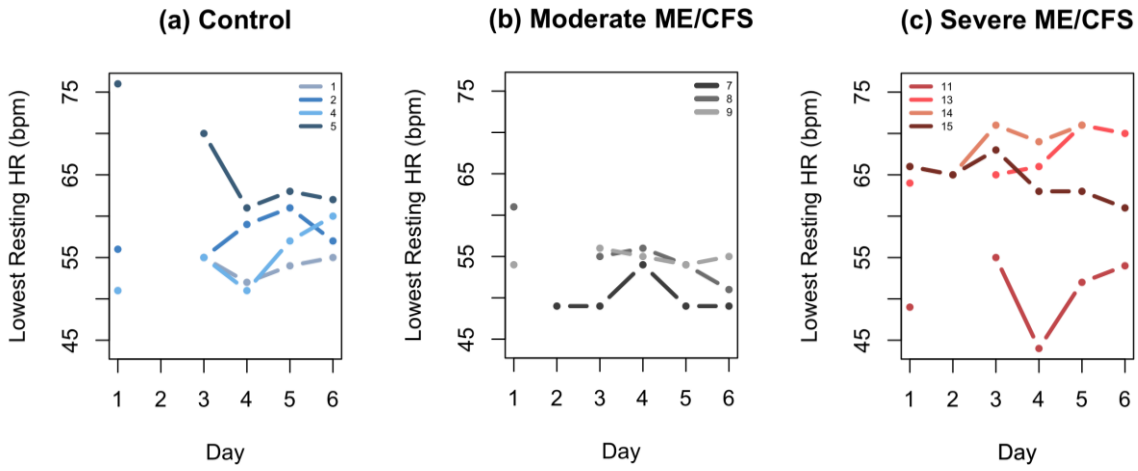

Figure 6.12. Lowest Resting HR separated by subject and sorted into groups.

Although we cannot pinpoint the exact reason disease levels are ordered in this unexpected manner, Lowest Resting HR does appear to differentiate between the moderate and severe ME/CFS groups. This distinction could explain more about the effects of ME/CFS as it increases in severity. Unfortunately, sample sizes are too small to make this claim with confidence. For this reason, it could be wise to focus on this measurement type during future investigations when funding allows for data collection from appropriately sized samples.

6.5.6 Activity Score

Another proprietary measurement of the Oura Ring is Activity Score. This measurement type gauges the user's recent activity, measured from 0 to 100%. Activity Score is unique because it is initially set at maximum, and then adjusted down to match the user's *actual* activity over time. Once the Activity Score has normalized to the user, this score will accurately measure the user's activity level. To some extent, this score normalization explains the downward trend seen in Figure 6.13(b) throughout the week; another portion of this downward trend is likely due to actual daily UpTime differences.

At first glance, a comparison of each group's Activity Score does not appear to show any significant differences (Figure 6.13(a)). However, a subtle difference can be seen when the groups are compared side-by-side (Figure 6.14(a)).

We believe that a combination of factors is at play here. For one thing, all data is normalizing from the first day's Activity Score (set at maximum in all cases) towards Activity Scores representing *actual* activity levels. To some extent, this explains the downward trends for all groups. We also know that the control group's UpTime

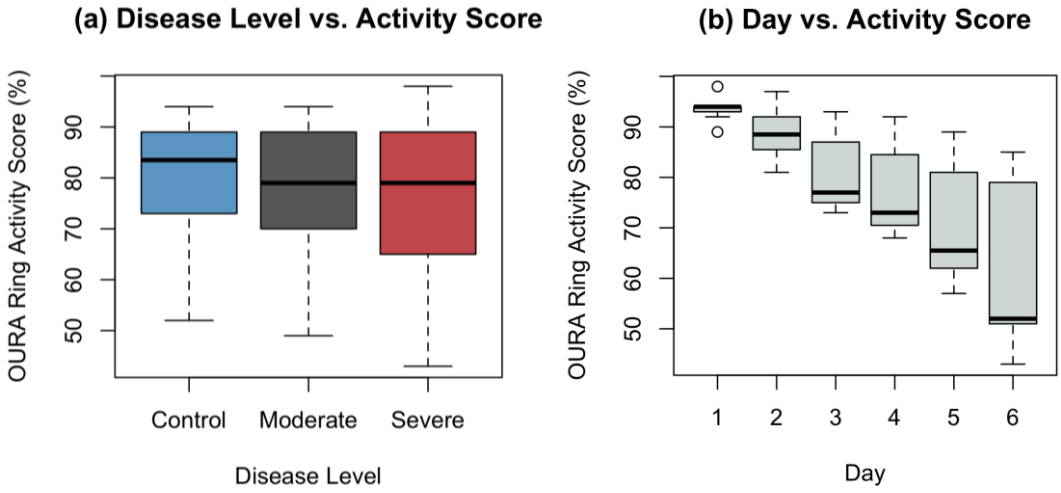

Figure 6.13. Mean plots of Oura Ring Activity Scores. Average group Activity Scores (a) are nearly identical. A weekly trend becomes clear when data are grouped by day (b).

decreases over the weekend (see control data in Figure 6.14(b)); this would explain the downward trend in Activity Score as well.

While we do not know the official number of days required to normalize Activity Score, we can see that both ME/CFS groups separate from the control group around day 3 or 4 (Figure 6.14(a)). By days 5 and 6, this separation is even more distinct. Future studies with longer data collection periods will allow researchers to compare Activity Scores more accurately. Oura Rings should be distributed to subjects well before the start date of the study to ensure that the Oura Ring's Activity Score normalization period has passed by the time data collection begins.

Evaluation of the Oura Ring's Activity Score was performed to see if UpTime could be approximated using an existing commercially available smart device. Figure 6.14 shows that the Oura Ring is not capable of efficiently capturing activity levels for individuals with ME/CFS. The IMU-based UpTime measurements developed in this thesis do not require this week-long normalization period before accurate activity

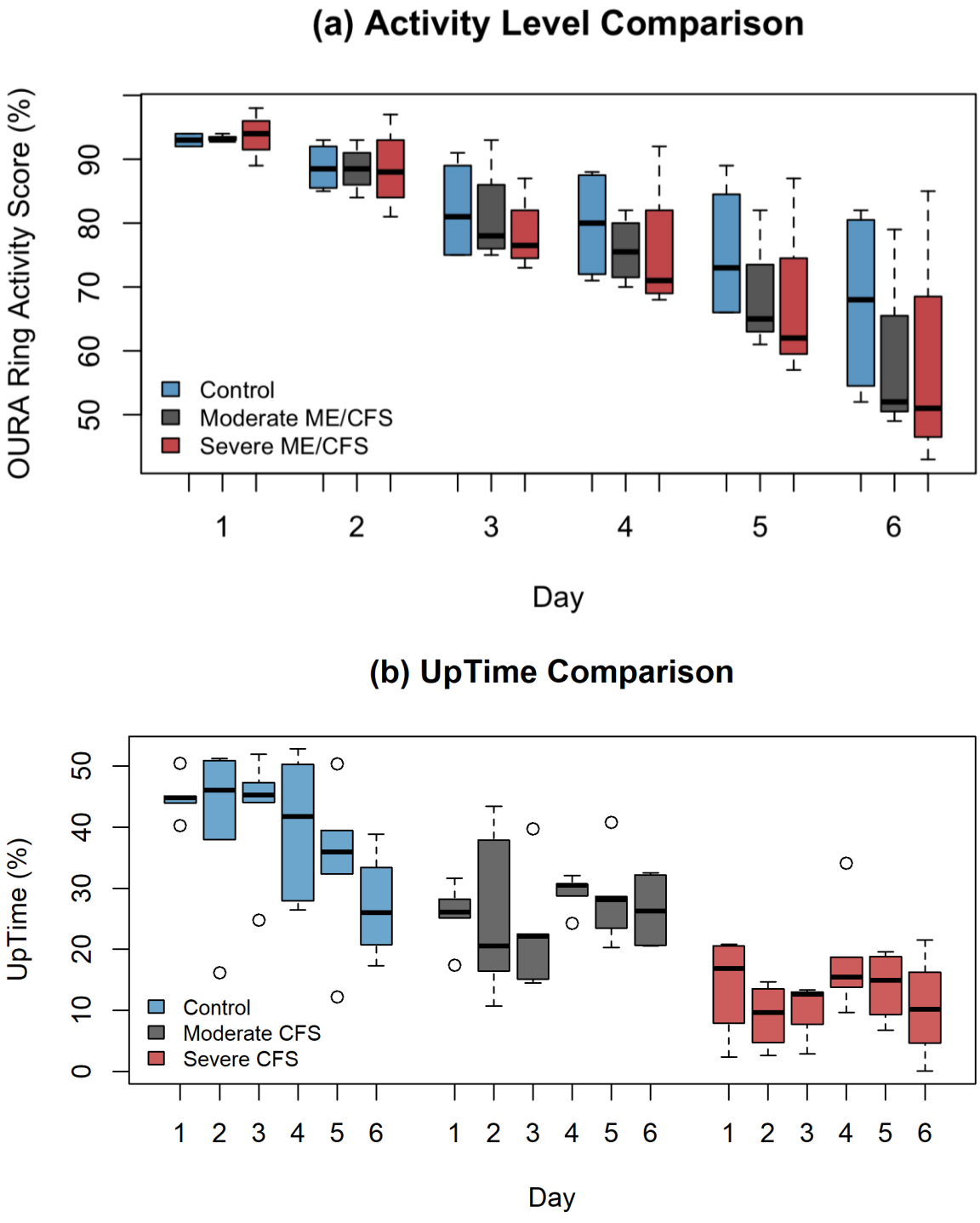

Figure 6.14. IMU-based UpTime compared to Oura Ring Activity Level. Activity Score separated by group and day (a). UpTime comparison—for trend comparison—also separated by group and day (b).

measurements can be recorded. Furthermore, placing an IMU on the lower leg (rather than on the hand) provides a more direct measurement upright activity.

6.5.7 HRV

The final Oura Ring measurement considered in this thesis is heart rate variability (HRV). This measurement quantifies the time-variation between heartbeats. HRV has been shown to correlate with stress levels and general health, with higher HRV indicating low stress [34] and good heart health [35]. A healthy heart tends to be a bit looser with the timing of its beating, thus increasing heart rate variation; when stress increases, the timing of the heart will become more precise, thus decreasing heart rate variation. The Oura Ring records HRV scores each night as the user sleeps, thereby ensuring consistent measurement circumstances and allowing for more comparable readings.

HRV varies widely from one person to another and is, therefore, most useful when evaluated as an individual trend. Due to its natural response to stress and illness, we expected average HRV to be lower for both ME/CFS groups. Moreover, we expected HRV to decrease for the ME/CFS groups after the NASA Lean Test due to the onset of PEM. Despite our expectations, HRV trends for almost all ME/CFS subjects were invariant throughout the week (Figure 6.15(b & c)). Figure 6.15(a) shows that the control group's HRV scores trended lower than we expected (Figure 6.15).

The balance of how stress and illness affect HRV could explain why the control group had lower HRV scores than we expected. For this group, a decrease in HRV scores (Figure 6.15(a)) makes sense on days 5 and 6, given the raised resting heart rates

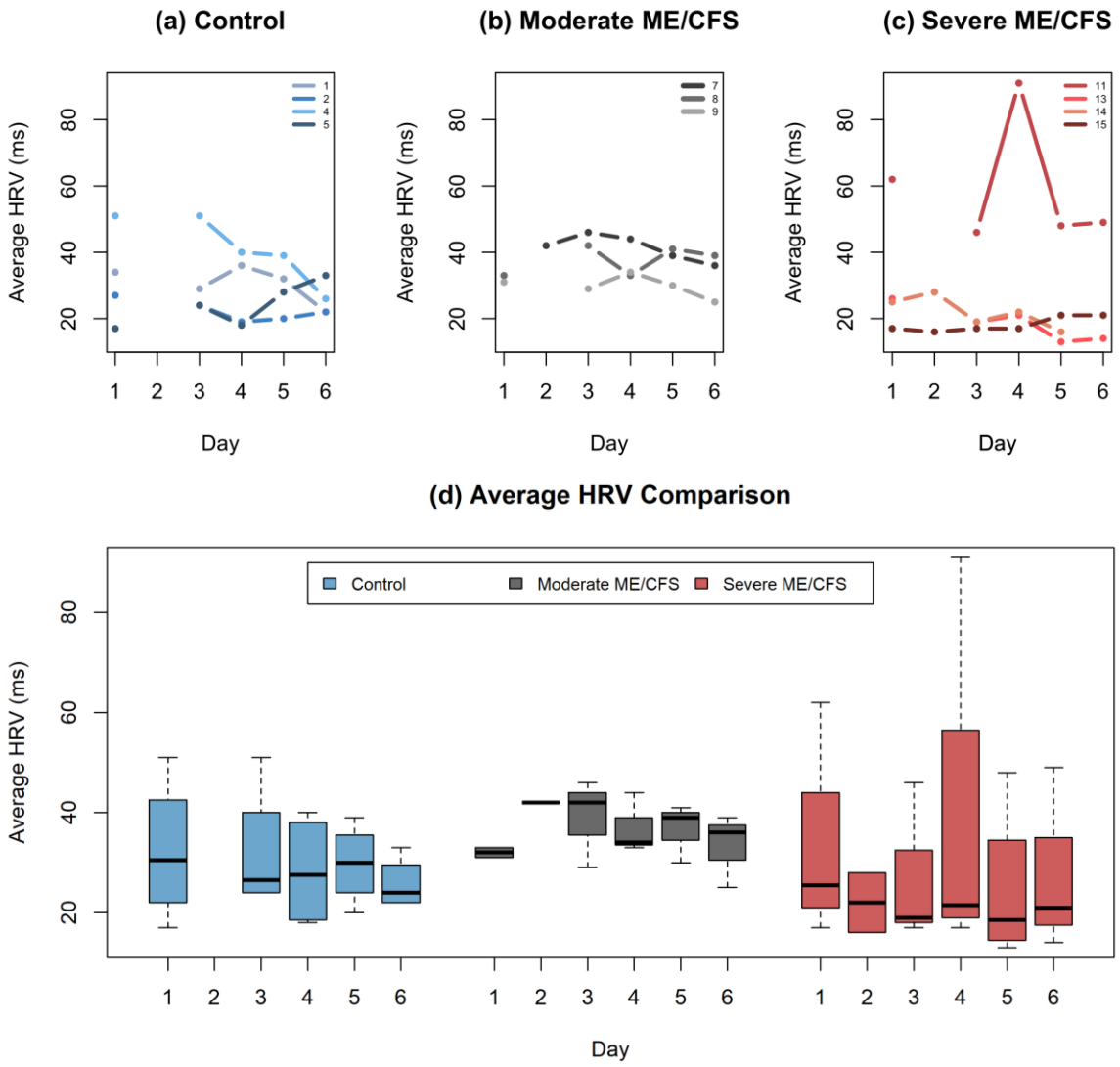

Figure 6.15. HRV measurement comparison. Subject Average HRV measurements (a-c), grouped according to disease level. Numerous gaps in the data exist, commonly occurring on the second day after equipping the Oura Ring. Subject 11's Oura Ring may have malfunctioned. The Oura Ring failed to collect data from subjects 3, 6, 10, and 12. HRV scores are also compared by disease level and day (d).

(Figure 6.11) and reduced sleep quality (Figures 6.9(a)) associated with the weekend. There does appear to be a difference in average HRV for the ME/CFS groups. Subjects with severe ME/CFS average HRV scores seven milliseconds lower than their moderate ME/CFS counterparts. Unfortunately, Oura Ring failures limited the sample size for this measurement type below what is needed to pursue a statistical analysis of this difference. Future studies should continue this investigation of HRV differences between ME/CFS groups.

Another trend we see in Figure 6.15(b-c) is invariant HRV scores from all but one ME/CFS subjects. Invariant HRV could be a side-effect of ME/CFS caused by the floor effect; HRV can only go so low, especially with the severe ME/CFS group. Alternatively, invariant HRV could be a valuable biomarker of ME/CFS disease severity, with lower scores corresponding with higher disease severity. As previously discussed, the stress-inducing demands of day one could have caused the onset of PEM for the diseased groups—another possible explanation for the ME/CFS groups' unchanging HRV. Again, we recommend a future study with a longer data collection period. Home visits may also help to reduce stressors associated with attaching devices at the beginning of the study.

Overall, the Oura Ring measures a multitude of clinically valuable data; however, it does so with intermittently occurring failures. In most cases, it is difficult to draw any robust conclusions from the Oura Ring data without a longer study and more samples. Another point worth noting is that the Oura Ring was calibrated on healthy people and was meant to be used as a tool to improve physical fitness. Score errors may have been produced simply by using this device to track diseased groups rather than the intended population. Regardless of the cause, we have seen numerous Oura Ring measurements

mark ME/CFS groups as “normal.” Additional measurements have shown the ME/CFS group to be unresponsive to the NASA Lean Test and ensuing PEM.

6.6 Limitations of Results

While results appear to be significant in some cases, they are not necessarily representative of the broader population. There are limitations on the generalizability of group UpTime ranges due to imperfect experiment design and insufficient sampling. Before beginning this study, we knew that our sample sizes were small enough to prevent us from generalizing our conclusions. However, we did not anticipate the device failures that caused our effective sample sizes to shrink even further. Both Shimmers and Oura Rings suffered from lost data, though individual device failure rates differed significantly.

Most often, Oura Rings failed to record data the second day after the device was equipped (Figures 6.9, 6.12, & 6.15). The Oura Ring’s Activity Score requires several days of use before adjusting to the user’s activity level. This adjustment period caused the relevant collected data to be skewed for days 1-3. We recommend that future studies equip each subject with an Oura Ring a full week in advance of the collection period to prevent either of these Oura Ring failures/shortcomings from impacting data analysis.

As for Shimmers, failures occurred randomly. The cause of gaps remains unknown. Fortunately, the validation of single-Shimmer data allowed us to salvage 72% of data that would have otherwise been lost as a result of these failures. It is recommended that researchers pursue in-house development of a custom IMU device for future UpTime-related studies. A custom device may function with a substantially lower

failure-rate than the Shimmers. Additionally, we recommend that researchers randomize the use of specific IMU devices to eliminate remaining biases in collected data.

The final limitation comes from the period selected for data collection. As we have identified, weekdays differ from the weekend in terms of activity, sleep, and other clinical outcomes. Regularly occurring trends made it difficult to accurately determine what portion of observed changes were caused by the NASA Lean Test. Future studies should randomize each subject's start day to remove day of the week as a confounding factor. This randomization was not done in our study because the BHC is only open Monday through Thursday, and it was vital to minimize study costs due to a lack of external funding.

6.7 Evaluation of UpTime Variability

In this section, we briefly address variations in UpTime measurements. When separated by group, the data collected in the case-control study has both day-to-day and subject-to-subject variability. Day-to-day variability describes the differences in UpTime scores caused by specific days of the week; this would include differences between weekday UpTime and weekend UpTime. Subject-to-subject variability describes differences in UpTime scores seen between the five subjects in each group. We can compare absolute UpTime variability using standard deviations as shown in Table 6.7.

For all disease groups, Table 6.7 shows that subject-to-subject variations exceed day-to-day variations. This hierarchy of variation is also shown in Figure 6.16, where daily UpTime means are more closely clustered together than subject UpTime scores for each day.

Table 6.7. Standards of deviation for collected UpTime data.

Variability Type	Standards of Deviation		
	Control	Moderate ME/CFS	Severe ME/CFS
Day-to-day	4.04	2.52	3.79
Subject-to-subject	10.98	8.08	7.03

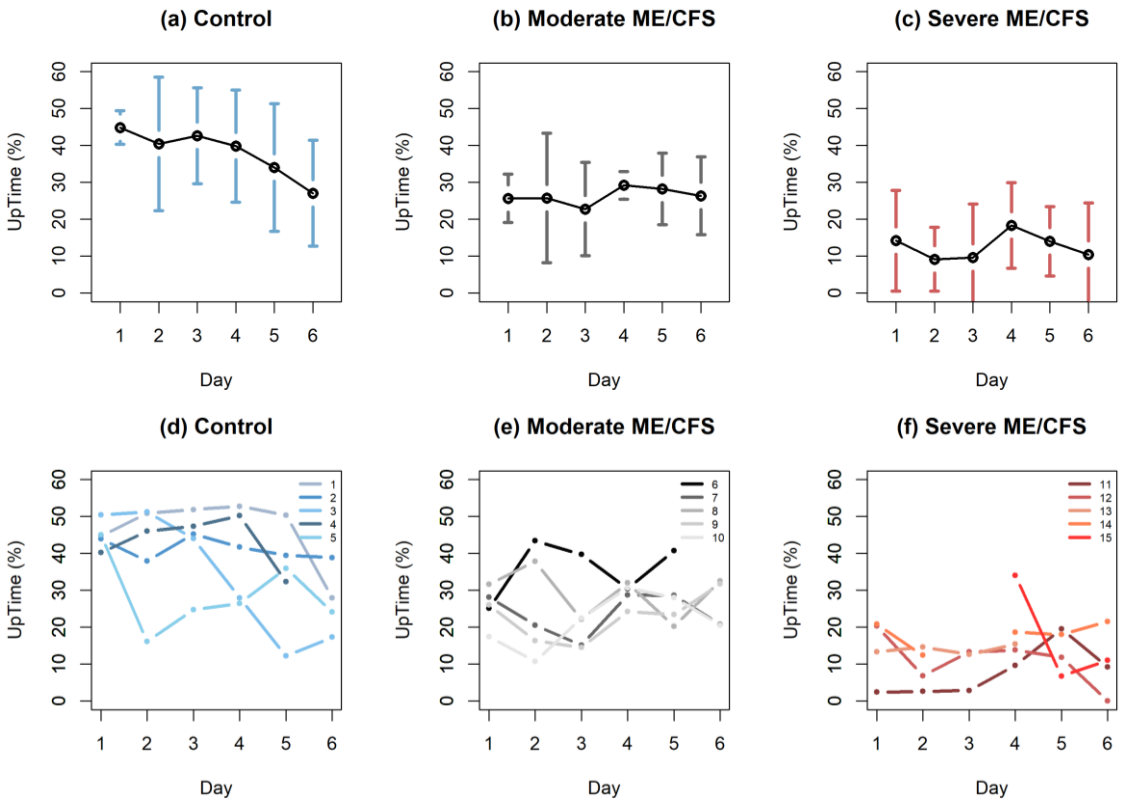

Figure 6.16. UpTime means and confidence intervals for each day of the week. Days correspond to Monday through Saturday. UpTime grouped according to disease level (a-c). Individual UpTime trends plotted by group (d-f).

In terms of day-to-day variation, Figure 6.16(a) shows that the control group's UpTime scores drop over the weekend. Furthermore, the control group has the greatest day-to-day UpTime variability of all three disease levels. Both ME/CFS groups have day-to-day variability, but this appears to be due to random variability rather than weekly trends (Figure 6.16(a & b)). Another day-to-day variability trend seen for the ME/CFS groups is that days one and four—the two days subjects were required to travel to the BHC during the study—correspond to local mean UpTime maximums. As was previously recommended, necessary subject interaction (e.g., IMU device attachment, NASA 10-minute Lean Test) could be done at each subject's home to minimize day-to-day UpTime variability for the ME/CFS groups.

We expected subject-to-subject UpTime variation to be greatest for the control group, given that they have the widest range of expected UpTime scores. This is confirmed by the results of Table 6.7. Figure 6.16(d-f) shows that the control group's UpTime scores have the widest spread of all three disease levels. Table 6.7 shows that subject-to-subject UpTime variation decreases as ME/CFS disease level increases in severity.

This leads us to another evaluation of UpTime variation that is especially useful in a clinical application. When evaluating treatment efficacy, it is necessary to distinguish between a patient's day-to-day UpTime variability and UpTime changes caused by a treatment; this can be done using a control chart. While control charts are generally used to monitor a process, they can also be used to evaluate performance improvement post-treatment. An I-Chart is a type of control chart with two phases. Phase I uses historical data (before a treatment) to create upper and lower control limits. Phase II data (after

treatment) can then compared against this historical baseline to see if the treatment signals additional variation beyond the controlled variation seen in Phase I. In future full-scale studies, this process can be used to evaluate treatment efficacy. If a change is identified, it can then be tested using a paired t-test. This same approach can be used to monitor patient improvement over time.

6.8 Conclusion

Regardless of device and experimental design weaknesses, many of the results of this preliminary study appear promising. While previous research showed that HUA differed by disease group, further analysis was required due to the weaknesses of HUA as an assessment of upright activity. In this research, we have used IMU-based UpTime measurements to conclusively show that upright activity differs significantly by disease group, thus validating UpTime as an efficacy endpoint for ME/CFS. Despite our expectations, the NASA Lean Test did not cause significant UpTime reductions, although we did see a clear decreasing trend over the weekend for controls. We suspect that the diseased groups' unchanging UpTimes pre- and post-Lean Tests are due to the unintentionally high stress caused at the start of the study; unchanged UpTimes could also be the result of subject medication.

By comparing HUA to UpTime, we have shown that the two are poorly correlated. Specifically, the control group was able to estimate UpTime with some level of accuracy, but both ME/CFS groups reported inaccurate levels of upright activity. This non-correlation further proves the value of UpTime as an objectively measured clinical outcome.

Oura Rings were used with varying rates of success. Despite device failures, we made several claims that could prove useful in guiding more extensive future studies. Analysis of lowest resting HR indicated that the moderate and severe ME/CFS groups were clustered separately, highlighting the differences between disease levels. All groups showed decreasing Activity Scores throughout the week. This is understood to be an effect of the method the Oura Ring uses to acclimate to each individual's activity level. Even so, when compared to the control group, the diseased groups appeared to have lower Activity Scores on days 5 and 6. Further research is required to draw any substantial conclusions regarding the Oura Ring's Activity Score. Lastly, the invariant HRV scores of ME/CFS groups could be due to the floor effect. Alternatively, PEM could have been induced on day 1, thereby preventing the possibility of change in HRV scores before and after the NASA Lean Test.

CHAPTER 7

CONCLUSION

7.1 Conclusions & Contributions

The main purpose of this research was to determine the accuracy of IMU-based UpTime measurements. Once verified, UpTime was used to evaluate two potential efficacy endpoints for ME/CFS. These endpoints are (1) the amount of time an individual spends upright and active, and (2) the extent to which individuals with ME/CFS experience unrefreshing sleep. HUA questionnaires are the current clinical standard for assessing time spent upright and active. Fundamentally, HUA is an inaccurate estimator of activity and is further limited by its low resolution—it is reported as a whole number of hours. To improve the method used to assess upright activity—our primary efficacy endpoint—we developed an IMU-based assessment of upright activity, which we call UpTime.

UpTime is an objective measurement of upright activity, which is based on a continuous assessment of lower leg angle. In this thesis, we have provided evidence supporting the accuracy of this measurement method. UpTime uses a critical angle of 39 degrees from vertical as a threshold to divide posture into two categories: (1) upright and (2) not upright.

As part of this research, we planned and executed a six-day study with 15 subjects

divided equally into three groups—control, moderate ME/CFS, and severe ME/CFS. Analysis of the UpTime data collected during this study indicates that these groups spend different proportions of their days upright and active. Healthy individuals are expected to have weekly UpTime scores above 30%, subjects with moderate ME/CFS are expected to have weekly UpTime scores between 20% and 30%, and subjects with severe ME/CFS are expected to have weekly UpTime scores below 20%. To ensure that these results have not been significantly confounded by activity trends relating to days of the week, future studies should randomize each subject's start day. Future studies should also cover an extended period beyond the six-day limit of our initial study.

Another objective of our study was to evaluate the effects of PEM brought on by the NASA Lean Test. However, our results showed no change in UpTime before and after the NASA Lean Test. Although this contradicts our expectations, we have confirmed that this test is humane; patients with ME/CFS do what they can to avoid stress-causing exertion, but we have seen that this test does not cause a drastic decrease in UpTime—indicating that the test does not substantially hurt them. While further investigation is beyond the scope of this thesis, future studies should incorporate home-visits to reduce the stress caused by participation, thereby ensuring that PEM is only induced during the Lean Test.

Due to an unexpectedly high Shimmer failure rate, 20% of all days had data from only one leg. After further analysis, we were able to validate the calculation of UpTime using only one Shimmer. The ability to measure UpTime using a single IMU device provides the possibility to reduce device-related costs associated with measuring UpTime. Furthermore, subjects could attach the device to the leg of their choice, thereby

increasing overall comfort.

We used Oura Rings to address our secondary efficacy endpoint—unrefreshing sleep. Among other less useful measurement types, the Oura Ring took daily measurements of Awake Time, HRV, and Hours of Sleep throughout the study. Analysis of these measurements showed that the NASA Lean Test had no significant effect on sleep quality for diseased groups—meaning that sleep was not shown to be any more or less refreshing while PEM was induced. In many cases, both healthy and diseased individuals had comparable measurements, indicating that sleep requirements may not differ by group. In terms of resting HR and HRV, Oura Ring measurements showed significant differences between the moderate and severe ME/CFS groups.

This research shows the value of UpTime as an objective replacement for HUA. While there are numerous smart devices with integrated IMUs available, most are designed to help healthy individuals achieve fitness goals. No existing smart device is specifically designed to be used on the lower leg for the evaluation of ME/CFS clinical indicators; hence the need to develop and validate an IMU-based assessment of ME/CFS disease severity, UpTime. If the conclusions of this research are confirmed in future full-scale studies, accurate UpTime measurements could become a valuable asset for healthcare providers in assisting ME/CFS patients. Furthermore, UpTime provides a method for pharmaceutical companies and independent researchers to prove the efficacy of their treatments—an important step towards receiving FDA-approval. The BHC's data shows that patients with severe ME/CFS are limited to a bed or reclining chair for all but five hours each day; increasing this number would be life-changing.

7.2 Future Work

This thesis has identified a host of pathways for directing future work. Most pertinent to this research is the need to use IMU-based UpTime measurements to evaluate treatment efficacy. The FDA recommends recording subject data for three months to avoid data skewing caused by weekly trends. The BHC has several treatment protocols that they have found to help aid recovery for patients with ME/CFS. Confirmation of treatment efficacy via UpTime would be a monumental step towards recovery for all people suffering from ME/CFS.

Although evaluating UpTime is a good approach to assessing ME/CFS disease severity, UpTime could be improved by adding depth and expanding its application. Beyond identifying body postures as upright or not upright, additional benefits could come from using machine learning algorithms to identify specific activities such as standing still, sitting in a chair with feet on the floor, driving a car, walking, running, etc. simply by using the lower leg angles provided by the method documented in this thesis.

So far, UpTime has only been used as a tool for evaluating disease symptom improvement. Future research could investigate the possibility of using UpTime, in conjunction with other measurements, to diagnose ME/CFS—a process that is continuously evolving as the scientific community's understanding of the disease improves.

Augmenting UpTime by adding sensors to other parts of the body could provide more detailed activity assessments. Another unexplored realm is the potential to improve device efficiency by changing the tools used to measure lower leg angle. Switching from a 9-axis IMU to an inclinometer could improve battery life for any future custom

UpTime-measuring devices.

The genesis of this research was the BHC approaching the Inertial Self-Powered Sensing Lab at the University of Utah about developing a measurement of upright activity. Now that this method has been developed, commercialization could benefit other disease groups, such as those with fibromyalgia and the geriatric segment of the general population. Furthermore, a commercialized UpTime-based product could also be an effective way for medical professionals to more closely and accurately monitor patients as they go through the process of rehabilitation from any number of diseases, injuries, or other procedures.

APPENDIX

A GAP CHECK (MATLAB)

In section 5.3.1 of this thesis, we describe the process used to check IMU data collected from a Shimmer for gaps. This was done using the MATLAB script below:

```
%% --- GAP CHECK ------------------------------------------------- %
%
% This MATLAB script compares the timestamps of consecutive
% measurements to make sure there are no gaps in the file. If gap(s)
% is/are found, the location(s) within the file(s) is/are displayed.

clear
clc


%% --- Import CSV file and isolate date/time data ------------------- %

    % Read .csv file to table (file location will vary)
    T = readtable('C:\Users\turne\OneDrive - University of Utah\School\
        Research\Final Study\Shimmer Data (Days)\01_D1_L.csv');

    % Isolate column with date/time data
        % Change T from table to array using specified format
        T = table2array(T(2:end,2));
            infmt = 'hh:mm:ss.SSS';
            outfmt = 'hh:mm:ss';
        % Change T from array to durations
        T = duration(T,'InputFormat',infmt,'Format',outfmt);

    % Break durations into pieces to isolate Seconds ('s')
    [h,m,s] = hms(T);


%% --- Check for gaps in seconds vector, i.e., look for an increase in
      's' greater than 1 ------------------------------------------- %

    % Iterate through entire 's' vector
    for i = 2:length(s)
        % Identify whether gap is greater than or less than 1
        if s(i) - s(i-1) > 1
            % Print error message if gap is greater than 1
            %(also state the gap's location within the trial)
            Message = "!!!!!!! GAP AFTER ROW("+ (i-1) + ") !!!!!!!"
        end
    end

    % Print message indicating that the check is complete
    Message = "Gap check = complete"

    % Notification to indicate script has completed
    load gong
    sound(y,Fs)
```

B ALIGN DAYS (MATLAB)

Section 5.3.1 also describes a preprocessing step where paired days are aligned. The MATLAB script which accomplishes this task is included below:

```
%% --- ALIGN TRIALS ------------------------------------------------- %
%
% This MATLAB script edits data from both leg to
% make sure that both datasets have the same number of rows.

clear;clc


%% --- Import files and prepare data -------------------------------- %
    filename_L = 'C:\Users\turne\OneDrive - University of Utah\School\
                  Research\Final Study\Shimmer Data (Days)\
                  01_D1_L.csv';
    filename_R = 'C:\Users\turne\OneDrive - University of Utah\School\
                  Research\Final Study\Shimmer Data (Days)\
                  01_D1_R.csv';

    % Read both .csv files to corresponding tables
    L = readtable(filename_L);
    R = readtable(filename_R);

    % Display first and last entries in each file
    L(2,:)
    R(2,:)
    L(end,:)
    R(end,:)

    % Define each file's length
    rL = height(L)
    rR = height(R)

    % Determine the difference in file lenghts (d).
    if rL > rR
        d = rL - rR
    else
        d = rR - rL
    end


%% --- Edit the file with more rows --------------------------------- %

    % Enter this if-statement if the Left file is longer
    if rL > rR
        while d ~= 0
            % L is longer --> Cut rows of L to match R
            i = [];i = rL:-ceil(rL/d):ceil(rL/d);
            L(i,:) = [];
            rL = height(L);
            % Display number of rows to cut out of L to match R
```

```
            d = rL - rR;
        end
        % overwrite existing left .csv file with edited table
        writetable(L,filename_L,'Delimiter','tab');

    % Enter this elseif-statement if the Right file is longer
    elseif rR > rL
        while d > 0
            % R is longer --> Cut rows of R to match L
            i = [];i = rR:-ceil(rR/d):ceil(rR/d);
            R(i,:) = [];
            rR = height(R);
            % Display number of rows to cut out of R to match L
            d = rR - rL;
        end
        % overwrite existing right .csv file with edited table
        writetable(R,filename_R,'Delimiter','tab');

    % Enter this elseif-statement if the files are the same length
    elseif rL == rR
        % Display the difference in file lengths, zero.
        d
    end


%% --- Print details of revised table(s) ---------------------------- %

    % Number of rows in each table
    rL = height(L)
    rR = height(R)

    % Logical comparison of table rows
    rL == rR

    % Display first and last entries in each file
    L(2,:)
    R(2,:)
    L(end,:)
    R(end,:)

    % Notification to indicate script has completed
    load gong
    sound(y,Fs)
```

C DAY PROCESSOR (MATLAB)

Section 5.3.2 describes the method used to process formatted CSV files. The main MATLAB file, “Day Processor,” calls four functions (included in the proceeding appendices) to convert sensor data to daily UpTime scores. The Day Processor script used to accomplish this task is included below:

```
%% --- DAY PROCESSOR ------------------------------------------------ %
%
% This MATLAB script converts preprocessed Shimmer files into daily
% UpTime scores. Only one subject's data can be processed at a time.
% However, any combination of days can be processed at once.

clear
clc


%% --- Define Days to be Processed ---------------------------------- %

    % Subject 01 - 15
    Subject = '01';

    % Days [1 2 3 4 5 6]
    Days = [1 2 3 4 5 6];

    % Workstation: Mac, PC, or Duffman
    Source = 'Duffman';


%% --- Read & Format Data ------------------------------------------- %

    % Build filenames based on selected Source
    switch Source
        case 'Mac'
            Source = '/Users/turnerpalombo/OneDrive - University of
                      Utah/School/Research/Final Study/Shimmer Data
                      (Days)/';
        case 'PC'
            Source = 'C:\Users\turne\OneDrive - University of
                      Utah\School\Research\Final Study\Shimmer Data
                      (Days)\';
        case 'Duffman'
            Source = 'D:\Turner\OneDrive - University of
                      Utah\School\Research\Final Study\Shimmer Data
                      (Days)\';
        otherwise
            error('Unrecognized source')
    end

    % Add to filenames based on selected Days
```

```
        % If Days includes 1
        if any(Days(:) == 1)
            Day = 1;
            [D1_L, D1_R] = fileimport(Source, Subject, Day);
        end
        % If Days includes 2
        if any(Days(:) == 2)
            Day = 2;
            [D2_L, D2_R] = fileimport(Source, Subject, Day);
        end
        % If Days includes 3
        if any(Days(:) == 3)
            Day = 3;
            [D3_L, D3_R] = fileimport(Source, Subject, Day);
        end
        % If Days includes 4
        if any(Days(:) == 4)
            Day = 4;
            [D4_L, D4_R] = fileimport(Source, Subject, Day);
        end
        % If Days includes 5
        if any(Days(:) == 5)
            Day = 5;
            [D5_L, D5_R] = fileimport(Source, Subject, Day);
        end
        % If Days includes 6
        if any(Days(:) == 6)
            Day = 6;
            [D6_L, D6_R] = fileimport(Source, Subject, Day);
        end


%% --- Calculate Angles (Kalman filter) ----------------------------- %

    % Define IMU (Shimmer) sample rate
    sample_rate = 30; % [Hz]

    % convert Subject from string to number
    subject = str2num(Subject);

    % Convert IMU data to Leg Angles
    %
    % For each day, this process includes three steps:
    % (1) Define Shimmer Gain, based on specific device used
    % (2) Using Kalman filter, calculate Left angles
    % (3) Using Kalman filter, calculate Right angles
        % If Days includes 1
        if any(Days(:) == 1)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 1);
            D1_L = kalmanfilter(D1_L,sample_rate,ShimmerGain_L);
            D1_R = kalmanfilter(D1_R,sample_rate,ShimmerGain_R);
        end
        % If Days includes 2
        if any(Days(:) == 2)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 2);
            D2_L = kalmanfilter(D2_L,sample_rate,ShimmerGain_L);
```

```
            D2_R = kalmanfilter(D2_R,sample_rate,ShimmerGain_R);
        end
        % If Days includes 3
        if any(Days(:) == 3)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 3);
            D3_L = kalmanfilter(D3_L,sample_rate,ShimmerGain_L);
            D3_R = kalmanfilter(D3_R,sample_rate,ShimmerGain_R);
        end
        % If Days includes 4
        if any(Days(:) == 4)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 4);
            D4_L = kalmanfilter(D4_L,sample_rate,ShimmerGain_L);
            D4_R = kalmanfilter(D4_R,sample_rate,ShimmerGain_R);
        end
        % If Days includes 5
        if any(Days(:) == 5)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 5);
            D5_L = kalmanfilter(D5_L,sample_rate,ShimmerGain_L);
            D5_R = kalmanfilter(D5_R,sample_rate,ShimmerGain_R);
        end
        % If Days includes 6
        if any(Days(:) == 6)
            [ShimmerGain_L, ShimmerGain_R] = shimmergain(subject, 6);
            D6_L = kalmanfilter(D6_L,sample_rate,ShimmerGain_L);
            D6_R = kalmanfilter(D6_R,sample_rate,ShimmerGain_R);
        end


%% --- Calculate UpTime ------------------------------------------- %

    % Define critical angle
    critical_angle = 39; %[degrees]

    % Initialize a column of zeros to be filled with UpTimes
    UpTime = zeros(6,1);

    % For each Day, calculate UpTime using paired Left and Right data
        % If Days includes 1
        if any(Days(:) == 1)
            [D1_UpTime] = UpTime(D1_L,D1_R,critical_angle);
            UpTime(1) = D1_UpTime;
        end
        % If Days includes 2
        if any(Days(:) == 2)
            [D2_UpTime] = UpTime(D2_L,D2_R,critical_angle);
            UpTime(2) = D2_UpTime;
        end
        % If Days includes 3
        if any(Days(:) == 3)
            [D3_UpTime] = UpTime(D3_L,D3_R,critical_angle);
            UpTime(3) = D3_UpTime;
        end
        % If Days includes 4
        if any(Days(:) == 4)
            [D4_UpTime] = UpTime(D4_L,D4_R,critical_angle);
            UpTime(4) = D4_UpTime;
```

```
    end
    % If Days includes 5
    if any(Days(:) == 5)
        [D5_UpTime] = UpTime(D5_L,D5_R,critical_angle);
        UpTime(5) = D5_UpTime;
    end
    % If Days includes 6
    if any(Days(:) == 6)
        [D6_UpTime] = UpTime(D6_L,D6_R,critical_angle);
        UpTime(6) = D6_UpTime;
    end

% Display UpTime vector for all days
disp('---------------------------------');
disp(UpTime)
disp('---------------------------------');

% Notification to indicate script has completed
load gong
sound(y,Fs)
```

D FILE IMPORT (MATLAB)

File Import is the first function called by the Day Processor script (Appendix C). This custom file-importing function was created to manage data processing, given the number of Shimmer failures which occurred during our study. File Import is shown below:

```
%% --- FILE IMPORT ------------------------------------------------- %
%
% This MATLAB function imports specified files. Once imported, these
% files are formatted to minimize RAM usage. If a requested file is
% marked as "too short" (due to Shimmer failure), a matrix of zeros is
% created in its place.

function [L, R] = fileimport(Source, Subject, Day)

% Convert Day from string to number
Day = num2str(Day);

% Define files to be imported
Filename_L = [Source Subject '_D' Day '_L.csv'];
Filename_R = [Source Subject '_D' Day '_R.csv'];

% Convert Subject from string to number
Subject = str2num(Subject);

% Check for files marked as too short due to Shimmer failure, aka BAD
    % If Left Data is BAD
    if (Subject==4  &&...
            (Day=='1'||Day=='2'||Day=='3'||Day=='4'||Day=='5'))||...
       (Subject==9  && (Day=='3'))||...
       (Subject==12 && (Day=='4'||Day=='5'))||...
       (Subject==14 && (Day=='4'||Day=='5'||Day=='6'))||...
       (Subject==15 && (Day=='5'||Day=='6'))
        R = readmatrix(Filename_R);
        % Delete unused columns of data
        R(:,[1,2,9]) = [];
        % Replace BAD file with matrix of zeros
        L = zeros(size(R));
    % If Right Data is BAD
    elseif (Subject==3  && (Day=='1'||Day=='2'||Day=='3'))||...
           (Subject==5  && (Day=='2'||Day=='3'))||...
           (Subject==12 && (Day=='6'))
        L = readmatrix(Filename_L);
        % Delete unused columns of data
        L(:,[1,2,9]) = [];
        % Replace BAD file with matrix of zeros
        R = zeros(size(L));
    % If both files are GOOD (sufficient length)
    else
        R = readmatrix(Filename_R);
```

```
        % Delete unused columns of data
        R(:,[1,2,9]) = [];
        L = readmatrix(Filename_L);
        % Delete unused columns of data
        L(:,[1,2,9]) = [];
    end

% Some Shimmer files exported with an extra column of NaN entries
    % Check for and delete these columns
    if isnan(L(1,1))
        L(1,:) = [];
    end
    if isnan(R(1,1))
        R(1,:) = [];
    end
end
```

E SHIMMER GAIN (MATLAB)

Shimmer Gain is the second function called by the Day Processor script (Appendix C). In our study, we used six different Shimmers. Each Shimmer has a unique device ID. This function was created as a repository containing information detailing which Shimmer was used for each day of the study. By identifying which Shimmer was used, we can then set the appropriate gain—as each Shimmer appeared to vary slightly in terms of measurement magnitudes. The Shimmer Gain function is shown below:

```
%% --- SHIMMER GAIN ------------------------------------------------- %
%
% This MATLAB function contains information detailing which Shimmer was
% used for each Subject on each Day. Corresponding Shimmer Gains are
% returned based on the queried days.

function [ShimmerGain_L, ShimmerGain_R] = shimmergain(Subject, Day)

% Identify Phase of study based on Day
if Day < 4
    Phase = 1;
else
    Phase = 2;
end

% Build reference tables for Shimmer IDs used on each days
    S = [repmat(1,2,1);  repmat(2,2,1);  repmat(3,2,1);...
         repmat(4,2,1);  repmat(5,2,1);  repmat(6,2,1);...
         repmat(7,2,1);  repmat(8,2,1);  repmat(9,2,1);...
         repmat(10,2,1); repmat(11,2,1); repmat(12,2,1);...
         repmat(13,2,1); repmat(14,2,1); repmat(15,2,1)];

    P = [1; 2];P = repmat(P,15,1);

    Shimmer_L = ['E11B';'E0EB';'E11B';'E0EB';'E11B';
                 'E0EB';'E0EB';'E0EB';'C71E';'E11B';
                 'E0EB';'E11B';'E11B';'C71E';'C71E';
                 'E0EB';'E0EB';'C71E';'E0EB';'E11B';
                 'C71E';'E0EB';'E11B';'C71E';'E11B';
                 'C71E';'E11B';'C71E';'?   ';'E11B'];

    Shimmer_R = ['DC5F';'DCFB';'DC5F';'DCFB';'DC5F';
                 'DCFB';'DCFB';'C722';'C722';'DC5F';
                 'DCFB';'DC5F';'DC5F';'C722';'DCFB';
                 'C722';'C722';'DCFB';'DCFB';'DC5F';
                 'C722';'DCFB';'DC5F';'C722';'DC5F';
                 'DCFB';'DC5F';'C722';'?   ';'DC5F'];
```

```
    sid.L = table(S, P, Shimmer_L);
    sid.R = table(S, P, Shimmer_R);

    if Phase == 1
        i = 2*Subject - 1;
    else
        i = 2*Subject;
    end

% Select ShimmerID from reference table
ShimmerID_L = sid.L{i,3};
ShimmerID_R = sid.R{i,3};

% Define Gains for each Shimmer
C71E = 0.86;
C722 = 0.83;
DC5F = 0.77;%
DCFB = 0.81;%
E0EB = 0.77;%
E11B = 0.76;%

% Set ShimmerGain based on ShimmerID
switch ShimmerID_L
    case 'C71E'
        ShimmerGain_L = C71E;
    case 'C722'
        ShimmerGain_L = C722;
    case 'DC5F'
        ShimmerGain_L = DC5F;
    case 'DCFB'
        ShimmerGain_L = DCFB;
    case 'E0EB'
        ShimmerGain_L = E0EB;
    case 'E11B'
        ShimmerGain_L = E11B;
    otherwise
        error('Unrecognized source')
end
switch ShimmerID_R
    case 'C71E'
        ShimmerGain_R = C71E;
    case 'C722'
        ShimmerGain_R = C722;
    case 'DC5F'
        ShimmerGain_R = DC5F;
    case 'DCFB'
        ShimmerGain_R = DCFB;
    case 'E0EB'
        ShimmerGain_R = E0EB;
    case 'E11B'
        ShimmerGain_R = E11B;
    otherwise
        error('Unrecognized source')
end
end
```

F KALMAN FILTER (MATLAB)

The Kalman filter function is the third function called by the Day Processor script (Appendix C). In this function, Shimmer data is converted to roll and pitch angles via a Kalman filter. These roll and pitch angles are combined into a single vector of leg angles representing the tilt from vertical using quaternions. The Kalman filter function is shown below:

```
%% --- KALMAN FILTER ------------------------------------------------- %
%
% This MATLAB function uses Phillip Salmony's Kalman filter to convert
% Shimmer data to lower leg angles.
%
% (http://philsal.co.uk/projects/imu-attitude-estimation)

function [angle] = kalmanfilter(Shimmer_data, Sample_rate,Shimmer_gain)

% --- Format Data & Create Estimations ------------------------------ %

    % Define timestep between each measurement
    dt = 1/Sample_rate; % [seconds]

    % Filter Accelerometer Data: 10 Hz second-order low-pass
    [b, a] = butter(2,10/( Sample_rate* 0.5), 'low');
    Shimmer_data(:,1:6) = filter(b,a,Shimmer_data(:,1:6));

    % Break up Shimmer_data
        % Accelerometer
        Ax = Shimmer_data(:,1);
        Ay = Shimmer_data(:,2);
        Az = Shimmer_data(:,3);
        % Gyroscope
        Gx = Shimmer_data(:,4);
        Gy = Shimmer_data(:,5);
        Gz = Shimmer_data(:,6);

    % Convert gyroscope data to radians
    Gx_rad = Gx*(pi/180);
    Gy_rad = Gy*(pi/180);
    Gz_rad = Gz*(pi/180);

    % Trigonometric estimations of roll and pitch using raw Acc data
    phi_hat_acc   = atan2( Ay, sqrt(Ax .^ 2 + Az .^ 2));
    theta_hat_acc = atan2(-Ax, sqrt(Ay .^ 2 + Az .^ 2));


% --- Kalman Filter ------------------------------------------------- %
```

```
LENGTH = length(Gx_rad);

% State-Space form of Kalman filter
A = [ 1 -dt   0   0;
      0   1   0   0;
      0   0   1 -dt;
      0   0   0   1];

B = [dt   0   0   0;
      0   0  dt   0]';

C = [ 1   0   0   0;
      0   0   1   0];

% P = Error Covariance Matrix
% Initially set by us, then updated by the Kalman Filter.
    % Large values = unsure if initial state is correct.
    % Small values = confident that initial guess is correct.
P = eye(4) * 1;

% Q = Process Covariance Matrix
% Tells filter how sure we are about model dynamics.
    % Large values = Model is inaccurate.
    % Small values = Model is accurate.
Q = eye(4) * 0.01;

% R = Measurement Noise Covariance Matrix
% Set depending on noise of the sensors used in the system.
% (based on details in IMU datasheet)
    % Large values = Greater sensor noise.
    % Small values = Minimal sensor noise.
R = eye(2) * 10;

% Initial value estimate
state_estimate = [pi/2 0 0 0]';

% Initialize vectors
phi        = zeros(1, LENGTH);
bias_phi   = zeros(1, LENGTH);
theta      = zeros(1, LENGTH);
bias_theta = zeros(1, LENGTH);

% Kalman filter loop
for i=2:LENGTH

    p = Gx_rad(i);
    q = Gy_rad(i);
    r = Gz_rad(i);

    phi_hat   = phi(i - 1);
    theta_hat = theta(i - 1);

    % Generate input vector based on Gyroscope
    phi_dot   = p...
```

```
                + sin(phi_hat) * tan(theta_hat) * q ...
                + cos(phi_hat) * tan(theta_hat) * r;

        theta_dot = cos(phi_hat) * q ...
                - sin(phi_hat) * r;

        % Predict
        state_estimate = A * state_estimate ...
                       + B * [phi_dot, theta_dot]';

        P = A * P * A' + Q;

        % Update
        measurement = [phi_hat_acc(i) theta_hat_acc(i)]';

        y_tilde = measurement ...
                - C * state_estimate;

        S = R + C * P * C';
        K = P * C' * (S^-1);

        state_estimate = state_estimate ...
                       + K * y_tilde;

        P = (eye(4) - K * C) * P;

        phi(i)        = state_estimate(1);
        bias_phi(i)   = state_estimate(2);
        theta(i)      = state_estimate(3);
        bias_theta(i) = state_estimate(4);

    end

% --- Format Angles to be returned by function ---------------------- %

    % Convert phi and theta to single angle output using quaternions
    roll  = pi/2 - phi;
    pitch = theta;

    qr =   cos(roll/2) .* cos(pitch/2);
    qi =   sin(roll/2) .* cos(pitch/2);
    qj =   cos(roll/2) .* sin(pitch/2);
    qk = - sin(roll/2) .* sin(pitch/2);

    angle = 2*atan2( sqrt(qi.^2 + qj.^2 + qk.^2 ), qr);

    % Convert angle from radians to degrees
    % (adjust for gain to match Shimmer to VICON)
    angle = angle*(180/pi)*Shimmer_gain;

    % Set Maximum angle to 180 degrees
    parfor i = 1:LENGTH
        if angle(i) > 180
            angle(i) = 180;
```

```
            end
        end
    end
```

G UPTIME (MATLAB)

UpTime is the fourth function called by the Day Processor script (Appendix C). This function calculates UpTime, as a percentage of the day spent upright, by comparing both leg angles to the critical angle at each point in time. This function is shown below:

```
%% --- UPTIME CALCULATOR ------------------------------------------- %
%
% This MATLAB function takes two equal-length vectors of angle
% measurements and compares them to the critical angle to calculate
% UpTime. The critical angle indicates which angles are considered
% "upright". All angles are measured as degrees from vertical.
%
% Vertical = 0 degrees
% Horizontal = 90 degrees

function [UpTime] = UpTime(L, R, critical_angle)

% Initialize UpTime_vec
UpTime_vec = zeros(1,length(L));

% For each angle measurement, determine the "uprightness" of each leg
for i = 1:length(L)

    % If only the Left leg is upright (elevated above critical angle)
    if L(i) > critical_angle && R(i) < critical_angle
        % Assign a value of 0.5, essentially averaging uprightness
        UpTime_vec(i) = 0.5;

    % If only the Right leg is upright (elevated above critical angle)
    elseif L(i) < critical_angle && R(i) > critical_angle
        % Assign a value of 0.5, essentially averaging uprightness
        UpTime_vec(i) = 0.5;

    % If both legs are upright (elevated above critical angle)
    elseif L(i) > critical_angle && R(i) > critical_angle
        % Assign a value of 1
        UpTime_vec(i) = 1;

    % If neither leg is upright (both legs below critical angle)
    else
        % Assign a value of 0
        UpTime_vec(i) = 0;

    end
end

% Calculate UpTime as a percentage of the day spent upright
UpTime = sum(UpTime_vec)/length(UpTime_vec)*100; % [%]

end
```

REFERENCES

[1] Naviaux, Robert; Naviaux, Jane; Li, Kefeng; Bright, Taylor; Alaynick, William; Wang, Lin; Baxter, Asha; Nathan, Neil; Anderson, Wayne; and Gordon, Eric. "Metabolic features of chronic fatigue syndrome." *Proceedings of the National Academy of Sciences of the United States of America* Vol. 113 No. 37 (2016): pp. E5472-E5480. DOI 10.1073/pnas.1607571113.

[2] Anonymous. "Beyond Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Redefining an Illness." *Military Medicine* Vol. 180 No. 7 (2015): pp. 721-723. DOI 10.7205/MILMED-D-15-00085.

[3] Lorusso, Lorenzo; Mikhaylova, Svetlana; Capelli, Enrica; Ferrari, Daniela; Ngonga, Gaelle; and Ricevuti, Giovanni. "Immunological aspects of chronic fatigue syndrome." *Autoimmunity Reviews* Vol. 8 No. 4 (2009): pp. 287-291. DOI 10.1016/j.autrev.2008.08.003.

[4] Prins, Judith; Bleijenberg, Gijs; and van Der Meer, Jos Wm. "Chronic fatigue syndrome and myalgic encephalomyelitis." *The Lancet* Vol. 359 No. 9318 (2002): pp. 1699-1699. DOI 10.1016/S0140-6736(02)08577-X.

[5] Murdock, Kyle; Wang, Xin; Shi, Qiuling; Cleeland, Charles; Fagundes, Christopher; and Vernon, Suzanne. "The utility of patient-reported outcome measures among patients with myalgic encephalomyelitis/chronic fatigue syndrome." *Quality of Life Research* Vol. 26 No. 4 (2017) pp. 913-921. DOI 10.1007/s11136-016-1406-3.

[6] Cairns, R and Hotopf, M. "A systematic review describing the prognosis of chronic fatigue syndrome." *Occupational Medicine* Vol. 55 No. 1 (2005): pp. 20-31. DOI 10.1093/occmed/kqi013.

[7] Fukuda, K; Straus, S; Hickie, I; Sharpe, M; Dobbins, J; and Komaroff, A. "The chronic fatigue syndrome: A comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group." *Annals of Internal Medicine* Vol. 121 No. 12 (1994): pp. 953-959.

[8] Chu, Lily; Valencia, Ian; Garvert, Donn; and Montoya, Jose. "Deconstructing post-exertional malaise in myalgic encephalomyelitis/ chronic fatigue syndrome: A patient-centered, cross-sectional survey." *PloS One* Vol. 13 No. 6 (2018): pp. E0197811. DOI 10.1371/journal.pone.0197811.

[9] Van Campen, Linda; Rowe, Peter; and Visser, Frans. “Blood Volume Status in ME/CFS Correlates With the Presence or Absence of Orthostatic Symptoms: Preliminary Results.” *Frontiers in Pediatrics* Vol. 6 No. 352 (2018): DOI 10.3389/fped.2018.00352.

[10] Jackson, M and Bruck, D. “Sleep Abnormalities in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis: A Review.” *Journal Of Clinical Sleep Medicine* Vol. 8 No. 6 (2012): pp. 719-728. DOI 10.5664/jcsm.2276.

[11] Esfandyarpour, R; Kashi, A; Nemat-Gorgani, M; and Wilhelmy, J; Davis, R. “A nanoelectronics-blood-based diagnostic biomarker for myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS).” *Proceedings of the National Academy of Sciences of the United States of America* Vol. 116 No. 21 (2019): pp. 10250-10257. DOI 10.1073/pnas.1901274116.

[12] Montoya, Jose; Holmes, Tyson; Anderson, Jill; Maecker, Holden; Rosenberg-Hasson, Yael; Valencia, Ian; Chu, Lily; Younger, Jarred; Tato, Cristina; and Davis, Mark. “Cytokine signature associated with disease severity in chronic fatigue syndrome patients.” *Proceedings of the National Academy of Sciences of the United States of America* Vol. 114 No. 34 (2017): pp. E7150-E7158. DOI 10.1073/pnas.1710519114.

[13] Gabriel, Pettee; Sidney, Stephen; Jacobs, David; Quesenberry, Charles; Reis, Jared; Jiang, Sheng-Fang; and Sternfeld, Barbara. “Convergent Validity of a Brief Self-reported Physical Activity Questionnaire.” *Medicine & Science in Sports & Exercise* Vol. 46 No. 8 (2014): pp. 1570-1577. DOI 10.1249/MSS.0000000000000278.

[14] “InvenSense Launches Sensor-Based SaaS Fitness Tracking Platform for Smartphone, Smartwatch Applications.” *M2 Pharma* (2015).

[15] Burton, A; Rahman, K; Kadota, Y; Lloyd, A; and Vollmer-Conna, U. “Reduced heart rate variability predicts poor sleep quality in a case–control study of chronic fatigue syndrome.” *Experimental Brain Research* Vol. 204 No. 1 (2010): pp. 71-78. DOI 10.1007/s00221-010-2296-1.

[16] Libman, Eva; Fichten, Catherine; Creti; Conrod, Kerry; Tran, Dieu-Ly; Grad, Roland; Jorgensen, Mary; Amsel, Rhonda; Rizzo, Dorrie; Baltzan, Marc; Pavilanis, Alan; and Bailes, Sally. “Refreshing Sleep and Sleep Continuity Determine Perceived Sleep Quality.” *Sleep Disorders* Vol. 2016 (2016): DOI 10.1155/2016/7170610.

[17] Choi, Seoyoung; Shin, Yong Beom; Kim, Soo-Yeon; and Kim, Jonghyun. “A novel sensor-based assessment of lower limb spasticity in children with cerebral palsy.” *Journal of NeuroEngineering and Rehabilitation* Vol. 15 No. 1 (2018): pp. 1-16. DOI 10.1186/s12984-018-0388-5.

[18] Yadav, Nagesh and Bleakley Chris. “Fast calibration of a 9-DOF IMU using a 3 DOF position tracker and a semi-random motion sequence.” *Measurement* Vol. 90 (2016): pp. 192-198. DOI 10.1016/j.measurement.2016.04.066.

[19] Novak, Domen and Riener, Robert. “A survey of sensor fusion methods in wearable robotics.” *Robotics and Autonomous Systems* Vol. 73 (2015): pp. 155-170. DOI 10.1016/j.robot.2014.08.012.

[20] Pedley, Mark. “Tilt sensing using a three-axis accelerometer.” *Freescale Semiconductor Application Note* No. AN3461 (2013). URL https://www.nxp.com/docs/en/application-note/AN3461.pdf.

[21] Narkhede, Parag; Joseph Raj, Alex Noel; Kumar, Vipan; Karar, Vinod; and Poddar, Shashi. “Least square estimation-based adaptive complimentary filter for attitude estimation.” *Transactions of the Institute of Measurement and Control* Vol. 41 No. 1 (2019): pp. 235-245. DOI 10.1177/0142331218755234.

[22] Fourati, Hassen. “Heterogeneous Data Fusion Algorithm for Pedestrian Navigation via Foot-Mounted Inertial Measurement Unit and Complementary Filter.” *IEEE Transactions on Instrumentation and Measurement* Vol. 64 No. 1 (2015): pp. 221-229. DOI 10.1109/TIM.2014.2335912.

[23] Kamali, C and Jain, S. “Multiplicative error state kalman filter vs nonlinear complimentary filter for a high performance aircraft attitude estimation.” *Defence Science Journal* Vol. 66 No. 6 (2016): pp. 630-637. DOI 10.14429/dsj.66.8838.

[24] Kottath, Rahul; Narkhede, Parag; Kumar, Vipan; Karar, Vinod; and Poddar, Shashi. “Multiple Model Adaptive Complementary Filter for Attitude Estimation.” *Aerospace Science and Technology* Vol. 69 (2017): pp. 574-581. DOI 10.1016/j.ast.2017.07.011.

[25] Bai, Thibault; Thibault, Jules; and Mclean, David. “Dynamic data reconciliation: Alternative to Kalman filter.” *Journal of Process Control* Vol. 16 No. 5 (2006): pp. 485-498. DOI 10.1016/j.jprocont.2005.08.002.

[26] Kalman, Rudolf. “A New Approach to Linear Filtering and Prediction Problems.” *Journal of Basic Engineering* Vol. 82 No. 1 (1960): pp. 35. DOI 10.1115/1.3662552.

[27] Bauer, Jeffrey and Andrisani, Dominick. *Estimating short-period dynamics using an extended Kalman filter.* National Aeronautics and Space Administration, Edwards (1990).

[28] Salmony, Philip. “IMU Attitude Estimation.” 2018. URL http://philsal.co.uk/projects/imu-attitude-estimation.

[29] Liu, H; Hou, C; Lin, Jh; Li, Yf; Shi, Qf; Chen, T; Sun, Ln; and Lee, Ck. “A non-resonant rotational electromagnetic energy harvester for low-frequency and irregular human motion.” *Applied Physics Letters* Vol. 113 No. 20 (2018). DOI 10.1063/1.5053945.

[30] Cho, Sanghee ; Grazioso, Ron ; Zhang, Nan ; Aykac, Mehmet ; and Schmand, Matthias. “Digital timing: Sampling frequency, anti-aliasing filter and signal interpolation filter dependence on timing resolution.” *Physics in Medicine and Biology* Vol. 56 No. 23 (2011): pp. 7569-7583. DOI 10.1088/0031-9155/56/23/015.

[31] OptiTrack. “Rizzoli Lower Body Protocol (26).” 2018. URL https://v21.wiki.optitrack.com/index.php?title=Rizzoli_Markersets#Rizzoli_Lower_Body_Protocol_.2826.29.

[32] Bateman, Lucinda. “NASA 10 Minute Lean Test | Instructions for Providers.” Bateman Horne Center, Salt Lake City, Utah. 2016. URL https://batemanhornecenter.org/wp-content/uploads/2016/09/NASA-Lean-Test-Instructions.pdf.

[33] Oura. “Oura Guide.” 2020. URL https://cloud.ouraring.com/edu.

[34] Saperova, Elena and Dimitriev, Dmitry. “The effects of mood state and psychological well-being on heart rate variability during mental stress.” *International Journal of Psychophysiology* Vol. 94 No. 2 (2014): pp. 232. DOI 10.1016/j.ijpsycho.2014.08.904.

[35] Thayer, Jf; Ahs, F; Fredrikson, M; Sollers, J; and Wager, T. “A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health.” *Neuroscience and Biobehavioral Reviews* Vol. 36 No. 2 (2012): pp. 747-756. DOI 10.1016/j.neubiorev.2011.11.009.