MODELLING AND CHARACTERIZATION OF FORCE PLATE MEASUREMENTS ON SUBACUTE POST-CONCUSSION SUBJECTS THROUGH MACHINE LEARNING

A Thesis
Presented to
The Academic Faculty

by

José Joaquín Casado Garrido

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

Georgia Institute of Technology
December 2021

MODELLING AND CHARACTERIZATION OF FORCE PLATE MEASUREMENTS ON SUBACUTE POST-CONCUSSION SUBJECTS THROUGH MACHINE LEARNING

Approved by:

Dr. Russell Gore, Advisor
School of Biomedical Engineering
Georgia Institute of Technology

Dr. Jason Allen
School of Medicine
Emory University

Dr. Cassie Mitchell
School of Biomedical Engineering
Georgia Institute of Technology

Dr. May Wang
School of Biomedical Engineering
Georgia Institute of Technology

Date Approved: December, 8, 2021

ACKNOWLEDGEMENTS

I firstly would like to thank Dr. Russell Gore and Dr. Jason Allen for giving me the opportunity to work in this project, and for all the help during the past year. I would also like to thank Dr. Jeremy Smith for his attentive responses and feedback during the development of this project. I am also thankful for my committee members Dr. May Wang and Dr. Cassie Mitchell that have guided me through the past months as I finished up my thesis and my stay at Georgia Tech. I would also like to thank the personnel at the Shepherd Center for their kindness and help during my visits, and for their great work in collecting the data used in this project. I am also grateful for my friends who have shown me their support and love even through the immense distance. Finally, I would like to thank my family for giving me the opportunity to study here, and for their continued support during my academic career.

.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

mTBI	Mild Traumatic Brain Injury
PCVD	Post-Concussion Vestibular Dysfunction
SOT	Sensory Organization Test
COP	Center of Pressure
AP	Anterior-Posterior
ML	Medial-Lateral
EMD	Empirical Mode Decomposition
IMF	Intrinsic Mode Function
ApEn	Approximate Entropy
SampEn	Sample Entropy
MSE	Multiscale Sample Entropy
CI	Complexity Index
RF	Random Forest

SUMMARY

Mild traumatic brain injuries (mTBI) are one of the leading causes of neurological disorders. Symptoms after a mTBI may include headache, dizziness, and balance issues, among others, with vestibular disorders observed in up to 80% of these patients. These symptoms generally resolve in the first few weeks after the injury, but some patients may develop persistent symptoms. Patients with Post-Concussion Vestibular Dysfunction (PCVD) may present alterations in the peripheral and central vestibular systems. These alterations may then affect postural control and stability, which coupled with visual motion sensitivity, cause the prolonged symptomatology. In this study, we evaluated postural control strategies in Healthy Controls (HC) and Subacute PCVD patients (ST) to identify underlying changes in the postural control system. Sensory Organization Test (SOT) was employed to measure Centre Of Pressure (COP) signals under different sensory conditions. Analysis of traditional linear metrics and entropy metrics of the COP signals demonstrated significant differences between groups. Complexity index was reduced for the ST group during “Eyes Closed” condition, with a median value of 7.93 vs 9.59 for the HC in the Medial-Lateral direction ($p=0.002$), and 5.17 vs 6.22 Anterior-Posterior direction ($p=0.0009$). Moreover, analysis of these metrics through machine learning, showed indications of interactions between these variables that may be predictive of the health condition of the patient. These results remark the potential of these metrics for evaluating changes in postural dynamics in patients with PCVD, and opens a new path for analysis of the COP signals with the support of machine learning models.

CHAPTER 1. INTRODUCTION

1.1 Motivation

Neurological disorders are characterized by changes in normal brain functioning that result from degenerative diseases, infections, stroke, or head trauma, among other causes. Changes in brain connectivity may drive network dysfunction and lead to cognitive, balance, and other functional impairments [1], [2]. Understanding the complexity of brain function and connectivity is a difficult task that generally involves correlating specialized brain imaging (e.g., functional magnetic resonance imaging) and/or electrical recordings (e.g., electroencephalogram) with clinical testing. However, clinical metrics may include subjective questionaries that do not offer the degree of precision that may be required for correlation to imaging data. In addition, while self-report questionnaires may help clinicians to assess a patient's symptomatology and condition, symptoms are often not specific to a single disease state and symptom reports do not always provide an unbiased, quantifiable measure. Establishing objective and repeatable metrics that offer insight into underlying brain function is fundamental for both appropriately assessing a patient's condition and to allow a better understanding of how the brain is organized in both healthy and disease states.

The applications of machine learning and data analytics to the healthcare and medical fields are rapidly growing and have been more recently applied to the question of brain function. There is an increasing demand for clinical decision support systems that can guide the diagnosis and treatment of a patient [3]. Moreover, the expanding availability of clinical instrumentation has led to the development of new clinical assessments that rely on

quantifiable patient biomedical metrics. Data from these clinical assessments are being evaluated through complex machine learning models capable of unraveling details and patterns that traditional methods are unable to detect [4].

MTBI or concussion is one of the most common neurological disorders. After an mTBI patients may report a variety of signs and symptoms including cognitive impairment, headaches, dizziness or imbalance, visual impairment, or sleep disruption. Post-concussion vestibular dysfunction, accounting for reports of dizziness and imbalance, is of particular interest because these symptoms are predictive of more long term recovery after injury [5] and there are well established treatments available when these patients are appropriately identified [6]. Although assessment of these patients is based on both subjective questionnaires and objective balance tests, current standard metrics may not be effective in revealing the full range of dysfunction in patients with post-concussion vestibular dysfunction.

The aim of this thesis is to analyze data derived from objective balance testing using force plate measurements of center of pressure signals. Specifically, evaluate both traditional linear metrics and entropy measures of patients with post-concussion vestibular dysfunction to:

1. Establish objective metrics that reveal insights into the underlying changes in postural control.
2. Identify a reduced set of metrics that more fully characterize the vestibular deficits of patients with PCVD than is currently possible with standard of care.

CHAPTER 2. BACKGROUND

2.1 Balance Control and Sensory Integration

Postural stability is achieved through complex processing and integration of sensory feedback. Integration of several sensory systems is fundamental for proper balance control including visual, vestibular, and proprioceptive feedback. The overall mechanism of postural stability can be defined as a closed-loop feedback control system, in which afferent signals from these sensory organs result in efferent control of muscles to achieve stability [7].

Deterioration or malfunctioning of one of these sensory systems can greatly affect the control strategy under different conditions. For example, ageing is one of the main causes of sensory degeneration, leading to lower visual acuity and reduction of vestibular hair cells [8]. Disturbances in the environment, such as walking in a dark room, coupled with impaired sensory function can elicit unreliable sensory feedback which may lead to imbalance and falls [9]. In order to avoid falls, the system may compensate for unreliable sensory feedback with a strategy that reweights sensory inputs, increasing the utilization of sensory inputs which are most reliable and decreasing the utilization of less reliable sensory inputs [10]. This compensatory mechanism has been evaluated in patients with diverse neurological disorders, including mild traumatic brain injuries [11].

2.2 MTBI and Persistent Post-Concussive Symptoms

MTBI is one of the most common neurological disorders, with an estimated 40 million people worldwide affected every year. Approximately, 10 to 25 percent of mTBI patients

suffer from persistent symptoms after injury, generally defined as symptoms persisting greater than 30 days post injury [12]. The diagnosis of mTBI is a clinical decision with a variety of accepted definitions. For example the WHO Collaborating Centre for Neurotrauma Task Force on mTBI requires that at least one of the following occur after an acute brain injury resulting from mechanical energy to the head: confusion or disorientation, loss of consciousness for 30 minutes or less, amnesia less than 24 hours, or transient neurological abnormalities not requiring surgery [13]. After the diagnosis of mTBI is established, assessment of symptoms and severity [14] may include a combination of self-report and objective measures including the Post-Concussion Symptom Scale (PCSS), a variety of cognitive tests, the Vestibular Ocular Motor Screen (VOMS) [15], and objective balance testing. The Balance Error Scoring System (BESS) and the Sensory Organization Test (SOT), among others, may be used to specifically evaluate balance and vestibular function.

2.2.1 Vestibular Dysfunction and Visual Compensation

Both the peripheral and central vestibular systems may be affected after injury, although recent studies suggest that central alterations may be more prevalent, at least in the subacute phase [16], [17]. Vestibular therapy is effective for recovering normal functioning and for resolving these discomforting symptoms [18]. Post-concussive vestibular dysfunction is common, present in up to 80% of patients with mTBI, and if vestibular symptoms persist, PCVD has been associated with sensory reweighting resulting in an overdependence on visual stimuli [19]. During the acute phase of a vestibular disorder, patients may overcome the lack of reliable vestibular feedback through higher dependence on visual or somatosensory cues. This compensatory behavior is a response to the inaccurate integration

of the vestibular system and may help patients maintain balance in the early period post-injury. However, this response may become maladaptive over time resulting in protracted symptoms of dizziness, imbalance, and disorientation [11], [20].

It is hypothesized that subjects who develop persistent symptoms may be suffering from this sensory maladaptation. These patients often suffer from visual motion sensitivity (VMS) and worsened symptoms when presented with complex visual environments. Detailed evaluation of these sensory systems is critical to understanding the underlying neuropathology behind each patient's symptomatology. Computerized Dynamic Posturography (CDP) is a clinical test used to analyze sensory integration in balance control and consists of a series of tests in which the patient must maintain standing balance while presented with different sensory conditions. One of the most widely used CDP assessments is the Sensory Organization Test (SOT).

2.2.2 Sensory Organization Test

The Sensory Organization Test is a battery of balance tasks that includes six different conditions as depicted on Figure 2-1. The sensory systems utilized or modified in each condition are shown at the bottom including visual, vestibular, and proprioceptive afferents. The subject stands on a force plate and maintains stable postural control for a controlled time period while confronted with differing test conditions, and the forces on the plate and the center of pressure are measured. These six conditions provide a structured method to analyze different sensory control systems during balance. The six tasks are depicted on the figure below and consist of: 1) Eyes Open, 2) Eyes Closed, 3) Eyes Open

with sway-referenced Vision, 4) Eyes Open with sway-referenced Support, 5) Eyes Closed with sway-referenced Support, and 6) Sway-referenced Vision and Support.

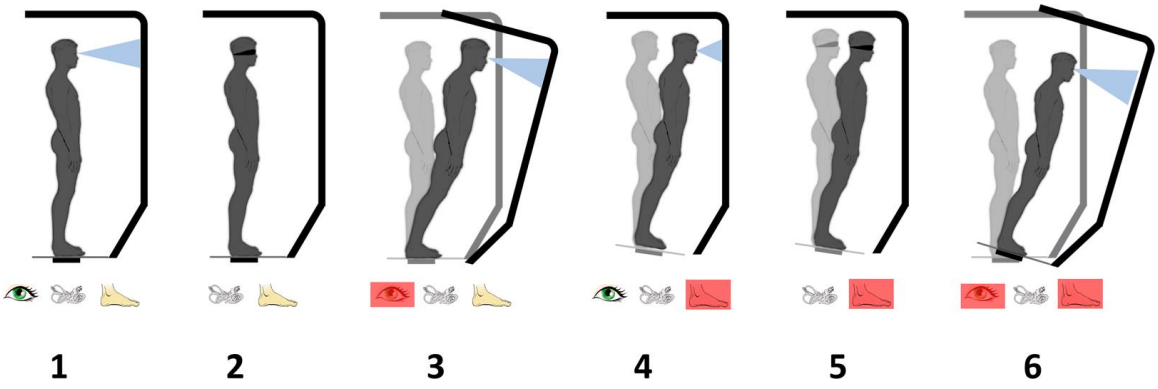


Figure 2-1. Sensory Organization Test illustration with conditions (1) Eyes Open, (2) Eyes Closed, (3) Eyes Open with sway-referenced Vision, (4) Eyes Open with sway-referenced Support, (5) Eyes Closed with sway-referenced Support, and (6) Sway-referenced Vision and Support. Sensory systems contributing on each condition are shown on the bottom: vision (eye), vestibular (ear canal), proprioception (foot). Unreliable feedback are highlighted in red (Image by Charkhkar H. et al. [21]).

The first condition serves as a baseline. Throughout the remaining five conditions, the patient must counteract the lack of reliable feedback from the different sensory systems in different combinations. During condition two, no visual feedback is available to the subject so maintaining balance depends on vestibular sensation and proprioception. On condition three, the subject is presented with sway-referenced visual feedback (red highlight), in which the visual environment moves in synch with the center of pressure, rendering visual feedback unreliable. On conditions four through six, the subject is presented with a sway-referenced proprioceptive feedback, in which the force plate platform moves in synch with the center of pressure, rendering proprioceptive feedback unreliable. Condition five removes visual feedback combined with unreliable proprioceptive feedback. Finally, condition six combines both sway-referenced visual and

proprioceptive feedback rendering both visual and proprioceptive feedback unreliable. An *equilibrium score* is generated for each condition based on balance sway. This score ranges from 0 (complete loss of balance) to 100 (no balance sway). Additionally, a *composite score* is calculated as a linear combination of the equilibrium scores of the six conditions. This test also provides a *sensory strategy score* that estimates how much a patient relies on each sensory system. Although these analyses have demonstrated reliability in diagnosing patients with different underlying conditions [22], composite scores are a less reliable balance assessment after mTBI and often fail to demonstrate meaningful changes among those with persistent post-concussive symptoms [23], [24].

2.3 Balance control and Center of Pressure

As mentioned above, there are different systems and tools to quantify balance control and postural stability in the clinic. Two of the main measures quantified for this purpose are the Center of Mass (COM) and Center of Pressure (COP). The COP is the resulting application point of the Ground Reaction Force (GRF) of a subject over a surface. The COM is typically measured with a motion-capture system, while the COP is assessed through a force plate. These two measures are tightly related but have been shown to provide distinct information [25]. The COM velocity is thought to be one of the most critical forms of sensory feedback to stabilize balance during quiet stance [26]. This feedback is integrated by the central nervous system which then produces a compensatory response through muscular activity. This muscular response is later reflected in the deviations of the COP. The COP is thus thought to show the nervous system's response to the COM movements, which can elucidate changes or disturbances in underlying sensory integration systems [27].

The COP has been studied for decades due to its easy acquisition process and its effective characterization of balance stability. During this thesis we will consider different metrics extracted from these signals to characterize patients with PCVD.

CHAPTER 3. DATA ACQUISITION AND SIGNAL PREPROCESSING

In this chapter we will introduce the data acquisition process as well as the signal preprocessing steps followed prior to feature extraction.

3.1 Background

3.1.1 Bertec Computerized Dynamic Posturography

There are multiple systems commercially available for CDP assessments. In this project, the Bertec FIT CDP/IVR system (see Figure 3-1) was employed for the SOT recordings acquisition. It consists of a visual surround immersive screen, with over 180-degree horizontal and 90-degree vertical fields of view, and a dynamic base with a dual-balance force plate for forces and center of pressure acquisition. The immersive screen provides a sway-referenced visual environment for SOT conditions 3 and 6, while the dynamic base allows for movement to create the sway-referenced support for SOT conditions 4 through 6.

The force plate contains a set of strain gaged load transducers that measure six components: three orthogonal forces and three moments along each axis. The coordinates of the center of pressure are then estimated as:

$$x_p = -\frac{M_y}{F_z}$$

$$y_p = \frac{M_x}{F_z}$$

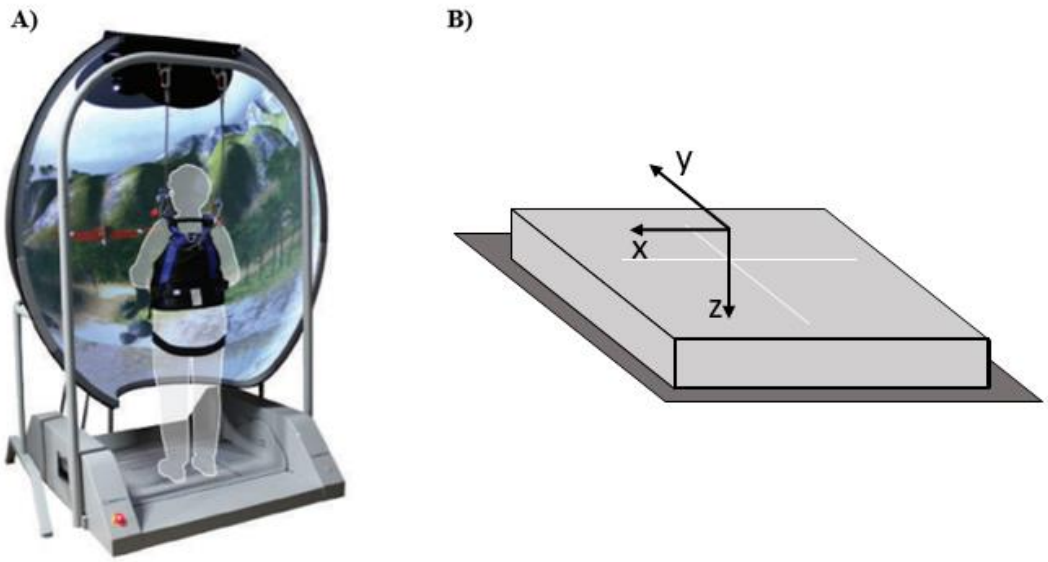


Figure 3-1. Bertec Immersive System. Depiction of the overall system with immersive screen (A, image from Bertec™ [28]) and illustration of the force plate coordinate system (B)

Where x_p and y_p are the x and y coordinates of the center of pressure respectively, M_y and M_x are the moment components along the y and x axes respectively and F_z is the force component along the z axis, as measured by the force transducer coordinate system. If we consider the two main components of the movement independently, x_p is the center of pressure movement along the medial-lateral (ML) direction and y_p is the center of pressure movement along the anterior-posterior (AP) direction.

3.1.2 Signal Pre-processing

Prior to analysis and feature extraction, it is fundamental to pre-process the signal to: 1) Eliminate any sources of noise that may have contaminated the signal, and 2) Downsample to a manageable sequence length for faster processing of the signal.

Filtering high frequency content is a key step to avoid aliasing when downsampling a signal. It is important to understand the concepts of sampling theory and aliasing to have a sense of how they can impact the resulting sequence. The sampling theory states that an analog signal may be completely recovered as long as the sampling frequency is at least double the maximal frequency contained in the original signal [29]. When this is not satisfied, the resulting signal will contain information from higher frequency components that would not be distinguishable from those of lower frequencies. This results in a distorted and inaccurate signal that does not reflect either the real high frequency or low frequency components of the original one. This is what we call aliasing.

When downsampling a signal we must ensure that aliasing is not present, so that the maximal frequency component is well characterized in the downsampled version. For this, we may have to perform lowpass filtering to remove high frequency components that we are not interested in keeping.

There have been multiple approaches in prior research regarding the filtering methods [30]–[32]. However, they have consisted mainly of two strategies: 1) Butterworth Low-pass filtering, and 2) Empirical Mode Decomposition (EMD). On this project we decided to employ the latter, as EMD is especially suitable for non-linear and non-stationary signals, as it is the case of the COP.

3.1.3 Empirical Mode Decomposition

EMD is an adaptive time-space method that operates in the time domain. This algorithm was originally proposed in 1998 by Huang et al. and consists of an iterative process in which the input signal $x(t)$ is decomposed into a set of independent components also

known as Intrinsic Mode Functions (IMFs) [33]. The IMF must satisfy two characteristics: 1) the number of extrema and zero-crossings must be equal or different by one, and 2) the signal is locally symmetric around the time axis [34]. The core element of this algorithm consists of *sifting* the signal $x(t)$ to obtain a function $r(t)$. This is performed in five main steps:

1. Set $r_{i,o} = x(t)$, where $r_{i,o}$ is the starting residual value r_o at iteration *i*
2. Find the local minima and maxima of $r_{i,o}$
3. Use the local maxima and minima to construct the lower and upper envelopes $e\_(t)$ and $e_+(t)$ respectively. Calculate the mean of the envelopes: $m(t)$
4. Subtract $m(t)$ from $r_{i,o}$ to obtain the *current* residual $r_{i,cur}(t) = r_{i,o} - m(t)$
5. Check if $r_{i,cur}(t)$ satisfies IMF characteristics:
    a. Not Satisfied: repeat steps 2-4 with $r_{i,cur}(t)$ as starting point until IMF characteristics are met
    b. Satisfied: $IMF_i(t) = r_{i,cur}(t)$, and $r_{i,cur}(t) = r_{i,o} - IMF_i(t)$. Then, find the next IMFs by repeating the process with $r_{i,cur}(t)$ as starting point ($r_{i,o}$)

This is iteratively performed until a stop criterion is satisfied and the decomposition is terminated. The stop criteria can be defined, as an example, by selecting the maximum number of IMFs to calculate or, by determining a minimum number of extrema that the residual should contain.

The original signal can then be reconstructed by:

$$x(t) = \sum_{i=1}^{N} IMF_i(t) + r_N(t)$$

Where N is the number of intrinsic mode functions obtained after the decomposition.

In the EMD signal, the lower order IMFs -the first extracted components- contain the higher oscillations contents (high frequencies), while the higher order IMFs contain the lower oscillations of the signal. This is the principle used to remove the higher frequency components of the signal [35]. To obtain a lowpass filtered signal $x'(t)$, the first order IMFs are subtracted from the original signal, keeping only the lower oscillations reflected in the lower order IMFs:

$$x'(t) = x(t) - \sum_{i=1}^{m} IMF_i(t)$$

Where m is the number of modes to subtract.

To select the appropriate number of IMF, it is necessary to estimate frequency contents of each IMF. EMD is usually coupled with the Hilbert-Huang transform (HHT), which allows for the time-frequency analysis of the resulting components.

3.2 Methods

Thirty-four healthy controls and twenty-one subjects with subacute post-concussive vestibular dysfunction were recruited. Subjects were consecutively recruited during clinical assessment at a referral concussion clinic. Age matched controls were recruited from the community. Informed consent was obtained according to procedures approved

by Shepherd Center, Emory University, and Georgia Institute of Technology Institutional Review Boards. Eligible subjects were greater than 16 years of age and sustained a concussion within the previous 6 months. Additionally, PCVD subjects required clinical evidence for vestibular impairment through PCSS and abnormal VOMS scores. All subjects underwent a variety of clinical assessments including balance testing with Bertec CDP. SOT testing was part of a larger protocol for the study of this subjects.

For each SOT condition, three trials of 20 seconds of duration were recorded, and the COP in the anterior-posterior (AP) and medial-lateral (ML) directions were extracted. During analysis of the SOT recordings, each signal was decomposed through EMD, and the frequency content of each IMF was visualized through HHT. The IMFs containing frequencies above 50 Hz were subtracted from the original signal prior to downsampling. Frequencies above this threshold are essentially removed from the resulting signal, preventing aliasing during downsampling. Filtered signals where were then downsampled to 100 Hz, in preparation for later metrics extraction.

3.3 Results

A sample of the EMD decomposition of the COP signal is shown in Figure 3-2. It is apparent that the higher-order IMFs (i.e., IMFs 6-10) contain the lower oscillations, while IMFs 1 and 2 appear as low amplitude noisy signals. However, it is not as evident which frequency ranges are contained on each of these components.

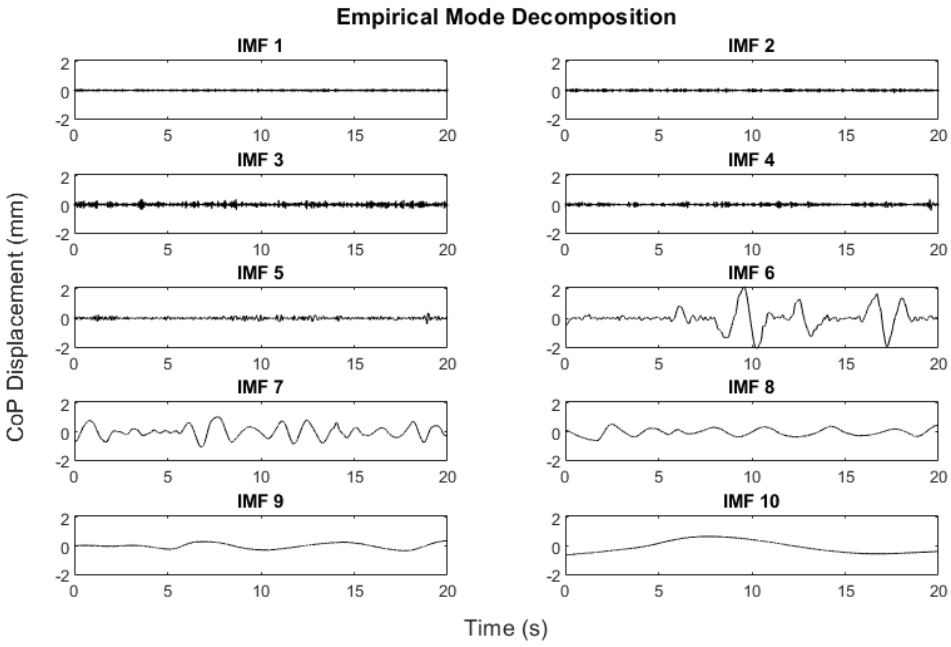


Figure 3-2. Empirical Mode Decomposition of COP in the ML direction. Ten first IMF are shown

When visualizing the HHT of the IMFs, the frequency contents are better described. This can be seen in Figure 3-3-A, where the Hilbert Spectrum of IMFs 1 and 2 are plotted. This image represents the instantaneous energy of all frequencies at each time instant. It is visible that these two IMFs contain high frequency components that could drastically affect the signal when performing downsampling. After inspection of all IMFs, we decided to subtract IMFs 1 through 3 from the original signal so that only low frequency contents were kept. The complete Hilbert Spectrum can be found on APPENDIX A. IMFS HILBERT SPECTRUM

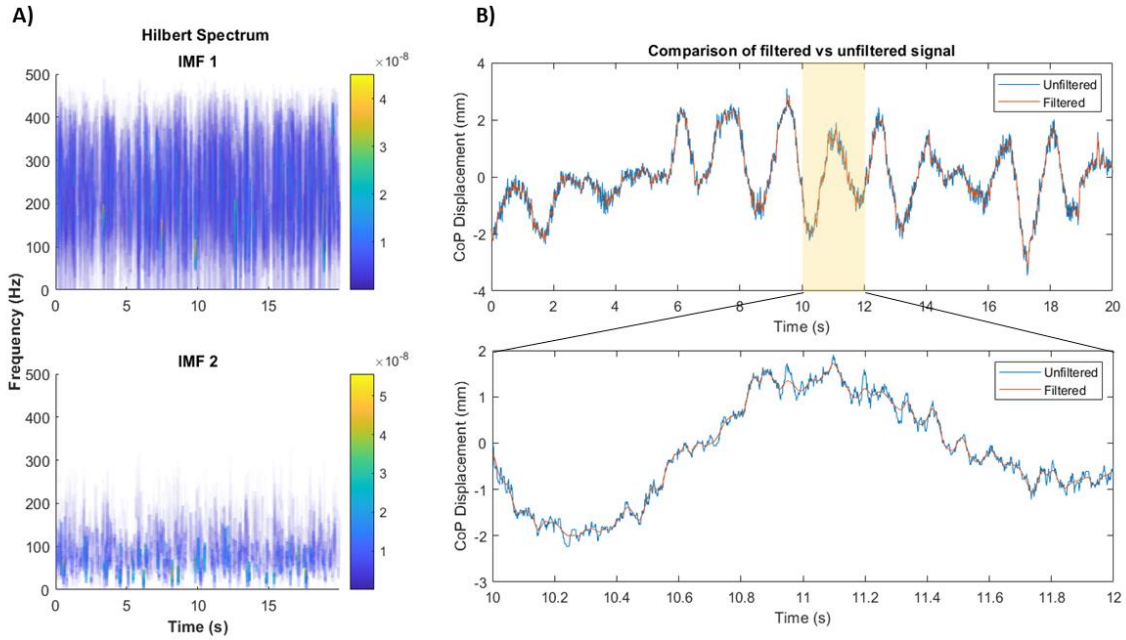


Figure 3-3. EMD Filtering. (A) Hilbert Spectrum of IMFs 1 and 2. Colobar indicates instantaneous energy. (B) Comparison between unfiltered COP (Blue) and filtered COP signal (Orange) after subtracting the first three IMFs. Bottom picture shows zoomed in version of the signals.

A comparison between the original signal and the filtered COP is shown in Figure 3-3-B. The bottom image is a zoomed-in visualization of two seconds of the signals. Both signals share the same global structure given by the low oscillations of higher order IMFs. The differences are more evident when we look at the finer details. The unfiltered signal presents more spikes and rapid changes, corresponding to the higher frequency contents.

3.4 Discussion

The filtering process through EMD was satisfactory, with low frequency contents and global dynamical changes preserved post-filtering. However, it is important to note the overlap in the frequency content between different IMFs. Looking at the frequency content of IMFs 1 and 2, it is evident that some lower frequency content is still contained in these components. This effect is called mode mixing, which has been reported before, and it is

known to be one of the major drawbacks of EMD [36]. Nonetheless, for our particular case, the energy content for these frequencies is one to two orders lower in IMFs 1 through 3 than on the preserved IMFs, meaning that the information loss may not be as substantial as compared to the retained signal. Alternatives to EMD have been proposed to solve this issue, one of the most popular ones being Ensemble EMD (EEMD) [37]. This algorithm is based on a noise-assisted version of the traditional EMD. For future work, more careful evaluation of these frequency contents and their impact on the final output should be considered.

Additionally, this project has focused on the 1-dimensional analysis of COP signals, with independent analysis of AP and ML directions. Future analysis could take into consideration a 2-dimensional analysis of COP, which may reflect better the interactions between both directions. An extension of EMD is the Multivariate EMD which could be appropriate for this type of study [38].

CHAPTER 4. FEATURE EXTRACTION AND ANALYSIS OF COP SIGNALS

4.1 Background

Characterizing changes in postural steadiness and control is fundamental to diagnose and evaluate the evolution of disorders and diseases affecting balance. As mentioned above, CDP is a widely used tool for evaluating patients suffering from balance issues after mTBI. Particularly, the analysis of COP signals is established as an effective means to evaluate balance and postural control [39]. Traditionally, these signals have been analyzed through linear metrics including the computation of mean displacements and variances, as well as the calculation of sway areas.

During the past decades, metrics based on the principles of information theory and entropy have been introduced into the analysis of medical signals. Specifically, Approximate Entropy (ApEn), Sample Entropy (SampEn), Multiscale Sample Entropy (MSE), and Complexity Index (CI) have been applied to the study of postural control [24], [40], [41]. In this study we analyzed a combination of linear and non-linear metrics for the evaluation of postural control in healthy and subacute concussed subjects.

4.1.1 Linear Metrics

As discussed previously, postural control is achieved through complex processing of sensory feedback, in which the integration of visual, vestibular, and proprioceptive cues plays a critical role. Alterations in the postural control system may disturb the dynamics of postural stability. Dynamic Posturography has allowed for the evaluation of these changes

under different sensory conditions. Traditional metrics based on the temporal patterns of the COP have been long studied for this purpose. Prieto *et al.* collected and proposed a set of measures that been the basis of most of the research in postural stability [42]. Here we focus on four of these metrics that reflect deviations in sway along the x-y plane.

The COP can be defined as a bivariate distribution constituted of the AP and ML coordinate functions. The recorded AP and ML time series are referenced on the basis of the force plate's coordinate system. Because each subject may be situated in slightly different positions off the center of the force plate, these signals are centered by subtracting the mean of the time-series:

$$AP[n] = AP_o[n] - \overline{AP_o}[n]$$

$$ML[n] = ML_o[n] - \overline{ML_o}[n]$$

Where $\overline{AP_o}[n]$ and $\overline{ML_o}[n]$ are the mean of the original AP_o and ML_o signals, and n are the discrete time points. Then, the distance vector (DIST) from the center of the COP can be estimated as:

$$DIST[n] = [AP[n]^2 + ML[n]^2]^{1/2} \quad n = 1, \dots, N \qquad [mm]$$

Where N is the number of time points of the signals. Then, the calculation for the mean distance or mean displacement (MDIST) from the COP is straightforward:

$$MDIST = \frac{1}{N}\sum_{n=1}^{N} MDIST[n] \qquad [mm]$$

The mean displacement along each direction can also be calculated. The mean distance along the AP direction is given by:

$$MDIST_{AP} = \frac{1}{N}\sum_{n=1}^{N}|AP[n]| \qquad [mm]$$

Another useful metric is the mean velocity (MVELO) of the displacement. This is calculated by dividing the total excursion, or total path, of the COP (TOTEX) by the total time *T*:

$$TOTEX = \sum_{n=1}^{N-1}[(AP[n+1]-AP[n])^2 + (ML[n+1]-ML[n])^2]^{1/2} \qquad [mm]$$

$$MVELO = \frac{TOTEX}{T} \qquad [mm/s]$$

Where T = 20 seconds, which is the duration of the SOT trials. The mean velocity in the AP and ML directions are calculated in an analogous way.

Another measure of stability is the sway area. This metric approximates the area enclosed by the COP per unit time. This can be calculated by adding up the area formed after triangulation of the COP time-series and dividing by the total time. The triangles are composed of each pair of consecutive points in the COP signal and the mean of the COP, as seen in Figure 4-1.Then, the calculation can be written as follows:

$$AREA_{SWAY} = \frac{1}{2T}\sum_{n=1}^{N-1}|AP[n+1]ML[n] - AP[n]ML[n+1]| \qquad [mm^2/s]$$

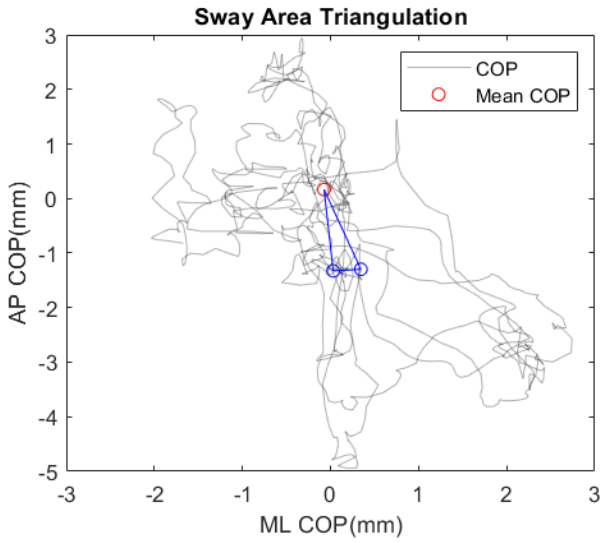


Figure 4-1. Triangulation of COP signals for Sway Area calculation. The blue dots indicate a pair of consecutive points. The red dot indicates the mean of the COP.

The last traditional metric examined in this study was the mean frequency (MFREQ). The frequency was approximated, in the temporal domain, as the rotational frequency of the COP if it had traveled its path in a circle of radius equal to MDIST:

$$MFREQ = \frac{TOTEX}{2\pi * MDIST * T} \quad [Hz]$$

In general, an increase of values for any of these metrics may be a sign of dysfunctional postural control or a deteriorated dynamical system.

4.1.2 Non-Linear Metrics: Entropy

Postural control is governed by a complex dynamic nonlinear system in which sensory feedbacks operate at different time scales. These nonlinear systems may reach equilibrium states, periodic or stable oscillations, or be susceptible to chaotic changes.

There is a new approach in the analysis of COP signals that considers that postural control is dominated by a chaotic but deterministic system. Alterations in this system may then be detected by measures of order or entropy. Cavanaugh *et al.* employed approximate entropy to detect abnormalities in postural control on athletes with concussion [43]. While traditional metrics of postural stability were not altered in these subjects, they exhibited reduced entropy in SOT conditions 1 and 2. Notably athletes post-concussion were found to have decreased entropy compared with control subjects.

Approximate entropy (ApEn) measures the complexity or randomness of a system. The framework for this metric is based on the field of information theory. Entropy is a measure of uncertainty or the amount of information we learn from a system given a set of observations. In this sense, entropy is a function of the probability distribution defined by the data series, not a function of the observations themselves. If this distribution is flat or uniform, all probabilities are equal, the entropy is high, and more uncertainty is generated [44].

ApEn has been applied to different research studies in the medical field [45], [46]. Biomedical signals are often of a nonlinear nature, for which measures of entropy may be more appropriate than other classical statistical metrics. ApEn essentially splits the signal in segments of equal size *m* and size *m+1* and checks whether these segments are repeated along the time-series within a tolerance of radius *r*. The algorithm behind ApEn works as follows:

1. Given a time-series $x = \{x_1, x_2, \dots, x_N\}$ with length N, select a window size *m*
2. Create the set of vectors $\vec{x} = \vec{x}_1, \vec{x}_2, \dots, \vec{x}_{N-m+1}$, where $\vec{x}_i = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$

3. Select a distance threshold or tolerance r
4. Calculate the distance between each pair of vectors $\{\vec{x}_i, \vec{x}_j\}$, and check if they are within the tolerance distance r:

$$\left\|\vec{x}_i - \vec{x}_j\right\|_m < r \quad \Leftrightarrow \quad \left|x_{i+k} - x_{j+k}\right| < r \quad \forall k \quad 0 \leq k \leq m-1$$

5. Given the tolerance r, the probability of vector $\vec{x}_i$ to be similar to vector $\vec{x}_j$ is given by:

$$C_i^m(r) = \frac{\sum_{j=1}^{N-m+1} \emptyset(i,j,m,r)}{N-m+1}$$

$$\emptyset(i,j,m,r) = \begin{Bmatrix} 1, & if \ \left\|\vec{x}_i - \vec{x}_j\right\|_m < r \\ 0, & otherwise \end{Bmatrix}$$

6. Then, with all vectors $\vec{x}_i$ we can calculate θ^m as:

$$\theta^m(r) = \frac{\sum_{i=1}^{N-m+1} \ln\left(C_i^m(r)\right)}{N-m+1}$$

7. Finally, ApEn is estimated as the difference of windows *m* and *m+1*:

$$ApEn(m,r) = \theta^m(r) - \theta^{m+1}(r)$$

For repetitive series with repeatable patterns, the approximate entropy tends to zero, as is the case of a perfect sinusoidal signal [47]. This value increases with more random or unpredictable time-series.

As shown in the above equations, this metric depends on two predefined parameters, the window length *m* and the tolerance distance *r*. These both parameters are highly dependent on the type of signal being evaluated, as well as the acquisition conditions and pre-processing steps. Traditionally, the window length is usually set to 2 or 3, and the tolerance is set between 0.1 to 0.25 times the standard deviation of the signal. However,

these values were initially determined for other types of physiological signals (e.g., respiratory time-series), and now there is ongoing research to establish more reliable parameters for COP signals [48].

One of the drawbacks of ApEn is that it counts self-matches in its calculation, so that the logarithm in the algorithm is always defined. This induces a bias towards a lower value of entropy, as the pattern becomes by default more repetitive. In an attempt to solve this issue, Sample Entropy (SampEn) was later defined [49]. SampEn is calculated as follows:

$$SampEn(r) = -log\frac{A(r)}{B(r)}$$

Where $A(r)$ is the number of pair of vectors of length *m+1* satisfying $\left\|\vec{x}_i - \vec{x}_j\right\|_{m+1} < r$, and $B(r)$ is the number of pair of vectors of length *m* satisfying $\left\|\vec{x}_i - \vec{x}_j\right\|_m < r$. In this calculation, self-matching vectors are not counted. By definition, *B* is always equal or greater than *A*, thus, the logarithm is always defined and of negative sign, making SampEn a positive number or zero for perfect repeatable patterns.

Multiscale Sample Entropy (MSE) is an extension of SampEn, that quantifies fluctuations at different time ranges or scales. MSE has two main steps:

1. Calculating the coarse-grained signal y^τ of the time-series x given a time scale τ:

$$y^\tau: \left\{y_1^\tau, y_2^\tau, \dots, y_{N/\tau}^\tau\right\} \quad with \quad y_j^\tau = \frac{1}{\tau}\sum_{i=(j-1)\tau+1}^{j\tau} x_i\,, \qquad 1 \le y_j \le N/\tau$$

2. Calculate SampEn of the coarse-grained signal y^τ: $SampEn(m, r, \tau)$

In essence, during coarse graining the signal x is split into segments of length τ, and the average value is calculated for each of those segments. These calculated values constitute the coarse-grained signal y^{τ}.

Generally, multiple time scales are evaluated simultaneously. It has been suggested that at least 200 data points should be present in the last time scale after coarse graining [50]. For this project, with a time series of length 2000 data points, the selected values for τ were 1 through 10, so that the condition mentioned above was satisfied.

An additional metric derived from the MSE, is the complexity index (CI). This metric is simply the summation of all SampEn values at each time scale τ:

$$CI = \sum_{\tau=1}^{10} SampEn(m, r, \tau)$$

Where $SampEn(m, r, \tau)$ is the SampEn of the coarse-grained signal at time scale τ.

4.2 Methods

An overview of the process followed for the signals analysis is shown in Figure 4-2. The first step, as discussed in the previous chapter, was to preprocess the signal to filter the high frequency components. The second step was to calculate the above-mentioned features on the COP signals, with special focus on the AP direction. The last step consisted of feature exploration, to visualize feature distributions among the two groups, and statistical analysis.

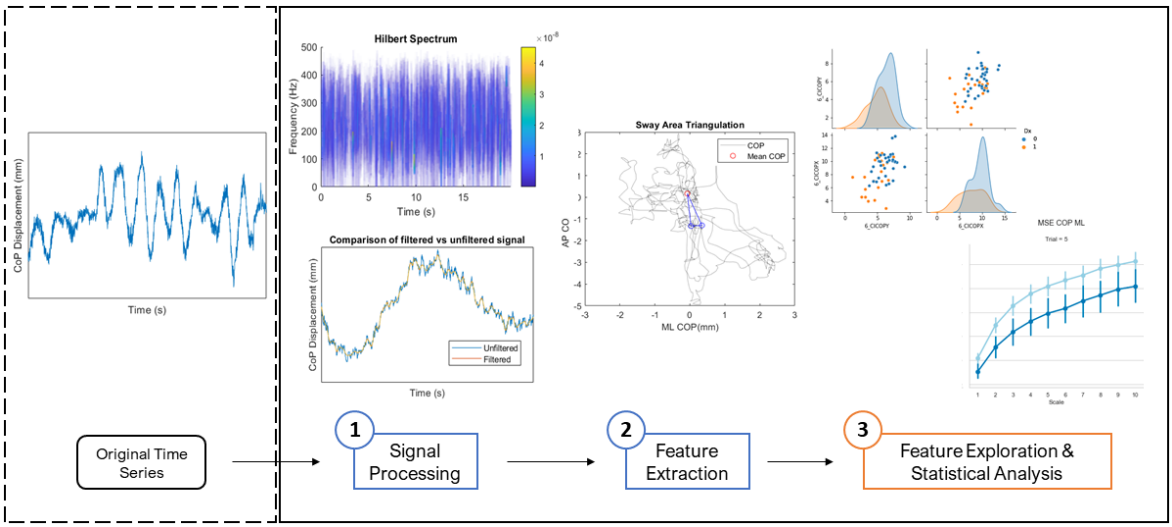


Figure 4-2. Schematic of Signal Analysis, with three main steps: 1) Signal Processing, 2) Feature Extraction and 3) Feature Exploration and Statistical Analysis. The Feature Extraction process includes both the calculation of traditional metrics and entropy metric

4.2.1 Calculation of Metrics

The above mentioned metrics were calculated for the COP signals of all 18 trials, and the average value for each SOT condition was then tested for significant differences between groups.

For the traditional linear metrics MDIST, MFREQ, MVELO and $AREA_{SWAY}$, as well as $MDIST_{AP}$, $MFREQ_{AP}$ and $MVELO_{AP}$ (mean distance, velocity and frequency in the AP direction respectively) were calculated. The non-linear metrics included $ApEn_{ML}$, $ApEn_{AP}$, CI_{ML} and CI_{AP}. The parameters for these metrics were set to $m=2$ and $r=0.15*STD$, where STD is the standard deviation of the signal being evaluated. Additionally, τ was set to $\tau = 1,2,\dots,10$ for the calculation of CI.

4.2.2 Statistical Analysis

Data distributions were tested to assess normality both visually and through Shapiro Wilk Tests. Because of the presence of outliers in certain variables, and the non-normality of the data, significant differences were evaluated through Wilcoxon Rank Sum tests. Bonferroni correction was employed to correct the *p-value* for multiple comparisons, resulting in a corrected *p-value* of $p < 0.008$.

4.3 Results

As mentioned above, a total of 21 ST subjects (12 Female, Age: 23.16 ± 4.67) were recruited for this study. This group presented on average with a history of two lifetime concussions, and a median of 35 days elapsed between the last injury and the time of examination. On the other hand, 34 HC subjects (20 Female, Age: 28.11 ± 4.46) were recruited, all with no history of previous lifetime concussion. A summary of the demographics and clinical scores is presented in Table 1.

4.3.1 Comparison with traditional linear metrics

After correction of the *p-value* for multiple comparisons, two of the linear metrics survived the threshold for significance. The sway area during condition SOT 2 showed significantly ($p=0.004$) greater values for ST with a median (IQR) value of 13.91 (11.24 to 22.55) mm^2/s vs the HC with a value of 10.62 (8.75 to 14.82) mm^2/s. Additionally, the mean frequency along the AP direction showed increased values for the HC group during SOT condition 2 ($p=0.005$), with a median value of 0.76 (0.65 to 0.88) Hz vs the ST group with 0.67 (0.47 to 0.74) Hz. It is also interesting to note a tendency for ST subjects to have greater mean distance - mean displacement from the center of the COP as seen in Figure 4-3-A. As an

example, for SOT condition 2, the ST subjects showed a median value of 3.37 (2.92 to 5.23) mm, as opposed to the HC group with 2.92 (2.43 to 3.49) mm, with *p=0.009*.

Table 1. Demographics and clinical scores of HC and subacute PCVD subjects. Data presented as median (IQR). Statistical significance tested through Wilcoxon Rank Sum Tests.

		CONTROL (HC)	**PCVD (ST)**	***P-VALUE***
N		34	21	n.a.
AGE		27 (24 - 30.0)	23 (19.8 - 27.3)	**0.003**
VOMS SCORE	**Baseline**	0 (0)	5.0 (3.75 - 6.25)	< **0.001**
	Smooth Pursuit	0 (0)	4.5 (4.00 – 7.25)	< **0.001**
	Horizontal Saccades	0 (0)	6.0 (4.75 – 9.25)	< **0.001**
	Vertical Saccades	0 (0)	6.5 (4.0 – 9.5)	< **0.001**
	Near-Point Convergence	0 (0)	6.0 (3.0 – 10.0)	< **0.001**
	Horizontal VOR	0 (0 – 0.75)	8.0 (5.0 -12.0)	< **0.001**
	Vertical VOR	0 (0)	9.0 (5.75 – 12.25)	< **0.001**
	VMS	0 (0 - 1)	9.0 (5.75 – 13.25)	< **0.001**
SOT EQUILIBRIUM SCORE	**Condition 1**	92.7 (91.7 - 94.0)	92.5 (89.6 - 94.4)	0.293
	Condition 2	92.4 (91.1 - 93.8)	91.9 (88.1 - 92.8)	**0.043**
	Condition 3	93.3 (92.1 - 94.3)	92.3 (86.6 - 93.8)	**0.039**
	Condition 4	74.4 (71.2 - 80.3)	75.8 (72.2 - 79.5)	0.441
	Condition 5	68.3 (63.6 - 73.8)	67.5 (54.8 - 74.8)	0.196
	Condition 6	71.4 (67.2 - 75.7)	69.9 (60.7 - 75.5)	0.186

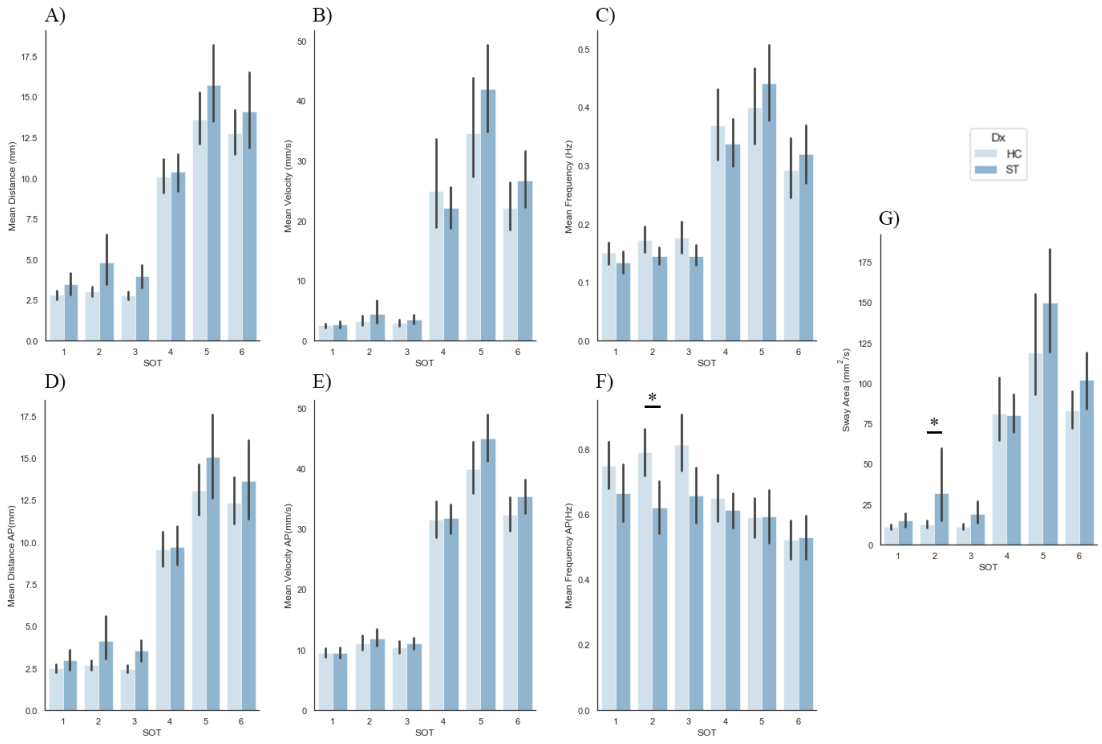


Figure 4-3. Comparison of linear metrics between HC and ST groups, with Mean Distance (A&D), Mean Velocity (B&E), Mean Frequency (C&F) and Sway Area (G). Data is averaged between every three trials for each SOT condition. Error bar displays 95% confidence interval. The asterisk (*) indicates significant metrics after Bonferroni correction ($p<0.008$)

4.3.2 Comparison with entropy metrics

In order to reduce the number of metrics assessed during the analysis, we opted to evaluate differences between groups based on ApEn and CI only, instead of using every independent temporal scale of the MSE. However, here we also take a deeper look at the different time scales for SOT condition 2, and its evolution with increasing trials, as seen in Figure 4-4. For all scales and trials, both SampEn in the AP and ML directions appear lower in the ST groups. These differences are more noticeable in the mid-scales (4 to 8) of trial 5. In terms of temporal evolution, if we look at the HC group from trial 4 to trial 6, which are recorded

sequentially, the SampEn generally increases for all time scales, as opposed to the ST groups.

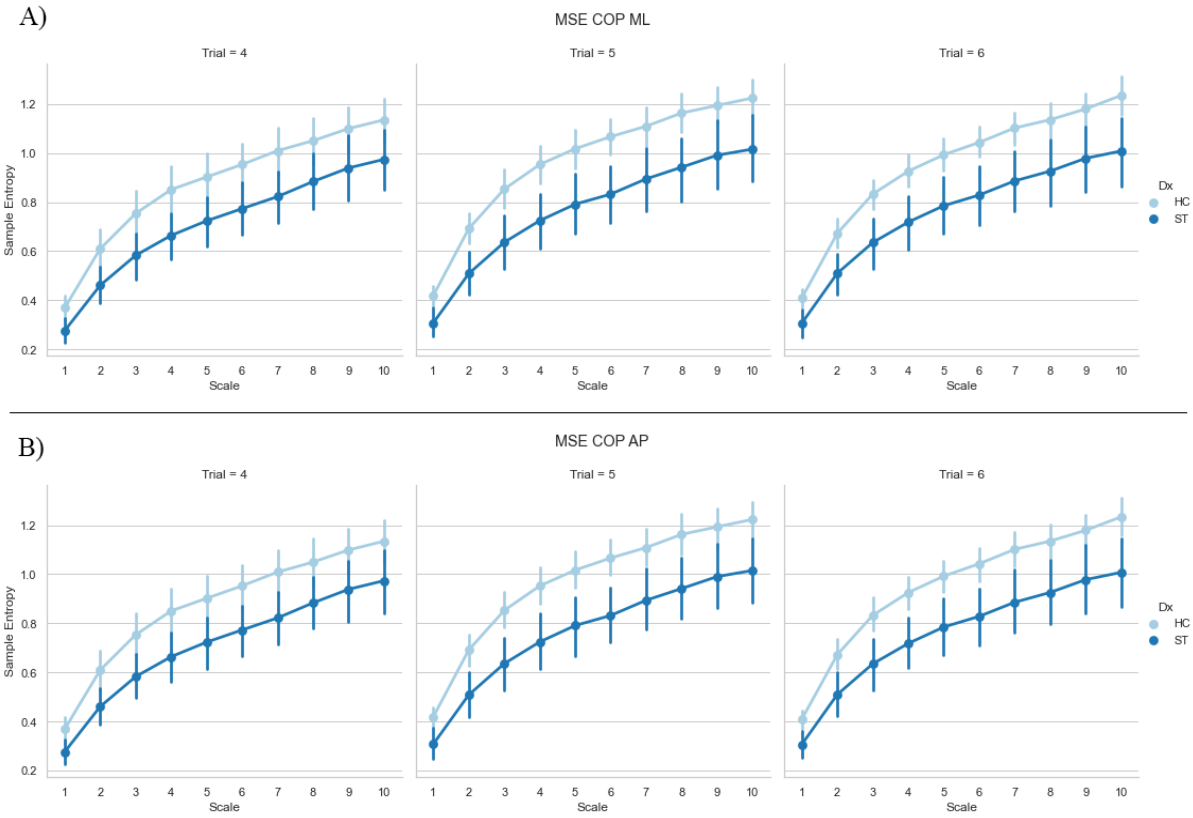


Figure 4-4. MSE of ML (A) and AP (B) COP signals for the three trials recorded during SOT condition 2. Point indicates mean value for each scale, and bars indicate 95% confidence interval.

These differences are also reflected in the CI values. As seen in Figure 4-5. All four non-linear metrics show significant differences between groups for the second SOT condition. In all cases, HC show higher entropy values as compared to ST, with 0.45 (0.39 to 0.51) vs 0.37 (0.25 to 0.44) for the ApEn in the ML direction (*p=0.004*), and 0.22 (0.19 to 0.26) vs 0.18 (0.12 to 0.20) for the ApEn in the AP direction (*p=0.004*). Complexity values were also statistically higher in HC for this SOT condition with 9.59 (8.55 to 10.39) vs 7.93 (5.93 to 9.16) for the CI in the ML direction (*p=0.002)*, 6.22 (5.48 to 7.30) vs 5.17 (3.67

to 6.11) for CI in the AP direction (*p=0.0009*). This behavior is also consistent for the rest of the SOT conditions, although statistical differences did not exceed the threshold.

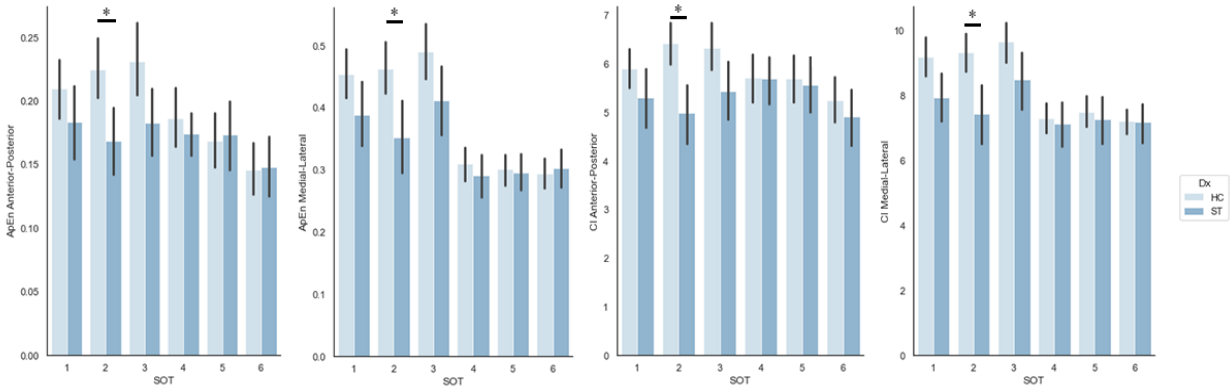


Figure 4-5. Comparison of nonlinear metrics between HC and ST groups in the six SOT conditions. Error bars indicate 95% confidence interval. Asterisk (*) indicates significant differences after Bonferroni correction (*p<0.008*).

4.4 Discussion

4.4.1 Traditional Linear Metrics for PCVD subjects' diagnosis

Linear metrics differentiated between the ST and HC groups including Mean Frequency in the AP direction, and Sway Area in condition 2 of the SOT. Although the mean displacement and mean velocity metrics did not show significant differences between subjects, there is a clear tendency for patients with PCVD to present higher values, which may be a direct response to changes in the postural control dynamics. This is also reflected in the sway area and the mean frequency metrics, especially for the second condition of the SOT. These results are consistent with prior studies evaluating balance after concussion in which linear measures often failed to discriminate concussion groups after the first week post injury [23], [24].

It is interesting to note that the mean frequency in the AP direction is significantly higher in the HC group. This can be interpreted as a rapid but small amplitude compensatory response as opposed to the ST, which may respond with higher amplitude but slower changes (i.e., lower frequency).

Additionally, this may have an interesting physiological interpretation. There has been research on the balance control system to understand the reweighting mechanism of the sensory input integration process. In the frequency analysis of balance control, frequency bands can be associated with different sensory afferent paths [51]. Friedrich M. et. al. estimated that the visual system operates in postural control in the lower frequency range of about 0.03 to 0.1 Hz [52], Oppenheim et al. estimated that vestibular feedback acts in the 0.1 to 0.39 Hz frequency range [53], while research by Kapoula Z. and Paillard T. indicated that the cerebellar system may operate from 0.4 to 1.5 Hz [54], [55]. Finally, Dietz V. and Diener HC. established that proprioceptive feedback and muscle activity may be concentrated in the higher frequencies between 1.5 and 5 Hz [56], [57].

Although our analysis of mean frequency is a simplified approximation of the mean frequency content, it is noteworthy to see higher values for mean frequency for the HC during SOT condition two. Given that eyes are closed in SOT condition two, this result may represent improved suppression of visual feedback among healthy controls with vision absent and a reweighting of other available sensory systems that operate in the lower time scales, shifting the mean frequencies towards higher values. However, deeper analysis of frequency contents would be necessary to establish more meaningful findings in this topic.

4.4.2 Entropy Metrics for PCVD subjects' diagnosis

The results obtained for this analysis are consistent with those reported by Cavanaugh et al. for post-concussion subjects [24]. Here we again demonstrate that ApEn values appear lowered after concussion, possibly indicating changes in the neurophysiological dynamical system. Head trauma can cause diffuse axonal injury that may disrupt neuronal interactions. This can in turn cause brain regions to become less coupled with an associated increment in the regularity of brain oscillations [58], [59], which could potentially affect the balance control system. Notably here we found prevalent changes in ApEn that go beyond the acute phase of the injury and that are sustained during the subacute phase an average of 35 days post injury, demonstrating changes in postural control dynamics in subacute PCVD subjects.

Similar results in Complexity Index were found in a research study by Purkayastha et. al. [60]. In this study, a cohort of concussed athletes showed reduced CI in the AP direction during eyes closed condition up to 90 days after injury, while linear metrics did not show significant differences with respect to controls. This is consistent with our findings, where both the ML and AP directions showed decreased CI. Lower complexity index may also be a direct indication of lower dynamic complexity in postural control. Sensory integration may then be affected, showing a system with lower dynamic interactions that produces less effective compensatory responses. This has already been shown in research on elderly populations as well as subjects affected by multiple sclerosis [40], [61].

These findings are in line with other studies evaluating neurophysiological changes after concussion, that are sustained even after clinical symptoms are resolved [62], [63]. For example, research by Pearce et. al. found evidence that subjects suffering from Post-Concussion Syndrome may show altered central information processing [64]. Proper

evaluation of these physiological changes should be further studied for improving clinical assessments. Specially athletes and professionals with higher risk of repetitive concussions should be carefully evaluated as continued head trauma may worsen their condition and negatively affect their recovery [65].

Accessible clinical assessments, such as the evaluations in this study, may be a good option to identify these physiological changes. However, further analysis of these metrics with bigger sample sizes may be necessary to establish clear correlations between COP dynamics alterations and neurophysiological abnormalities caused by mTBI.

CHAPTER 5. FEATURE IMPORTANCE THROUGH MACHINE LEARNING

5.1 Background

The application of machine learning models to the biomedical research has become more popular in the recent years [66]. The study of complex physiological processes becomes more intricate when evaluating multiple parameters or metrics and finding relationships between those may be critical to understand the overall function of the studied system.

Machine learning can be used as a tool for studying sets of data without much knowledge of the underlying patterns between the metrics or features analyzed. A machine learning model can be trained to learn relationships and functions between the input variables to solve a specific task. Some of the most common tasks are classification, regression, clustering, and dimensionality reduction.

Classification and regression tasks are part of a subset of machine learning applications called supervised learning. Supervised models are fed with samples that include the desired target or label to be predicted. In this way, the model can learn the optimal functions between the input features that generate most accurately the target output value.

On the other hand, dimensionality reduction and clustering are applications of the unsupervised learning domain. Unsupervised models are uniquely fed with the set of features to analyze, with no label or target output information. These models are trained to find patterns or relationships inherent to the feature set with no additional information.

In this project, we used a combination of both types of models to find patterns within the data that could reveal information about the features and their relationship with the two populations of subjects here studied. Specifically, we implemented a Random Forest classification model for feature importance, and an agglomerative clustering model for unsupervised grouping of the subjects. Additionally, Principal Component Analysis was employed for dimensionality reduction of the feature space.

5.1.1 Random Forest

Random Forests (RF) are one of the most commonly used classification models in machine learning. It is an ensemble method, meaning that it is composed of multiple individual models, also called weak learners, that work in conjunction to solve the specified task. In the case of RF, these individual models are called decision trees.

Decision trees are a type of non-parametric supervised algorithm. The model is structured into a flow-chart or tree-like graph that consists of three main elements: root node, decision nodes, and leaf nodes [67]. Nodes are organized by layers, with the root node being on the first layer, and the rest of the nodes expanding into subsequent layers. Each node can be thought as a test based on an attribute of the dataset. Each possible response to this test, constitutes a path or branch connecting to the next node. The root node constitutes the first test and, thus, determines the first splits on the dataset. Decision nodes represent subsequent tests that draw the possible paths within the tree. Finally, leaf nodes are the terminal nodes of the tree in which a final decision on the output value is taken. In this sense, decision trees are a sequence of questions posed to the studied sample, with each response determining the likelihood of belonging to a specific target output.

During training, the set of optimal attributes, thresholds and splits must be determined to reach the best performance of the model. Optimization of the model is generally based on one of two metrics: entropy and Gini Index, which try to estimate the inhomogeneity or impurity of the splits. Assuming we want to classify a set of samples *A* into *m* classes, we can define p_i $(i = 1, \dots, \mathrm{m})$ as the fraction or probability of samples in *A* that correspond to class *i*. Then, the entropy (*S*) for the probability distribution of set *A* is expressed as:

$$S = -\sum_{i=1}^{m} p_i \log p_i$$

The lower the entropy value *S*, the higher the homogeneity, and the purer the set *A* is. Alternatively, we can define the Gini index as:

$$Gini = 1 - \sum_{i=1}^{m} p_i^2$$

For a set with samples from one unique class, p_i is only non-zero (and equal to 1) for one class. *S* and *Gini* become zero, showing that the set is completely homogeneous. These metrics are used to determine the best splits on the tree, by minimizing inhomogeneity at each step, and ensuring that the splits are as pure as possible [68].

Once an impurity metric *I* is selected, the weighted average of impurities after a split can be estimated as:

$$I_{split} = \sum_{j=1}^{k} \frac{|A_j|}{|A|} I(A_j)$$

Where *k* is the number of children nodes after the split. Then, we try to find the test at the given node, that minimizes the impurity of the children nodes.

One of the main disadvantages of decision tress, is that they are prone to overfitting. This means that they can perfectly adjust the predictions to the training set with the downside that their potential for generalization in an external set is greatly diminished. RF deal with this problem by training multiple decision trees simultaneously and aggregating their predictions to one final result.

On the positive side, decision trees and RF are both interpretable models, as they allow us to look into the tree structures and take note of the features or attributes that best separated the dataset.

5.1.2 Principal Component Analysis

Principal Component Analysis (PCA) is the most widely used technique for dimensionality reduction. PCA consists of the transformation of a given dataset to a new feature space defined by uncorrelated variables, also called Principal Components (PCs), in which the variance of the data is maximally preserved, and the information loss is minimized.

PCA is an unsupervised, linear technique which can be directly solved through Singular Value Decomposition (SVD). PCA is calculated in three steps: 1) Mean-centering of the Data Matrix, 2) Calculation of Covariance Matrix, 3) Application of SVD on the Covariance Matrix [69].

The mean-centered matrix *Y* is calculated from the original matrix *X* by subtracting the mean of each feature column from *X*. Then, the Covariance Matrix *C* of *Y* is calculated as:

$$C = Y^T Y$$

Where Y^T, is the transpose of matrix *Y*. Then, the SVD of *C* can be expressed as:

$$C = V \Delta V^T$$

Where *V* is the matrix of eigenvectors, and Δ is the matrix of eigenvalues. Finally, the PCs can be calculated by:

$$T = YV$$

Where *T* represents the projected dataset *X* into the new defined space by the PCs.

One of the main characteristics of the PCs, is that they are orthogonal to each other. Moreover, they are organized in order of explained variance, this is, the first PC contains the most amount or percentage of variance of the original dataset, the second PC contains the second most, and so on. This is a critical characteristic for dimensionality reduction, as we can select a subset of PCs that will explain most of the variance of the original dataset without losing much information.

5.1.3 Agglomerative Clustering

Hierarchical clustering is an unsupervised learning method that consists of defining a hierarchy of groups (or clusters) within the data. Agglomerative Clustering is a specific type of hierarchical clustering, in which the hierarchy is built with a bottom-up approach, this is, each sample starts as a unique cluster and pairs of clusters are merged together moving up in the hierarchy[70].

Clusters are defined based on affinity or similarity between samples. Two main parameters define the algorithm for agglomerative clustering, the distance metric, and the linkage criterion. The distance metric is the measure employed to estimate the similarity between two specific samples. Examples of distance metrics include the Euclidian distance, Manhattan distance or Mahalanobis distance.

On the other hand, the linkage criterion specifies the way in which subset of observations are compared, this is, how the distances between samples of different clusters are weighted. Examples of linkage criteria include the Weighted Pair Group Method with Arithmetic Mean (WPGMA), or the Ward's criterion.

The Ward's criterion is a method that tries to minimize the increase of within variance of the new merged clusters. For two initial clusters *A* and B, this can be expressed as follows:

$$W_{AB} = SSE_{AB} - (SSE_A + SSE_B)$$

Where SSE_{AB} is the Sum Square Error (SSE) of the cluster *AB* (union of clusters A and B), SSE_A is the SSE of cluster A, and SSE_B is the SSE of cluster B. The clusters *A* and *B* that minimize W_{AB} would be merged together.

5.2 Methods

Two independent analyses were performed with this dataset. On the one hand, an analysis of feature relevance or importance for predicting health condition (i.e., HC vs ST) via RF. On the other hand, clustering was performed to find possible subpopulations within the studied populations. A diagram showing the complete analysis on this project is shown in

Figure 5-1. This figure shows the three machine learning models implemented in this section, and their respective preprocessing steps.

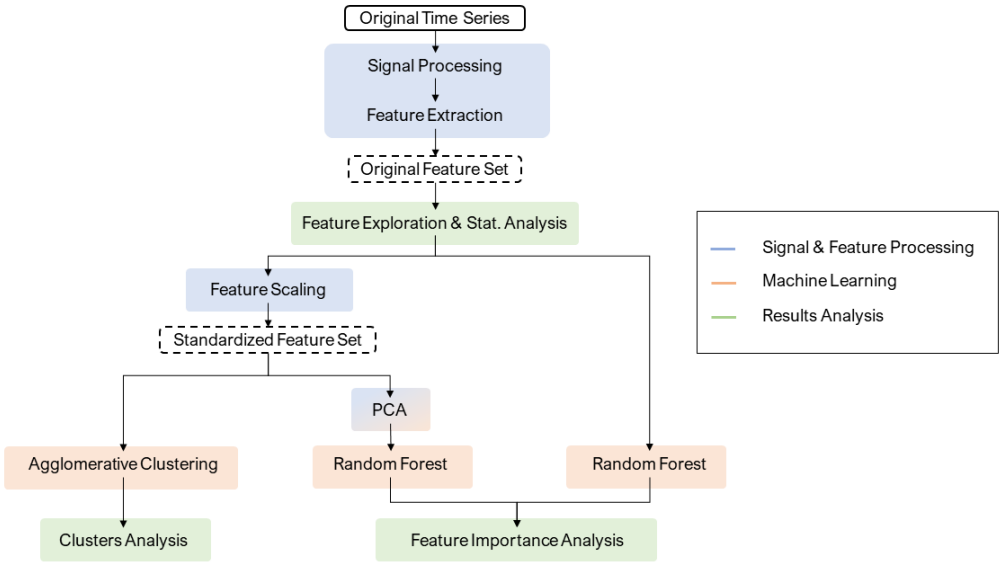


Figure 5-1. Schematic of the complete analysis including: 1) Signal and Feature Processing (Blue), 2) Machine Learning Techniques (Orange), 3) Analytic Evaluation of Results (Green). Feature Processing includes the calculation of features as discussed in Chapter 4.

5.2.1 Random Forests: Feature Importance

The dataset was split into training and testing set with an 80:20 ratio. Two random forests were trained, one with the original set of features, and another with a PCA decomposition of the features. RF can deal with non-normalized data, so the original features did not require any scaling prior to being inputted to the model.

Alternatively, PCA requires standardized data so that the scale of each feature does not influence the amount of explained variance. Features were scaled with a standard scaler (z-score) on the training set, and later PCA was applied. These transformations were then applied to the testing set to project into the same dimensional space.

In both cases, hyperparameter tuning was performed through 5-fold cross-validation with a grid-search on the training test. Performance was evaluated on the testing set through accuracy, ROC-AUC and F1-score to compensate for the imbalance nature of the dataset. Feature importance as estimated from the RF were then visualized.

The scikit-learn python package version 0.23.22 was employed for PCA and training of the RF [71]. The hyperparameters tuned for the RF during optimization were the number of estimators, and the maximum depth of the trees.

5.2.2 Agglomerative Clustering: Identification of subpopulations

An agglomerative clustering model was fit with the scaled original features to identify any possible groupings within the data. Agglomerative clustering is sensitive to the scaling of the inputted features, so the variables were standardized to have zero mean and unit variance. The Euclidean distance metric and the Ward's linkage criterion were selected. The number of clusters was set to none, to expand the hierarchy completely into individual clusters for each subject. A cluster-map with the hierarchical organization of subjects and features after clustering was drawn to visualize the estimated groupings with the Seaborn python package version 0.11.0 [72].

5.3 Results

5.3.1 Random Forest: Feature Importance

5.3.1.1 Random Forest with Original Features

The random forest trained with the original features led to an accuracy of 81%, ROC-AUC of 0.80 and F1-score of 0.75. The optimal RF contained six estimators with a maximum depth of two. The predictive features as estimated by the random forest are shown in Figure 5-2.

The three top selected features were the CI on the ML direction during SOT condition 6, the CI in the AP direction during SOT condition 2, and the ApEn on the AP direction during SOT condition 3.

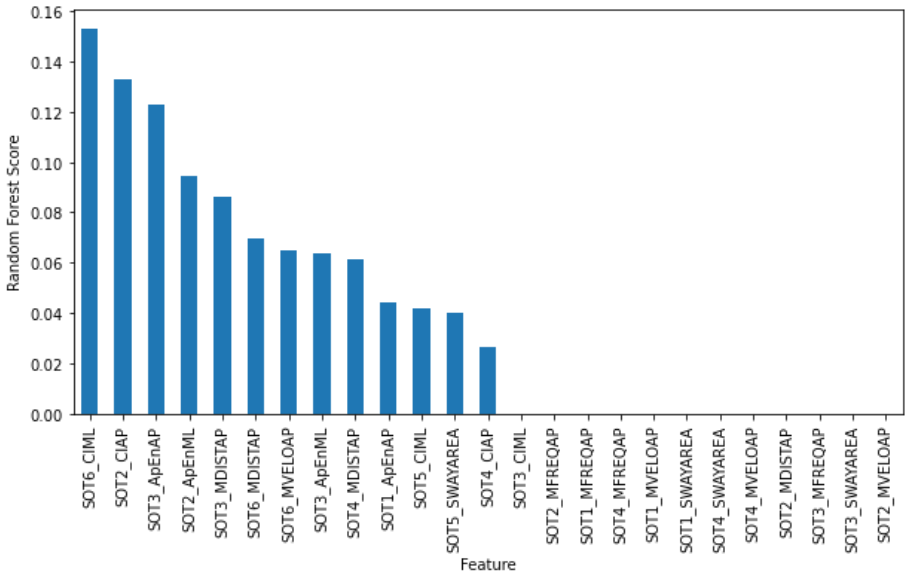


Figure 5-2. Feature Importance Scores calculated through Random Forest with original features. The higher the score, the higher the predictive power.

5.3.1.2 Random Forest with PCs

The explained variance for each PC, and the cumulative explained variance are shown in Figure 5-3. The first PC accounts for 36.6% of the variance, and the first 14 PCs account for around 95% of the total variance.

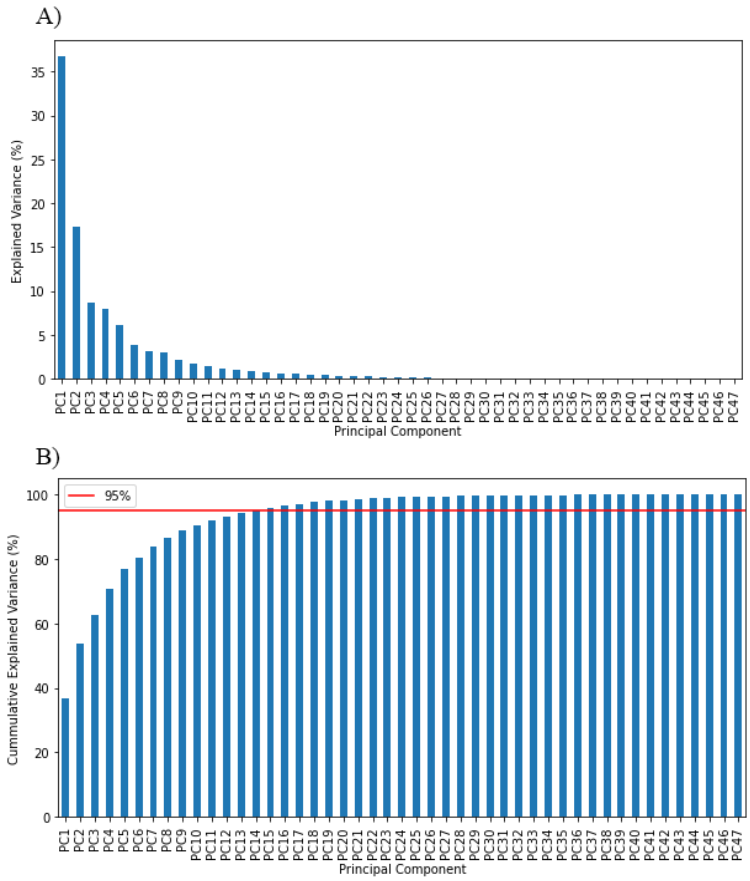


Figure 5-3. Principal Component Analysis: Variance Explained by each PC (A) and Cumulative Explained Variance (B) in percentages. Red horizontal line indicates 95% of cumulative explained variance

All PCs were fed as features to the random forest during training. The final performance on the testing set was 63% accuracy, 0.71 ROC-AUC and 0.66 F1-score. The optimal random forest consisted of seven estimators with a maximum depth of two. The predictive features for this random forest are shown in Figure 5-4. There are three PCs that appear

with the highest predictive power out of the original 47 PCs. These PCs are PC8, PC15 and PC24. These PC account for 3%, 0.79% and 0.13% of the variance of the feature space respectively.

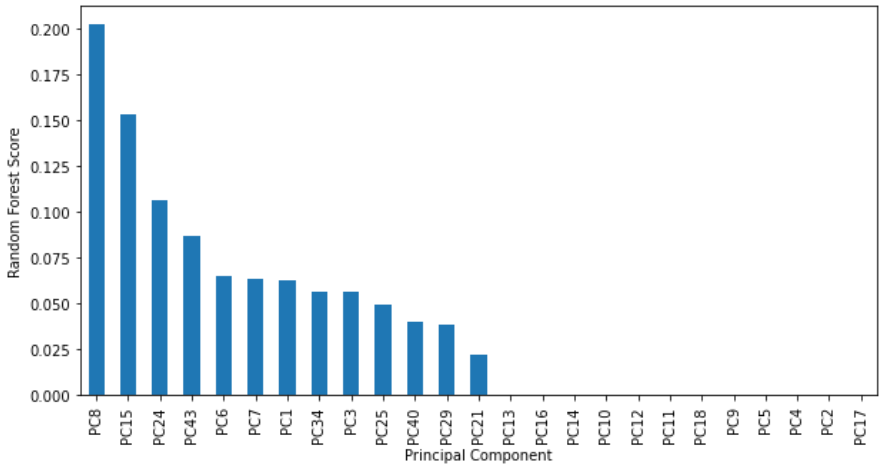


Figure 5-4. Feature Importance Scores calculated through Random Forest with PCs. The higher the score, the higher the predictive power.

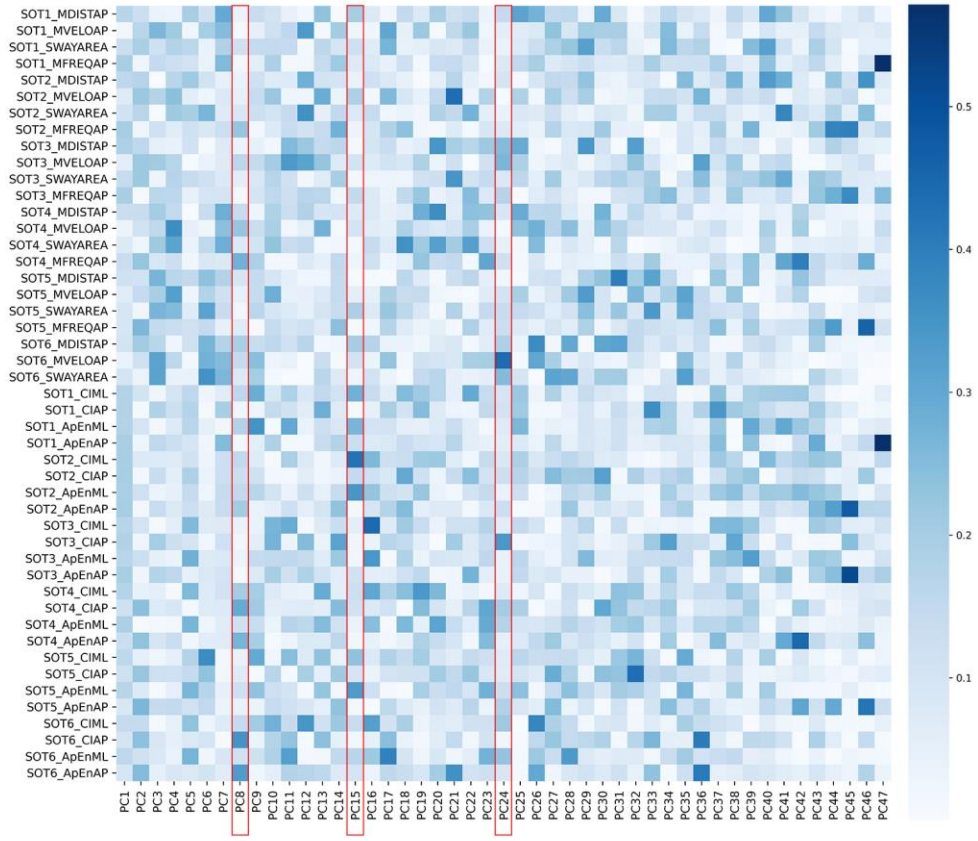


Figure 5-5. Contribution of original features to the Principal Components. Color-bar indicates absolute value of the correlation between feature and PC. The three PC with highest predictive power, as estimated by the random forest, are highlighted in red.

5.3.1.3 Agglomerative Clustering: *Identification of subpopulation*

The hierarchical organization found by this model is shown in the clustering heatmap in Figure 5-6. This heatmap represents the standardize value for each sample and variable. The first column (Dx) represents the diagnosis of the subject, colored in light blue for the HC and in dark blue for the ST. The clustering of the subjects is shown by the dendrogram on the left vertical axis, while the clustering on the feature space is represented by the dendrogram on the top horizontal axis.

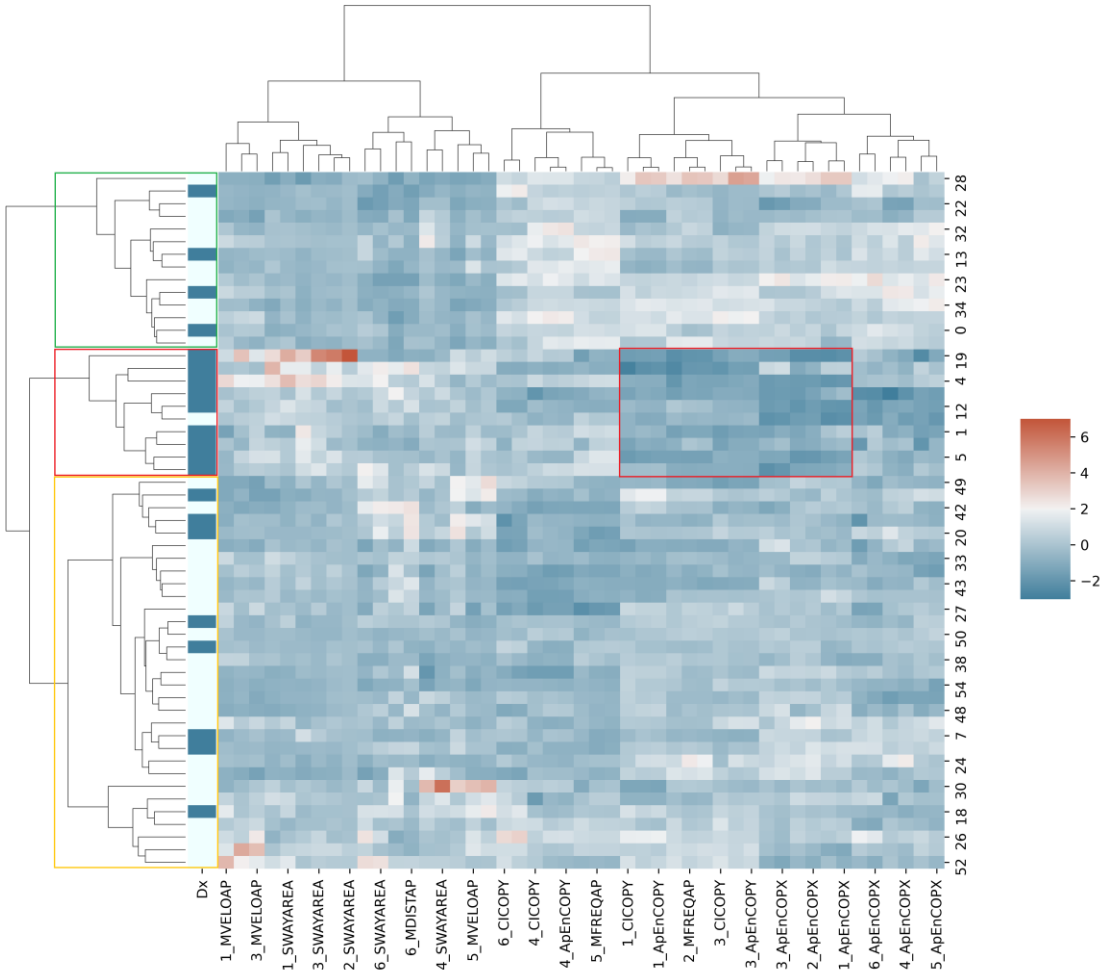


Figure 5-6. Hierarchically clustered Heatmap Organization of the dataset determined through Ward's criterion on the subjects (Vertical axis) and the features (Horizontal axis). The first column represents the diagnosis (Dx) of the subject, with a light blue for HC and dark blue for ST. A subset of ST identified by the clustering model is highlighted in red.

5.4 Discussion

5.4.1 Feature Importance

5.4.1.1 Feature Importance on Original Features

The set of features selected by the RF is in partial agreement with the most significant features found on the statistical analysis in the previous chapter. Metrics like the ApEn and

the CI in SOT condition 2 appear as predictive features helpful in discerning between HC and ST. These metrics showed clear differences, with a tendency for lower values in the ST group.

However, it is also remarkable the appearance of metrics extracted from SOT condition 6, particularly the CI in the ML direction, which appears as the most predictive feature. This variable, which alone did not show meaningful differences between groups, may have good predictive power in combination with the rest of the variables selected. This may imply the existence of complex relationships between variables, with non-linear behaviours, or it may be an indicative of the heterogeneity of the population, with subgroups of subjects showing different sets of metrics with abnormal values.

It is notable that metrics extracted from conditions 6, 2, and 3 emerge as predictive variables and are represented in the PCA and clustering analysis. In each of these SOT conditions, vision is manipulated with eyes closed in condition 2 and sway-referenced vision in conditions 3 and 6 resulting in unreliable visual sensory feedback. These conditions may best discriminate between ST and HC groups because PCVD results in an over-reliance on visual feedback.

5.4.1.2 Feature Importance on PCA transformation

The selection of PC8, PC15 and PC24 as predictive features, with very low explained variance, may indicate that there are interactions between the original features that are more relevant to the classification task than the original features themselves. If we take a look at the correlation between PCs and the original features (see Figure 5-5), we find that the CI and ApEn of SOT condition 6 in the AP direction, as well as the CI and ApEn of SOT

condition 2 in the ML direction are contributing the most to PC8 and PC15 respectively. The entropy metrics for SOT condition 2 already appeared in the statistical analysis as significantly different between groups. On the other hand, CI in the SOT condition 6 did not appear as significantly different but may be a key contributor in this complex dynamical system.

This is in accordance with the results found in the previous subsection. Although there are features that clearly show differences between populations, there may be hidden interactions between variables that may be more valuable in explaining changes between groups.

One of the main limitations of the analysis with RF is the small sample size. Generally, machine learning models require an extensive amount of data to generate generalizable models. In this case, these models have been utilized as a measure for feature importance, which may be indicative and supportive of predominant metrics withing our analysis. However, these results should not be understood as a definite statement on which metrics are most valuable within the complete analysis. We acknowledge the potential of these models and propose further analysis with more extensive datasets in the future.

5.4.2 Agglomerative Clustering

Regarding the feature clustering, it is interesting to note how this unsupervised model is capable of grouping features of the same type into the same clusters. If we focus on ApEn on the ML direction ("ApEnCOPX"), it appears clustered on the right side of the heatmap. We can also notice how SOT conditions 1 through 3 are clustered on one side and conditions 4 through 6 are clustered apart on a different grouping. This is consistent with

the nature of the SOT tests where, starting with condition 4, the proprioceptive feedback becomes unreliable provoking more disturbances to the subject and, thus, more changes into the entropy of the recorded signals.

Additionally, there are three main clusters of subjects found by the model, two of which present mixed subjects from both groups studied. However, it is interesting to note how this completely unsupervised model was capable of clustering a set of 9 ST subjects, shown on the red box on Figure 5-6. This may be indicative that our cohort of subjects may present different degrees of PCVD, with this specific cluster being more negatively affected. This can also be appreciated when looking at the red box highlighted on the heatmap. These are a set of features that appear more distinctive, for this set of subjects, with respect to the rest of the population. Specifically, these subjects present lower values in ApEn and CI on conditions 1 through 3. Furthermore, they present more positive values on the traditional metrics, particularly in Sway Area and MVELO, as seen on the first columns of the heatmap.

Machine learning has already been applied for phenotyping or subtype definition of other diseases like cancer, or cardiovascular diseases like heart failure and atrial fibrillation [73], [74]. An example is the research presented by Ahmad et. al. who utilized cluster analysis to define chronic heart failure phenotypes from clinical variables and biomarkers [75]. One of the major advantages of these methods is the definition of subpopulations within the disease that may be generally disregarded. These subgroupings may present specific characteristics that could be valuable for the clinician, to determine the health condition of the patient and provide treatments specific to the patient's needs [76]. However, most of

the limitations encountered in these studies are the limited clinical data available for analyses.

Agglomerative clustering is very informative and easy to interpret and visualize. However, one of the biggest limitations is susceptibility to noise and outliers. As with the case of RF, more extensive datasets with bigger sample sizes may be helpful to build more robust models with better defined clusters that may more faithfully represent subpopulations within the dataset.

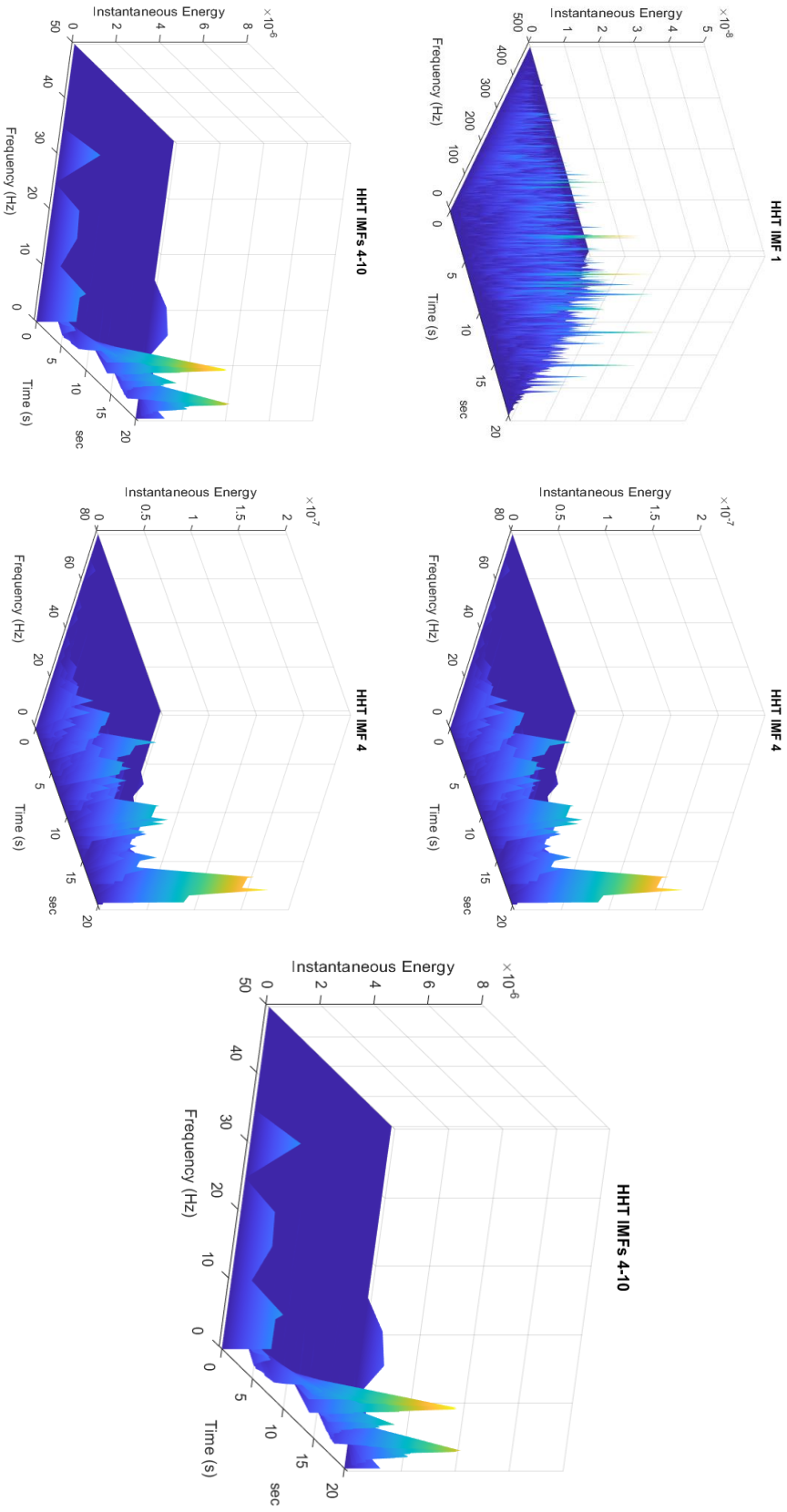
CHAPTER 6. CONCLUSION

This work has demonstrated the applicability of entropy measures and machine learning methods for detecting changes in postural control in patients suffering from PCVD. Moreover, it has been proved that changes in entropy, already present in the acute phase of injury, may extend into the subacute phase after mTBI. These findings are consistent with the traditional linear metrics, where deviations in the sway area may be directly related with postural complexity and the ability of a patient to adequate postural changes in different time scales. However, entropy metrics have demonstrated more sensitivity to these alterations during the subacute phase.

The use of machine learning shows potential for investigating patterns within the studied features. Complex dynamical systems, as the case of postural control, require great understanding of the underlying pathology and their associated manifestations. Machine learning may be a powerful tool to examine these extracted metrics, for which we have an interpretation into the associated pathology, to find complex relationships that may not be as apparent through classical analysis. However, it is recommended to reevaluate these models in future studies with larger sample sizes, as this may be a key factor for obtaining generalizable models. Additionally, machine learning may offer a way into personalized care, by analyzing subpopulations within the affected patients that may require individualized treatment for their specific underlying conditions.

In conclusion, this work establishes a new way to evaluate PCVD and demonstrates the potential for machine learning to provide decision support in clinical settings.

APPENDIX A. IMFS HILBERT SPECTRUM



APPENDIX B. TABLE TRADITIONAL METRICS

Table 2. Traditional metrics comparison between HC and ST. Results are presented as Median (IQR). Statistical differences were assessed through Wilcoxon Rank Sum tests. Significant p-values (p < 0.008, Bonferroni correction) are highlighted in bold.

METRIC	SOT CONDITION	CONTROL	PCVD	*P-VALUE*
MDIST (MM)	1	2.765 (2.273 - 3.241)	2.694 (2.487 - 3.632)	0.164
	2	2.924 (2.426 - 3.494)	3.372 (2.921 - 5.229)	0.009
	3	2.744 (2.224 - 3.299)	2.811 (2.460 - 5.143)	0.018
	4	9.583 (7.825 - 12.506)	10.257 (7.823 - 11.901)	0.368
	5	13.343 (10.752 - 15.413)	14.969 (11.19 - 20.054)	0.132
	6	11.483 (10.240 - 15.549)	13.475 (10.079 - 16.847)	0.299
MDISTAP (MM)	1	2.497 (1.926 - 2.928)	2.466 (2.116 - 3.311)	0.200
	2	2.625 (2.044 - 3.294)	2.966 (2.535 - 4.292)	0.018
	3	2.480 (1.969 - 2.917)	2.637 (2.269 - 4.814)	0.023
	4	8.906 (7.444 - 11.611)	9.473 (7.371 - 11.084)	0.394
	5	12.609 (10.348 - 14.943)	13.864 (10.655 - 19.169)	0.147
	6	11.061 (9.636 - 15.264)	12.688 (9.781 - 16.302)	0.293
MVELO (MM/S)	1	2.244 (1.724 - 3.121)	2.189 (1.918 - 3.483)	0.281
	2	2.657 (2.142 - 3.781)	3.123 (2.578 - 3.923)	0.042
	3	2.567 (2.038 - 3.621)	3.153 (2.592 - 3.781)	0.063
	4	19.800 (16.133 - 26.215)	21.56 (16.486 - 26.937)	0.361

	5	26.227 (19.711 - 33.703)	35.823 (28.869 - 56.895)	0.009
	6	20.009 (14.001 - 27.400)	24.83 (19.218 - 29.661)	0.028
MVELOAP (MM/S)	1	9.161 (7.899 - 10.837)	9.178 (8.448 - 10.754)	0.408
	2	10.172 (8.998 - 12.759)	10.766 (9.709 - 12.314)	0.105
	3	9.797 (8.363 - 12.045)	11.209 (9.788 - 12.088)	0.069
	4	30.129 (26.410 - 36.161)	31.997 (27.224 - 33.992)	0.252
	5	35.538 (31.484 - 42.393)	43.765 (38.682 - 51.291)	0.011
	6	31.219 (26.111 - 34.322)	35.508 (31.463 - 37.295)	0.029
MFREQ (HZ)	1	0.140 (0.098 - 0.188)	0.119 (0.109 - 0.160)	0.164
	2	0.152 (0.127 - 0.192)	0.138 (0.121 - 0.173)	0.069
	3	0.141 (0.121 - 0.223)	0.136 (0.121 - 0.169)	0.160
	4	0.339 (0.219 - 0.477)	0.312 (0.263 - 0.406)	0.317
	5	0.345 (0.253 - 0.500)	0.421 (0.311 - 0.565)	0.118
	6	0.265 (0.175 - 0.345)	0.301 (0.234 - 0.370)	0.114
MFREQAP (HZ)	1	0.755 (0.571 - 0.842)	0.659 (0.527 - 0.800)	0.101
	2	0.766 (0.647 - 0.878)	0.670 (0.469 - 0.737)	**0.005**
	3	0.772 (0.655 - 0.940)	0.660 (0.500 - 0.809)	0.011
	4	0.636 (0.505 - 0.847)	0.620 (0.511 - 0.715)	0.342
	5	0.567 (0.449 - 0.724)	0.566 (0.437 - 0.760)	0.421
	6	0.506 (0.369 - 0.624)	0.543 (0.385 - 0.622)	0.448
SWAY AREA (MM^2)	1	9.855 (7.791 - 12.041)	11.572 (8.450 - 16.501)	0.160

	2	10.615 (8.747 - 14.821)	13.916 (11.244 - 22.556)	**0.004**
	3	9.907 (7.737 - 12.739)	12.569 (9.574 - 22.529)	0.021
	4	69.563 (47.855 - 95.760)	76.895 (58.606 - 96.145)	0.151
	5	92.087 (76.074 - 118.683)	132.548 (96.031 - 201.677)	0.018
	6	76.570 (62.507 - 94.602)	99.796 (59.876 - 133.385)	0.084

APPENDIX C. TABLE NON-LINEAR METRICS

Table 3. Non-Linear Metrics Comparison between HC and ST. Results are presented as Median (IQR). Statistical differences were assessed through Wilcoxon Rank Sum tests. Significant p-values (p < 0.008, Bonferroni correction) are highlighted in bold.

METRIC	SOT CONDITION	CONTROL	PCVD	*P-VALUE*
APEN AP	1	0.212 (0.149 - 0.243)	0.183 (0.133 - 0.230)	0.101
	2	0.219 (0.178 - 0.263)	0.179 (0.118 - 0.206)	**0.004**
	3	0.220 (0.173 - 0.273)	0.181 (0.128 - 0.236)	0.012
	4	0.180 (0.134 - 0.247)	0.175 (0.138 - 0.216)	0.355
	5	0.162 (0.117 - 0.217)	0.162 (0.119 - 0.234)	0.394
	6	0.135 (0.097 - 0.180)	0.148 (0.100 - 0.177)	0.448
CI AP	1	6.132 (4.858 - 6.507)	5.463 (4.484 - 5.983)	0.061
	2	6.218 (5.481 - 7.298)	5.174 (3.669 - 6.110)	**0.001**
	3	6.228 (5.179 - 7.304)	5.043 (4.348 - 6.957)	0.026
	4	5.818 (4.615 - 6.862)	5.479 (5.066 - 6.732)	0.476
	5	5.248 (4.782 - 6.956)	5.364 (4.547 - 6.724)	0.381
	6	5.296 (4.128 - 6.318)	4.849 (3.971 - 5.873)	0.177
APEN ML	1	0.455 (0.385 - 0.493)	0.390 (0.338 - 0.445)	0.036
	2	0.453 (0.396 - 0.510)	0.370 (0.252 - 0.442)	**0.004**
	3	0.491 (0.414 - 0.571)	0.442 (0.291 - 0.489)	0.047
	4	0.296 (0.256 - 0.352)	0.277 (0.226 - 0.353)	0.226
	5	0.294 (0.258 - 0.341)	0.305 (0.255 - 0.341)	0.455
	6	0.281 (0.245 - 0.329)	0.294 (0.266 - 0.338)	0.191

CI ML	1	9.052 (8.313 - 9.987)	8.299 (7.263 - 8.696)	0.014
	2	9.595 (8.555 - 10.389)	7.928 (5.926 - 9.156)	**0.002**
	3	9.690 (8.753 - 10.875)	9.228 (6.396 - 10.191)	0.038
	4	7.173 (6.253 - 8.091)	6.821 (6.120 - 8.085)	0.361
	5	7.572 (6.449 - 8.289)	7.735 (6.464 - 8.325)	0.497
	6	7.194 (6.438 - 8.037)	7.286 (6.704 - 8.205)	0.374

REFERENCES

[1] W. Li, Y. Li, W. Zhu, and X. Chen, "Changes in brain functional network connectivity after stroke," *Neural Regeneration Research*, vol. 9, no. 1, p. 51, 2014, doi: 10.4103/1673-5374.125330.

[2] H. J. van der Horn *et al.*, "Brain network dysregulation, emotion, and complaints after mild traumatic brain injury," *Human Brain Mapping*, vol. 37, no. 4, pp. 1645–1654, Apr. 2016, doi: 10.1002/HBM.23126.

[3] A. E. W. Johnson, M. M. Ghassemi, S. Nemati, K. E. Niehaus, D. Clifton, and G. D. Clifford, "Machine Learning and Decision Support in Critical Care," *Proceedings of the IEEE*, vol. 104, no. 2, pp. 444–466, Feb. 2016, doi: 10.1109/JPROC.2015.2501978.

[4] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Zeitschrift für Medizinische Physik*, vol. 29, no. 2, pp. 102–127, May 2019, doi: 10.1016/J.ZEMEDI.2018.11.002.

[5] B. C. Lau, A. P. Kontos, M. W. Collins, A. Mucha, and M. R. Lovell, "Which on-field signs/symptoms predict protracted recovery from sport-related concussion among high school football players?," *The American journal of sports medicine*, vol. 39, no. 11, pp. 2311–2318, Nov. 2011, doi: 10.1177/0363546511410655.

[6] K. J. Schneider *et al.*, “Cervicovestibular rehabilitation in sport-related concussion: a randomised controlled trial,” *British journal of sports medicine*, vol. 48, no. 17, pp. 1294–1298, Sep. 2014, doi: 10.1136/BJSPORTS-2013-093267.

[7] R. J. Peterka, “Sensory integration for human balance control,” *Handbook of Clinical Neurology*, vol. 159, pp. 27–42, Jan. 2018, doi: 10.1016/B978-0-444-63916-5.00002-1.

[8] L. Fernández, H. A. Breinbauer, and P. H. Delano, “Vertigo and dizziness in the elderly,” *Frontiers in Neurology*, vol. 6, no. JUN, p. 144, 2015, doi: 10.3389/FNEUR.2015.00144/BIBTEX.

[9] J. H. Pasma, D. Engelhart, A. C. Schouten, H. van der Kooij, A. B. Maier, and C. G. M. Meskers, “Impaired standing balance: The clinical need for closing the loop,” *Neuroscience*, vol. 267, pp. 157–165, May 2014, doi: 10.1016/J.NEUROSCIENCE.2014.02.030.

[10] J. H. Pasma, D. Engelhart, A. C. Schouten, H. van der Kooij, A. B. Maier, and C. G. M. Meskers, “Impaired standing balance: The clinical need for closing the loop,” *Neuroscience*, vol. 267, pp. 157–165, May 2014, doi: 10.1016/J.NEUROSCIENCE.2014.02.030.

[11] A. M. Bronstein, J. F. Golding, and M. A. Gresty, “Vertigo and Dizziness from Environmental Motion: Visual Vertigo, Motion Sickness, and Drivers’ Disorientation,” *Seminars in Neurology*, vol. 33, no. 03, pp. 219–230, Sep. 2013, doi: 10.1055/S-0033-1354602.

[12] R. C. Gardner and K. Yaffe, “Epidemiology of mild traumatic brain injury and neurodegenerative disease,” *Molecular and Cellular Neuroscience*, vol. 66, no. PB, pp. 75–80, May 2015, doi: 10.1016/J.MCN.2015.03.001.

[13] L. J. Carroll, J. D. Cassidy, L. Holm, J. Kraus, and V. G. Coronado, “Methodological issues and research recommendations for mild traumatic brain injury: the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury,” *Journal of rehabilitation medicine*, no. 43 Suppl, pp. 113–125, Feb. 2004, doi: 10.1080/16501960410023877.

[14] S. Polinder *et al.*, “A Multidimensional Approach to Post-concussion Symptoms in Mild Traumatic Brain Injury,” *Frontiers in Neurology*, vol. 0, p. 1113, 2018, doi: 10.3389/FNEUR.2018.01113.

[15] M. A *et al.*, “A Brief Vestibular/Ocular Motor Screening (VOMS) assessment to evaluate concussions: preliminary findings,” *The American journal of sports medicine*, vol. 42, no. 10, pp. 2479–2486, Oct. 2014, doi: 10.1177/0363546514543775.

[16] J. W. Allen *et al.*, “Altered Processing of Complex Visual Stimuli in Patients with Postconcussive Visual Motion Sensitivity,” *AJNR. American journal of neuroradiology*, vol. 42, no. 5, pp. 930–937, May 2021, doi: 10.3174/AJNR.A7007.

[17] A. Trofimova, J. L. Smith, V. Ahluwalia, J. Hurtado, R. K. Gore, and J. W. Allen, “Alterations in Resting-State Functional Brain Connectivity and Correlations with Vestibular/Ocular-Motor Screening Measures in Postconcussion Vestibular

Dysfunction," *Journal of neuroimaging : official journal of the American Society of Neuroimaging*, vol. 31, no. 2, pp. 277–286, Mar. 2021, doi: 10.1111/JON.12834.

[18] H. J. Marcus *et al.*, "Vestibular dysfunction in acute traumatic brain injury," *Journal of Neurology 2019 266:10*, vol. 266, no. 10, pp. 2430–2433, Jun. 2019, doi: 10.1007/S00415-019-09403-Z.

[19] A. Mucha, S. Fedor, and D. DeMarco, "Vestibular dysfunction and concussion," *Handbook of Clinical Neurology*, vol. 158, pp. 135–144, Jan. 2018, doi: 10.1016/B978-0-444-63954-7.00014-8.

[20] J. M. Gurley, B. D. Hujsak, and J. L. Kelly, "Vestibular rehabilitation following mild traumatic brain injury," *NeuroRehabilitation*, vol. 32, no. 3, pp. 519–528, Jan. 2013, doi: 10.3233/NRE-130874.

[21] H. Charkhkar, B. P. Christie, and R. J. Triolo, "Sensory neuroprosthesis improves postural stability during Sensory Organization Test in lower-limb amputees," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–13, Apr. 2020, doi: 10.1038/s41598-020-63936-2.

[22] A. Ruhe, R. Fejer, A. Gänsslen, and W. Klein, "Assessing postural stability in the concussed athlete: what to do, what to expect, and when.," *Sports health*, vol. 6, no. 5, pp. 427–33, Sep. 2014, doi: 10.1177/1941738114541238.

[23] K. M. Guskiewicz, S. E. Ross, and S. W. Marshall, "Postural Stability and Neuropsychological Deficits After Concussion in Collegiate Athletes," *Journal of Athletic Training*, vol. 36, no. 3, p. 263, Jul. 2001, doi: 10.17615/4t5p-9y62.

[24] J. T. Cavanaugh, K. M. Guskiewicz, C. Giuliani, S. Marshall, V. Mercer, and N. Stergiou, “Detecting altered postural control after cerebral concussion in athletes with normal postural stability,” *British journal of sports medicine*, vol. 39, no. 11, pp. 805–811, Nov. 2005, doi: 10.1136/BJSM.2004.015909.

[25] S. B. Richmond, B. W. Fling, H. Lee, and D. S. Peterson, “The assessment of center of mass and center of pressure during quiet stance: Current applications and future directions,” *Journal of Biomechanics*, vol. 123, p. 110485, Jun. 2021, doi: 10.1016/J.JBIOMECH.2021.110485.

[26] J. Jeka, T. Kiemel, R. Creath, F. Horak, and R. Peterka, “Controlling Human Upright Posture: Velocity Information Is More Accurate Than Position or Acceleration,” *https://doi.org/10.1152/jn.00983.2003*, vol. 92, no. 4, pp. 2368–2379, Oct. 2004, doi: 10.1152/JN.00983.2003.

[27] R. M. Palmieri, C. D. Ingersoll, M. B. Stone, and B. A. Krause, “Center-of-Pressure Parameters Used in the Assessment of Postural Control,” *Journal of Sport Rehabilitation*, vol. 11, no. 1, pp. 51–66, Feb. 2002, doi: 10.1123/JSR.11.1.51.

[28] “CDP/IVR — Bertec.” https://www.bertec.com/products/cdp-ivr (accessed Nov. 23, 2021).

[29] L. Tan and J. Jiang, “Signal Sampling and Quantization,” *Digital Signal Processing*, pp. 15–56, Jan. 2013, doi: 10.1016/B978-0-12-415893-1.00002-0.

[30] J. J. Koltermann, M. Gerber, H. Beck, and M. Beck, “Validation of Various Filters and Sampling Parameters for a COP Analysis,” *Technologies 2018, Vol. 6, Page 56*, vol. 6, no. 2, p. 56, Jun. 2018, doi: 10.3390/TECHNOLOGIES6020056.

[31] C.-W. Huang, P.-D. Sue, M. F. Abbod, B. C. Jiang, and J.-S. Shieh, “Measuring Center of Pressure Signals to Quantify Human Balance Using Multivariate Multiscale Entropy by Designing a Force Platform,” *Sensors (Basel, Switzerland)*, vol. 13, no. 8, p. 10151, 2013, doi: 10.3390/S130810151.

[32] M.-S. Chen and B. C. Jiang, “Resistance Training Exercise Program for Intervention to Enhance Gait Function in Elderly Chronically Ill Patients: Multivariate Multiscale Entropy for Center of Pressure Signal Analysis,” *Computational and Mathematical Methods in Medicine*, vol. 2014, no. 21, 2014, doi: 10.1155/2014/471356.

[33] H. Liang, S. L. Bressler, R. Desimone, and P. Fries, “Empirical mode decomposition: a method for analyzing neural data,” *Neurocomputing*, vol. 65–66, no. SPEC. ISS., pp. 801–807, Jun. 2005, doi: 10.1016/J.NEUCOM.2004.10.077.

[34] H. Li, X. Qin, D. Zhao, J. Chen, and P. Wang, “An improved empirical mode decomposition method based on the cubic trigonometric B-spline interpolation algorithm,” *Applied Mathematics and Computation*, vol. 332, pp. 406–419, Sep. 2018, doi: 10.1016/J.AMC.2018.02.039.

[35] A. O. Boudraa and J. C. Cexus, “EMD-based signal filtering,” *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, Dec. 2007, doi: 10.1109/TIM.2007.907967.

[36] A. Stallone, A. Cicone, and M. Materassi, “New insights and best practices for the successful use of Empirical Mode Decomposition, Iterative Filtering and derived algorithms,” *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–15, Sep. 2020, doi: 10.1038/s41598-020-72193-2.

[37] Z. Wu and N. E. Huang, “ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOISE-ASSISTED DATA ANALYSIS METHOD,” *https://doi.org/10.1142/S1793536909000047*, vol. 1, no. 1, pp. 1–41, Nov. 2011, doi: 10.1142/S1793536909000047.

[38] X. Lang *et al.*, “Fast Multivariate Empirical Mode Decomposition,” *IEEE Access*, vol. 6, pp. 65521–65538, 2018, doi: 10.1109/ACCESS.2018.2877150.

[39] A. Ruhe, R. Fejer, and B. Walker, “Center of pressure excursion as a measure of balance performance in patients with non-specific low back pain compared to healthy controls: a systematic review of the literature,” *European Spine Journal*, vol. 20, no. 3, p. 358, Mar. 2011, doi: 10.1007/S00586-010-1543-2.

[40] M. A. Busa and R. E. A. van Emmerik, “Multiscale entropy: A tool for understanding the complexity of postural control,” *Journal of sport and health science*, vol. 5, no. 1, pp. 44–51, Mar. 2016, doi: 10.1016/J.JSHS.2016.01.018.

[41] D. W. Powell and D. S. B. Williams, “Athletes trained using stable compared to unstable surfaces exhibit distinct postural control profiles when assessed by traditional and nonlinear measures,” *Human movement science*, vol. 44, pp. 73–80, Dec. 2015, doi: 10.1016/J.HUMOV.2015.08.013.

[42] T. E. Prieto, J. B. Myklebust, R. G. Hoffmann, E. G. Lovett, and B. M. Myklebust, "Measures of postural steadiness: differences between healthy young and elderly adults," *IEEE transactions on bio-medical engineering*, vol. 43, no. 9, pp. 956–966, Sep. 1996, doi: 10.1109/10.532130.

[43] J. T. Cavanaugh, K. M. Guskiewicz, C. Giuliani, S. Marshall, V. Mercer, and N. Stergiou, "Detecting altered postural control after cerebral concussion in athletes with normal postural stability," *British Journal of Sports Medicine*, vol. 39, no. 11, p. 805, Nov. 2005, doi: 10.1136/BJSM.2004.015909.

[44] A. Delgado-Bonal and A. Marshak, "Approximate Entropy and Sample Entropy: A Comprehensive Tutorial," *Entropy 2019, Vol. 21, Page 541*, vol. 21, no. 6, p. 541, May 2019, doi: 10.3390/E21060541.

[45] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American journal of physiology. Heart and circulatory physiology*, vol. 278, no. 6, 2000, doi: 10.1152/AJPHEART.2000.278.6.H2039.

[46] D. Caldirola, L. Bellodi, A. Caumo, G. Migliarese, and G. Perna, "Approximate entropy of respiratory patterns in panic disorder," *The American journal of psychiatry*, vol. 161, no. 1, pp. 79–87, Jan. 2004, doi: 10.1176/APPI.AJP.161.1.79.

[47] G. Manis, "Fast computation of approximate entropy," *Computer Methods and Programs in Biomedicine*, vol. 91, no. 1, pp. 48–54, Jul. 2008, doi: 10.1016/J.CMPB.2008.02.008.

[48] L. Montesinos, R. Castaldo, and L. Pecchia, “On the use of approximate entropy and sample entropy with centre of pressure time-series,” *Journal of NeuroEngineering and Rehabilitation*, vol. 15, no. 1, pp. 1–15, Dec. 2018, doi: 10.1186/S12984-018-0465-9/FIGURES/4.

[49] A. Delgado-Bonal and A. Marshak, “Approximate Entropy and Sample Entropy: A Comprehensive Tutorial,” *Entropy 2019, Vol. 21, Page 541*, vol. 21, no. 6, p. 541, May 2019, doi: 10.3390/E21060541.

[50] M. A. Busa and R. E. A. van Emmerik, “Multiscale entropy: A tool for understanding the complexity of postural control,” *Journal of Sport and Health Science*, vol. 5, no. 1, pp. 44–51, Mar. 2016, doi: 10.1016/J.JSHS.2016.01.018.

[51] I. S. Lin *et al.*, “Reweighting of the sensory inputs for postural control in patients with cervical spondylotic myelopathy after surgery,” *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, pp. 1–12, Jul. 2019, doi: 10.1186/S12984-019-0564-2/FIGURES/7.

[52] M. Friedrich *et al.*, “Influence of pathologic and simulated visual dysfunctions on the postural system,” *Experimental Brain Research*, vol. 186, no. 2, pp. 305–314, Mar. 2008, doi: 10.1007/S00221-007-1233-4/FIGURES/6.

[53] U. Oppenheim, R. Kohen-Raz, D. Alex, A. Kohen-Raz, and M. Azarya, “Postural characteristics of diabetic neuropathy.,” *Diabetes Care*, vol. 22, no. 2, pp. 328–332, Feb. 1999, doi: 10.2337/DIACARE.22.2.328.

[54] T. Paillard, C. Costes-Salon, C. Lafont, and P. Dupui, “Are there differences in postural regulation according to the level of competition in judoists?,” *British Journal of Sports Medicine*, vol. 36, no. 4, pp. 304–305, 2002, doi: 10.1136/BJSM.36.4.304.

[55] Z. Kapoula, E. Matheron, E. Demule, C. Fauvel, and M. P. Bucci, “Postural Control during the Stroop Test in Dyslexic and Non Dyslexic Teenagers,” *PLOS ONE*, vol. 6, no. 4, p. e19272, 2011, doi: 10.1371/JOURNAL.PONE.0019272.

[56] H. C. Diener, J. Dichgans, B. Guschlbauer, and H. Mau, “The significance of proprioception on postural stabilization as assessed by ischemia,” *Brain research*, vol. 296, no. 1, pp. 103–109, Mar. 1984, doi: 10.1016/0006-8993(84)90515-8.

[57] V. Dietz, K. H. Mauritz, and J. Dichgans, “Body oscillations in balancing due to segmental stretch reflex activity,” *Experimental brain research*, vol. 40, no. 1, pp. 89–95, Aug. 1980, doi: 10.1007/BF00236666.

[58] S. M. Pincus, “Quantifying complexity and regularity of neurobiological systems,” *Methods in Neurosciences*, vol. 28, no. C, pp. 336–363, Jan. 1995, doi: 10.1016/S1043-9471(06)80040-6.

[59] B. T. Dunkley *et al.*, “Low-frequency connectivity is associated with mild traumatic brain injury,” *NeuroImage: Clinical*, vol. 7, pp. 611–621, Jan. 2015, doi: 10.1016/J.NICL.2015.02.020.

[60] S. Purkayastha *et al.*, “Balance Testing Following Concussion: Postural Sway versus Complexity Index,” *PM & R : the journal of injury, function, and*

rehabilitation, vol. 11, no. 11, pp. 1184–1192, Nov. 2019, doi: 10.1002/PMRJ.12129.

[61] B. Manor *et al.*, "Physiological complexity and system adaptability: Evidence from postural control dynamics of older adults," *Journal of Applied Physiology*, vol. 109, no. 6, pp. 1786–1791, Dec. 2010, doi: 10.1152/JAPPLPHYSIOL.00390.2010/ASSET/IMAGES/LARGE/ZDG0121093630003.JPEG.

[62] A. D. Wright, J. D. Smirl, K. Bryk, S. Fraser, M. Jakovac, and P. van Donkelaar, "Sport-related concussion alters indices of dynamic cerebral autoregulation," *Frontiers in Neurology*, vol. 9, no. MAR, p. 196, Mar. 2018, doi: 10.3389/FNEUR.2018.00196/BIBTEX.

[63] D. Esterov and B. D. Greenwald, "Autonomic Dysfunction after Mild Traumatic Brain Injury," 2017, doi: 10.3390/brainsci7080100.

[64] A. J. Pearce, M. Tommerdahl, and D. A. King, "Neurophysiological abnormalities in individuals with persistent post-concussion symptoms," *Neuroscience*, vol. 408, pp. 272–281, Jun. 2019, doi: 10.1016/J.NEUROSCIENCE.2019.04.019.

[65] T. McAllister and M. McCrea, "Long-Term Cognitive and Neuropsychiatric Consequences of Repetitive Concussion and Head-Impact Exposure," *Journal of Athletic Training*, vol. 52, no. 3, pp. 309–317, Mar. 2017, doi: 10.4085/1062-6050-52.1.14.

[66] J. Goecks, V. Jalili, L. M. Heiser, and J. W. Gray, “How Machine Learning Will Transform Biomedicine,” *Cell*, vol. 181, no. 1, pp. 92–101, Apr. 2020, doi: 10.1016/J.CELL.2020.03.022.

[67] G. B. Berikol and G. Berikol, “Predictive models in precision medicine,” *Artificial Intelligence in Precision Health*, pp. 177–188, Jan. 2020, doi: 10.1016/B978-0-12-817133-2.00007-0.

[68] C. Kingsford and S. L. Salzberg, “What are decision trees?,” *Nature biotechnology*, vol. 26, no. 9, p. 1011, Sep. 2008, doi: 10.1038/NBT0908-1011.

[69] H. Abdi and L. J. Williams, “Principal component analysis,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, Jul. 2010, doi: 10.1002/WICS.101.

[70] M. Z. Rodriguez *et al.*, “Clustering algorithms: A comparative approach,” *PLOS ONE*, vol. 14, no. 1, p. e0210236, Jan. 2019, doi: 10.1371/JOURNAL.PONE.0210236.

[71] F. Pedregosa FABIANPEDREGOSA *et al.*, “Scikit-learn: Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011, Accessed: Nov. 28, 2021. [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html

[72] M. L. Waskom, “seaborn: statistical data visualization,” *Journal of Open Source Software*, vol. 6, no. 60, p. 3021, Apr. 2021, doi: 10.21105/JOSS.03021.

[73] A. Banerjee *et al.*, “Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility,” *BMC Medicine 2021 19:1*, vol. 19, no. 1, pp. 1–14, Apr. 2021, doi: 10.1186/S12916-021-01940-7.

[74] G. Guzzetta, G. Jurman, and C. Furlanello, “A machine learning pipeline for quantitative phenotype prediction from genotype data,” *BMC Bioinformatics*, vol. 11, no. SUPPL. 8, pp. 1–9, Oct. 2010, doi: 10.1186/1471-2105-11-S8-S3/FIGURES/4.

[75] T. Ahmad *et al.*, “Clinical Implications of Chronic Heart Failure Phenotypes Defined by Cluster Analysis,” *Journal of the American College of Cardiology*, vol. 64, no. 17, pp. 1765–1774, Oct. 2014, doi: 10.1016/J.JACC.2014.07.979.