Acceleration Signals in Determining Gait-Related Difficulties and the Motor Skill of Walking in Older Adults

by

Pritika Dasgupta

Bachelor of Science, Cornell University, 2012

Master of Public Health, University of Michigan – Ann Arbor, 2014

Master of Health Informatics, University of Michigan – Ann Arbor, 2016

Master of Science, University of Pittsburgh, 2017

Submitted to the Graduate Faculty of

the School of Medicine in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2021

UNIVERSITY OF PITTSBURGH

SCHOOL OF MEDICINE

This dissertation was presented

by

Pritika Dasgupta

It was defended on

April 23, 2021

and approved by

Ervin Sejdić, Associate Professor, Department of Electrical and Computer Engineering

Xinghua Lu, Professor, Department of Biomedical Informatics

Jessie VanSwearingen, Associate Professor, Department of Physical Therapy

Mark Redfern, Professor, Department of Electrical and Computer Engineering

Sofia Triantafillou, Assistant Professor, Department of Biomedical Informatics

Dissertation Director: Ervin Sejdić, Associate Professor, Department of Electrical and Computer Engineering

Acceleration Signals in Determining Gait-Related Difficulties and the Motor Skill of Walking in Older Adults

Pritika Dasgupta, PhD

University of Pittsburgh, 2021

In adults 65 years or older, falls or other neuromotor dysfunctions are often framed as walking-related declines in motor skill; the frequent occurrence of such decline in walking-related motor skill motivates the need for an improved understanding of motor skill in walking. Simple gait measurements, such as speed, do not provide adequate information about the quality of the translation of the body motion during walking. Furthermore, there is a great need in the clinical literature and clinical practice for more accurate measures of the loss of the motor skill of walking, so that clinical practice can provide better therapeutic interventions to improve the motor skill of walking. This dissertation suggests a consensus on what the motor skill of walking is and dissects it into seven interrelated characteristics and traits. Subsequently, we purport that these characteristics of the motor skill of walking cannot be represented by simple gait measurements or raw sensor measurements alone. Gait measures from accelerometers placed on the lower trunk, or trunk-acceleration gait measures, can enrich measurements of walking and motor performance. To support our claim, we will map these acceleration gait measures (AGMs) to the various aspects of the motor skill of walking. Additionally, influential AGMs will be elected through feature selection methods. Various machine learning algorithms ranging from logistic regression, non-linear regression, evolutionary algorithms, and ensemble methods will be used to make predictions on age-related gait-related difficulty outcomes (such as fall risk). Overall, we hope to find that the combination of high-fidelity artificial intelligence algorithms and acceleration gait measures derived from low-cost sensors can fulfill the severe and crucial need for the clinical measurement of the motor skill of walking in older adults.

Keywords: motor control, motor skill, movement control, acceleration, wearables, gait, clinical informatics, machine learning, feature specification.

Table of Contents

List of Tables

List of Figures

1.0 Introduction

1.1 Motivation

1.1.1 Increasing Global Life Expectancy of Older Adults

In 2019, according to the World Population Prospects 2019 report by the United Nations, 1 in 11 people were over the age of 65; predictions indicate that the population of older adults may increase to 1 in 6 people being over the age of 65, by 2050 [237]. Furthermore, life expectancy for older adults has grown over the years, with those aged 65 years or over expecting to live an average of an additional 19.1 years, and those who are 80 years or over may live an average of a further eight years [263]. However, increased life expectancy can correspond to an increased number of patients in the health system; moreover, older adults typically have higher rates of chronic disease and other pre-existing conditions [263].

1.1.2 Gait Difficulties in Older Adults

Falls are one of the common causes of injury in older adults [33]. Some of the factors that may lead to falls are gait and balance disorders [33]. Typically, walking is acquired through motor learning [321]. The hallmark of the motor skill of walking is a smooth and efficient movement that requires minimal attention [321]. Among older adults, the motor skill of walking varies widely [99, 208, 308] with declines in motor skill being among the most significant causes of falls [13], morbidity [118], and low quality of life [57, 119, 100]. Age-related decline in sensorimotor function further increases motor decline and detrimental changes to gait [9].

1.2 Gait

1.2.1 Anatomy and Human Movement

In this work, we present and study the human body in its anatomical state. There are three anatomical planes that body is partitioned in; these planes will define how we describe human movement. We define three anatomical planes of the human body: coronal plane, the transverse plane, and the saggital plane. We will also define the anterior-posterior (AP), mediolateral or medial-lateral (ML), and vertical (V) directions, which are the directions in which the body moves. A reference figure can be found in Figure 5 in [111].

The coronal plane partitions the body into the anterior (front) and posterior (rear) sections. Anterior locations are on the front, while posterior locations are on the back of the body. This plane is also called the "frontal" plane. Movement about this plane occurs in the AP direction. Example movements in the coronal plane are abduction and adduction. Abduction is defined to be the movement that is moving away from the midline of the body or from an adjacent part of the body. Adduction is to move inward toward the median axis of the body or toward an adjacent part of the body.

The transverse plane partitions the the body into superior (upper) and inferior (lower) sections. Movement about this plane is on the longitudinal or V axis. Example movements in this plane include internal and external rotations, such as twisting.

The saggital plane partitions the body into left and right halves. Movement about this plane occurs around the ML axis. Example movements in the saggital plane are called flexion and extensions.

1.2.2 The Gait Cycle

Walking can be described as the way humans move across space, using their legs in an alternating manner for propulsion and support. The period in which walking occurs is represented by: the gait cycle or a stride. The gait cycle, or a stride, can be divided into a variety of several events. Broadly speaking, it can be divided into the stance phase and the swing phase. Each leg occupies about 60% and 40% of the gait cycle in the stance and

swing phase, respectfully.

1.2.2.1 Stance Phase The stance phase can be further divided into the initial contact phase, the loading response phase, the midstance phase, the terminal phase, and the pre-swing phase [241].

A leg enters the stance phase when its foot is in contact with the ground (this is called initial contact). Both legs can be in this position during the gait cycle, and that is called the double support period [241].

If we start with the left foot, initial contact begins when the left foot first makes contact with the ground. In normal walking, this is when the heel first strikes the ground (i.e., heelstrike). This is also the beginning of the double support event.

The loading response, which accounts for about 10% of the gait cycle, is next and it is considered the period in which the left foot rocks from its heel to midfoot and accepts the weight of the body. This phase occurs all the way up to the toe-off event of the right foot.

Midstance follows the loading response of the left foot. The weight of the body is on the left leg (single support), while the right food swings from its toe-off towards its next heelstrike. Midstance accounts for about 25% of the gait cycle.

Terminal stance (about 20% of the gait cycle) follows midstance, and it is still in single support. This phase marks the moment the heel of the supporting foot rises off the ground until the footstrike of the swinging ipsilateral leg.

The pre-swing sub-phase (about 10% of the gait cycle) follows next, where the body is again in double support. However, the weight shifts from the left leg to the right leg and the left foot continues to rock from midfoot to toe-off. This phase positions the limb for the swing phase.

1.2.2.2 Swing Phase The swing phase follows the pre-swing sub-phase. The swing phase allows the left foot to advance forward and position the foot in preparation for the next stance phase. The swing phase can be subdivided into the sub-phases: initial swing, midswing, and terminal swing [241].

The initial swing (about 13% of the gait cycle) happens when the left foot leaves the

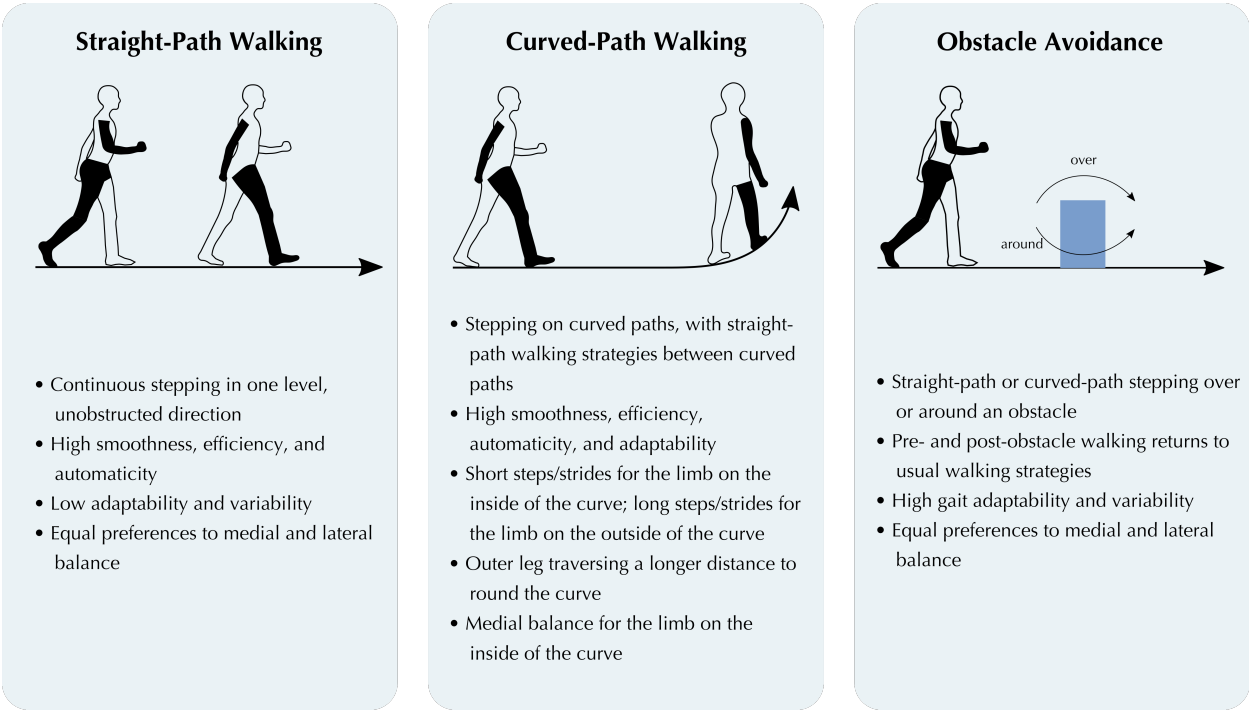
ground until the swing foot is next to the stance foot. Midswing occurs when the feet are adjacent to each other, and it continues until the tibia of the swing leg is vertical. The terminal swing comes next, and it is the final sub-phase of the gait cycle. The terminal swing begins when the tibia of the swing leg is vertical and lasts until the foot makes contact with the floor.

1.2.3 Definition of Walking

Walking is defined as gait with intent, specifically, the control of the body's center of mass and the continuation of movement; it involves multiple aspects of motor skill, which we call "the motor skill of walking" [321, 49]. In the most general sense, walking can be thought of as moving the body through space by repetitive stepping (i.e., gait cycle) while maintaining postural stability and balance [337]. Postural stability refers to the inter-segmental coordination during locomotion, including pelvic, torso, and head control, as well as arm swing coordination. Balance is the ability to remain upright while walking. Thus, walking requires complex coordination to be successful [337].

The motor skill of walking is the set of learned coordinated actions that result in the translation of the body through space while maintaining postural control and balance [321, 49]. In various real-world environments (e.g., indoor, outdoor, in crowded malls, uneven or littered ground), motor skill needs to be tractable. For example, this tractability can be defined for three general paths of walking: a straight path, a curved path, and an obstacle avoidance path (Figure 1.1) [321, 35, 65, 43]. In each case, changes in foot placement and postural adjustments are superimposed upon the gait cycle. Kinematic measurements during walking are used to quantify aspects of gait in order to evaluate the motor skill of walking. Several metrics can be calculated from these characteristics, which focus on particular aspects of the motor skill of walking. Aligning the right metrics to the particular aspect of the motor skill of walking is imperative in defining healthy walking and impairments.

Figure 1.1: For each of the three walking tasks: straight-path, curved-path, and obstacle avoidance, we define "good" motor skill of walking.



1.2.4 Characteristics of the Motor Skill of Walking

The intact motor skill of walking results in a smooth and efficient translation of the body over the surface. A decline in motor skills often leads to coordination loss, haphazard timing of stepping, postural instability, and asymmetries in the phases of gait during walking. Each of these aspects of motor skill is important in the evaluation of locomotion towards defining impairments and guiding rehabilitation. Based on the literature search, we defined seven critical characteristics of the performance outcome of the motor skill of walking:

- **Smoothness** is the consistent forward progression and regular, repeatable pattern of steps during walking [44, 215, 146]. Specifically, the smoothness of walking refers to the acceleration and deceleration of the trunk during walking. An interruption of the gait cycle events, such as heel strike and toe-off, can lead to uneven walking, characterized by

an extended deceleration of the "the leading limb at heel strike and altered accelerations of the trunk to advance the trailing limb [321, 44, 215]."

- **Efficiency** is inversely related to the energy expenditure during walking; the higher energy cost of walking, the lower the efficiency [321, 330].
- **Automaticity** is the reproducibility of walking motor skill with limited attentional, central nervous system resources for guidance [321, 66].
- **Adaptability** is the set of accommodations to walking based on the response before or after the loss of postural balance (due to obstacles or biomechanical defects) [122].
- **Variability** (or Regularity) is the change or fluctuation in walking from one stride to the next [108]. Multiple metrics claim to measure gait variability, leading to many ambiguous definitions; however, gait variability refers to stride to stride fluctuations [200].
- **Stability** is the notion that when a body, in and out of motion, goes out of equilibrium, forces or moments will restore the original condition of the body [341, 142, 122]. Furthermore, stability refers to the capability of maintaining an upright balance during walking [158].
- **Symmetry** is the agreement between the actions and behavior of the lower limbs during walking [271, 254]. While smoothness and variability may include some aspects of symmetry, symmetry is more focused on the concordance of contralateral motion while walking [203, 44, 173].

The above characteristics can be evaluated in various locomotor tasks. For example, in straight-line walking, good motor skill is indicated by clinical measures of low gait variability. In contrast, for curved-path and obstacle-avoidance walking, excellent performance is indicated by clinical measures of high gait adaptability, particularly in step lengths and widths. Furthermore, in curved-path walking, a good motor skill can be indicated by high gait variability [35]. Hallmarks of poor straight-path and curved-path walking are a decrease in walking speed, a decrease in stride length, a reduction in trunk movement, decreased strength and flexibility, and decreased balance [70]. Signs of poor obstacle-avoidance walking are decreased swing velocity, rapid stepping to maintain balance, shorter step lengths, shorter obstacle-heel strike distance, and freezing/stopping in motion.

1.2.5 Impact of Gait Difficulties in Older Adults

In those with gait-related difficulties, the limited motor skill of walking or a mobility disability can lead to exorbitant healthcare costs (e.g., 31.3 - 49.5 billion US dollars were spent on treating fatal and nonfatal fall injuries in the hospital, ED, and outpatient settings in 2014-2015 [120]). The motor skill of walking is affected by age- and disease-related metabolic, cardiovascular, musculoskeletal, and neurological changes. Thus the altered motor skill of walking can be a functional indication of the aging system decline or disease states. For example, for those who have Parkinson's, walking in a straight path is more accessible than walking on a curved path or through/over obstacles [116]. Even in the presence of adequate muscle strength and endurance, and pain-free, the greater difficulty navigating curved-path walking and obstacle avoidance illustrate the disease-related altered basal ganglia to cortical communication impact on the timing and coordination, and adaptability of walking necessary for these walking tasks [230, 38].

While there are therapeutic options such as exercise and physical therapy, exercise therapies only make a small dent in reducing gait difficulties (e.g., less than 0.07 m/s improvements). Motor learning physical therapies make a slightly more significant dent in reducing these gait difficulties (e.g., an increase in walking speed and gait efficiency).

1.2.6 Evaluation of Gait Difficulties

In order to improve the motor skill of walking, there is a scientific imperative to quantify the loss of the motor skill of walking through gait measures. Gait measures, such as gait speed, step length, and step temporal variability [57, 243], are used to characterize specific aspects of motor skill; however, these measures are somewhat limited. For example, some older adults may walk slowly with adapted optimal motor skill, while others may walk slowly with poor motor skill. Older adults with or without diagnosed disease may walk at clinically normal speeds with altered control [129, 321]. Other measures of walking that are a better match to specific aspects of motor skill may prove to be useful when evaluating the gait of older adults.

The evaluation of the motor skill of walking considers multiple environmental factors.

Evaluating walking in the clinic, while useful, is limited and may not capture the multiple dimensions of skills in everyday mobility. The recent emergence of wearable technology can capture numerous gait characteristics in a variety of settings (e.g., clinical facilities, community settings, and in the home). Indeed, the amount of physical activity and human movement data that can be collected from wearables is virtually unlimited; however, much of the data are not analyzed or used in a meaningful manner [182]. One way of making better use of this new source of data is to develop metrics that match the motor skills of interest in older adults. This endeavor will require a collaborative effort between researchers in geriatrics of mobility and experts in engineering and data analytics.

1.3 Directions and Goals

One wearable technology that has gained prominence and has great potential to match with motor skills of gait is accelerometry. Acceleration gait measures (AGMs), which are derived or calculated from the raw acceleration values acquired with accelerometer wearables, capture the motion of body segments. Researchers have proposed that AGMs, particularly those derived from accelerations in the lower trunk, can be global indicators of the motor skill of walking [164, 288, 167, 133, 228]. AGMs are not only widely used [111] but can be proxies for center-of-mass dynamics [253, 336].

It is crucial to investigate motor skill in walking concerning aging and illness. AGMs can help with the clinical measurement of the motor skill of walking. Trunk acceleration measurements have been used in the evaluation of normal aging [216], Parkinson's disease [210], the impact of Alzheimer's disease [207], and numerous other impacts on gait and balance [164, 124]. Previous studies found that older adults adopt more conservative gait patterns than younger adults, potentially to compensate for degeneration in physiological systems such as those associated with vision, sensation, and lower limb strength [216, 169]. These conservative gait patterns result in reduced walking velocity and accelerations, accompanied by reduced step length and increased step width [216].

Since AGMs are more robust than simple gait measures and are associated with the

motor skill of walking, the goals of this dissertation are to fortify AGM use in future gait literature and to demonstrate how AGMs can descry longitudinal changes in mobility in older adults.

1.4 Dissertation Scope

The goal of this dissertation is to establish and evolve the understanding and the use of feature selection methods and machine learning algorithms using accelerometer data. The limitation of machine learning is the quality of the input data and whether or not the right feature is selected. Since gait patterns differ from human to human due to disease, age, and other bio-demographic factors, this research will use machine learning to investigate how well AGMs as features aid in predicting different modes of walking in older adults.

1.5 Main Contributions

This research will elucidate the methods by which raw accelerometer signal data gathered from various points on the human body can be processed to produce simple and acceleration gait measures. Further, this research will demonstrate the usage of these acceleration gait measures in measuring aspects of the motor skill of walking, using statistical analysis and machine learning modeling.

- We will investigate whether we can perform symbolic regression to produce non-linear symbolic models of human locomotion using accelerometer data.
- We will predict the spatiotemporal location of the foot by determining the step length and width based on the acceleration signals.
- From a large dataset of older adults, aged 65-91 years, we will perform feature extraction and selection methods to determine the most influential AGMs for predicting gait speed.
- The AGMs mentioned above will be used to classify fall-risk in older adults.

- From a large dataset of older adults in the Emergency Department, we will use deep learning to detect and classify past and future falls, respectively.

1.6 Dissertation Organization

Chapter 2 will define the terms and methods to understand aspects of the motor skill of walking, accelerometry, data acquisition, acceleration gait measures, signal processing methods, and machine learning. Chapter 3 will describe the datasets used in this research and address areas of study where there are open questions in the gait literature from both the clinical and engineering domains. These open areas are:

- Is it possible to model the human body as a network using accelerometers?
- Can we predict the spatiotemporal location of the foot by determining the step length and width based on the acceleration signals?
- Is it possible to model and classify fall-risk in a large dataset of older adults walking, using accelerometers?
- Is there a concordance (and validation) of the smoothness of walking measures in the literature?
- Using accelerometers, is it also possible to model and classify falls in older adults who present themselves to the Emergency Department?

Chapters 4-8 will be proposed papers that address the open questions highlighted in Chapter 3. Finally, Chapter 9 will detail future work and the proposed research for further study.

2.0 Background

Various parts of Chapter 2 are part of a peer-reviewed review paper [78]. Co-authors of this review paper include: Jessie Vanswearingen, Alan Godfrey, Mark Redfern, Manuel Montero-Odasso, and Ervin Sejdić.

2.1 Acceleration Gait Measures (AGMs)

2.1.1 Categories of AGMs

In this dissertation, AGMs are grouped by methodologies and metrics: 1) gait cycle event timings, 2) statistical features, 3) signal-frequency features, 4) time-frequency features, and 5) information-theoretic features. In Table 2.1, we define each of the categories and compare/contrast the differences between them. For the following attributes, we compare the strengths and weaknesses across AGM categories: 1) "Ease of calculation" refers to the difficulty of calculation of the AGMs, 2) "Directly applicable to clinical problems" refers to how contextually relevant the AGMs are without further explanation or back-calculation, 3) "Popular across literature" is how prevalent these set of AGMs are, 4) "Reduce complexity and dimensionality" is the extent to which AGMs capture a wide amount of information, and 5) "Tied to multiple aspects of walking" refers to how well the AGMs relate to walking elements (Table 2.1).

Table 2.1: Qualitative attributes of the different categories of acceleration gait measures.

Attributes	**Categories of AGMs**				
	Gait Cycle Event Timings	Statistical Features	Signal Frequency Features	Time-Frequency Features	Information-Theoretic Features
Ease of Calculation	-	●	-	-	-
Directly Applicable to Clinical Problems	●	●	-	-	-
Popular Across Literature	●	●	●	-	-
Reduce Complexity and Dimensionality	-	●	●	●	●
Tied to Multiple Aspects of Walking	●	●	●	●	●

● = strength; - = weakness

2.1.1.1 Gait Cycle Event Timings The coordinated trajectories of each leg define the gait cycle into the swing and stance phases of each leg during single support and double support [270, 80, 334]. Specific events of particular interest are heel contact, foot flat, heel off, mid-swing, and toe-off (see [227] for details on gait cycle parameterization) (Figure 2.1). Using AGMs to measure gait cycle characteristics often requires knowing these events and how often they occur (i.e., the number of strides).

Combining the knowledge of the gait cycle with simple measures, such as gait speed and speed variability [352], we can estimate the times of each of these events and phases (e.g., lengths and times of stance, swing, step, and strides [131, 140, 222, 291, 294, 295, 293, 350, 351, 353, 247, 292]). Cadence [222, 291, 295, 344, 353, 162, 292, 16, 90, 201], stride frequency [19, 109, 191, 199], step or stride irregularity [76], stride or step amplitude [209], and stride frequency variability [287, 265] can be calculated with AGMs. Other articles may provide some good insight into how gait cycle AGMs are measured and used [145, 312, 345, 346, 347].

2.1.1.2 Statistical Features In the majority of studies, statistical summaries are performed on different gait cycle metrics over a time period [59]. These include:

- The mean trends of the acceleration signal amplitudes in the mediolateral (ML), vertical (V), and anterior-posterior (AP) directions from different gait phases and activities can be used to characterize a feature of the specific gait phase or activity [117, 79].
- Correlation and covariance of the acceleration signal amplitudes between the pairs of ML, V, and AP directions (ML-V, ML-AP, V-AP) can help elucidate differences among activities that involve translation in just one dimension [79]. For example, in Dasgupta et al. and Sejdić et al., correlations and covariances were calculated as a basic statistical features [288, 79]. Additionally, the auto-covariance can also be used to estimate stride time and stride regularity [265, 227]. Cross-correlations are a measure of similarity between signals [115]. The autocorrelation coefficient of the acceleration signal is the cross-correlations of a signal with itself [227, 316].
- The coefficient of variations of gait parameters, like gait speed and stride time, can help us see the fluctuations between strides, steps, or other gait phases [126].

- Skewness and kurtosis of the acceleration signal amplitude in the ML, V, and AP directions respectively describe the lack of symmetry and whether the amplitude signals are peaked or flat relative to a normal distribution[219, 288, 103].
- The maximum and minimum acceleration signal amplitude in the ML, V, and AP directions can be used to estimate stride time variability [265].
- Mean, standard deviations, median, percentiles, correlation, covariance, coefficient of variation, skewness, and kurtosis of variables and gait events/phases, which are not raw acceleration signals, can also be done to reduce the complexity of those features [265].
- The root mean square of the acceleration signal amplitudes can measure the magnitude of the acceleration signal and can be correlated with walking speed [224, 223, 132, 290] and can be used to estimate the variability of the signal [215, 15].

2.1.1.3 Signal-Frequency Features Signal-frequency features are those acquired by the frequency spectra of the acceleration signals. Signal-frequency features include:

- The peak frequency is the frequency when the maximum spectral power takes place, and it can be an indicator of gait cycle events [288, 289, 219, 304, 239]. The spectral centroid, like peak frequency, is another marker of spectral change in signal [288, 289, 219].
- The bandwidth of the signal expresses the spectral spread of the signal and is often used to differentiate between walking tasks [288, 219].
- The index of harmonicity is the spectral power of the basic harmonic divided by the sum of the power of the first six harmonics [161, 160]. Harmonic ratios can evaluate the harmonic balance and content for a stride, step, or gait phase [89, 288]. The total harmonic distortion of the signal is represented by how much distortion the signal-frequency has from other factors [17]. All these measures are known to assess motor skill of gait [252].

2.1.1.4 Time-Frequency Features Time-frequency features are features from the information from signal and time dimensions, using time-frequency functions [285], such as short-time Fourier transform and wavelet transformations. While some of the time-frequency features in this section may fit into the other AGM categories, they are specifically grouped

here by the way they are extracted from the acceleration signals. Time-frequency features include:

- Estimation of initial contact/final contact of the foot can be done with wavelet transformations, bi-orthogonal spine wavelet, and jerk cost functions [112, 56].
- Wavelet transforms are used to determine the relative energy from each time-frequency band [288]. Overall, wavelet transforms of acceleration signal data are useful in segmenting gait cycle events, gait surveillance and monitoring, and many advanced statistical calculations such as principal components analysis and independent components analysis are used in conjunction with wavelet transformation [242].
- Wavelet entropy measures the "degree of time-frequency based order-disorder of the acceleration signal," and it is computed by wavelet decomposition [288].

2.1.1.5 Information-Theoretic Features Information-theoretic features measure the amount of variability and uncertainty in the information context of a signal [311, 288]. Many of these features can be measured for each direction or a gait event (i.e., a stride). Information-theoretic features include:

- Entropy can measure the uncertainty of a specific variable, which means that the entropy rate can be used to measure the regularity of the acceleration signal [288]. With regards to gait, it can be used the assess the predictability of acceleration signals (from one step/stride to the next step/stride) [5].
- Cross-entropy rate measures the entropy rate between data points from two different signals [288]. Cross-entropy between acceleration signals were used to differentiate walkers with Parkinson's disease and without Parkinson's disease [18].
- Lempel–Ziv complexity can be used to assess discrete time signals, in terms of their dynamics, complexity, bandwidth, and variability [2]. Lempel-Ziv complexity has been used as a measure of complexity in the non-linear analysis of gait patterns in those with Parkinson's disease [95].
- Fractal dynamics can capture the non-linear aspects of gait, typically through analysis like the Detrended Fluctuation Analysis [306, 128]. Hausdorff et al. studied stride-to-

stride changes using fractal dynamics [127].

- Lyapunov exponents can measure disruptions or perturbations in signals, which can expose new insights about gait variability [50, 306, 338, 319].
- Local dynamic stability (LDS) is measured by evaluating the maximal finite time Lyapunov exponents [50, 86]. LDS can measure the gait stability and quality [265] .

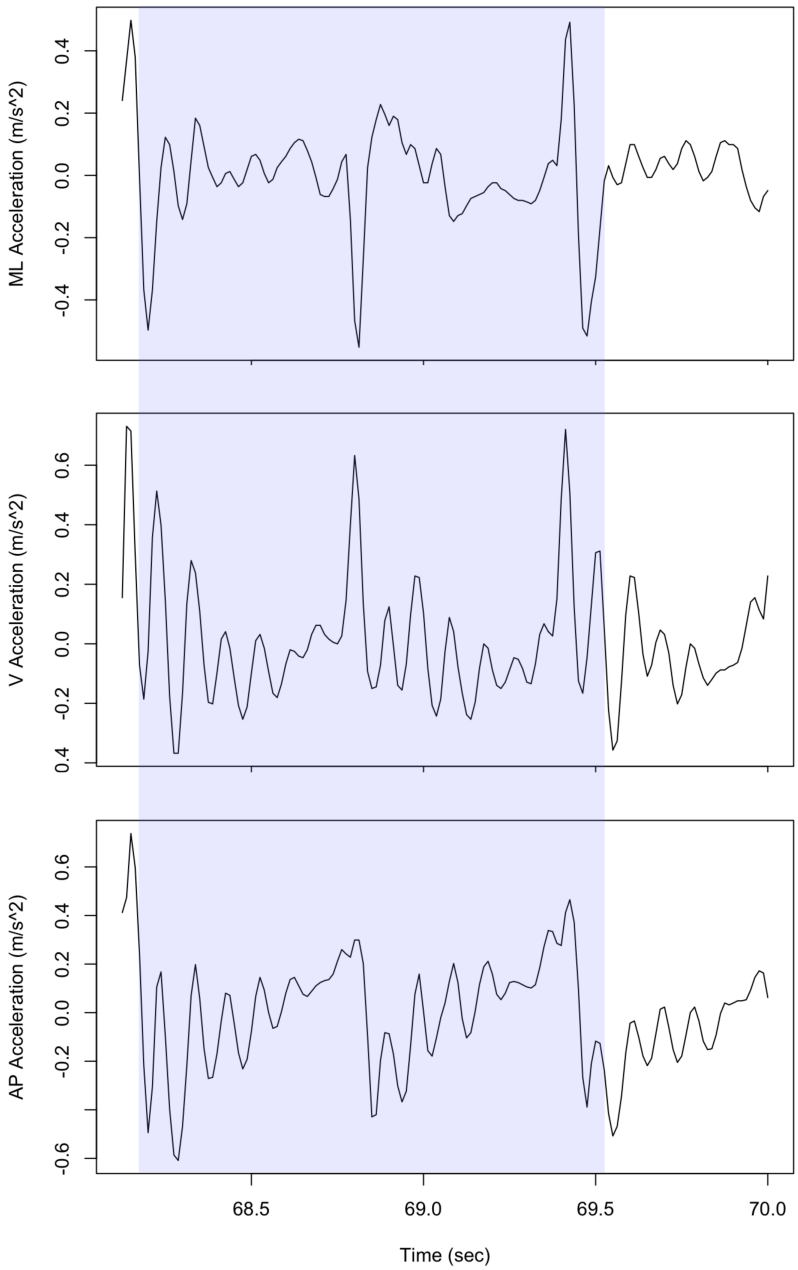
2.2 Data Collection

Accelerometers are used to study age- and illness-related changes in walking [72]. Accelerometers measure the accelerations of objects in motion along three orthogonal axes, often generally aligned with anatomical coordinates (e.g., ML, superior-inferior or V, and AP [343]; these accelerations are time-series, and an example is shown in Figure 2.1. Inertial measurement units (IMUs) and accelerometers are preferred over other wearables because the acceleration measurements can be used to validate the velocity of walking, distance walked, and the intensity of movement [12, 343]. Accelerometers can also predict the rotation and orientation of objects. These characteristics allow accelerometers to determine body postures [343].

Accelerometers are often located at the level of the L3-L5 vertebrae [44], and they estimate the accelerations of the center of mass of the body and can measure aspects of motor skill, such as smoothness and symmetry. From a clinical perspective, lower-trunk placement succeeds because the trunk segment covers over half the mass of the body, and so is prioritized by the nervous system [167].

Data collection also involves handling technical issues [217], such as designating the sampling rates used, frequency response requirements for different tasks, placement, and alignment of the accelerometer on the trunk [124], and how the sensors are attached for long-term and short-term use.

Figure 2.1: Example of acceleration signals (mediolateral, vertical, and anterior-posterior) from an accelerometer placed on the lower back. A full gait cycle of the right foot (starting from a heel strike) is shaded (data and gait extraction done by [79]).



2.3 Signal Pre-Processing

To derive AGMs, there are several pre-processing steps that can be used to prepare the signal data [219, 317]. These techniques involve formatting, cleaning, removing errant/redundant/missing data, resampling, and normalizing the signal data. Some specific tasks include filtering or extracting noise from the signals [113, 335, 41], event detection and labeling [305, 244, 287, 265, 227], wavelet analysis and decomposition [328, 58, 270], integration [317, 11, 42], and tilt correction [219].

There are many sources of noise in a signal: human-made, artifacts of the measurement procedure, and random noise. Often noise is erratic and incalculable. Filtering signals refer to the reduction of noise and other artifacts through smoothing filters (e.g., low-pass filter, Kalman filter, and median filter) [113, 335, 41].

Raw signals are often resampled (or normalized) to ensure that the time intervals between the acceleration sample timepoints are the same [322, 236]. These signals can either be upsampled or downsampled through linear interpolation, cubic interpolation, or spline interpolation [236]. Information on tilt, or the angle of the wearable on the individual, can help researchers get closer to the actual signal. Tilt correction is done by measuring the rotation matrix component of the signal, relative to the gravitational field. Another way to process the signal is to perform integration of acceleration values concerning time to calculate velocities and positional displacement.

One of the ways that signal processing can be approached is through finding cycles of gait events in the raw acceleration; these gait events are then manually and algorithmically labeled. The most typical components of the gait cycles that are captured are stance time, swing time, step time, and stride time. Signal segmentation is the temporal or event-based segmentation of the signal data.

Data points are often removed to eliminate any start-up, pausing, and ending effects that were caused by the subject becoming familiar with the exercise. Outliers based on average stride time are removed using the interquartile range (IQR) rule [314]. An Analysis of variance (ANOVA) test (null hypothesis: no statistical difference in means of stride features between strides) can be done before and after errant data removal.

2.3.1 Stride Extraction

Acceleration signal based stride extraction is a clinically useful tool to evaluate cognitive changes while walking concerning the gait cycle [46]. A gait cycle is defined to start when one foot's heel strikes the floor and ends when the same foot's heel next strikes the floor. For both the right and left foot, successive heel strikes and toe-offs on the same side have been used to define stride, stance, and swing intervals [105]. Heel strike and toe-off events were extracted from the lower back sensor using the algorithm outlined in Sejdić et al. [287]. Local minima points in the V direction correspond to toe-offs, and local minima points in the AP direction were related to heel strikes [287]. The order of which foot initially made the first step is found by taking the average of the first 10 milliseconds of acceleration in the ML direction [287]. If there is a positive mean value, the right foot came first; otherwise, the left foot came first [287].

2.3.2 Window Extraction

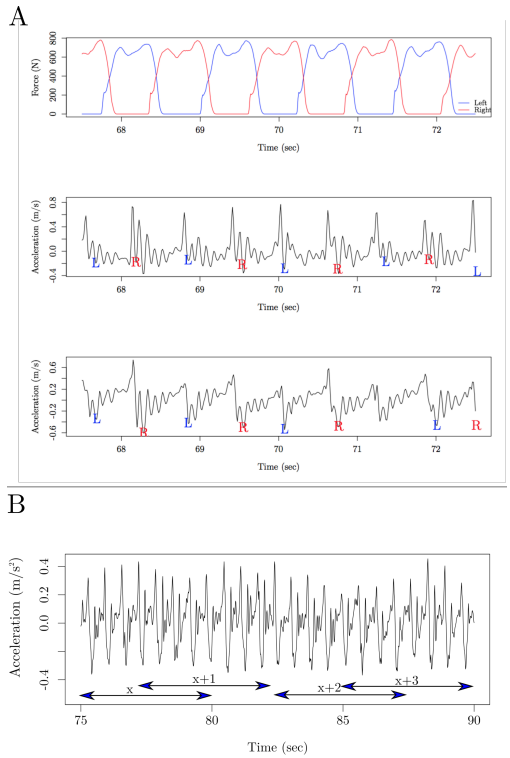
A simpler alternative to stride extraction is the creation of sliding observation windows, by partitioning sensor signals into smaller time segments [259]. Sliding observation windows are commonly used in the activity monitoring and machine learning analysis of accelerometer data [259]. For each of the subject's four raw datasets, one additional dataset was formed: acceleration signals split up in 5-second intervals with a fifty percent overlap.

2.4 Machine Learning for Accelerometer Signal Analysis

Mapping AGMs to the aspects of motor skill walking can aid in differentiating gait-related outcomes through machine or statistical learning. The overall goal of machine learning is to predict or classify unknown outcomes based on a set of past exposures/features. Often, these predictions or classifications are made based on probability measures. Machine learning algorithms classify these outcomes via a learning process in which the algorithm models on a training dataset and then apply the model mentioned above on a testing dataset.

In machine learning, there are two tasks: supervised learning and unsupervised learning. In the field of motor skill research, the goal of supervised learning is to learn a function from labeled data and approximate the relationship between the observable exposure and outcome variables in the data; in unsupervised learning, walking tasks, other gait-related, or motor decline outcomes are not labeled, and the goal is to deduce the relationships within the data.

Figure 2.2: **A.** Stride extraction sample: subject 1's toe-offs and heel-strikes, which were determined from foot pressure recordings during treadmill walking and acceleration signal recorded using the lower back sensor. The top graph shows the pressure readings from the treadmill, the middle graph are the V acceleration readings, and the bottom graph are the AP acceleration readings. Red and blue colored lines and labels depict the right and left foot respectively. **B.** Acceleration signal extraction during walking: subject 3's AP acceleration signals from the back sensor. 5 second observation windows with 50% overlap are shown.



Among the paradigms of classifiers for recognizing gait-related outcomes, regression, Naïve Bayes, support vector machines, decision trees, k-nearest neighbors, Hidden Markov Models, neural networks, and deep learning are the most popular. Typically, the pipeline for machine learning with acceleration signals follows the following steps: 1) signal acquisition (Section 2.2) and signal pre-processing (Section 2.3), 2) derive AGMs through feature extraction and selection, 3) label the outcomes (if performing supervised learning), 4) use a single classifier or a combination of classifiers to test and validate the model, and 5) applying models to test data to predict probabilities of class assignments. This dissertation will primarily consider supervised learning methods.

However, with the use of machine learning and AGMs, it can be challenging to determine which selected features (AGMs) are less significant than others. Mechanistically, there are feature selection methods, such as forward or backward or recursive methods. However, it can be more clinically useful to pick out relevant AGMs that fit the context of the clinical problem.

This section is not a comprehensive overview of machine learning, but rather, it is a summary of the most well-developed and well-used algorithms and techniques used in machine learning for accelerometer signal analysis. The section is organized in the following sub-sections: feature extraction methodology, feature selection methodology and reducing dimensionality, a brief description of various categories of supervised learning models, between-subject modeling, within-subject modeling, leave-one-subject-out modeling, and validation techniques.

2.4.1 Feature Extraction

A feature is a characteristic of the events being observed in the sample space. The feature matrix is the collection of all the features extracted from the set of exposures and information of each of the samples or observations. The feature matrix typically contains rows and columns, which correspond to the observations and features, respectively. These features include AGMs as well as other bio-demographical data. In machine learning, the aim is to have non-redundant features that can directly aid in the learning of the model. The

following sub-sections include statistical calculations and transforming methods (Laplace, Fast Fourier, and Discrete Wavelet Transforms).

2.4.1.1 Statistical Calculations These statistical calculations include performing the calculations laid out in Section 2.1.1.2. Many of these statistical calculations include basic summary statistics and measures of statistical spread (e.g., standard deviation, root-mean-square, and cross-correlation).

2.4.1.2 Fast Fourier Transform The Fourier series is the expansion of a periodical function, which is the additive combination of sine and cosine waves [302]. When a particular function has a period that is very magnanimous and has a discrete set of frequencies, a Fourier transform can be used on that function to produce a continuous spectrum. The Fourier transform (and similar transformations) are used to fit simpler, algebraic models (such as sinusoidal waves) to processes that would otherwise need to be modeled with complex calculus or differential equations. The Fourier transform is ideal for assessing continuous, steady signals.

The discrete Fourier transform (DFT) is used for converting a discrete function from the time domain with equally spaced samples to the frequency domain [302]. The fast Fourier transform (FFT) is a technique to compute the DFT [302].

2.4.1.3 Laplace Transform Related to the Fourier transform, the Laplace transform is used to determine changes in the signal - particularly those moments when the stepping changes during walking [302]. It is beneficial with the use of discrete-time signals.

2.4.1.4 z-Transform The z-transform is very similar to the Laplace transform, except that it is used for discrete temporal datasets, instead of continuous ones.

2.4.1.5 Discrete Wavelet Transform Wavelets can be used, in addition to transforms, to decompose signals into waves - these waves can be assessed for their symmetry or regu-

larity. This is just another way of reducing the complexity of the signal and enable AGM or feature extraction.

The Fourier transform cannot calculate when the frequency components of a signal occur even though it finds the frequency components. The discrete wavelet transform (DWT) can capture both the frequency and when the frequency occurs. Wavelet transforms are made up of a wavelet basis function and the integral of the product of the two functions that make up the signal; some conventional wavelets are the Haar wavelet, Meye wavelet, and the Morlet wavelet.

2.4.2 Feature Selection and Reducing Dimensionality

High dimensionality may lead to over-fitting in machine learning analysis, and it is advantageous to reduce the number of features. Feature selection is performed to produce a final feature set that can be inputted in the machine learning model; this feature set is the subset of the entire feature matrix that was previously extracted. Some of the reasons for conducting feature selection are: allowing for a more straightforward interpretation of the models, reducing model training time, reducing overfitting, increasing model generalization, and reducing dimensionality. There are many ways to perform feature selection, such as principal components analysis, independent components analysis, linear discriminant analysis [23], generalized discriminant analysis, correlation feature selection, and feature selection methods within learning algorithms. Examples of some of the latter methods are: a correlation matrix between each feature is constructed, and highly correlated features (r ¿ 0.75) are eliminated; logistic regression models can run step-wise variable selection via a likelihood ratio test, which selects the model with the lowest Akaike information criterion (AIC) value [6].

2.4.2.1 Principal Components Analysis and Independent Components Analysis

Principal components analysis (PCA) is typically used to reduce dimensions in a dataset without compromising on the information, by maximizing the variance and identifying orthogonal linear combinations with the most substantial covariance. Assumptions of PCA

include having multiple continuous variables, and there is a linear relationship between the variables. This assumption is essential because PCA is based on Pearson correlation coefficients.

We assume that we may not need every predictor in the dataset and that a weighted combination of predictors may be more beneficial. PCA does this by "projecting the entire dataset onto a different feature (sub)space" in the hopes that the new subspace can filter out the potential noise in the original dataset and reveal hidden dynamics. Accurately, principal components (PCs) are sequentially captured the maximum variability among the columns of the dataset, thus guaranteeing minimal information loss; principal components are also uncorrelated. However, the downside to PCA is that each PC is a linear combination of all p variables, and the loadings are typically nonzero, which means there could be other types of interactions between variables.

The way that the PCA is performed is as follows: first, the data is standardized, the eigenvectors and eigenvalues are computed from the covariance matrix or the correlation matrix, the eigenvalues are sorted, eigenvectors are chosen, the projection matrix is created from the chosen eigenvectors, and the original dataset is transformed through the projection matrix.

Independent Component Analysis (ICA) finds the independent components in the data. With PCA, PCA attempts to compress the data to reduce dimensionality globally; ICA aims to separate the data into independent elements. This means that we do not have the same amount of independent components (ICs) as we have dimensions (which we see in PCA). It is not necessary for independent components (ICs) to be orthogonal, and they are not ranked or ordered (as is with PCs). The purpose of ICA is to separate linearly mixed sources, by maximizing non-Gaussian elements in the dataset. This means that normally distributed sources will be hard to be separated. Non-independent sources can still have ICA performed on them because ICA will still try to find a subspace where they are maximally independent.

2.4.3 Supervised Learning Models

2.4.3.1 Linear Regression Algorithms Linear regression algorithms include linear regression and logistic regression. Regression often predicts and computes the outcome from a set of covariates or features. Linear models assume and ascertain a linear relationship between these covariates. Logistic regression is often commonly used as a simple machine learning classifier because it is often used in a predictive analysis where there is a binary dependent variable. The logistic regression model gives the relationship between this binary dependent variable and the predictor variables [339].

2.4.3.2 Non-Linear Regression Algorithms Many biological systems are not linear, but many of our tools rely on assumptions of linearity because of the ease of calculation. Nonlinear techniques can help give meaning to randomness in the data because it can fit multiple polynomial or algebraic equations to the unknowns in the data.

Symbolic regression is similar to other regression techniques, in that it searches for parameters for a mathematical model to fit data; however, unlike linear regression, symbolic regression also frames the structure and operators within the model. In other words, symbolic regression is capable of generating minimally constrained nonlinear models, while linear regression can only find a linear combination of parameters to scale regressors. Symbolic regression is not restricted to a linear format.

2.4.3.3 Classification and Regression Methods One of the popular classification and regression methods is known as a support vector machine (SVM). SVMs are kernel-based, where an optimal hyperplane is found for linearly (or radially or quadratically, depending on the kernel function used) separable patterns in the dataset. SVMs aim to maximize the geometric margin on the training set and minimizes the training error.

2.4.3.4 Evolutionary Algorithms Evolutionary algorithms (EA) mimic how biological evolution works; one such EA is genetic programming (GP). Although all EAs are slightly different from one another, they typically follow the same structure: 1) the initialization

phase, 2) the fitness evaluation phase, 3) the selection of candidate solutions phase, and 4) the application of the genetic operators phase. The initialization phase creates a population of chromosomes/candidate solutions. The fitness evaluation phase, which is often the most computationally expensive part of the EA, determines how effective each of the candidate solutions are at solving the given problem; the effectiveness of the candidate solution is known as the fitness of the given candidate solution. In this study, symbolic regression is performed. Thus the fitness value is the mean squared error of the regressed model to the actual recorded signal (the typical fitness measure when performing symbolic regression). The ultimate goal of the evolutionary search is to find a candidate solution with a high-quality fitness value.

The selection and application of genetic operators phases work together. Once every candidate solution's fitness is calculated, a tournament selection is performed by 1) choosing two random candidate solutions, 2) selecting the better candidate solution, and 3) repeating step 1-2 with another two random candidate solutions. This tournament selection then produces two reasonably good candidate solutions to be propagated into the next generation. Once two reasonably satisfactory candidate solutions are selected, there is a chance the genetic operators will be applied to them before they enter the next generation's population. There are two types of genetic operators: crossover and mutation. The idea behind crossover is to take the two selected candidate solutions and mate them and place their offspring into the next generation. The motivation is that mating two reasonably effective candidate solutions may result in even better candidate solutions. Mutation acts on a single candidate solution at a time and will result in some random change that may add diversity to the population and improve results over time. This selection and application of genetic operators repeat until the next generation's population is full. Once the new population is full, the algorithm returns to the fitness evaluation phase, and the whole process continues to repeat until some stopping criteria are met. These criteria may be after some number of generations have occurred, or until some fitness threshold is met. Typically, the evolutionary algorithm will run for a high number of generations as it will take some time before the algorithm can converge on a high-quality solution.

The GP implementation used for this dissertation is heavily inspired by Schmidt et al.'s work [281]. The major strengths of this system are in the use of subpopulations, an acyclic

graph representation, and the use of fitness predictors. Unlike traditional GP systems where the chromosomes are tree-based S-expressions, this study's application uses an array-based acyclic graph representation of the underlying mathematical expressions. The benefits of such representation include faster runtimes, easy reuse of subexpressions, and the reduction of bloat (a GP phenomenon where the trees grow arbitrarily complex with no meaningful improvement in the quality of the effectiveness).

2.4.3.5 Other Algorithms Other algorithm categories are regularization algorithms, ensemble methods, Bayesian algorithms, clustering algorithms, and deep learning algorithms. Three common regularization algorithms are elastic net, decision tree, and classification and regression trees. Decision trees are often used in the clinical space; these trees have a node/root, leaves, and internal nodes, where the internal nodes utilize particular features to split the instance space into two or more sub-spaces, and the leaves refer to the outcome classes. Decision trees can be ensembled to form the random forest algorithm. In the random forest algorithm, the nodes are split using the best split among a random subset of features; generally, random forests are very robust and provide high accuracies in classification. Other ensemble algorithms include AdaBoost, other boosting approaches, and other bagging approaches. Bayesian algorithms include Naive Bayes and Bayesian networks. Clustering algorithms include k-means and artificial neural networks (ANN) (e.g., perceptron, back-propagation). ANNs can lead to deep learning, which refers to a whole set of neural networks, specifically convolutional neural networks (CNN). CNNs have a large number of parameters, often with a magnanimous amount of observations.

2.4.4 Between-Subjects and Within-Subject Modeling

For the within-subjects model, each subject's dataset is appended to each other. For the between-subjects model, all subjects' datasets were appended.

2.4.5 Leave One Subject Out Modeling

A leave one subject out cross-validation can be done to capture how well other subjects' features and data can predict an individual's gait-related outcome. This was done by combining all subjects' datasets. Each model is trained on n-1 out of the n subjects, and the model is tested on the remaining subject.

2.4.6 Validation

2.4.6.1 Metrics For a binary outcome, metrics from a confusion matrix is typically used to evaluate the performance of a classifier or algorithm 2.2. As shown in [123], the confusion matrix contrasts the results from the prediction with the true classification labels; these result in the following sub-samples: true positives, true negatives, false positives, and false negatives. Many metrics can be derived from the confusion matrix such as accuracy, true positive rate (TPR) or sensitivity, true negative rate (TNR) or specificity, false-positive rate (FPR), positive predictive value (PPV) or precision, negative predictive value (NPV), and area under the curve (AUC). The formulas for these metrics are shown below:

Table 2.2: Example of a confusion matrix, which gives the performance of a classifier algorithm.

		Actual	
		Positive	**Negative**
Predicted	**Positive**	True Positive	False Positive
	Negative	False Negative	True Negative

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.1}$$

$$TPR = Sensitivity = \frac{TP}{TP + FN} \tag{2.2}$$

$$TNR = Specificity = \frac{TN}{TN + FP} \tag{2.3}$$

$$FPR = \frac{FP}{TN + FP} \tag{2.4}$$

$$PPV = Precision = \frac{TP}{TP + FP} \tag{2.5}$$

$$NPV = \frac{TN}{TN + FN} \tag{2.6}$$

The AUC refers explicitly to the area under the receiver operating characteristic (ROC) curve; the ROC curve is formed by plotting the TPR against the FPR at different thresholds. The AUC is the density of the area underneath the ROC curve. The AUC can help compare the performance of classifiers; a "good" classifier has an AUC of 0.9-1.0 while a "poor" classifier has an AUC of 0.5-0.6.

For binary or multiple outcomes, accuracy is a widely used measure of the performance of a classifier. Accuracy is defined as the rate of correct predictions (Equation 2.1).

2.4.6.2 Cross-Validation and Holdout Validation Overall, validation techniques are used to evaluate the generalization of a trained classifier on a test set. During the learning or training process of the classifier, the accuracy is often estimated from a portion of the dataset that is left out of the training.

Conventional validation techniques are k-fold cross-validation and holdout validation. For the k-fold cross-validation, the dataset is randomly partitioned into k subsets of equal size. The model is trained for k iterations, where each iteration uses one k subset for testing and the remainder k-1 subsets for training. The final accuracy is the average of these k iterations. For the holdout validation technique, the dataset is randomly partitioned into a training set and a test set (often, this partition is the Pareto split of 80%/20%, respectively). The model is trained on the training set, and then the accuracy is evaluated on the testing set. Multiple runs of the holdout validation technique may be run with different random training-testing splits.

2.5 AGMs in Action

2.5.1 Motor skill and AGMs

Understanding the use of AGMs as proxies for the aspects of the motor skill of walking will provide better clinical features for models that can potentially predict the motor skill of walking. Clinically, mapping motor skill characteristics (Section 1.2.4) to categories of AGMs (Table 2.1) may be capable of providing relevant and accurate measurements. In Table 2.3, we summarized a selection of references for each of the aspects of motor skill–AGM mappings. By doing so, we also identify the existing gap by seeing how researchers have combined multiple features extracted from gait accelerometry signals into a derived AGM that could potentially be a marker for walking-related changes in physical function.

Table 2.3: Literature citations that depict the mapping between the seven aspects of the motor skill of walking and acceleration gait measures.

Aspects of Motor Skill	**Categories of AGMs**				
	Gait Cycle Event Timings	Statistical Features	Signal-Frequency Features	Time-Frequency Features	Information-Theoretic Features
Variability	[216, 124, 128, 174, 165, 36]	[288, 228, 265, 227, 191, 201, 174, 159]	[288, 108, 19, 316, 153]	[288, 54]	[288, 153, 54]
Stability	[341, 19, 159]	[265, 159, 191, 201]	[219]	-	[307]
Smoothness	[192]	[290, 293, 21, 88, 342]	[44, 89, 224, 223, 90, 201, 293, 342, 288, 342]	[146, 56]	-
Automaticity	[321, 294, 293, 351, 201, 342, 20, 102, 166, 296]	[321, 296, 191, 201, 20, 176]	[294, 293, 351, 201, 342, 20]	-	[5]
Efficiency	[219, 291, 295, 191, 199]	-	[191, 199, 295, 219, 291]	-	-
Adaptability	[131, 140, 222, 291, 294, 295, 293, 350, 351, 353, 247, 292, 134, 272, 84, 226]	-	[288]	-	-
Symmetry	[133, 271, 140, 344, 19, 109, 191, 174, 21, 20, 156, 301]	[219, 288, 103, 344, 19, 191, 174, 21, 20, 301]	-	-	[2, 174]

2.5.1.1 Variability Typically, gait variability is calculated through simple measures (and by simple methods), such as step/stride length/time [165]. Because accelerometers can collect massive amounts of data over time, they are especially useful in assessing stride-to-stride or step-to-step variability of walking [174]. Some common AGMs describing variability presented are:

- Standard deviation and coefficient of variation of the gait cycle events can directly measure variability [174].
- The median of the modal frequencies for the V, ML, and AP directions and the strength of the relative fluctuations in the phase progression can determine step/stride frequency [265].
- The autocorrelation coefficient of the signal can capture inter-stride variability [228, 174].
- The peak values of the first and second dominant periods of the autocorrelation function, simple statistical features, individual curve estimates, and adaptive peak thresholds can determine step/stride variability [316, 108, 19].
- Root mean square of the acceleration signal can be a measure of variability. For example, Rispens et al. define "movement intensity" as the root mean square of the acceleration [265, 159, 191, 201].
- Entropy, entropy rate, and Lyapunov exponents are indicators of variability [288, 5].

While many gait cycle events are used for variability, step duration is a much better measure than step length when investigating loss of balance in older adults [124, 216, 128, 36]. Statistical summaries of step length, in conjunction with a low root mean square value are often indicative of a typical gait pattern during walking. On the other hand, autocorrelation coefficient of the signal and other signal-frequency features can better pick up characteristics of overall walking patterns. Finally, information-theoretic features can provide some insight into variability if other motor skill aspects are also being investigated [288]; for example, the regularity of a time series can be captured via entropy or entropic features [54].

Some specific examples in the literature have shown that measuring variability via AGMs is helpful to differentiate between classes of older adults. Older adults who have neuromotor difficulties have one or more of the following: lower step/stride variability, lower step/stride frequency, and higher movement intensity in all forms of walking [108, 216]. In [153], the authors use linear (mean velocity, the peak-to-peak amplitude of accelerations, root mean square, and frequency dispersion) and non-linear AGMs (Lyapunov exponent and entropy) instead of simple footfall data, to measure the gait variability in patients with multiple sclerosis. In [213], the authors posit that gait variability AGMs can be part of a clinical screening method for the locomotive syndrome since AGMs provide a complete, accurate,

and personalized measurement of locomotive disorder in older patients with or without the musculoskeletal disease. Gait irregularities and variability can also be measured to create a reference database, investigate outcomes in patients with gait disorders, and study rehabilitation for those with limited knee function [293, 291, 294, 295]. Similarly, other articles directly assess gait variability through trunk AGMs [347, 21, 292, 91].

2.5.1.2 Stability To measure how people maintain gait stability, many papers test a strategy of changing walking speeds or measuring accelerations. However, raw trunk acceleration data could enrich the measure of stability. V accelerations can show the moments when toe-offs and heel strikes occur - decreased moments and low acceleration at heel contact, foot flat, mid-swing, and initial push-off are more prevalent in older adults [341, 19, 159]. Additionally, measures such as root mean square [265, 159, 191, 201], standard deviations, and coefficient of variations of the acceleration signals can provide a better depiction of stability.

Even so, root mean square, standard deviations, and coefficient of variations of the acceleration signals have limitations in estimating the non-linear features of human movements, such as stability. LDS can address those limitations because it is measured with the maximal Lyapunov exponent from chaos theory. LDS is often used to assess gait stability and falling risk [307]. The LDS per stride can be measured by dividing LDS by the stride frequency. A high LDS is indicative of good motor control and dynamically-stable gait. Other non-linear measures of stability are the step stability index and the harmonic ratio. The step stability index is a function of standard deviations of the intrinsic mode functions (which are derived from acceleration signals from the vertical direction).

2.5.1.3 Smoothness Walking smoothness is a high indicator of fall-risk in older adults. The most common way to measure smoothness is through root mean square [290, 293, 21, 88, 342], indices of harmonicity, or harmonic ratios (estimated for each of the three directions as the index of harmonicity) [44, 89, 224, 223]. Larger harmonic ratios can indicate a smoother gait pattern, while a lower ratio is found in older adults and older adults with unsteady gaits [44, 90, 201, 293, 342, 288]. During, most modes of walking, the most significant impact on

the harmonic ratio, due to increased age, is in the ML direction. Another way to measure smoothness is to measure the jerk-cost function [146, 56] from gait movement. Lower jerk indicates higher smoothness in gait and higher motor control [146]. In [176], they used power spectrum entropy of the acceleration signals to differentiate persons likely to fall and persons not likely to fall, by their gait.

2.5.1.4 Automaticity Automaticity often goes hand in hand with variability/regularity [66]. Many of the features that measure inter-step or inter-stride variability in walking can be indicative of automaticity. For instance, the coefficient of variation of stride velocity, coefficient of variations of the axial directions of accelerations, and swing time variability are measures of automaticity [321, 102, 166]. Other useful AGMs include the periodicity of accelerations [20, 201, 294, 293, 342, 351], and measures of efficiency [296]. For example, in patients who freeze or momentarily stop walking, a sign of Parkinson's disease, these measures are particularly useful [333, 87, 249, 22]. Moreover, automaticity becomes an important motor skill to investigate when studying cognitive impairment or load within aging adults [326].

2.5.1.5 Efficiency Efficiency, the inverse of energy expenditure, can also be used to assess gait and evaluate balance in older adults [320, 330]. In [296], they measured energy expenditure along with the center of mass accelerations in all forms of walking to come up with guidelines on how older adults can improve their walking. Another way to measure efficiency is through measuring periodicity, specifically constant acceleration periods and changes [191, 199, 295, 219, 291]. While these AGMs are effective in measuring efficiency, validation methods such as measuring the oxygen rate during walking are often used [320, 196].

2.5.1.6 Adaptability Adaptability is directly tied to stability since people try to increase their stability in the ML direction to maintain an upright posture. Similarly, adaptability can be tied to variability/regularity, since people adapt back into their regular gait pattern when they are perturbed [226]. Thus, measures of stability and variability can apply

to adaptability. Statistical features of gait cycle events and the harmonic ratio can be used to measure gait adaptability [288]. In obstacle avoidance studies [134, 272, 84], gait pattern adaptations were measured via step length variability. Step length variability is measured in the following studies: [131, 140, 222, 291, 294, 295, 293, 350, 351, 353, 247, 292, 292].

2.5.1.7 Symmetry Similar to variability, fractal dynamics [174] and autocorrelation coefficient of the signal [174], the mean, standard deviation, coefficient of variation, and correlation of the gait cycle events [174, 156, 140, 21, 191, 301] are used to determine symmetry. In [19, 20, 344], symmetry was derived from the autocorrelation function of the vertical acceleration signal. There are more metrics of symmetry [271]: step asymmetry [109], symmetry ratio, symmetry index, gait asymmetry, and symmetry angle using step length, swing time, stance time, double support time and an intra-limb ratio of swing: stance time.

Gait symmetry is often a necessary clinical measure when older patients are in recovery from falls, stroke, amputation, osteoarthritis, and hip replacement [133]. In [133], the authors measured trunk movement regularity and stride asymmetry to characterize asymmetry among persons with history of chronic stroke.

2.5.2 Example Uses of Motor Skill–AGM Mapping for Gait-Related Outcomes

Adverse motor skill and slowing down could be caused by peripheral muscle fatigue and can be measured by multiple AGMs. The amount of walking, changes in walking speed, and the amount of physical activity can be derived from velocity, statistical measurements, and variability measurements. Another use for motor skill–AGM mapping is to measure energy expenditure and pelvic displacement, pelvic rotation, pelvic tilt, and stance knee flexion – it has been shown that vertical displacement of the pelvis during walking can be a predictor of oxygen consumption, and in turn, energy expenditure [330].

2.6 Selection and Use of AGMs

The literature is overpopulated with multiple AGMs, and very few researchers can say they measure specific aspects of motor skill. Thus, there are several issues to be addressed to move the field of gait and rehabilitation forward.

Extracting AGMs from raw acceleration values is a natural step in biomedical informatics research. With the increased use of artificial intelligence, feature selection and specification is necessary for scientists to build statistical models in order to make predictions in the context of their problem. Clinical researchers in rehabilitation and physical-activity sciences may find utility and insight from conducting more studies in observational and clinical trials with AGMs in order to further the field.

However, the current selection and use of AGMs in research have limited value because of a lack of gold-standard information from acceleration measurements. Only a few studies have compared various AGMs within the same sample or dataset, let alone in different study designs. Moreover, there is a discrepancy in how AGMs are used between age, sex, gender, and disease groups. Further, previous research is limited to comparing AGMs to common simple gait measurements [202]. Collectively, research has a minimal consensus on the validity of using many of these AGMs.

There is little consensus on the most useful AGMs to use for analyzing locomotion in general, and particularly with an accelerometer located on the lower back. There are very few studies that examine more than one AGM from one dataset [242]. Most of the current single AGMs studies only differentiate generalized populations (e.g., older adults vs. young adults) as opposed to more specific groups (e.g., older adults who are more prone to falling vs. older non-fallers). To improve the accuracy of the AGMs for detection of gait impairment, future research needs to combine multiple AGMs through modeling [242]. Analyzing AGMS collected pre- and post-intervention can examine discriminative ability, responsiveness and construct validity for various AGMs [309, 242].

2.7 Issues in validity and interpretation of AGMs

There are multiple construct validity issues with the use of AGMs, because of the various methods for the derivation of an AGM from gait accelerometry and no known means to compare across the derived AGMs. It is not certain if various AGMs represent the same findings of the motor skill of walking, or if differences in the ability of various AGMs to distinguish the level of physical functioning in daily life.

In the studies that we have identified that investigate the impact of aging and illness on specific walking tasks, older adults adopt more conservative and compensatory gait patterns [169]. Older adults typically have reduced walking velocity and trunk-accelerations accompanied by reduced step length; these reduced accelerations are possibly induced to compensate for degeneration in vision, sensation, and lower-limb strength [216]. Notably, in straight path walking and curved-path walking, older adults have increased sub-movements, deceleration, and hesitancy [146].

Furthermore, few studies have researched how multiple AGMs within the same sample can effectively improve a statistical model. Several investigators report individually defined indexes of the acceleration signal, derived by proprietary algorithm methods [274, 53]. Little replication of AGMs in the same target population exists, including by the same investigator in subsequent studies of a similar sample. As a result, the clinical investigator has little to base an informed decision about the usefulness of derived AGMs to describe, detect and monitor walking abnormalities, and determine intervention outcomes. Therefore, there is an obligation for further study into comparing AGMs in a more standardized way.

AGMs can also be used in combination with other measures to investigate gait measures. For example, Salarian et al. developed a Timed Up and Go test using from five to seven accelerometer sensors; which had good psychometric properties at a pilot study for Parkinson's patients; main features that demonstrated association with the Unified Parkinson's disease rating scale, extracted from instrumented Timed Up and Go are step counting, seconds, peak arm velocity, cadence, stride and turning and among the sub-elements of the instrumented Timed Up and Go test, gait, turning, and turn-to-sit were the most reliable [273].

2.8 Other Limitations

Even though stride extraction is more clinically relevant for each subject, window extraction is often done to compare how machine learning algorithms perform without respect to gait cycle events. Accuracy values between windows and strides are often comparable. Due to the time and memory consuming task of signal processing, window extraction is preferable [329].

The data collection process for producing AGMs has multiple limitations. Even though there may be low generalizability of the gait data acquired by subjects, this can be remedied by future work including a higher sample size and a more extensive set of acceleration signals. Furthermore, statistical signal processing is a tedious task, which involves not only implementing the algorithm outlined in [287], but also manual checks to ensure biological accuracy. While stride extraction was not validated in this dissertation, a consistent number of strides with a small standard deviation was found amongst study participants per trial [79]. This limitation could be further addressed with video or pressure report validation [313].

Cognitive task abandonment and lack of recording cognitive mistakes are other limitations. Some participants could have either abandoned or incorrectly performed the cognitive task for short periods during the trials, and these periods were difficult to identify and remove. For example, these cognitive abandonment may correlate with walking pauses because the participant may pause when they have strayed away from task or they may not correlate with walking pauses, because pauses may represent better focus on the cognitive task.

3.0 Datasets and Aims

This chapter summarizes the three datasets used in this paper and lays out the aims that Chapters 4-8 will address.

3.1 Healthy Subjects Dataset

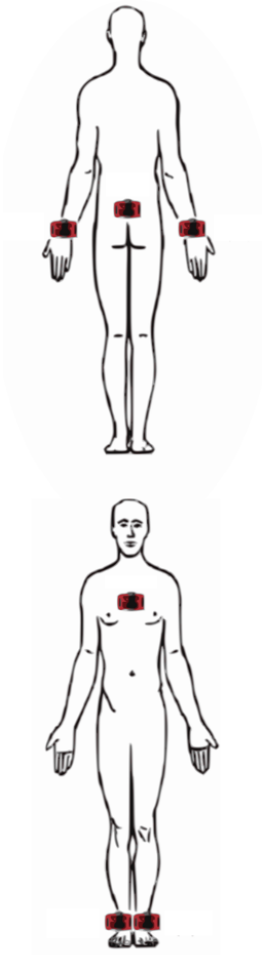
3.1.1 Study Design and Procedures

This data collection is of a repeated-measurements design, where subjects were their own controls. Participation in this data collection consisted of two sessions in which each session had five stages: affixing the sensors (10 to 20 minutes), first walking trial (10 minutes), a rest period (10 to 20 minutes), second walking trial (10 minutes), and sensor detachment (10 minutes). The two sessions of this study for each participant occur with at least 48 hours between each session. Each session lasted for about 75 to 90 minutes.

The walking trials were performed on a treadmill which rested on top of a flat non-compliant surface. The treadmill captured pressure data (in units of Newton force (N)). The treadmill was set at a steady pace of 2.2 mph, which is 0.98m/s which is less than the usual adult walking speed of 1.2-1.3 m/s. This treadmill speed was chosen so that all participants would be at a comfortable, and slower than usual walking speed to be closer to a speed common among community-dwelling older adults and persons with walking difficulties. The first walking trial consisted of walking, while the second walking trial consisted of walking while counting backward from 10,000 in increments of 7, an arithmetic task that is mentally involved. We will refer to the first walking trial as normal walking while referring to the second walking trial as walking under cognitive load.

Participants were affixed with six wGT3X-BT triaxial accelerometer sensors (produced by ActiGraph LLC, Fort Walton Beach, Florida, USA) located on their chest, bilateral ankles, wrists, and lower back (Figure 3.1). These sensors captured linear accelerations (in

Figure 3.1: A wGT3X-BT triaxial accelerometer sensor was placed on each of the following six locations: the participants' chest, bilateral ankles, wrists, and lower back.



units of $\frac{m}{s^2}$) at a frequency of 80 Hz from the x, y, and z directions which correspond to the ML, V, and AP directions. Overall, each sensor relayed 48,000 data points for each walking trial. The sensors were all clinically accepted monitoring devices and presented minimal risk to the subjects. All recorded data did not include any personal identifying information.

3.1.2 Subjects

Volunteers were recruited from the Pittsburgh area in Pennsylvania, USA. Participation in the study was entirely voluntary and could be discontinued at any time. Acceleration and treadmill pressure data were collected from ten healthy volunteers (18-35 years old) at the University of Pittsburgh. Other than age and a subjective perception of healthiness, no other screening criteria were included in the volunteer recruitment. Demographic characteristics for the participants, for the variables of age (mean (sd)), gender (N (%)), height (mean (sd)), weight (mean (sd)), and BMI (mean (sd)) were: 21.40 (4.38), 5 (50%), 1.72 m (0.09), 66.36 kg (8.41), and 22.87 (1.65).

3.2 PRIMA Dataset

This dataset is from the Program to Improve Mobility in Aging (PRIMA), and thus, this dataset will be known as the "PRIMA dataset". PRIMA is supported by The National Institute on Aging (AG045252).

3.2.1 Study Design, Procedures, and Subjects

PRIMA is a five-year "randomized single-blind two arm intervention trial to compare the effects on mobility, activity and participation of a standard strength, endurance, and flexibility program to the standard plus timing and coordination program in 248 community-dwelling older adults walking slower than the desired gait speed of 1.2 m/s." Participants were asked to come into the lab in the University of Pittsburgh for a baseline assessment, a 12-week post-intervention assessment, a 24-week post-intervention assessment, and a 36-week post-intervention assessment. Overall, this study planned to enroll "community-dwelling older adults who walk faster than 0.60 m/s and slower than the desired gait speed of 1.2 m/s. We plan to enroll approximately 124 subjects who walk slowly (i.e. gait speed > 0.60 and < 1.0 m/s) and approximately 124 faster walkers (i.e. gait speed ≥ 1.0 and < 1.20 m/s) for a total of 248 subjects."

Accelerometers were attached to participants on the skin over the L3 segment of the lumbar spine, left ankle, right ankle, left wrist, and right wrist. These trunk accelerations were sampled at 200 Hz. Participants are roughly between 65 – 91 years of age. The participants were surveyed on their fall history and screened for depression, mobility, pain, sleep, and confidence in walking. Blood pressure and heart rate measurements were taken from these participants. Subsequently, they were asked to perform in triplicate a short walk from an orange cone to another orange cone. A six minute walk test was done with the accelerometers. Afterwards a treadmill and strength test were done.

3.3 Elder Falls in Emergency Room Dataset

3.3.1 Study Design and Procedures

Participants were affixed with one triaxial accelerometer sensor on their lower back. This sensor captured linear accelerations (in units of $\frac{m}{s^2}$) at a frequency of 100 Hz from the x, y, and z directions which correspond to the ML, V, and AP directions. Participants were asked to sit and then do 5 tasks consecutively: stand up, walk 10 steps forward, turn around 180 degrees, walk 10 steps back to their original location, and sit down. These tasks are coded 1, 2, 3, 4, and 5 respectively.

3.3.2 Subjects

Volunteers were recruited from the Pittsburgh area in Pennsylvania, USA. Participation in the study was entirely voluntary and could be discontinued at any time. Survey data were collected from 166 volunteers (60-96 years old) who were also patients in the Emergency Room at the University of Pittsburgh. Of these 166 patients, 153 patients had acceleration data collected. A table of the demographic characteristics and responses to physical activity questions are shown in Table 3.1.

Table 3.1: Demographic and survey data from older adult patients in the emergency room.

Variables	**n = 166**
Age (mean (SD))	69.92 (8.46)
Identified Gender = Woman (%)	71 (42.8)
Race (%)	
Black	48 (28.9)
White	116 (69.9)
Other	2 (1.2)
Hispanic or Latino = No (%)	164 (98.8)
Marital Status (%)	
Married	58 (34.9)
Single	74 (44.6)
Separated	2 (1.2)
Widowed	32 (19.3)
Living Arrangement (%)	
Live by Self	63 (38.0)
Live with another, age ¿=65 years	37 (22.3)
Live with another, age ¡65 years	39 (23.5)
Live with multiple family members	27 (16.3)
Current Mobility (%)	
I have no problems walking	118 (71.1)
I have slight problems walking	24 (14.5)
I have moderate problems walking	16 (9.6)
I have severe problems walking	8 (4.8)
Mobility Prior to Injury (%)	
I had no problems walking	129 (77.7)
I had slight problems walking	19 (11.4)
I had moderate problems walking	13 (7.8)
I had severe problems walking	5 (3.0)
Self-care Ability (%)	
I have no problems washing or dressing myself	156 (94.0)
I have slight problems washing or dressing myself	6 (3.6)
I have moderate problems washing or dressing myself	2 (1.2)
I have severe problems washing or dressing myself	2 (1.2)
Routine Activity Ability (%)	
I have no problems doing my usual activities	150 (90.4)
I have slight problems doing my usual activities	7 (4.2)
I have moderate problems doing my usual activities	4 (2.4)
I have severe problems doing my usual activities	5 (3.0)
Pain during Walking (%)	
I have no pain or discomfort	91 (54.8)
I have slight pain or discomfort	28 (16.9)
I have moderate pain or discomfort	23 (13.9)
I have severe pain or discomfort	17 (10.2)
I have extreme pain or discomfort	7 (4.2)
Fall(s) in the Last Year? (%)	
Fallen in the Last Year	72 (43.4)
Have Not Fallen in the Last Year	93 (56.0)
Don't Remember if Fallen in the Last Year	1 (0.6)
Fall Frequency (%)	
Once	41 (56.9)
A few times	28 (38.9)
Monthly	1 (1.4)
Weekly	2 (2.8)
Fall Injuries? = No (%)	42 (58.3)
Use of Cane = No (%)	138 (84.7)
Unsteady during Walking = No (%)	105 (63.3)
Hold Furniture at Home during Walking = No (%)	116 (69.9)
Fall Worry = No (%)	118 (71.1)
Push with Hands to Stand Up = No (%)	102 (61.4)
Trouble Stepping onto Curb = No (%)	122 (73.5)
Feel Rushed to the Toilet = No (%)	80 (48.5)
Loss of Feeling in Feet = No (%)	128 (77.1)
Medication that Makes you feel Light-headed or More Tired than Usual = No (%)	129 (77.7)
Medication that Helps Sleep or Improve Mood = No (%)	120 (72.3)
Drink More than a couple of Alcoholic Drinks in One Day = No (%)	134 (80.7)

3.4 Dissertation Aims

There are various aims of this dissertation.

Chapter 4 will discuss how, while hypothesizing that accelerometer data can model this BSN, we collected accelerometer signals from six body areas from ten healthy participants performing a cognitive task. Chapter 4 will also discuss how machine learning based on genetic programming was used to produce a collection of non-linear symbolic models of human locomotion.

Chapter 5 also uses the healthy subjects dataset. In this chapter, we estimate the step length and step width using raw acceleration signals, using two methods: the acceleration-peak finding method and the inverted pendulum method. These two methods are supplemented by an algorithm (with the aim of improving the spatiotemporal estimates) called the piecewise aggregate approximation method.

Chapter 6 and 7 use the PRIMA dataset. Chapter 6 discusses how we use the PRIMA dataset's acceleration datasets to create AGMs. We calculated around 213 AGMs. These AGMs were used to assess whether we could predict gait speed. The aforementioned AGMs were also used to classify and recognize falls in older adults.

Chapter 7 discusses how concordant other smoothness of walking measures are with a validated measure, the harmonic ratio. A concordance analysis was done, using Bland-Altman plots.

Chapter 8 uses the dataset with older adults in the Emergency Department. Using the collected data, we implemented a hybrid-convolutional recurrent neural network. With this deep learning framework, results may indicate that clinicians do not have to conjecture which feature(s) to include in a machine learning model.

4.0 Is Human Walking a Network Problem?: An Analysis Using Symbolic Regression Models with Genetic Programming

Co-authors include: James Alexander Hughes, Mark Daley, and Ervin Sejdić. This article has been accepted in Computer Methods and Programs in Biomedicine.

4.1 Introduction

The human body is a network of moving parts, and the concept of "network medicine" can investigate how these moving parts interact. Network medicine is the biometric concept of modeling and identifying a person by using their body attributes [168, 349, 348, 27, 163, 26, 233, 62, 29]. This branch of clinical analysis overlaps with the goals of the fields such as "personalized medicine" or "systems biology." Specifically, it refers to the concept of considering the relationship between the unique traits or features of a specific individual to make diagnostic decisions.

In particular, the topics of network medicine and human gait analysis have been investigated by multiple medicine-adjacent fields, such as biomechanics, computer science, and robotics [349, 348]. Using model-based approaches, researchers have modeled the human body's motion through the individual's body structures [168, 29]. Walking requires nodes of the body, such as the arms, legs, chest, and lower back, to move the body through space while preserving stability and balance [321]. While walking, the body's center of gravity vacillates between the right and left sides; the coordination of all these nodes makes walking a complex task, where all the nodes of the body must work in concert [349, 321, 268, 269, 157]. Viewing walking as a network can help create individual-specific gait models based on walking patterns [205].

In walking-related network medicine studies, how walking can be modeled using the current statistical and analytical approaches is the open research question. In conjunction with signal processing algorithms and wearable devices, machine learning algorithms may

provide a way to monitor walking as a network [114]. Gait data collection is often done through non-invasive, inexpensive accelerometers, which can be placed in various locations on the body as a body sensor network (BSN), allow for continuous monitoring, and are used in clinical settings [229, 83, 79]. The resulting datasets often result in large time-series datasets per subject, which has been shown to give favorable results in studies with low sample sizes [79, 73, 327, 32]. While these datasets can sometimes be "noisy," where some of the data are artifacts of the accelerometer, signal processing methodology analysis of these datasets can provide reliable measurements and the creation of clinical gait variables.

Symbolic regression (SR) networks are machine learning models that may be more adaptable and robust than other statistical approaches in modeling walking behavior [151]. SR is similar to other regression techniques. SR searches for parameters for a mathematical model to fit data; however, unlike linear regression, SR also frames the structure and operators within the model. SR has been widely used in medicine and physiology to study various topics, such as transcriptomics, metabolomics, the dynamics of the human gut microbiome, and the cardiovascular signals in sleep apnea patients [197, 61, 206]. SR has also been used for signal modeling in modeling bipedal locomotion and physiological signals [245, 206]. In gait studies, genetic programming (GP) can be used to search through clinical gait variables for the best predictors to put into SR models of human locomotion [1, 179, 180, 181].

We hypothesize that human locomotion can be considered as a network of well-connected nodes, or areas of the body. The objective of this study is to use GP to perform SR on accelerometer data to generate human BSN models for each individual in the study. By doing so, a "network of walking" will emerge where the models are a representation of distributed nodes, proxied by the placement of accelerometer on different parts of the body, and their interactions. Explicitly, this study records and uses signal processing methods to transform gait accelerations from six separate accelerometers on the body, which are referred to as the nodes in the human network. The novelty of this approach is that these SR models can be used for gaining insight into the underlying kinematics and can also be used as a predictive model, without the use of other machine learning algorithms, such as artificial neural networks or support vector machine, since the models should represent the system's metastable state.

4.2 Methods

4.2.1 Participant Demographics

This study included ten human volunteers (five male and five female) from the University of Pittsburgh in Pittsburgh, Pennsylvania, USA. The ages of the participants ranged from 18-35 years, with a mean of 21.40 and a standard deviation of 4.38. The mean height was 1.72 m (sd = 0.09) and the mean weight was 66.36 kg (sd = 8.41). Based on these basic measurements, the participants were considered "healthy" individuals. Further information about the participants is found in Table 1 of Dasgupta et al. [79].

4.2.2 Materials

Six wGT3X-BT triaxial accelerometer sensors (produced by ActiGraph LLC, Fort Walton Beach, Florida, USA) were placed on participants' chest (C), bilateral ankles (left ankle (LA) and right ankle (RA), wrists (left wrist (LW) and right wrist (RW), and lower back (LB) (a figure of these locations can be found in [79]). These sensors are known to be accepted monitoring sensors, and they present minimal risks to participants. Each sensor collected linear accelerations (meters per second squared) at a frequency of 80 Hz over from the mediolateral (ML), vertical (V), and anteroposterior (AP) directions. Treadmills were set at a constant speed of 2.2 miles per hour (or 0.98 m/s).

4.2.3 Data Collection

Data were collected from ten participants from two sessions. These two sessions were wholly identical and were scheduled at least 48 hours or more apart.

For each session, there were five phases: 1) the six sensors were fastened on the participants, 2) the first walking trial, consisting of walking for 10 minutes, 3) the participants were asked to rest for a period of 10 to 20 minutes, 4) the second walking trial, consisting of walking for 10 minutes with a counting activity, and 5) then the sensors were unfastened from the participant.

The first trial consisted of walking on the treadmill, and the second trial involved walking while under an arithmetic cognitive load counting backward from 10,000 in increments of 7. During each trial, each of the six sensors transmitted 48,000 time points for each of the mediolateral (ML), vertical (V), and anteroposterior (AP) directions, resulting in 18 streams of data for each participant for each trial.

4.2.4 Experimental Design

First, PCA was performed on each of the six sensors for all recording sessions. Each sensor recorded three dimensions or axes (ML, V, and AP); however, variations in the physical orientation of the sensor may have had negative consequences as the dimensions may be askew. Thus, for each of the sensors, the ML, V, and AP values were transformed into three principal components. Even if the ML, V, and AP axes were misaligned, the PCA will linearly transform and order the axes based on the amount of variation in the recorded data. The first dimension on all sensors had the most variation, the second orthogonal dimension had the second most variation, and the third orthogonal dimension had the least variation.

Evolutionary searches were used to find a candidate solution with a high-quality fitness value. Typically, the evolutionary algorithm will run for a high number of generations as it will take some time before the algorithm can converge on a high-quality solution. EAs typically follow the same structure: 1) the initialization phase, 2) the fitness evaluation phase, 3) the selection of candidate solutions phase, and 4) the application of the genetic operators phase. In this study, symbolic regression is performed, and thus the fitness value is the mean squared error of the regressed model to the actual recorded signal (the typical fitness measure when performing symbolic regression).

The GP implementation used for this work is heavily inspired by Schmidt et al.'s work [281]. The symbolic regression models are $\hat{y} = f(X_1, X_2, \ldots, X_m)$, where the function f is some combination of the independent variables and linear and nonlinear basis functions defined by the GP system's language.

Four broad sets of experiments (described below) were separately performed on the same data, which includes both walking regularly and walking under cognitive load. For every

individual, session, and trial, each sensor's data (48,000 data points) were broken into ten collections (or batches) of 4,800 time points. For statistical significance and, given the stochastic nature of GP, to increase the chance of generating a high-quality model, 100 models were generated for each of the batches of 4,800 time points. A total of 40,000 models were generated for each of the four broad experiment sets described above (160,000 in total).

For generating these models, one of the directions/axes from a single accelerometer was selected to be fit to the dependent variable ($\hat{y}$), and all other axis/features were used as independent variables of the equation (X). The goal is to develop a model of $\hat{y}$ in terms of X, which accurately approximates y; $y \approx \hat{y} = f(X)$.

A summary of the GP system settings for these runs is presented in Table 4.1. Throughout the evolutionary process, the mean squared error was used as the fitness metric. The number of mating events is low, and the mutation rate of the GP is relatively high compared to traditional GP settings. Thus, preliminary results with longer run times showed significantly better results but provided minimal absolute improvements in the means. The higher mutation rate is a consequence of the acyclic graph representation as not all genes will be represented in the final expression, resulting in some mutations having no impact.

4.2.4.1 Experiment 1 The first experiment set was generated using only data collected from an accelerometer located on the subjects lower back due to it being the most typical accelerometer placement in the existing literature; however, both lower back and hip accelerometer placements have been touted to be the "best" location for providing data to detect an array of human activities [214, 39, 67]. The dependent variable (y) was the axis with the most variance.

4.2.4.2 Experiment 2 The second experiment set included data recorded from the subjects' left and right ankles, left and right arms, and chest in addition to the lower back data. The dependent variable was the axis with the most variance from the lower back (same as the first set's).

4.2.4.3 Experiment 3

The third experiment set was the same as the second (included all axis from all six accelerometers), except the dependent variable was the axis with the most variance from the right ankle.

4.2.4.4 Experiment 4

The fourth experiment set was the same as the second, except the models were fit to only a tiny subset (100) of time points per data set. Testing was two-fold: first, we tested the models against only 100 time points, and then we applied the models to all the data from each batch. These types of models are desirable as fitting to fewer data points will speed up run times, require less computing resources, and may be less susceptible to overfitting.

4.2.5 Computational Resources

Models were generated on a desktop computer with two quad-core (8 cores) i7-4770 CPUs at 3.4GHz with 8GB of RAM. Runtimes for models fit 4,800 data points took between 7-11s each, and the models fit only 100 data points (experiment set four) took between 0.4-0.5s. Evolutionary searches have an element of stochasticity, and even though these runtimes are slow for linear regression, they are fast for a typical EA.

Table 4.1: Parameter settings for GP System. The values for migrations and generations per migration to 100 each when performing the analysis on models fit to a subset of data.

Elitism	1
Population	101
Subpopulations	7
Migrations	5
Generations	100 per migration (500 total)
Crossover	80%
Mutation	10% (x2 chances)
Trainers	15
Predictors	20
Predictor Pop. Size	10% of Dataset
Max # Graph Nodes	16
Fitness Metric	Mean Squared Error: $\frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2$
Language	$+, -, *, /, exp, abs, sin, cos, tan$

4.3 Results

For each experiment, from the 100 models generated for each subject, the one with the lowest training error was selected as the top model (no rigorous model selection strategy was done). Additionally, in our initial data analysis between the trials, many clinical aspects of the participants' gait had no noticeable difference (t-test p-value < 0.05). Some of these aspects include stride lengths/times and cognitive behavior. This is with respect to the idea that there would be a different manifestation of the action depending on if the subject's brain was under a load.

Table 4.2 presents the median training errors over all top models, for each subject's recording, (400 total) along with median testing error calculated by applying all top models to every batch from the same subject, excluding the batch the model was fit to (3,600 total); for each model (400 total), there were 9 unseen batches. For these, the mean absolute error (MAE) and mean absolute percentage error (MAPE) was calculated for all cases, and then the median was presented here. This shows both the MAE and MAPE. Note that this percentage error is related to the mean relative error presented in the time series figures below (Figure 4.1). Although the MAE is reported, the MAPE is a more balanced view of how the models performed relative to one another because not all datasets were within the same range. Observe that although no rigorous model selection was performed and some overfitting is likely present, the models still generalized very well; the difference in errors from training and testing is small.

Table 4.2: Median errors and interquartile range over all subject, session, and trial on all of the broad experiment sets. The median mean absolute error and mean absolute percentage error are presented on both the data the models were fit to (training) and all batches from the same subject (from each trial), excluding the batch the model was fit to (testing). For the training, the models were averaged over all 400 instances, and for the testing, the models were averaged over all batches excluding the batch the model was fit to.

	Training		Testing	
	Median MAE (IQR)	Median MAPE (IQR)	Median MAE (IQR)	Median MAPE (IQR)
Set 1 (Lower Back Only)	0.099 (0.019)	150.404 (48.503)	0.100 (0.020)	151.967 (54.269)
Set 2 (Lower Back)	0.078 (0.017)	114.341 (38.322)	0.082 (0.023)	117.979 (42.960)
Set 3 (Right Leg)	0.230 (0.071)	205.378 (87.952)	0.241 (0.081)	208.221 (96.844)
Set 4 (Lower Back 100tp)	0.074 (0.023)	103.597 (46.987)	0.096 (0.037)	123.036 (73.992)
Set 4 (Lower Back 100tp on All)	0.094 (0.032)	130.367 (56.122)	0.098 (0.035)	132.107 (60.642)

Table 4.3: Probability values obtained with a Mann-Whitney U test comparing the MAPE testing results from each of the experimental sets to each other. Note that this matrix of information is symmetric and the diagonal compares each set to itself. Be aware that all p-values for set 3 are 0. This is due to precision errors and should be interpreted as very small, but not exactly 0.

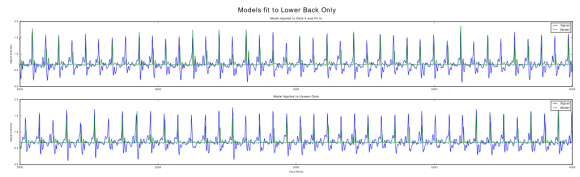
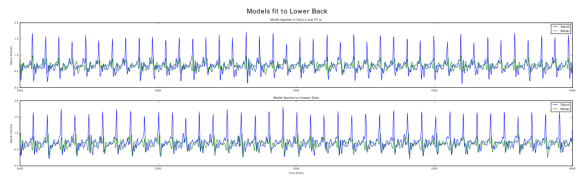
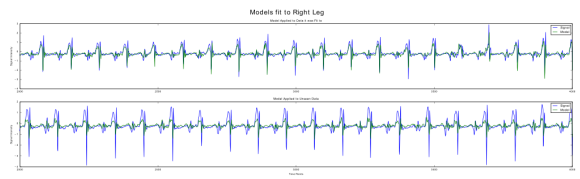
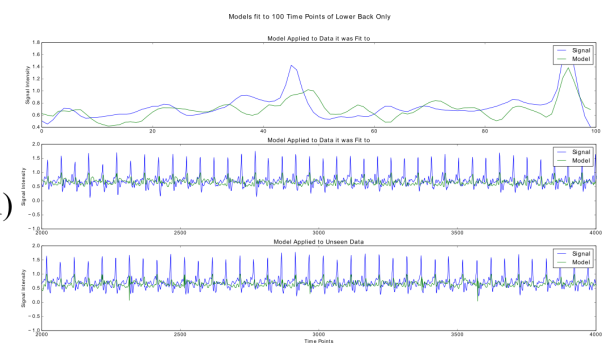
	Set 1	Set 2	Set 3	Set 4	Set 4 (on All)
Set 1 (Lower Back Only)	0.5	$4.331 * 10^{-258}$	0.000	$4.168 * 10^{-95}$	$6.690 * 10^{-61}$
Set 2 (Lower Back)	$4.331 * 10^{-258}$	0.5	0.000	$5.388 * 10^{-8}$	$9.426 * 10^{-53}$
Set 3 (Right Leg)	0.000	0.000	0.5	0.000	0.000
Set 4 (Lower Back 100tp)	$4.168 * 10^{-95}$	$5.388 * 10^{-8}$	0.000	0.5	$1.125 * 10^{-13}$
Set 4 (Lower Back 100tp on All)	$6.686 * 10^{-61}$	$9.426 * 10^{-53}$	0.000	$1.125 * 10^{-13}$	0.5

Table 4.3 presents a symmetrical matrix of probability values from a Mann-Whitney U test comparing the MAPE values obtained in testing each of the broad experimental sets. The diagonal represents the case where the experimental set was compared to itself. The Mann-Whitney U test's p-values comparing the training and testing errors on the median MAPEs are 0.243, 0.006, 0.025, $1,785 * 10^{-18}$, and 0.085 respectively (Table 4.3).

Figure 4.1 shows excerpts of the models plotted against the data they were fit to (training data), and the models alongside unseen data from the same subject, session, and trial they were applied to (testing data). In part **A**, example times series of a model fit with lower back only. Top graph plots the model against the data it was fit to (MAE of 0.136 and mean relative error of 0.247), and the bottom plots the model against unseen data from the same subject (MAE of 0.135 and mean relative error of 0.203). This particular example came from the second and third segment of 4800 time points of Subject 10's first trial of session 1. In part **B**, example time series of a model fit to lower back but fit with all features. Top graph plots the model against the data it was fit to (MAE of 0.131 and mean relative error of 0.210), and the bottom plots the model against unseen data from the same subject (MAE of 0.147 and mean relative error of 0.188). This particular example came from the second and ninth segment of 4800 time points of Subject 10's first trial of session 1. In part **C**, example time series of a model fit to right leg but fit with all features. Top graph plots the model against the data it was fit to (MAE of 0.187 and mean relative error of 1.04), and the bottom plots the model against unseen data from the same subject (MAE of 0.255 and mean relative error of 0.781). This particular example came from the tenth and seventh segment of 4800 time points of Subject 3's first trial of session 2. In parts **D1 and D2**, example timeseries of a model fit to lower back but fit with all features and only on 100 time points. Top graph plots the model against the data it was fit to (MAE of 0.129 and mean relative error of 0.173), middle graph plots the model against all data points from the same data segment the 100 time points came from (MAE of 0.157 and mean relative error of 0.221), and the bottom plots the model against unseen data from the same subject (MAE of 0.163 and mean relative error of 0.225). This particular example came from the seventh and eighth segment of 4800 time points of Subject 10's first trial of session 1. These particular models were selected as they were the ones with the best (lowest) error when applied to unseen data from their own experimental set (no rigorous model selection was performed). In other words, these models were not necessarily the ones with the lowest testing error. The errors reported here are MAE and mean relative error (multiply this by 100 to obtain the MAPE).

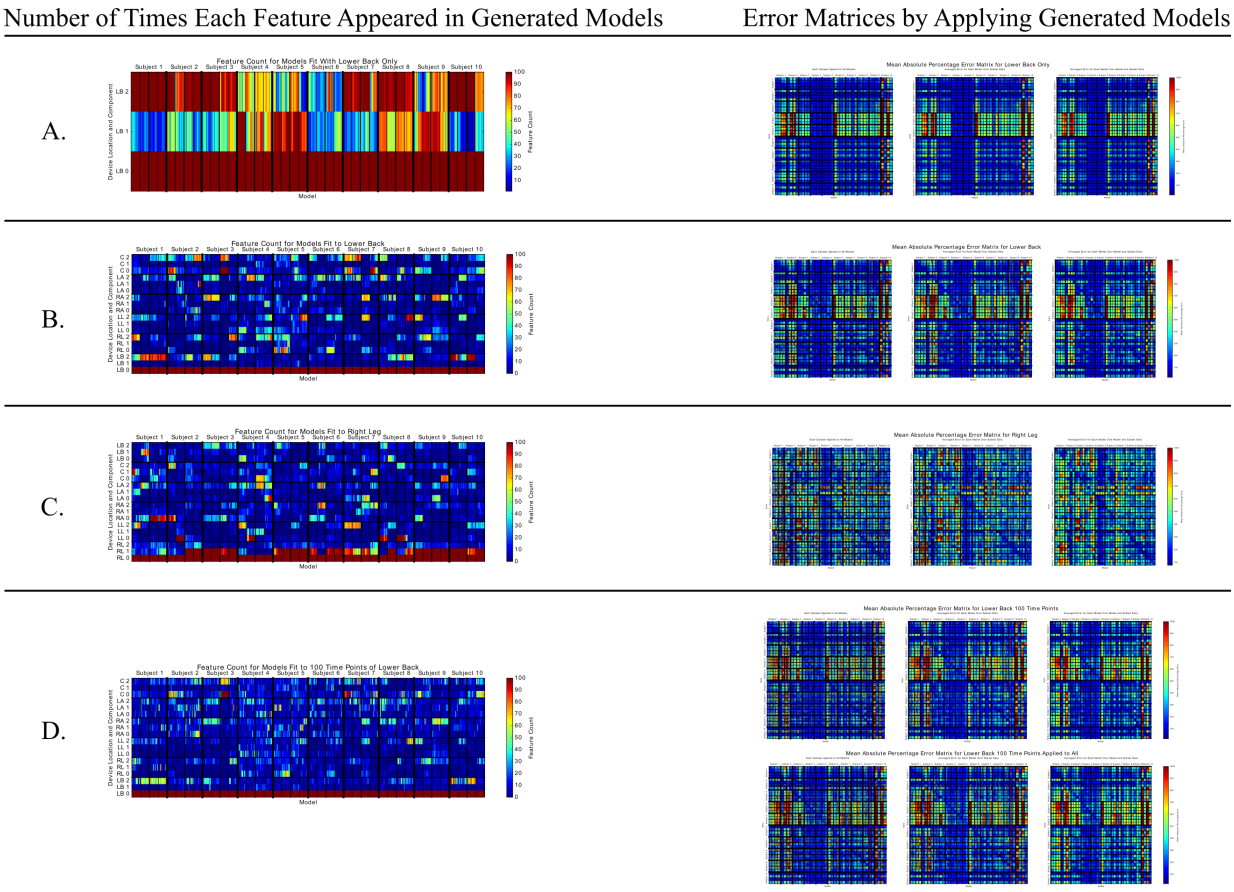
On the right side of Figure 4.2, the left most matrix is all models applied to all data from each subject, session, trial, and batch instances. In the left matrix, rows correspond to

Figure 4.1: Example models from each of the 4 experimental sets (respectively). In these models, the variable corresponds to the component from the accelerometer the data came from, 0 being the component with the highest variance. Additionally, these models correspond to the models presented in the time-series.

Example Models for Each Experiment	Example Time Series of Model Fit
A. Experiment 1 (Lower Back Only) $LB_0 = LB_2 - \cos(LB_2 * 5.0612 + 13.870) * 0.198$	
B. Experiment 2 (Lower Back) $LB_0 = \cos(RA_2 * \cos(RA_0 + LB_2))$	
C. Experiment 3 (Right Leg/Ankle) $RL_0 = \frac{RA_2}{-3.688} - RL_1 * \tan(\cos(LB_2))$	
D1. Experiment 4 (Lower Back 100 tp) $LB_0 = RA_2 + (e^{C_0} * (LA_1 + LB_2))$ D2. Experiment 4 (Lower Back 100 tp on All) $LB_0 = RA_2 + (e^{C_0} * (LA_1 + LB_2))$	

the specific instance and columns to the top model created for that instance and diagonal represents the case when the top generate model for a given instance is being applied to the instance it was fit to. The center matrix contains the averaged errors each model obtained on all batches from the same subject, session, and trial combination. The right most matrix

Figure 4.2: Each row corresponds to the following experiments: A corresponds to experiment 1, B corresponds to experiment 2, C corresponds to experiment 3, and D corresponds to experiment 4. **On the left:** Number of time each feature appeared in the models generated. Rows correspond the feature and columns specific subject, session, trial, and batch instances. The percentage of models each feature appeared in each collection of data (every row and column combination) is represented by a color, as laid out in the color bars. The first matrix corresponds to the first set of experiments (lower back only), second matrix corresponds to the second set, and so on. Note that the last row in each matrix is solid red as it was forced to be in each model since it was the left-hand side of the equation. **On the right:** Three error matrices generated by applying all generated models to every collection of data for the first set of experiments.



averages the errors of all models generated for a given subject, session, and trial combination for all batches from the same subject, session, and trial combination. In all matrices, the MAPE is represented by color.

On the left side of Figure 4.2, these error matrices give an idea of how well these models generalize to other subjects. The best model in this case was the one with the lowest training error. The errors reported here are the MAPEs, not the MAEs. Although these matrices are of error values, they can occasionally be seen as a means of quantifying how similar subjects are. For example, if the subject A's model obtained a similar error when applied to itself and subject B's data, then the two subjects' walking manifestations may be intrinsically similar. However, note that this will not be the case in general. For example, note subject 4 and 5 obtain similar error values on each other, but they also obtain similar error values on all other subject's data (Figure 4.2). Although there are similarities in the feature counts, it is by no means obvious that these subject's models produce these results (Figure 4.2).

4.4 Discussion

This study created models that portray the network relationship between six different areas of the human body. While studies typically use only the accelerometer placed on the lower back, by using accelerometers placed in various places of the body and treating the system as having a network relationship between these body parts, not only do the models fit their data better, but models are far more able to generalize to unseen data.

4.4.1 Symbolic Regression with Genetic Programming

This study produced symbolic regression models with GP. Symbolic regression is similar to other regression techniques, in that it searches for parameters for a mathematical model to fit data; however, unlike linear regression, symbolic regression also frames the structure and operators within the model. The major strengths of this particular implementation of producing symbolic regression models with GP are in the use of subpopulations, an acyclic

graph representation, and the use of fitness predictors. Unlike traditional GP systems where the chromosomes are tree-based S-expressions, this study's application uses an array-based acyclic graph representation of the underlying mathematical expressions. The benefits of such representation include faster runtimes, easy reuse of subexpressions, and the reduction of bloat (a GP phenomenon where the trees grow arbitrarily complex with no meaningful improvement in the quality of the effectiveness [278]).

Typically, the evaluation of a candidate solution involves applying all data points to it. For example, when performing symbolic regression, all data points are applied to the expression and the error the candidate solution is measured from the actual measured data. Since this evaluation is the most computationally expensive part of the evolutionary search, fitness predictors were used to alleviate the computational load. Fitness predictors are a small, always changing subset of data points selected for their ability to reasonably represent the whole data set while also selecting data from areas of the search space the models have less consensus on (if the models have little consensus on an area of the search space, then the training should focus on that area). This approach has been shown to improve runtimes, produce higher quality results, and even reduce overfitting [279, 280].

4.4.2 Comparison of Experiments

As mentioned previously, lower back and hip accelerometer placements are ubiquitous in the literature. After fitting the data collected from the lower back accelerometer only to the dependent variable (the axis from the lower back accelerometer with the most variance) (experimental set 1), the resulting models generalized to all other subjects far better than the other experiments. The error matrices for this experiment also indicate that these models could generalize the best, and the diagonal is not particularly well pronounced, indicating little intrasubject overfitting (Figure 4.2). Figure 4.1, showing the model for set 1, has the second-best testing MAPE at 0.203.

Adding the data from the subjects' left and right ankles, left and right arms, and chest, as well as the lower back (experimental set 2), increased the performance of the models when considering the testing results (Table 4.2). Table 4.3 also indicates that these results

outperformed other experiments by a significant amount. Figure 4.1 showing a model for set 2 had the best testing MAPE (0.188), which does align with the overall results seen in Table 4.2.

In our next experiment, we changed the dependent variable to the axis with the most variance from the right ankle, while keeping the data or features the same (experiment set 3). This change allows for an analysis of the network model of human locomotion but from a different perspective. Despite having all the same data to fit with, these models had very distinct features when compared to the previous experiment (Figure 4.2). This finding also corresponds to the observations made when discussing Table 4.2; the network arising depends on the dependent variable. Unfortunately, these models appear not to generalize well; however, some of the observations made about subject outlier errors (subject's 1, 2, 6, and 10 discussed above) from the other matrices cannot be seen here, suggesting that the choice of dependent variable is crucial (Figure 4.2).

Next, we tried altering the second experiment by fitting the models to only 100 time points and then subsequently applied to all 4,800 time points (experimental set 4). These models were still performing very well and fit significantly better than those fit to only the lower back's accelerometer. It seems that very few data points are required to create a useful model of this complex system. This finding is advantageous in a real-world application as it would require much less data gathering, less pre-processing, and less computational power to generate the models.

Even though these models were able to fit its data the best (experimental set 4), it should be noted that this testing error was calculated on only 100 time points as opposed to the 4,800 time points all other sets used. For this reason, one should focus on the second version of experimental set 4, where the models fit to 100 time points were applied to the full 4,800 time points to calculate the errors.

When comparing the second and fourth experimental sets, the error matrices all look similar with a well pronounced diagonal (Figure 4.2). Both sets appear to have overfit nuances in the few data points they saw (Figure 4.2).

Regardless of the experimental set, one can see that the models track the signals well. The most considerable deviations seem to come from the spikes in the signals; some of the

higher frequencies can be found near the middle of the signal. However, the example shown in time series corresponding to experimental set 4 did not fit the signal as well as those seen for sets 1 and 3 when considering the testing error (Figure 4.1).

In the error matrices in Figure 4.2, we can see some common trends between all four experiments, which may be due to the experimental setup, rather than any biological explanation. For example, subject 10's trial one in both sessions appear to be markedly worse. A similar thing can be seen in session two for subjects 1 and 2 and on subject 6's trial one of session one. Perhaps these subjects' devices were in markedly different locations, or there were problems with accelerometer calibration. Additionally, the high MAPEs are likely due to significant spiking events and other artifacts in the signal.

While the third experiment also treated the system as a network relationship, but those models performed poorly and significantly worse than those fit to just the lower back's accelerometer. It seems that this network relationship is necessary for describing the system, but the choice of the dependent variable is critical. For example, the experiment fitting the lower back as the dependent variable substantially outperformed the experiment having the right ankle as the dependent variable.

4.4.3 Distinguishing Cognitive Load

We expected that cognitive load would affect these models. However, in this study, we found no noticeable differences between the trials (walking regularly vs. walking under cognitive load). This is in direct contrast to the study by Dasgupta et al., where they could distinguish between the features from the accelerometer data from the two trials, using machine learning models [79]. This result could be due to the small sample size in conjunction with a sample of solely healthy adults whose gait only changes slightly with an added cognitive load. Additionally, perhaps symbolic regression models are better at distilling noise; in this case, noise could be the inconsistent differences caused by cognitive load [150]. Hypothetically, symbolic regression models may only capture high level trends, and other modelling techniques, such as those used in Dasgupta et al., may be better methods [79].

4.4.4 Network of Walking

One of the main, interesting results of this study is that a network relationship was found between the different body parts. The network models fit to the lower back's accelerometer were able to describe subject-specific data the best when compared to all other models by a significant amount, suggesting the network relationship is essential in describing the system. This is as expected since there was more information to fit with, and the testing results show that overfitting was not a significant problem. The quality of the network models depended on what feature the models were fit to — as one would expect based on the literature, the lower back appears to be critical for developing high-quality models. The second best performing models were network models fit to only 100 time points, as opposed to the 4,800 all others were fit to. Although these models did not perform as well as the lower back network models, it shows that with high-quality data, very few data points are required to generate effective models.

The features arising in the models had little consistency between subjects, but there was an emphasis on the network relationship within the same subject, session, and trial. We suspect the importance of the feature in describing the system is dependent on the location and orientation of the accelerometer.

Although the network models fit to the lower back were the most effective intrasubject, the models fit with only the lower back were the best at generalizing between subjects. However, subjects 4 and 5's models were still beneficial at generalizing to other subjects across all experimental sets. This is not surprising as these models were fit to data with fewer features, which would result in smaller errors.

No apparent trends can be seen in the feature counts in Figure 4.2 with respect to which features arose in the models. The largest trend that can be seen is the features arising in all batches from the same experimental set (from all the subjects), or sometimes between the same subject in the same experimental set. Unsurprisingly, the models' fit for experimental sets 2 and set 4 (both fit to lower back with all features) have a lot in common. In general, the same features arise, but to a lesser extent in experimental set 4's models.

Because there is little consistency in features for each subject, it seems that this sys-

tem (human locomotion) can be explained with a large variety of models, or perhaps the usefulness of a feature in a model is very dependent on the precise physical location of an accelerometer. Although PCA would align the axes, variations in the physical location of the accelerometer between subjects, or even between sessions or trials, cannot be corrected. More focus is concentrated on the latter as the commonality in counts between set 2 and 4's models show consistency. This does provide more evidence for the necessity of a network relationship for describing the system; however, some between-subject generality may be lost.

With the results presented here, little consistency arose in the features between subjects, and it seems that these models are very subject specific. This could be a result of different physiology (weight, size, and shape) between individuals or our modeling strategy. Given this, it would be challenging to develop a general-purpose network relationship model with the presented strategy as is, however, alterations to the modeling strategy may be able to generate a more general model. Thus, this network of modeling can be an additional data perspective to existing biological databases that target cognitive or neuromotor disease.

4.4.5 Conclusions

Overall, we presented a collection of nonlinear models of a complex dynamic system: human locomotion. These models were developed to analyze nonlinearities in the system and explore the network relationship between body parts during locomotion. These models were also developed to understand how the presence of a cognitive load (in the form of a cognitive task) affects these models, although no significant trends were noticed.

5.0 Spatiotemporal Estimations using Accelerometers

Co-authors include: Mark Redfern and Ervin Sejdić.

5.1 Introduction

Spatiotemporal measures of gait help evaluate the progression of diseases and related conditions, ranging from Parkinson's disease and Alzheimer's disease to obesity and arthritis. For example, people with Parkinson's disease often walk with both a reduced gait speed and step length [81]. Spatiotemporal gait parameters have also been used to identify people at high risk for falls and older adults with altered cognitive states [232, 138, 231]. Changes in these parameters have been linked to increased falls, comorbidities, and injuries [125]. Thus, the evaluation of human locomotion is now recognized as an important clinical measure.

Current, non-exhaustive spatiotemporal estimation techniques include the use of laboratory-grade gait analysis equipment (e.g., GAITRite, Vicon) or clinical, observation-based screenings [82, 92]. However, these gait analysis methods have limitations: space, decreased portability, high financial cost, and a limited number of gait cycles collected [12, 92]. Other methods of clinical gait analysis rely on clinical observation and subjective evaluation (e.g., the functional gait analysis [283]); however, these methods are qualitative and often have only moderate interrater reliability among health professionals. One recent measurement system that holds much promise is the use of inertial measurement units (IMUs) that contain linear accelerometers, gyroscopes, and magnetometers. IMUs are low-cost, and set-up is easily managed in either research, clinical, or home environments [82]. Low-cost estimation of spatiotemporal estimation may even be performed using acceleration signals from IMUs integrated into smartphones [351].

Processing IMU data to obtain gait parameters has been a challenge to obtain the most clinically valuable measures from various environments. Accuracy depends on several factors, including the location of the IMU on the body, walking speed, and navigation (i.e., straight

path or turning). Some common approaches to spatiotemporal estimation have included thresholding [240] and peak detection techniques on accelerometer and gyrometer signals [246]. However, the efficacy and accuracy of these methods under various conditions are not fully known [97].

In this study, we focus on two established methods to estimate spatiotemporal gait parameters from acceleration signals using IMUs and propose a method to improve them for slower gait speeds and make them less dependent upon the IMU location. These two methods use an acceleration peak-finding [79] and an inverted pendulum model [112]. In 2003, Zijlstra and Hof did a set of experiments, where they estimated spatiotemporal gait parameters including stride time, step length, and walking speed from lower-trunk accelerations of healthy subjects during overground walking using an inverted pendulum approach [352]. The acceleration peak-finding algorithm chosen was one we have used previously [79]. However, in each case, the methods did not investigate accelerations from other body-sites other than the lower back, nor during slower gait speeds. Expanding to other sites on the body and for slower speeds can have tremendous value clinically [97]. Further, step width was not able to be reliably estimated in either technique.

We focus on spatiotemporal aspects of a person's gait, expressly the number of steps, step length, and step width [79, 251, 112]. We hypothesize that the acceleration peak-finding method and the inverted pendulum algorithms will accurately estimate these spatiotemporal measures. Moreover, we use an instrumented treadmill to establish a 'gold standard for these measurements during data collection.

5.2 Methods

5.2.1 Participants and Data Collection

Ten healthy volunteers (5 male, 21.4 (4.2) years, 1.72 (.09) m height, 66.4 (8.4) kg weight) were recruited. All participants were in good overall physical health and reported no current medical conditions that could affect gait. Participants were equipped with six

wGT3X-BT triaxial accelerometer sensors (ActiGraph LLC, Fort Walton Beach, Florida, USA) positioned on their chest, bilateral ankles, bilateral wrists, and lower back (Figure 5.2) [79]. These sensors captured 3 orthogonal linear accelerations at a sampling frequency of 80 Hz. The axes were aligned approximately with the mediolateral (ML), vertical (V), and anteroposterior (AP) directions when in standard anatomical position (Figure 5.2) [79]. In Figure 5.2, we show exemplar data from the ML, V, and AP acceleration signals for each sensor across two consecutive steps.

The protocol consisted of walking on an instrumented treadmill (10 minutes), a rest period (10-20 minutes), and walking again (10 minutes). This was performed twice, once in each of two sessions with at least 48 hours between each session. In this study, we used the data from the first session's first trial only. For further information on the data collection methods see Dasgupta et al. [79]. An institutional review board at the University of Pittsburgh approved the data collection (No. PRO14060107), and all subjects participating in the study provided informed consent [79].

The instrumented treadmill (Noraxon USA, Scottsdale, Arizona, USA) captured foot force data for later calculations of foot contact times, step length, and step width [79]. The treadmill speed was approximately 1.0 m/s, which is less than the usual adult walking speed of 1.3 m/s [79]. This treadmill speed was chosen so that all participants would be at a comfortable and slower than usual walking speed, which is closer to a speed common among community-dwelling older adults and persons with walking difficulties [79].

5.2.2 Algorithms for Spatiotemporal Estimation

The following methods will highlight the peak-finding and the inverted pendulum methods used. From these two methods, we extracted the number of steps (#), step length [m], and step width [m], which correspond to the distance between the ipsilateral and contralateral heel strikes [51]. We applied these two methods to acceleration signals from all six body sites. We compared these measures with measures recorded with the treadmill. Note that stride and step duration [s] was not analyzed due the constant speed set by the treadmill.

5.2.2.1 Method 1: Acceleration Peak-Finding The peak-finding method of stride and step extractions used was based on the methodology used in Dasgupta et al. and Sejdic et al., stride and step extractions were done using a peak-finding methodology from the back accelerometers and other sites [79, 286]. We also applied this methodology to all the accelerometers for the six body sites. The steps in this algorithm implementation include: (1) pre-processing gait accelerometry signals via median filters, (2) determine which foot came first by calculating the average of the ML acceleration signals in the first 10 ms, (3) identify heel strike events via the local minima points in the AP signals, and (4) identify toe-off events via the local minima points in the V signals. (Extractions of heel strike events are shown in dotted lines in Figure 5.2.) Heel strike and toe-off events were used to identify stride and step timings. After stride extraction, the number of steps were calculated. Step length was found by multiplying gait speed by step time. Step width was not found through this method.

5.2.2.2 Method 2: Inverted Pendulum Model This method is based upon an inverted pendulum model which predicts the body's center of mass trajectory during walking. (See [112] for details.) This model leads to the step length estimation via Equation 1, where h is the double integration of the V acceleration (Equation 5.2), and l is height from which the back accelerometer from the ground. (This is approximately the leg length, which can be estimated to be 45% of the height of the individual) [12, 212, 352, 185, 112]. In Equation 5.2, $t0$ and $t1$ are the two consecutive instants of foot-flat during the gait cycle, and $c1$ and $c2$ are the two offsets in the integration process [12]. Step width was estimated by using a similar equation (Equation 5.3) as for Method 2, except with h as the double integration of the ML acceleration instead (Equation 5.3). Equation 5.3 was found using triangulation estimates using the Law of Sines (estimating 30 degree separation between steps) on Equation 5.1.

$$2\sqrt{2lh - h^2} \tag{5.1}$$

$$\begin{aligned} &\int_{t_0}^{t_1} (a_V + c_1)\, dt = 0 \\ &\int_{t_0}^{t_1} \left(\int_{t_0}^{t} (a_V + c_1)\, dt \right) c_2 dt = 0 \end{aligned} \tag{5.2}$$

$$\sqrt{2lh - h^2} \tag{5.3}$$

5.2.3 Determining Step Metrics from Instrumented Treadmill

The gold standard for step metrics were derived from treadmill underfoot force data. Briefly, the forces from the loadcells within the treadmill are used to calculate a center or pressure across the surface. This center of pressure (CoP) time series is then used to identify heel strike and toeoff event timing, number of steps, step length and step width. (A description of metric extraction from the treadmill is given in Dasgupta et al. and Truong et al. [79, 313].)

5.2.4 Overall Validation and Statistical Analysis

Means and standard deviations were calculated for each metric (number of steps, step length, and step width). The normal distribution for each metric were tested with a Kolmogorov-Smirnov test (all metrics were normally distributed). An equivalence test (two one-sided test (TOST) for paired data) was performed to determine significant similarities between the three metrics using the acceleration peak-finding method, the inverted pendulum methods, and the treadmill gold standard results [189]. For TOST, the level of significance was set at $p<0.05$ for similarity.

The TOST procedure specifies an upper (ΔU) and lower equivalence bound (ΔL) based on the smallest effect size of interest. Two null hypotheses are tested: H01: $\Delta \leq -\Delta L$ and H02: $\Delta \geq \Delta U$ [189]. When both the one-sided tests can be statistically rejected, then we can conclude that the observed effect is close enough to zero to be considered equivalent (or if the observed effect falls between the equivalence bounds: $-\Delta L<\Delta<\Delta U$) [189].

All analyses (extraction and statistical) were done in RStudio version 1.3.1073 via R version 4.0.3 [262].

5.3 Results

Using our gold-standard technique, treadmill observations, the number of steps per trial across subjects was 996 (±47). The mean step length was 0.59 (±.03) m and the mean step width was 0.12 (±.03) m. Estimates made for the low-back accelerometer for the acceleration peak-finding method and the inverted pendulum method were similar to the treadmill results. However, there were significant differences between the methods when using some of the other body sites (Figure 5.1). The acceleration peak-finding method did not find gait metrics similar to the treadmill observations for most non-low back accelerometers.

The analysis of step width was applied to the inverted pendulum model. We found that using the inverted pendulum method, we were able to produce similar mean step widths for each subject for all body sites (Figure 5.1). The back accelerometer site was not significantly different than the treadmill gold standard for all three spatiotemporal estimates.

Figure 5.1: Means (standard deviation) for method 1 and 2's step length, number of step results, and step width for each accelerometer for each body site. For all values, one significant figure is shown. Green highlighted values show a p-value of <0.05, which indicates significant similarity to the gold standard.

Method 1: Acceleration Peak-Finding

	Back	Chest	LA	RA	LW	RW
Number of Steps	1000 (60)	1000 (40)	1000 (60)	2000 (50)	1000 (50)	1000 (50)
Step Length	0.60 (0.03)	0.60 (1.0)	0.40 (1.0)	0.30 (0.04)	0.50 (0.03)	0.50 (0.03)

Method 2: Inverted Pendulum

	Back	Chest	LA	RA	LW	RW
Number of Steps	1000 (50)	1000 (60)	1000 (40)	1000 (40)	1000 (40)	1000 (40)
Step Length	0.60 (0.07)	0.60 (0.06)	0.50 (0.7)	0.50 (0.06)	0.60 (0.07)	0.60 (0.06)
Step Width	0.10 (0.05)	0.10 (0.06)	0.10 (0.06)	0.10 (0.06)	0.10 (0.1)	0.10 (0.01)

5.4 Discussion

This study demonstrates the two methods of extracting the number of steps and step length from accelerometers. We also demonstrate that it is possible to extract average step width from accelerometers. The step extractions from the accelerometer (methods 1 and 2) and treadmill (gold-standard) are very similar. For instance, the step length means are comparable. However, the standard deviations for these means are much tighter from the treadmill measurements, suggesting that the accelerometer measurements may have wider inconsistencies from one measurement to the next. Our findings conclude that, on average, nearly all the number-of-step estimates were lower or slightly lower than the treadmill measurements. Feng et al. corroborate this result, who found that accelerometers may underestimate step counts in slow speeds [97].

There may be two reasons that there is a decreased accuracy of average step counts from accelerometers: an intrinsic fault in the algorithms or decreased sensitivity for step detection at lower speeds [97]. Further research needs to be done on the sensitivity of accelerometer data in capturing changes in various gait speeds and those of a variety of ages.

Our data and results reinforce the use of the inverted pendulum model (method 2) and the assumption that the left-to-right trunk accelerations are correlated with the left and right foot positions as one walks (i.e. step widths). Clinically, the left foot support phase shows right trunk accelerations, while the right foot support phase shows left trunk accelerations [352]. These accelerations are not periodic during the support phase, and they have small amplitudes. It is useful to check the lateral accelerations around the heel strike gait event; there is often a large lateral acceleration spike during this gait event. With lower gait speeds, these spikes may seem small, but they are significant. With higher gait speeds, these spikes are much larger and more noticeable. Thus, these lateral acceleration spikes may indicate that further clinical investigation is needed; even though there is a large variability between subjects, the subjects' ML accelerations reflect individual walking patterns' idiosyncrasies.

5.4.1 Limitations

This study's primary limitations are the population (healthy adults) and the sample size (n=10). This study seeks to improve the techniques to capture basic measures of gait to be applied to larger datasets in the future. While our study's population size is only ten, the amount of data collected within-subject was extensive (10 minutes of continuous walking). Furthermore, data was collected on a treadmill, as opposed to free-ranging walking. Thus, the measures captured overall characteristics of walking. The participants' "healthy" status allows us to explore the effect of the treadmill's slow speed on the device's accuracy in spatiotemporal estimation. However, it may prevent us from interpreting abnormal gait results.

Additionally, while we used some common approaches to spatiotemporal extraction, such as filtering, thresholding, and peak detection techniques on accelerometer signals [246], limitations of these approaches are that these methods depend on multiple heavily-tuned computational parameters, such as width of the bandpass filter, window length, thresholds [246], many of which may have significant impact in a clinical environment.

Limitations to current feature extraction algorithms that estimate spatiotemporal gait elements are measured in either healthy subjects or a particular population (i.e., Parkinson's patients). There needs to be concerted research on a wide variety of people so that the algorithms can generalize. Furthermore, few data correction procedures were done due to the accelerometer's small sampling rates compared to other studies in the literature [124]. These data correction procedures include measuring the tilt of body parts, gravity artifacts, and skin motion artifacts.

5.4.2 Clinical Application and Future Directions

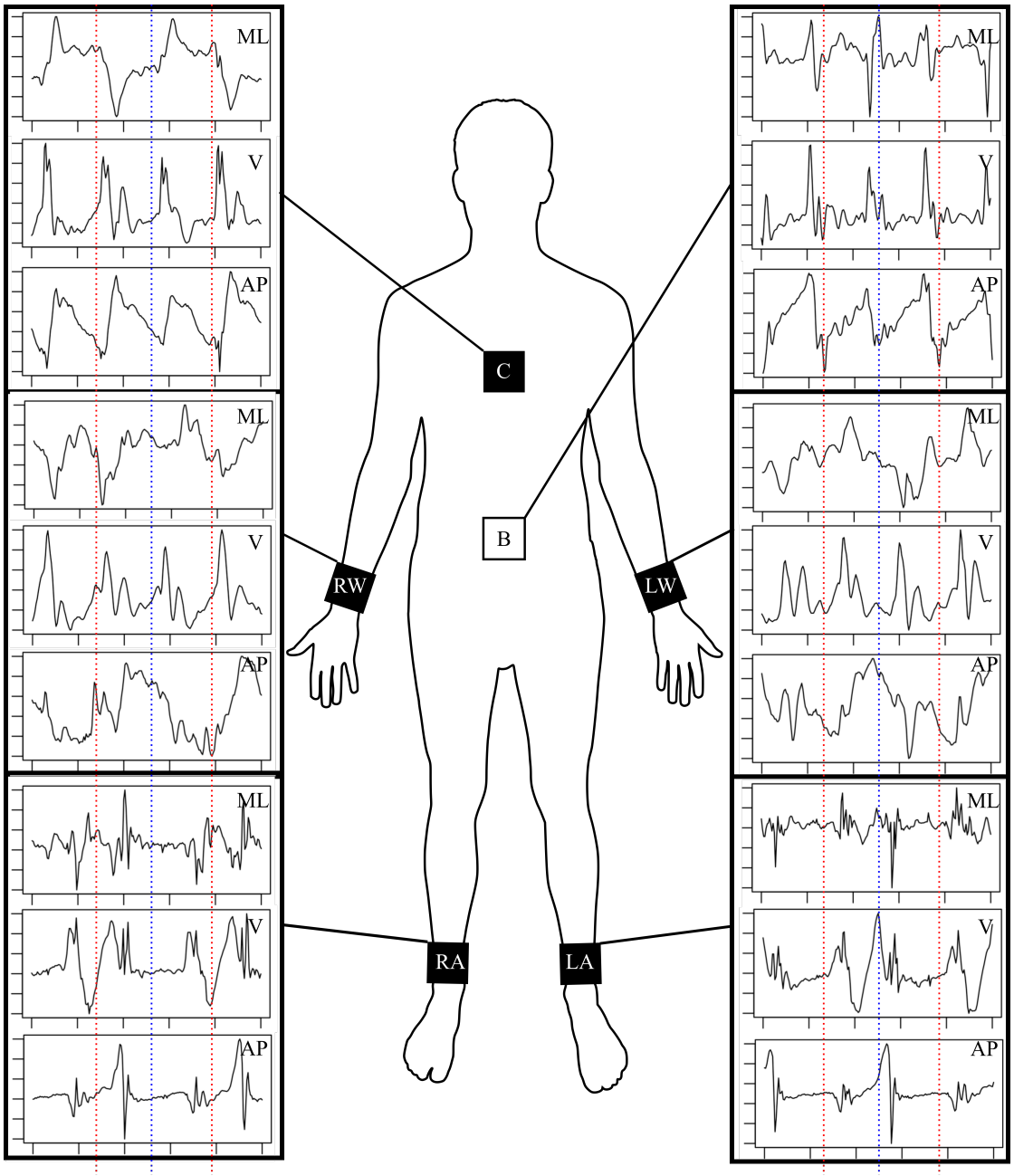
The assessment and interpretation of spatiotemporal gait metrics are essential for patient evaluation, predicting fall risk and determining gait implications of altered cognitive states [92]. Many healthcare professionals and therapists often rely on their observational skills alone or extensive lab setups. However, with the use of accelerometers, there is the potential to validate and improve the reliability of those observations [325]. For example, Wall et al.

state that "it is useful to gather data on spatial gait variables of step length and step width in patients with a locomotor disability" in addition to observations or video-recordings of a person's gait [325]. This extends to older adults who have a history of falling have step length and step width abnormalities [325].

5.5 Conclusion

Our study found that accelerometer signals with accelerometer-based methods of calculating spatiotemporal estimates of walking when applied to a number of accelerometer body site locations were successful. This improvement is shown for slower walking, where estimating spatiotemporal parameters are more difficult. These results have the potential for improving clinical measures for patients and older adults. Further development and testing of models will allow application of these algorithms to a more generalizable population with varying gait speeds.

Figure 5.2: Schematic of the accelerometer locations and the corresponding acceleration signals (ML, V, and AP) over two steps from each of the accelerometer positions from subject 10. The red and blue dotted lines correspond to right and left heel strikes, respectively. Note the locations of the accelerometers: C = chest, B = back, LW = left wrist, RW = right wrist, LA = left ankle, RA = right ankle.



6.0 Using Acceleration Gait Features Associated with Gait Speed to Detect Fall-risk in Older Community-Dwelling Adults

Co-authors include: Jessie VanSwearingen, Jennifer Brach, and Ervin Sejdić.

6.1 Introduction

Multiple factors, including clinical neurological developments, can affect aging older adults' motor skill of walking [40]. In particular, older adults' walking difficulty is of concern due to the immense global healthcare cost of falls and fall-related injuries [130, 220, 120, 256]. The motor skill of walking is complex, skilled movements coordinated with muscle activation patterns on both sides of the body [193, 321]. Aspects of the motor skill of walking include adaptability, symmetry, smoothness, efficiency, and automaticity [321]. The continued, effective use of the aspects of the motor skill of walking in older adults promotes walking, with decreased risk of falls [321].

Gait analysis research often uses gait speed, a standard, widely used outcome metric for mobility, and a good predictor for "walking difficulty" in older adults [258, 37]. Older adults exhibit decreased gait speed, stride length, and cadence. Additionally, they also experience an increased stance and double support phase duration. However, gait speed alone isn't enough to predict whether someone is going to fall, on a case by case basis. There have been studies on determining the relationship between gait speed and fall-risk [94, 261]. While a relationship exists: low gait speeds are associated with indoor fall-risk and fast gait speeds are associated with a outdoor fall-risk [261], multiple aspects of mobility along with gait speed is required to make a determination of whether someone is at fall-risk (e.g, adaptability and stability, aspects of the motor skill of walking, was quantified by Lyapunov exponents to aid in calculating fall-risk) [121].

Accelerometers can commonly measure motor skill factors and aspects [78, 303]. Acceleration gait measures (AGMs), features that are calculated from the time-series accelerometer

data, are crucial to proxying aspects of the motor skill of walking [78]. This proxying practice can help differentiate gait-related outcomes through machine or statistical learning. Machine and statistical learning's overall goal is to predict or classify unknown outcomes based on past exposures/features. However, it can be computationally and mentally demanding to determine which selected AGM features are less significant than others.

AGMs may be able to investigate fall-risk behaviors more closely. In the literature, it is shown that the following AGMs are helpful in classifying fall-risk behavior in machine learning studies: descriptive statistics, cadence, stride time, Fast Fourier Transforms to produce a percentage of acceleration frequencies in the first quartile of the frequency plot, harmonic ratios (the ratio of even to odd harmonics), maximum Lyapunov exponents, local dynamic stability, discrete wavelet transform, sample entropy, and power spectral density [144, 264, 93, 78].

There exists no gold standard of specific AGMs for a particular aspect of motor skill, with respect to gait speed and fall-risk. It can be difficult to decide which AGM to pick to evaluate gait speed and determine who is at fall-risk. Often, researchers pick AGM features that are common in use in their particular field. Clinically, we could pick out relevant AGMs that fit the context of the clinical problem. Mechanistically, there are many feature selection methods, such as forward or backward or recursive methods. In this paper, we take the mechanistic approach to finding AGMs that are most descriptive or influential for gait speed in older adults. The argument is that the AGMs that are influential for gait speed are also influential for fall-risk, as gait speed is associated with the fall-risk in older adults [94, 71, 238].

This study seeks to find AGM features most influential for gait speed in older adults; since gait speed is associated with fall-risk, doing so will help differentiate those who are a fall risk. Secondarily, the clinical relevance of the chosen AGM features will be explored and analyzed. We hypothesize that while multiple AGMs could be used in modeling gait speed, there can be differences in which AGMs prove to be more influential. We will calculate four categories of AGMs: statistical, signal-frequency, time-frequency, and information-theoretic features. Subsequently, we will implement a host of prominent wrapper and filter feature selection techniques. We will perform feature selection methods from a wide array of ma-

chine and deep learning algorithms. High-performing AGM features will be scored, and the clinical implications of including these AGM features in a model will be discussed. Using a multitude of various machine and deep learning methods, the AGM classes mentioned above will accurately classify whether someone has fallen (with baseline data).

6.2 Materials and Methods

6.2.1 Data Collection

This study uses the dataset from the "program to improve mobility in aging" (PRIMA) study [47]. The PRIMA study is a randomized, single-blind two-arm intervention trial of 249 community-dwelling older adults who walk > 0.60 m/s and < 1.2 m/s [47]. Thus, the inclusion into the PRIMA study was based on gait speed and age (65 years or older) [47]. Additionally, inclusion criteria also required participants to be ambulatory without an assistive device or the assistance of another person and have physician clearance to participate in a moderate-intensity exercise program [47].

Participants were recruited through the Pittsburgh Pepper Center Registry of over 2200 older adults who have signed consents to be directly contacted by participating researchers about mobility research studies. Each participant was measured at several time points: baseline, 12 weeks, 24 weeks, and 36 weeks. At each visit, we placed Actigraph (Actigraph, LLC; Pensacola, FL, USA.) accelerometers at the level of L3 on the lumbar spine. This study exclusively uses the overground files for the baseline and 12-week visits for the L3 location (of the 249 only 196 had measures at both time points and were considered eligible for the analyses in this paper). The demographics of these 196 participants are shown in Table 7.1.

Table 6.1: Demographic information and a selection of survey data from older adult patients in the PRIMA study. Note: SD = standard deviation.

Variables	**n=196**
Age (mean (SD))	77.69 (6.54)
Identified Gender = Male (%)	68 (34.7%)
Height [m] (mean SD))	1.67 (0.09)
Weight [kg] (mean SD))	79.51 (17.70)
Race (%)	
Asian	1 (0.5%)
Black or African American	18 (9.2%)
Native Hawaiian or Other Pacific Islander	1 (0.5%)
White	172 (87.8%)
Other	1 (0.5%)
Refused to Answer	3 (1.5%)
Live Alone = No (%)	110 (56.1%)
Last Grade Completed in School (%)	
High School/Equivalent	32 (16.3%)
College	71 (36.2%)
Post-Graduate	85 (43.4%)
Other	8 (4.1%)
Fall(s) within a year before baseline visit (%)	
Fallen	135 (68.9%)
Have Not Fallen	61 (31.1%)

6.2.2 Data Processing

For each participant's accelerometer data during an overground walk, the acceleration signals for the ML, V, and AP directions were pre-processed with the following steps: 1) we re-sampled the signals to 100 Hz, and 2) we filtered the signals through a fifth-order, low pass Butterworth filter [141]. For this Butterworth filter, we performed a harmonic analysis to determine an optimal cut-off frequency [337]. These pre-processing steps were crucial in reducing external noise sources from the accelerometer signal [110].

6.2.3 AGM Calculation Methods

We used window segmentation, where we segmented the processed acceleration signals in 10-second windows/vectors, each with a 50% overlap between two consecutive windows. For each window, we calculated popular AGMs in each of the four AGM categories (statistical features, signal-frequency features, time-frequency features, and information-theoretic features) [78]. We had 42 statistical features, 15 information-theoretic features, 99 time-frequency fea-

tures, and 57 signal-frequency features. Overall, we had 213 features for each window. These features and their descriptions are listed in the Appendix (Table 6.4, 6.5, 6.6, 6.7). These gait measures, and variations of them, are often used in machine learning modeling, and we will compare the performance of the popular AGMs to gait speed features.

6.2.4 Feature Selection and Importance Methods

The goal of feature selection is to find the subset of the calculated features that are most "important" to predict the value of the outcome. Several approaches attempt to rank individual features and to choose the most significant variables. However, there are disadvantages such as (1) choosing more variables than necessary because it does not take redundancy into account and (2) missing variables that are relevant together although useless individually.

Feature selection methods are also preferred over projection and compression techniques [171]. In order to contextualize the features clinically, we do not want to alter the features. There are three overall approaches to feature selection: filter, wrapper, and embedded. For example, wrapper feature selection methods were used to decrease the feature space for fall-risk prediction [141].

Apart from the removing correlated features, the approaches that we perform in this paper are random-forest based: intrinsic or model-based, filter methods, and wrapper methods. We restricted these methods to random forest, because random forests are also a useful technique for estimating feature importance [74].

6.2.4.1 Intrinsic or Model-Based Approaches Random forests do computation from the whole training data, and they compute feature importance based on out-of-bag error. In general, a random forest is composed of multiple decision trees, and each decision tree has a set of internal nodes and leaves. In each internal node, one of the features is used to decide on the dataset. In the internal nodes, the feature with the highest decrease in variance is chosen. In this way, we can assess how each feature decreases the impurity of the split. For each feature, the average of how it decreases the impurity over all the forest's

decision trees measures feature importance. In this paper, we include the total decrease in node impurities, or the mean decrease Gini, (IncNodePurity) and mean decrease accuracy (%IncMSE) as measures of feature importance, where the highest of these metrics indicate how influential the feature is [74, 184].

The IncNodePurity metric measures the split's quality for each node of the tree via the Gini Index. Higher IncNodePurity values represent "purer" nodes and higher variable importance. The %IncMSE is produced by "permuting" the values of each variable of the test data and contrasting it with the non-permuted test data. In regression, this is represented by the average increase in the test data's squared residuals when the variable is permuted. A higher %IncMSE value shows higher variable importance.

6.2.4.2 Highly Correlated Features Redundant features can be assessed with unsupervised methods. Highly correlated features, a filter method, can increase computing time, introduce bias, or mask effects on a machine learning model's outcome. Removing one of a pair of highly correlated features is an unsupervised method of feature selection.

The first method we used to eliminate highly correlated features was to create a correlation matrix of all the numerical features. We removed features with a 0.80 Pearson correlation or higher (Pearson for continuous variables); this correlation metric measures the linear correlation between two continuous variables, so it may not detect any non-linear correlation between the AGM features.

6.2.4.3 Wrapper Methods The wrapper method we use is the Recursive Feature Elimination (RFE); RFE operates by wrapping itself around a machine learning algorithm [77]. We use the random forest as the machine learning algorithm. RFE operates in a way that is different from other filter-based feature selection methods that score each feature and selects the feature with the most significant score (whether that is the largest or smallest) [183]. RFE operates by starting with all the features in the feature set and then removing features until a specific number (typically designated beforehand) of features is reached [183]. In this study, we set this number to be the number full feature set, and we evaluated each number of possible features. This is achieved by fitting the given machine learning algorithm used

in the core of the model, ranking features by importance, discarding the least important features, and re-fitting the model [183]. Features are then scored using the given machine learning model or using a statistical method [183]. In this study, we picked the set of features and the number of features by the lowest root mean squared error.

6.2.5 Machine Learning Methods and Validation for Gait Speed

The primary outcome for gathering influential features is mean gait speed because it is an indicator of morbidity and mortality. Mean gait speed was collected as part of the PRIMA study.

We produced model fits, using random forest [74], by using these methods to fit the training data (a random sample of 80 % of the data) using all the features and the different chosen feature sets. We then fit the resulting model on test data (20% rest of the data). The metric for comparing the model fits the test data was the Root Mean Square Error (RMSE), which is the residuals' standard deviation (prediction errors). We also reported the R-squared values, which represents the proportion of the variance for the outcome that's explained by the AGM features, and the mean absolute error (MAE) values for comparing the model fits [218]. The mtry values, which define the number of variables randomly sampled at each split, for each model fit are also included.

6.2.6 Machine Learning Methods and Validation for Falls

For classifying and predicting falls (a binary outcome), we used a host of machine learning algorithms such as logistic regression, random forest, XGBoost, support vector machine (with linear and radial kernels), simple neural networks, and multi-layer perceptron models. We chose these algorithms because they are common and standard in many hypothesis-driven gait studies [52, 155, 318, 297, 143, 315].

For logistic regression, random forest, support vector machine, and simple neural networks, we implemented the machine learning models under default settings using the caret package in R [184].

To perform XGBoost, this paper used the xgboost package in R, created by Chen et

al. [60]. XGBoost (eXtreme Gradient Boosting) was implemented by Friedman et al. [101]. The package includes an efficient linear model and tree learning models, and this package is primarily used for regression, classification, and ranking. XGBoost is very popular and is known for its speed and high performance [60].

Using the keras and tensorflow packages in R, we used two different configurations for the multi-layer perceptron network model: MLP-1 and MLP-2 [63]. MLP-1 used two dense layers, with 8 hidden units in one layer and two units for the last layer. MLP-2 also used two dense layers, except we used 28 hidden units in the first layer. For these models, we used an adam optimizer and a binary cross-entropy loss function [63].

We performed both hold-out sampling (with 20% of the data as test data) and leave-one-out sampling (with one participant's data as test data) techniques. We reported the leave-one-out sampling results due to its clinical use. Accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and F1-scores are reported.

6.3 Results

The mean gait speed for the sample at baseline was 0.96 m/s. The association between mean gait speed and falls, at baseline, had a Pearson correlation estimate of -0.04. A t-test between gait speed and falls, at baseline, was performed where the null hypothesis is the true correlation is equal to zero: the t-statistic of this test was -0.65 and the p-value is 0.52, indicating that we cannot reject the null hypothesis.

From the feature selection methods, intrinsic approach, removing highly correlated features method, and the recursive feature elimination method, we determined the recursive feature elimination method produced the best result in modeling mean gait speed with the lowest amount of features (Table 6.2). In Table 6.2, we determined the best feature selection result by performing the random forest algorithm.

There were 10 features that were identified through recursive feature elimination. These features were primarily from the groups: statistical features and time-frequency features, which include mean, mode, quantiles, and minimum of the vertical signals, biorthogonal

spline wavelet energy, entropy, and log energy of the acceleration signals, and the root mean square of the mediolateral signals.

Table 6.2: Random forest results with AGM feature sets from different feature selection methods.

Feature Set (# of features)	**RMSE**	**R-squared**	**MAE**
All Features (213)	0.03	0.98	0.02
Features from Intrinsic RF method - IncNodePurity (27)	0.03	0.98	0.02
Features from Intrinsic RF method - %IncMSE (30)	0.05	0.95	0.03
After removing highly correlated features (45)	0.07	0.91	0.05
Features from RFE (10)	0.03	0.99	0.02

Using the ten features we found through feature selection (with gait speed as an outcome), we classified past falls with baseline data using a variety of machine learning algorithms (Table 6.3). The main finding is the that the MLP-1 and MLP-2 models, compared to other common machine learning models, provided a 100 % accuracy using the overall leave-one-out model with the ten features described above.

Table 6.3: Table of classification and prediction validation metrics for each model for the LOO sampling method, using logistic regression, random forest, XGBoost, SVM (support vector machine), simple neural network, and multi-layer perceptron network models.

Classifying Past Falls						
Model	**Accuracy (95% CI)**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**F1 Score**
Logistic Regression	0.46 (0.30 - 0.63)	0.52	0.33	0.64	0.24	0.57
Random Forest	0.69 (0.52 - 0.83)	1	0	0.69	NA	0.82
XGBoost	0.69 (0.52 - 0.83)	1	0	0.69	NA	0.82
SVM (Linear Kernel)	0.56 (0.40 - 0.72)	0.78	0.08	0.66	0.14	0.71
SVM (Radial Kernel)	0.69 (0.52 - 0.83)	1	0	0.69	NA	0.82
Simple Neural Network	0.69 (0.52 - 0.83)	1	0	0.69	NA	0.82
Multi-Layer Perceptron 1	0.73 (0.57 - 0.86)	0.93	0.18	0.76	0.5	0.84
Multi-Layer Perceptron 2	0.44 (0.29 - 0.59)	0.65	0.15	0.51	0.23	0.57

Classifying Past Falls (LOO)						
Model	**Accuracy (95% CI)**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**F1 Score**
Logistic Regression	0.52 (0.44 - 0.59)	0.53	0.48	0.69	0.32	0.6
Random Forest	0.69 (0.62 - 0.75)	1	0.03	0.69	0.5	0.81
XGBoost	0.69 (0.62 - 0.75)	1	0	0.69	NA	0.82
SVM (Linear Kernel)	0.46 (0.39 - 0.54)	0.58	0.2	0.62	0.18	0.6
SVM (Radial Kernel)	0.69 (0.62 - 0.75)	1	0	0.69	NA	0.82
Simple Neural Network	0.69 (0.62 - 0.75)	1	0	0.69	NA	0.82
Multi-Layer Perceptron 1	1	1		1	NA	1
Multi-Layer Perceptron 2	1	1		1	NA	1

6.4 Discussion

Using data from an accelerometer on the L3 location on the trunk, we found a set of AGM features that would be most influential for older adults' gait speed and use those features to predict and classify falls in those older adults. We also implemented several standard machine learning and deep learning algorithms to predict two binary outcomes: the occurrence of falls within the last year of the baseline visit. Out of the 213 features calculated, only a fraction was identified to be influential. Our main finding was that the ten features from the recursive feature selection proved to be most successful in modeling gait speed, which was the outcome chosen to proxy mobility in older adults. Using these ten features, we were able to achieve 100% accuracy in classifying and predicting falls by using multi-layer perceptron deep learning algorithms using a leave-one-out model.

AGMs are ubiquitous and varied through human gait-focused literature. In the machine-learning analysis of gait tasks, using all or a subset of the available AGMs 1) may not be appropriately correlated to specific aspects of the motor skill of walking and 2) could lead to disastrous overfitting, redundant information, and dimensionality issues, and 3) may lead to being vulnerable to adversarial data points. Our results indicate that simple, easy-to-calculate AGMs such as statistical features AGMs or time-frequency AGMs can be sufficient in modeling gait speed and predicting fall-risk [78].

Fall risk is determined by many intrinsic factors relating to mobility, gait, and balance. These factors not only include gait and balance, but they also include audio-visual factors, medication, depression, cognitive impairment, diabetes, and arthritis [315]. Tunca et al. report that understanding, gathering, and analyzing these intrinsic risk factors are quite complex (e.g., many of the factors require numerous medical visits and amassing many medical reports) [315]. These risk factors can be overcome by fewer visits with wearable technology, such as accelerometers. Combining deep learning algorithms, which outperforms most standard machine learning techniques, and clinical assessment of fall-risk can be highly beneficial, while misclassifying patients, especially if they are at high fall risk, can have severe consequences. Thus, our result of complete 100% accuracy with the multi-layer perceptron models is desired. Our results confirm that an individual's gait characteristics can

be differentiated from other individuals' gait characteristics.

Clinically, detecting fall-risk in patients could lead to an increased quality of life for patients as well as decreasing the costs to the healthcare system. While there are other risk factors for falls, such as cognitive decline, medication use, weak muscles, and other signs of aging, early detection based on unsteady gait or other abnormal gait patterns can reduce the risk of future falls.

6.5 Conclusion

We investigated whether we could capture the features that were influential for older adult mobility and whether or not we could detect past and future falls using machine or deep learning approaches applied to three-axis accelerometer data collected during overground walking among older community dwelling adults. With the leave-one-out model, the multi-layer perceptron model algorithm achieved the highest accuracy for both detecting past falls and predicting future falls. This is highly advantageous in clinical environments, where the combination of deep learning models with the leave-one-out approach could help health care professionals determine who is at risk for future falls over repeated visits.

6.6 Appendix

6.6.1 Feature Descriptions

Table 6.4: Table of Statistical Features - Acceleration Gait Measures. Note: ts represents the vector of acceleration signals at the frequency, which represents the frequency of the signal (100 Hz).

AGM Group	Feature Name	Feature Description
Statistical Features	mean_x, mean_y, mean_z	$Meanoftheaccelerationsignal(ML, AP, V)$
Statistical Features	median_x, median_y, median_z	$Medianoftheaccelerationsignal(ML, AP, V)$
Statistical Features	mode_x, mode_y, mode_z	$Modeoftheaccelerationsignal(ML, AP, V)$
Statistical Features	min_x, min_y, min_z	$Minimumoftheaccelerationsignal(ML, AP, V)$
Statistical Features	max_x, max_y, max_z	$Maximumoftheaccelerationsignal(ML, AP, V)$
Statistical Features	q25_x, q25_y, q25_z, q75_x, q75_y, q75_z	$25\%and75\%percentilesoftheaccelerationsignal(ML, AP, V)$
Statistical Features	Iqr_x, iqr_y, iqr_z	$IQRoftheaccelerationsignal(ML, AP, V)$
Statistical Features	sd_x, sd_y, sd_z	$StandardDeviationoftheaccelerationsignal(ML, AP, V)$
Statistical Features	skewness_x, skewness_y, skewness_z	$Skewnessoftheaccelerationsignal(ML, AP, V)$
Statistical Features	kurtosis_x, kurtosis_y, kurtosis_z	$Kurtosisoftheaccelerationsignal(ML, AP, V)$
Statistical Features	correlation_xy, correlation_yz, correlation_xz	$Correlationsbetweentheaccelerationsignals(ML, AP, V)$
Statistical Features	covariance_xy, covariance _yz, covariance _xz	$Covariancebetweentheaccelerationsignals(ML, AP, V)$
Statistical Features	rms_x,rms_y,rms_z	$Rootmeansquaretheaccelerationsignals(ML, AP, V)$

Table 6.5: Table of Information-Theoretic Features - Acceleration Gait Measures.

AGM Group	Feature Name	Feature Description
Information-theoretic features	crossentropy_xy, crossentropy_yz, crossentropy_xz	$CrossEntropybetweentheaccelerationsignals(ML, AP, V).Note: ThisistheKullback-Leibler(KL)divergenceoftwodistributions$
Information-theoretic features	dfa_x, dfa_y, dfa_z	$Detrendedfluctuationanalysis(fractaldynamics)alphavaluesoftheaccelerationsignal(ML, AP, V)[117, 300]$
Information-theoretic features	entropy_x, entropy_y, entropy_z	$Shannonentropyoftheaccelerationsignal(ML, AP, V)$
Information-theoretic features	lzc_x, lzc_y, lzc_z	$Lempel-ZivComplexityoftheaccelerationsignal(ML, AP, V)$
Information-theoretic features	mle_x, mle_y, mle_z	$MaximumLyapunovExponentoftheaccelerationsignal(ML, AP, V)$

Table 6.6: Table of Time-Frequency Features - Acceleration Gait Measures.

AGM Group	Feature Name	Feature Description
Time-Frequency features	dwt_array_x_energy1, dwt_array_x_energy2, dwt_array_x_energy3, dwt_array_x_energy4, dwt_array_x_energy5, dwt_array_x_energy6, dwt_array_x_energy7, dwt_array_x_energy8, dwt_array_x_energy9, dwt_array_x_energy10, dwt_array_y_energy1, dwt_array_y_energy2, dwt_array_y_energy3, dwt_array_y_energy4, dwt_array_y_energy5, dwt_array_y_energy6, dwt_array_y_energy7, dwt_array_y_energy8, dwt_array_y_energy9, dwt_array_y_energy10, dwt_array_z_energy1, dwt_array_z_energy2, dwt_array_z_energy3, dwt_array_z_energy4, dwt_array_z_energy5, dwt_array_z_energy6, dwt_array_z_energy7, dwt_array_z_energy8, dwt_array_z_energy9, dwt_array_z_energy10	$EnergyoftheDaubechieswavelet(family: DaubExPhase)transform(fortenfilters)oftheaccelerationsignal(ML, AP, V)$[68].
Time-Frequency features	dwt_array_x_logenergy1, dwt_array_x_logenergy2, dwt_array_x_logenergy3, dwt_array_x_logenergy4, dwt_array_x_logenergy5, dwt_array_x_logenergy6, dwt_array_x_logenergy7, dwt_array_x_logenergy8, dwt_array_x_logenergy9, dwt_array_x_logenergy10, dwt_array_y_logenergy1, dwt_array_y_logenergy2, dwt_array_y_logenergy3, dwt_array_y_logenergy4, dwt_array_y_logenergy5, dwt_array_y_logenergy6, dwt_array_y_logenergy7, dwt_array_y_logenergy8, dwt_array_y_logenergy9, dwt_array_y_logenergy10, dwt_array_z_logenergy1, dwt_array_z_logenergy2, dwt_array_z_logenergy3, dwt_array_z_logenergy4, dwt_array_z_logenergy5, dwt_array_z_logenergy6, dwt_array_z_logenergy7, dwt_array_z_logenergy8, dwt_array_z_logenergy9, dwt_array_z_logenergy10	$Log-energyoftheDaubechieswaveletttransform(fortenfilters)oftheaccelerationsignal(ML, AP, V)$[68]
Time-Frequency features	dwt_array_x_entropy1, dwt_array_x_entropy2, dwt_array_x_entropy3, dwt_array_x_entropy4, dwt_array_x_entropy5, dwt_array_x_entropy6, dwt_array_x_entropy7, dwt_array_x_entropy8, dwt_array_x_entropy9, dwt_array_x_entropy10, dwt_array_y_entropy1, dwt_array_y_entropy2, dwt_array_y_entropy3, dwt_array_y_entropy4, dwt_array_y_entropy5, dwt_array_y_entropy6, dwt_array_y_entropy7, dwt_array_y_entropy8, dwt_array_y_entropy9, dwt_array_y_entropy10, dwt_array_z_entropy1, dwt_array_z_entropy2, dwt_array_z_entropy3, dwt_array_z_entropy4, dwt_array_z_entropy5, dwt_array_z_entropy6, dwt_array_z_entropy7, dwt_array_z_entropy8, dwt_array_z_entropy9, dwt_array_z_entropy10	$ShannonentropyoftheDaubechieswaveletttransform(fortenfilters)oftheaccelerationsignal(ML, AP, V)$[68]
Time-Frequency features	bsw_energy_x, bsw_energy_y, bsw_energy_z	$Biorthogonalsplinewaveletenergyoftheaccelerationsignal(ML, AP, V)$
Time-Frequency features	bsw_logenergy_x, bsw_logenergy_y, bsw_logenergy_z	$Biorthogonalsplinewaveletlogenergyoftheaccelerationsignal(ML, AP, V)$
Time-Frequency features	bsw_entropy_x, bsw_entropy_y, bsw_entropy_z	$Biorthogonalsplinewaveletentropyoftheaccelerationsignal(ML, AP, V)$

Table 6.7: Table of Signal-Frequency Features - Acceleration Gait Measures.

Feature Group	Feature Name	Feature Description
Signal-frequency features	peak_freq_x, peak_freq_y, peak_freq_z	$Numberoffrequencypeaksoftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	mean_peak_freq_x, mean_peak_freq_y, mean_peak_freq_z	$Meanofthepeakamplitudesoftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	centroid_x, centroid_y, centroid_z	$Centroidofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralprec_x, spectralprec_y, spectralprec_z	$Frequencyprecisionofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralmean_x, spectralmean_y, spectralmean_z	$Meanofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralsd_x, spectralsd_y, spectralsd_z	$Standarddeviationofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralsem_x, spectralsem_y, spectralsem_z	$Standarderrorofthemeanofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralmedian_x, spectralmedian_y, spectralmedian_z	$Medianofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralmode_x, spectralmode_y, spectralmode_z	$Modeofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralq25_x, spectralq25_y, spectralq25_z	$25\%percentileofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralq75_x, spectralq75_y, spectralq75_z	$75\%percentileofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectraliqr_x, spectraliqr_y, spectraliqr_z	$IQRofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralskewness_x, spectralskewness_y, spectralskewness_z	$Skewnessofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralkurtosis_x, spectralkurtosis_y, spectralkurtosis_z	$Kurtosisofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralsfm_x, spectralsfm_y, spectralsfm_z	$Spectralflatnessmeasureofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralentropy_x, spectralentropy_y, spectralentropy_z	$Spectralentropyofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	spectralbandwidth_x, spectralbandwidth_y, spectralbandwidth_z	$Bandwidthofthespectralsignaloftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	HR_array_x, HR_array_y, HR_array_z	$Harmonicratiooftheaccelerationsignal(ML, AP, V)$
Signal-frequency features	ih_x, ih_y, ih_z	$Indexofharmonicityoftheaccelerationsignal(ML, AP, V)$

7.0 Concordance of Acceleration Gait Measures that Measure Smoothness of Walking in Older Adults

Co-authors include: Jessie VanSwearingen, Jennifer Brach, and Ervin Sejdić

7.1 Introduction

The smoothness of walking is not only an important aspect of the motor skill of walking, but it is an important aspect of walking to assess and inform physical rehabilitation [45]. However, along with multiple definitions of smoothness, there are also various metrics that quantify walking smoothness. The smoothness of walking definitions, such as the level of the regularity or intermittency of walking [45], all indicate the continuousness of walking. The measurements that quantify smoothness capture the irregular, or intermittent, movements during the gait cycle [30]. It is difficult to know how smoothness is defined in the field and how it is measured.

There are multiple acceleration gait measures (AGMs), measures derived from accelerometer use, that reflect smoothness, but in this paper, harmonic ratios from trunk accelerometers (estimated for the mediolateral (ML), vertical (V), and anterior-posterior (AP) directions) are considered the validated, gold standard for measuring the smoothness of walking [24, 45]. The harmonic ratio is a well-validated trunk-AGM for the smoothness of walking in a stride. Harmonic ratios are shown to capture the complex acceleration trajectories for each direction of motion (ML, V, and AP) as a single calculable value for a given stride by quantifying the departure from an ideally smooth and symmetrical acceleration pattern [203]. Higher harmonic ratio values represent increased gait stability, improved symmetry, smoothness, and rhythmicity [277]. With increased age, we see a significant impact on the harmonic ratio in the ML direction when walking. Larger harmonic ratios may show a smoother gait pattern while a lower harmonic ratio is shown in those with unsteady gait patterns [44, 90, 201, 293, 342, 288]. In Pau et al., lower harmonic ratios were seen in the

direction of motion (the AP direction) in people with Multiple Sclerosis [255].

Despite being the gold standard, harmonic ratios are inconsistently calculated across the gait literature (Table 1 of [252]), making it difficult to interpret. If proper stride segmentation is not done, harmonic ratios are greatly affected by gait variability [252]. There is no standard as to how harmonic ratios are to be calculated concerning how many harmonics should be included in the indices of harmonics estimation. While the literature agrees upon the number of 20 harmonics, this number seems arbitrary and random (Table 1 of [252]) [299]. Additionally, it is not easy to know how many strides the harmonic ratio needs to be calculated to represent one's walking [252]. The harmonic ratio is also difficult to calculate and requires more calculation steps (e.g., Fourier transform) than the other smoothness measures. Lastly, harmonic ratios may measure step-to-step symmetry rather than smoothness, according to some authors [31, 25, 176, 252]. Smoothness and symmetry are different aspects of the motor skill of walking, and this inconsistency in the literature makes the use of harmonic ratios confusing. These reasons make harmonic ratios very inconsistent and hard to compare amongst different gait studies.

Comparing and contrasting the efficacy of measuring smoothness with other smoothness of walking AGMs may give us more insight into how to evaluate the smoothness of walking. The other smoothness AGMs (mean, maximum, jerk-cost, root mean square, and index of harmonicity) chosen in this paper are derived from AGM categories from Dasgupta et al. [78]. From Dasgupta et al.'s AGM and motor skill of walking categories, the AGM categories that are represented in this paper are statistical features, signal frequency features, and information-theoretic features [78]. However, many of these AGMs are related to more than just smoothness of walking (i.e., in addition to smoothness, these AGMs can also be tied to symmetry, stability, automaticity, efficiency, adaptability, and variability). In this paper, we have various expectations for each of these measures (Appendix Table 7.3), with regard to how they measure smoothness of walking. Each of these AGMs have associations with other aspects of the motor skill of walking as well as having finer associations with different aspects of the smoothness of walking. This study's innovation is to challenge how well other smoothness measures agree with the harmonic ratio since it can be an inconsistently calculated measure.

The aim of this study is to assess other smoothness of walking AGMs' concordance with the harmonic ratio, derived from data collected from a clinical study conducted at the University of Pittsburgh [47]. Typically, with Bland-Altman plots, a concordance analysis is done where a measure is deemed to be the gold standard. In this case, we have determined that the harmonic ratio is that current gold standard [45]. We hypothesize that the various AGMs will show signs of concordance with the harmonic ratio measure (See "Expectations" column in Appendix Table 7.3). We expect that this concordance analysis will give more insight into smoothness and how comparable other AGMs are to harmonic ratios.

7.2 Methods

7.2.1 Procedures

The dataset for this work is from the "program to improve mobility in aging" (PRIMA) study [47]. The PRIMA study consists of 249 community-dwelling older adults (65 years or older) who walk > 0.60 m/s and < 1.2 m/s [47]. Additionally, inclusion criteria also required participants to be ambulatory without an assistive device or the assistance of another person and have physician clearance to participate in a moderate-intensity exercise program [47].

Older adults were recruited from the Pittsburgh Pepper Center Registry of over 2200 older adults, who have consented to be contacted about their interest in participation in mobility research studies. We measured each participant at several time points: baseline, 12 weeks, 24 weeks, and 36 weeks. This study includes analysis on only those participants that have been measured at each of the four timepoints ($n_{baseline}$=203, $n_{12weeks}$=195, $n_{24weeks}$=166, and $n_{36weeks}$=148 for each visit.). Since gait can be altered under observation (Hawthorne effect), we included all results of each visit [75]. Additionally, in order to have a full set of data across all timepoints, we restricted our sample size to the 148 patients who had data from all four timepoints.

7.2.2 Data Collection

We used the acceleration data collected using the Actigraph accelerometers secured on the skin at the level of L3 vertebra of the lumbar spine (Actigraph, LLC.; Pensacola, FL, USA). The acceleration signals were sampled at 100 Hz. The acceleration signals were collected during overground walking. These accelerometers recorded data for the three directions of movement (vertical (V), anterior–posterior (AP), and medial–lateral (ML)).

7.2.3 Data Processing

For each subject, the accelerometer placed on the L3 region produced four variables of data: accelerations in the ML, V, and AP directions ($\frac{m}{s^2}$)) and the time-stamps (s) for each measurement. Harmonic ratios are typically measured via strides [31]. Thus, we portioned each of the acceleration signals for each direction into gait strides, using the method outlined by Sejdić et al. [286].

7.2.3.1 Smoothness of Walking Measures For each of the strides, in each subject and for each visit, we calculated the following five smoothness of walking measures (further descriptions of these measures can be found in Appendix Table 7.3).

- The mean value of the acceleration signals in a stride [324, 79].
- The maximum value of the acceleration signals in a stride [323, 324].
- The jerk-cost metric in a stride is related to the time integral of squared jerk, where jerk is defined as the rate of change in acceleration, or the third time derivative of one's physical position [137]. Jerk-cost is the area under the jerk curve [147]. The formula for total jerk-cost can be found in Schneider et al. [282]. We did not use the magnitudinal or directional definitions of jerk-cost.
- The root mean square (RMS) of the acceleration signals in a stride [225, 290].
- The index of harmonicity (IH) of the acceleration signals in a stride [190, 152]. While doing spectral analysis, we first perform a discrete Fourier transform to estimate the power spectrum of the signal from the raw acceleration data. From the power spectrum, the

index of harmonicity is defined to be the power spectral density of the first harmonic multiplied by the cumulative sum of the power spectral density of the fundamental frequency and the first five superharmonics.

7.2.3.2 Harmonic Ratios of Acceleration Values We will be comparing the above measures with the harmonic ratios of the acceleration signals in a stride. The harmonic ratios for each direction of the acceleration signals were calculated by performing a fast Fourier transform, where we generated the first twenty harmonics of the signal in the stride [265]. The ratio of the sum of the amplitudes of the even harmonics to the sum of the amplitudes of the odd harmonics make up the harmonic ratio for each acceleration signal and for each direction [152]. This calculation has been implemented by many in gait literature [215, 45]. Commonly, we see that higher harmonic ratios (in either of the three directions) represent smoother walking; this finding is corroborated with the fact that higher harmonic ratios mean that the accelerations are in-phase with the stride's oscillations [194].

7.2.4 Data Analysis

To perform data processing, signal processing, the creation of the Bland-Altman plots, and all other mathematical and programmatic analyses, we used R version 4.0.3 within the RStudio integrated development environment (Vienna, Austria) [262]. The R (or "GNU S") programming language and software environment is upheld by the R Foundation for Statistical Computing [262]. The official R software environment is written in C, Fortran, and R, and it is freely available to use under the GNU (GNU's Not Unix!) General Public License [262]. R is used primarily for statistical computing and graphics, and many R packages are available on the Comprehensive R Archive Network (CRAN) [262].

7.2.4.1 Bland-Altman Plots We created a Bland-Atlman plot for each individual and for each smoothness measure. The Bland-Altman plots show the differences between two measures (a smoothness measure and the harmonic ratio measure) vs. the means of those two measures [34]. Thus, the x-axis of the plot are the means of both the measures, and

the y-axis of the plot is the differences between the two measures. Each Bland-Altman plot has 3 horizontal lines; the middle horizontal line is the the mean-of-differences line, while the other two lines are 1.96 standard deviations (95% confidence interval) away from the mean-of-differences line [34]. The two measures are said to be concordant if the majority of the points fall between these the standard deviation lines, and it is even more advantageous to have a mean-of-differences close to 0.

From investigating each Bland-Altman plot for each individual, we created a table to depict how many of the smoothness measures had a mean-of-differences close to zero (which would indicate high concordance).

Bland and Altman recommended that, under a clinically meaningful context, the placement of the limits of agreement can be determined; we used the upper and lower limits of $\pm$ 1 standard deviations [34]. We also tested each mean-of-difference with a one-sample t-test to test whether the mean-of-difference value was significantly different than zero.

We did not perform correlation analysis in this study, because the Bland-Altman analysis sufficiently shows deviation between a smoothness measure and the harmonic ratio [186]. Furthermore, in typical correlation analysis, a high correlation does not necessarily mean that the two measures are highly concordant or in agreement (i.e., a sample could be ubiquitous and extensive and could result in a high correlation).

7.3 Results

Our sample consists of older adults with an average age of 77.29 years and is predominantly female. We also noted that the sample consists of predominantly "White" race (89.9%) with the majority of the sample having up to some form of post-graduate degree (51.4%). The anthropometrics and demographics of this study's 148 participants are shown in Table 7.1.

Table 7.1: Demographic information and a selection of survey data from older adult patients in the PRIMA study. Note: SD = standard deviation.

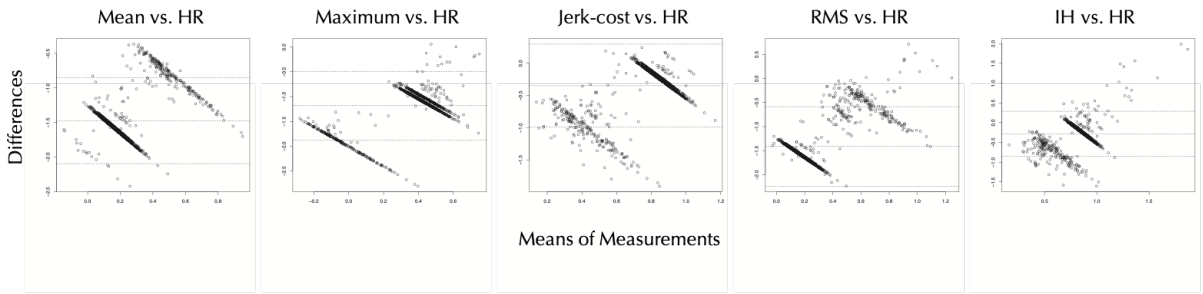
Variables	**n=148**
Age (mean (SD))	77.29 (6.69)
Identified Gender = Male (%)	46 (31.1%)
Height [in] (mean SD))	65.08 (3.99)
Weight [lbs] (mean SD))	170.50 (37.96)
Race (%)	
Black or African American	11 (7.4%)
Native Hawaiian or Other Pacific Islander	1 (0.7%)
White	133 (89.9%)
Refused to Answer	3 (2.0%)
Live Alone = No (%)	75 (50.7%)
Last Grade Completed in School (%)	
High School/Equivalent	21 (14.2%)
College	48 (32.4%)
Post-Graduate	76 (51.4%)
Other	3 (2.0%)

Each subject had a different trajectory of their smoothness measures from baseline to the other timepoints. Following the measurement of all the smoothness measures for each subject, Bland-Altman plots for each smoothness measure were examined (Figure 7.1). Diagonal trends seen on the graphs indicate a significant bias, not necessarily concordance, between the two measures [107]. This bias stems from a constant or average result arising from specific values - similar trends between the two measures may be seen [107]. Diagonal trends can also mean that there is a relationship between the two measures. The Bland-Altman plot only quantifies the bias and a range of agreement.

For the jerk-cost of the AP signals, consistently more than 80% of the mean-of-differences were close to zero for the following timepoints: baseline, 12 weeks, 24 weeks, and 36 weeks (Table 7.2). For the jerk-cost of the ML signals there was moderate concordance with the AP signals of the harmonic ratio than with the ML signals with the harmonic ratio (Table 7.2). For the jerk-cost of the V signals, there was moderate concordance with the AP signals of the harmonic ratio than with the V signals with the harmonic ratio (Table 7.2).

Other notable smoothness measures that showed signs of medium concordance were the mean of the AP signals, maximum of the ML signals, and the maximum of the AP signals (Table 7.2). No concordance demonstrated for the other AGMs with the harmonic ratio of smoothness.

Figure 7.1: Bland-Altman plots for each smoothness measure for a subject at baseline. Each plot's x-axis is the mean of the two measures, while each plot's y-axis is the differences between the two measures. Note: the ML directions for each measure, including the harmonic ratio, are shown.



7.4 Discussion

The accelerometer signal's harmonic ratios have been a well-validated metric to represent smoothness in walking in gait literature [104, 299, 225, 215]. However, harmonic ratios are inconsistently calculated in the literature. We hypothesized that the other smoothness AGMs would be all concordant with the harmonic ratio (Appendix Table 7.3). However, we found that out of all these smoothness measures, using concordance analysis, only the jerk-cost metric, in the AP direction, showed high concordance (mean of difference value close to zero according to the Bland-Altman plot) with the harmonic ratio of the accelerometer signal, for all directions. Other smoothness measures that showed medium concordance were the mean of the AP signals, maximum of the ML signals, and the maximum of the AP signals. This corroborates findings in the gait literature, where mean and maximum of the AP signals were shown to be associated with smoothness, because they were able to describe gait variability and stride-to-stride fluctuations [30]. A notable finding was that the concordance was not for the same direction of the acceleration measure for jerk-cost and the harmonic ratio. While there may be learning or improvement in performance over the different timepoints, that would not affect the concordance of the measures. Thus, the concordance differed from one

time point to another due to the stochasticity of repeated measures over a wide range of individuals.

Table 7.2: Percentage of participants whose mean-of-difference value is close to zero within the cut-points of within one standard deviation from their mean-of-differences (Bland-Altman plot) and is not significantly different than zero (t-test) for the comparison between the various smoothness measures and the Harmonic Ratio for the ML, V, and AP signals (respectively for each column at each visit). High percentages indicate high concordance of those measures, over all 148 patients, with the harmonic ratio.

Smoothness Measure	**Baseline**			**12 weeks**			**24 weeks**			**36 weeks**		
Harmonic Ratio	**ML**	**V**	**AP**	**ML**	**V**	**AP**	**ML**	**V**	**AP**	**ML**	**V**	**AP**
Mean of the ML signals	5%	12%	12%	5%	12%	33%	5%	7%	38%	7%	12%	38%
Mean of the V signals	14%	14%	14%	12%	12%	21%	7%	10%	14%	19%	17%	31%
Mean of the AP signals	52%	52%	52%	38%	40%	43%	33%	33%	36%	43%	45%	50%
Maximum of the ML signals	17%	21%	29%	12%	38%	67%	12%	31%	64%	7%	21%	60%
Maximum of the V signals	14%	14%	14%	12%	12%	24%	10%	12%	19%	19%	21%	38%
Maximum of the AP signals	55%	60%	57%	40%	48%	52%	33%	43%	48%	45%	48%	50%
RMS of the ML signals	7%	10%	21%	14%	29%	55%	10%	10%	38%	7%	12%	33%
RMS of the V signals	2%	5%	2%	5%	7%	14%	5%	10%	19%	5%	10%	21%
RMS of the AP signals	0%	0%	2%	5%	12%	24%	2%	17%	21%	2%	7%	24%
Jerk-cost of the ML signals	52%	60%	74%	31%	43%	71%	43%	50%	76%	19%	33%	74%
Jerk-cost of the V signals	50%	52%	71%	62%	57%	88%	60%	74%	95%	52%	64%	88%
Jerk-cost of the AP signals	83%	86%	90%	90%	86%	98%	98%	93%	95%	100%	98%	100%
Index of Harmonicity of the ML signals	0%	0%	5%	5%	12%	45%	0%	7%	52%	0%	7%	50%
Index of Harmonicity of the V signals	0%	12%	21%	7%	29%	57%	2%	17%	67%	0%	14%	62%
Index of Harmonicity of the AP signals	0%	12%	24%	7%	33%	40%	0%	26%	57%	0%	14%	55%

Jerk-cost in the AP direction was found to be most concordant with all directions of the harmonic ratio. Moderate concordance can be seen in the jerk-cost in the ML and V signals. As mentioned by Fazio et al. and other authors, the repetitive patterns of the body during walking, which include information about the smoothness and variability of one's walking, can be captured by jerk analysis [136, 96]. Typically, from a physics standpoint, jerk-cost

will increase, if there is unsmooth movement [136]. Jerk has been used in other areas, besides gait, to great effect (e.g., swimming, chewing, head movement) [170, 106, 221, 48].

There are a variety of other jerk measures used in studies of walking and human movement. While they each purport to characterize and quantify smoothness, they have widely different results. In Hogan et al., they point out studies in patients with Parkinson's disease and Huntington's disease reported statistically significant abnormalities of jerk measures, but in studies with stroke patients and patients with cerebellar dysfunction did not report statistically significant results with jerk measures [137]. Hogan et al. state that jerk measures may be scaled differently with movement amplitude, duration, and intervals of arrest [137].

Dimensionless measures are more suited to measuring smoothness, because smoothness is an aspect of the motor skill of walking that is different than speed and distance [137]. Jerk-cost, which is derived from jerk, is not dimensionless, and jerk-cost may have some of the biases of not picking up brief periods of arrest, as noted by Hogan et al. [137]. Dimensionless jerk-derived measures may be better suited for measuring smoothness. For example, Brodie et al. demonstrated that the dimensionless ratio, ML/V jerk, demonstrated 89% accuracy in differentiating older from younger people, when measuring dynamic stability. Dynamic stability is associated with smoothness, but it's not exactly the same [78]; however, harmonic ratios have been associated with stability as well [266].

Since harmonic ratio is a dimensionless measure, and if jerk-cost is highly concordant with the harmonic ratio, then we must consider what is being measured: smoothness [45], step-to-step symmetry [31], or dynamic stability [266]? It's possible that all three are being measured to some degree. In order to understand how jerk-cost of the AP signals and the harmonic ratio (in all directions) are similar, we may need to look at granular aspects of smoothness. Perhaps, smoothness of walking when measuring with the harmonic ratio, only captures the rhythmicity of the acceleration patterns. The even harmonics for the AP and V axes from the harmonic ratio calculation represents the in-phase components of the acceleration signal, while the odd harmonics, in those directions, represent the the out-of-phase components [164]. Healthy walking seeks to minimize these out-of-phase components [164]. In a similar way, the AP direction of the jerk-cost may measure rhythmicity (as noted

by Palmerini et al. and Hubble et al., harmonic ratios, RMS, and jerk analyses may measure smoothness and rhythmicity [248, 149]).

Jerk-cost and other jerk analyses of the trunk accelerations may be capturing a measure of whole body movement, just like the harmonic ratios and RMS from trunk accelerations. However, a surprising result was the discordance of the index of harmonicity and RMS measures with the harmonic ratio measure. Index of harmonicity is shown to be discordant with the harmonic ratio, and the potential rationale for this discordance may be due to inconsistencies in the literature of where inertial measuring units or other measurement techniques are used. One of the primary uses of index of harmonicity comes from Lamoth et al. who placed accelerometers on ankles [190]. In a more recent article, Caramia et al. measured index of harmonicity from hand-held smartphones [55]. Another reason for the discordance could be due to differences in measuring symmetry versus smoothness. In Huijben et al., the harmonic ratio results and the index of harmonicity result differed; there was an increased index of harmonicity in the V direction with decreases in the ML and AP directions while the harmonic ratio had an increases in all directions [152]. While Huijben et al. were measuring symmetry with harmonic ratio and smoothness with the index of harmonicity, we believe that these differing results indicate that index of harmonicity does not follow the same trends as harmonic ratio [152].

Similarly, RMS was also discordant with the harmonic ratio in all directions. Menz et al. describes normalized RMS as a measure of dispersion of the data relative to zero [215]. Many studies, including Mizuike et al., have used this description to characterize smoothness, especially with RMS of the trunk-acceleration signals, because they purport that RMS reveals the magnitude of trunk fluctuations without a dependence on the gait speed during walking [222, 14]. This quantification of trunk fluctuations are shown here not to be concordant with the ratio of in-phase to out-of-phase harmonics. Trunk fluctuations may present as fluctuations in the acceleration signal but may not sync up to the transform of the signal (as is used to calculate the harmonic ratio). RMS may be more prone to peaks in the signal data, which wouldn't necessarily corroborate with peaks in the transform of the signal. In this way, RMS may be more associated with the mean and maximum measures, which describe gait variability.

While we purport that all the AGMs chosen in this paper represent smoothness in a general way, the implications of these results are that there may be more granular aspects of smoothness that are not captured by the harmonic ratio or jerk-cost but by the other smoothness measures in this paper. Jerk-cost is also very different than the other measures of smoothness in this paper. For example, it's an information-theoretic AGM, which represents the amount of variability and uncertainty in the information context of a signal [78]. Moreover, mean, maximum, and RMS could represent smoothness but only in interaction with symmetry and variability. IH may need to be calculated alongside the harmonic ratio to make conclusions on a person's smoothness as it pertains to rhythmicity.

RMS may be tied to symmetry and variability in a more impactful way than it is tied to smoothness. Smoothness is also related to other motor skill of walking aspects, such as symmetry; gait metrics that reflect symmetry may also reflect smoothness [78]. However, the smoothness of gait is not equivalent to the symmetry of gait. According to Balasubramanian et al., a person can walk with a symmetric gait but an unsmooth or very intermittent gait pattern [25]. Additionally, according to Dasgupta et al., gait smoothness and symmetry have different definitions. Symmetry is concerned with the agreement of contralateral motion while walking [78]. In contrast, gait smoothness can also be considered the consistent forward progression and regular, repeatable pattern of steps during walking [78]. Thus, we might use RMS as a measure of symmetry more than we use it as a measure of smoothness.

7.4.1 Applications

In instances where health professionals need to measure or quantify smoothness of walking (e.g., stroke recovery), jerk-cost or one of the other moderately concordant measures may be used as a feature or variable in decision-making models or machine learning models. Statistical models, such as these, require informative features. If other features are not as informative as (or concordant with) the harmonic ratio, then it is dubious that the rhythmicity of walking is captured in the model; however, other aspects of the motor skill (or even finer aspects of smoothness) may be captured.

7.4.2 Limitations of the Study

Jerk-cost is a well-known smoothness metric for movement, but it is not often used for various reasons. Jerk-cost can be challenging to calculate because it may include corrections to jerk with respect to distance, and jerk may be influenced by noise in the acceleration signal. Jerk-cost has also been a metric that physical therapists haven't used due to time and environmental constraints [340]. However, we purport that jerk-cost has a more consistent mathematical calculation than harmonic ratio; for example, there is no need to do spectral analysis.

One of our other main limitations is that we did not compare every measure of smoothness in the literature. We limited our analyses to measures that were common in the past ten years of gait literature. For example, Kojima et al. in 2008 proposed a novel measure for smoothness: the power spectrum entropy from acceleration signals during movement [177]. However, this method was not as standard as other statistical, time-frequency, signal-frequency, and information-theoretic acceleration measures. Other methods that are commonly used but were not included in this study is the general entropy measure. While entropy has been reported to have a link to smoothness (large entropies indicate low smoothness because the increase in high-frequency components indicates more fine adjustments of movements), entropy is often more associated with adaptability and variability in walking [340, 78].

Future studies should critically investigate new and old measures of smoothness in the literature with validated measures, such as the harmonic ratio. Other research can include investigating why discordant measures still measure smoothness but do not reveal similar trends and patterns as the harmonic ratio.

7.5 Conclusion

In this study, we acknowledge that the harmonic ratio from trunk-accelerations is a validated measure of the smoothness of walking, and we used concordance analysis techniques

to confirm whether other measures of smoothness used in the literature were concordant with the harmonic ratio. Only jerk-cost showed a level of high concordance. Future work includes the investigation of AGMs in all areas of the motor skill of walking; it will be important to explain the varied associations of the AGMs with each other and attempt to define an approach to the application of AGMs in recognition of human movement control in walking.

7.6 Appendix

Table 7.3: Descriptions of each smoothness of walking measure. Each measure also includes authors' expectations of concordance with the harmonic ratio and the results from this paper.

Measure	Description	Expectations	Results
Mean	The mean of the acceleration signal is often tied to the variability, symmetry, and smoothness of walking. Taking the average of the signals at a given frequency over a stride takes into account of the spread of the acceleration signals. Since mean is tied to the spread of the signals, it may capture gait variability (or how gait fluctuates from one stride to the next) and symmetry (concurrence between the actions and behavior of the lower limbs during walking). Higher mean acceleration values in the AP, ML, and V directions indicate higher levels of walking smoothness [45, 225, 215, 137].	High concordance with the Harmonic Ratio for all directions.	Low concordance with the mean of the ML and V signals with the Harmonic Ratio's ML, V, AP signals. Moderate concordance with the mean of the AP signals with the Harmonic Ratio's ML, V, AP signals.
Maximum	The maximum of the acceleration signal is associated with smoothness. Higher maximum acceleration values in the ML and V directions indicate higher levels of smoothness; opposite trends can be seen in the AP direction [323, 324]. With maximum, the higher the maximum acceleration signals in a stride can indicate high peaks in the signal from which we can hypothesize that this measure will have moderate concordance with the harmonic ratio.	Medium concordance with the Harmonic Ratio for all directions.	Low concordance with the mean of the ML signals with the Harmonic Ratio's ML, V, AP signals. Moderate concordance with the mean of the AP signals with the Harmonic Ratio's ML, V, AP signals.
RMS	The root mean square (RMS) is used for measuring variability, smoothness, symmetry, automaticity, and stability. The root mean square (RMS) of the acceleration signals characterizes gait abnormality (according to Sekine et al.) [290]. Low RMS values in the ML direction are seen in people with impaired balance, while, generally, low RMS is seen in abnormal, unsmooth gait [225, 290]. RMS can represent variability because it captures movement intensity [265]. RMS is also associated with walking speed [223]. Due to these reasons, RMS most likely represents automaticity and smoothness – higher RMS occurs when individuals walk faster with more smoothness.	Medium concordance with the Harmonic Ratio for all directions.	Low concordance with the mean of the ML, V, and AP signals with the Harmonic Ratio's ML, V, AP signals.
Jerk-cost	The jerk-cost metric is associated with stability and smoothness in not just trunk-acceleration studies, but other areas of human movement (e.g., head jerk) [48]. In trunk-AGM analysis, jerk-cost analysis is very similar to jerk analysis and RMS of jerk analysis, because all three measure smoothness. All three analyses are derived from jerk (rate of change in acceleration). Jerk quantifies the accelerations experienced by the trunk during walking; jerk also captures the cyclical movement of the trunk as it decreases and increases speed, rises and falls, and moves from side to side [96, 72]. Lower values of jerk indicate smoothness, and since jerk-cost is the area under the jerk curve, the same association applies: overall, decreased jerk-cost values indicate higher smoothness. Furthermore, maximizing smoothness in body movements may be modelled by minimizing the mean-square jerk, according to Hogan [135]. Jerk RMS, as shown by Fazio et al. and Angellini et al., may have use in assessing rigidity, an aspect of smoothness, while walking; Fazio et al. showed that their healthy subjects had higher jerk RMS values compared to those with ataxia and Parkinson's disease [96, 14]. Jerk RMS has an inverse relationship with jerk (and jerk-cost), as Fazio et al. also states that the time-integrated square jerk decreases with increased smoothness of movement for those with ataxia and Parkinson's disease [96]. In those with ataxia and Parkinson's disease, jerk RMS were lower and more stagnant, and they hypothesize this is due to rigidity [96]. Jerk RMS quantifies the spread of jerk. The RMS of jerk is calculated differently from jerk-cost, but it characterizes the source of jerk by taking the square root of the mean jerk values. In walking, while lower jerk-cost values are preferred, increasing trends of jerk values in the ML and V directions indicate higher levels of smoothness; opposite trends can be seen in the AP direction [96].	Medium to high concordance with the Harmonic Ratio for all directions.	Medium concordance with the mean of the ML and V signals with the Harmonic Ratio's ML, V, AP signals. High concordance with the mean of the AP signals with the Harmonic Ratio's ML, V, AP signals.
IH	The index of harmonicity (IH) is similar to harmonic ratio, because, through spectral analysis, it evaluates the oscillating components to the observed coordination patterns [190]. Lower index of harmonicity values in the ML direction are seen with higher levels of smoothness [190]. A higher index of harmonicity in AP direction has been seen in those with a history of falls [318]. Lower index of harmonicity has also been seen in the V direction in older adults (e.g., persons with a history of stroke and at high risk for falls). [260].	High concordance with the Harmonic Ratio for all directions.	Low concordance with the mean of the ML, V, and AP signals with the Harmonic Ratio's ML and V signals. Medium concordance with the mean of the ML, V, and AP signals with the Harmonic Ratio's AP signals.

8.0 Deep Learning for Fall Detection Using Smartphone Accelerometer Data Among Older Community-Dwelling Adults

Co-authors include: Brian Suffoletto, Adam Frisch, and Ervin Sejdić

This paper is under review at Intelligence-Based Medicine.

8.1 Introduction

Older adults are prone to falls, which can lead to an inordinate amount of injuries, deaths, and financial costs to healthcare systems [33, 275]. According to Florence et al., in 2015 in the United States, the medical costs that were derivable from falls was about 50 billion USD [98]. In healthcare settings, questionnaires and subjective clinical scales are typically used to assess fall-risk behaviors. Rarely, healthcare providers make observations about the facility of movements by asking patients to perform actions such as standing up, walking a few steps, and sitting down. These tests provide some demonstrative evidence of identifying symptoms of cognitive and motor control disabilities [257]. However, these tests are limited by sub-optimal discrimination and unreliable inter-rater variability.

Wearable sensors that measure movement patterns could reduce the subjectivity of fall-risk assessments [229, 83, 257]. In particular, accelerometry used in conjunction with these clinical scales could help provide more objective data to evaluate patients' fall risk [331]. For example, the Timed Get Up and Go test, often abbreviated as the timed GUG test, helps classify binary fall-risk with an accuracy of 63%, but adding accelerometry to the timed GUG test raises the classification accuracy of fallers to 87% [257, 331].

Accelerometry can provide researchers either raw acceleration signals or processed acceleration gait measures (AGMs). Traditionally, the analysis of acceleration data is often a complex process. This data type is time-series data and goes through multiple transformations, such as smoothing and normalization. Moreover, for machine learning processes, this transformed (or "cleaned-up") time-series data goes through segmentation and feature ex-

traction; subsequently, machine learning models use these AGM features as variables. These features are often statistical (e.g., mean) or directly signal-based (e.g., Fourier transforms). Conventional machine learning algorithms use these AGMs to train the models. However, with deep learning, features can be "abstracted" automatically from the raw or processed acceleration signals.

Deep learning is an umbrella term referring to a family of neural network machine learning algorithms [110]. Eponymously, neural networks operate by collecting neurons connected with a set of synapses [110]. Various researchers have used neural networks to abstract features from acceleration signals. Neural networks typically have three fundamental components: the input layer, one or more hidden layers, and the output layer. If the neural network has more than one hidden layer, the neural network becomes a deep network, hence the term "deep learning" [110]. Hidden layers detect "complicated patterns [110]." Neural networks also perform forward- and back-propagation, which are non-exclusively, standard methods to adjust the weights in a neural network. The former applies a set of weights to the input data, and the latter calculates the margin of error of the output and adjusts the input weights accordingly [110]. Neural networks repeat both forward- and back-propagation until the weights are calibrated to predict an output accurately [110]. These weights are adjusted by gradient descent optimization methods, such as Stochastic Gradient Descent (SGD), Adam, or the RMSprop optimizer.

Convolutional neural networks (CNN) are a type of neural network that reduces the parameters from a typical neural network [7]. Notably, CNNs can abstract features as the input propagates through the layers [7]. These reasons are primarily why CNNs are preferred for complicated deep machine learning tasks, such as image or signal classification [7]. For example, Lee et al. show that CNNs outperform a standard machine learning approach, random forest [195]. Additionally, Santos et al. came up with an accelerometer-based framework using CNNs on three different datasets to improve fall detection in clinical spaces [275]. CNNs are particularly effectual with images [298]; their effectiveness may reach other data types. Addabbo et al. also used convolutional neural networks, specifically temporal convolutional neural networks (TCNN), to identify individuals based on the gait characteristics [4]. Incorporating different forms of deep learning tactics like CNN on biomedical data can

help tackle common biomedical dataset problems, such as imbalanced class labels, missing data, and high-dimensional low sample data [175].

Using acceleration signals from an accelerometer with deep learning may be able to pick up on human movements that conventional machine learning models cannot. In the literature, human gait has been considered a biometric trait - the identity and behavior of individuals based on their walking can be identified and classified, using deep learning models [110]. According to Horst et al., deep learning models, such as artificial neural networks, have been used to analyze human walking based on time-continuous gait patterns [139]. Applications of artificial neural networks have featured the following insights of deep learning gait analysis: gait patterns are different for individual people, emotional states can be identified, and gender and age-specific gait patterns can be classified [139]. In fact, many studies have successfully identified and categorized specific gait patterns into contextual, clinical groups [139, 58, 284].

We were motivated to use CNNs, RNNs, and ultimately CRNNs because of prior research stating that that gait patterns were detected by CNNs and LSTMs [10]. CNNs are capable of handling 1D, 2D, or 3D data, and CNNs are able abstract features from large datasets by implementing a convolution operation to the inputs. LSTMs (a type of RNN) are really good for time-sequence data because order is important (for the recursive process). Specifically, we wanted to see patterns of the inertial gait time series that were indicative of low motor skill of walking that might lead to a fall. With a CNN, there is a sequence of 1-D kernels that guarantees the output convolution features retain the properties of the input (1-second window time series data), then a RNN could be used process the resulting features for gait recognition [354].

This study aimed to evaluate a deep learning framework applied to inertial (i.e., accelerometry) data from smartphone sensors worn during a brief walking task to detect a history of falls and predict future falls in older adults receiving care in the emergency department. We hypothesized that hybrid CNNs could detect past and predict future falls over a 90 day period. We did not design the study to examine prediction of short-term falls (e.g. 7 days) our data did not provide such a granularity. Our experiments evaluated the performance of the following deep learning models: CNNs, recurrent neural networks,

hybrid convolutional recurrent neural networks, and conventional machine learning models (support vector machine, logistic regression, and random forest) using processed AGMs and accelerometer data. Both processed AGMs and accelerometer signals were used as inputs to these machine learning models, for comparison purposes.

8.2 Materials and Methods

8.2.1 Data Collection and Processing

We recruited volunteers for this study from patients from an urban emergency department in Pittsburgh, Pennsylvania, USA. Participation in the study was entirely voluntary, completed informed consent, and patients were allowed to discontinue at any time. We included a convenience sample of 166 older adults (60-96 years old) who were patients at the emergency department between May 2019 and October 2019. For each participant, we performed a 90-day follow-up phone call to ascertain if there was a fall within the 90-days post-enrollment. Table 1 shows the patients' demographic characteristics and a selection of the patients' responses to the baseline assessment. Of these 166 patients, 153 patients had acceleration data collected. Of these 153 patients, only 134 patients had data for both past falls (self-reported at the baseline assessment) and future falls (self-reported at the 90-day follow-up phone call) and thus, were included in the analyzed cohort. This leads us to two outcomes for patients: past falls and future falls.

During their emergency department visit, we affixed a smartphone on the lower back of the participants using an elastic belt. This smartphone captured linear accelerations (in units of $\frac{m}{s^2}$) at a sampling frequency set to 100 Hertz from the x, y, and z directions, which correspond to the mediolateral (ML), vertical (V), and anterior-posterior (AP) directions (Figure 8.1-A and B). While the sample frequency was set to 100 Hz, because of sampling jitter, the actual sampling frequency may be lower. In those cases, we re-sampled the signal to 100 Hz.

After the smartphone's placement, we asked the participants to complete a walking trial:

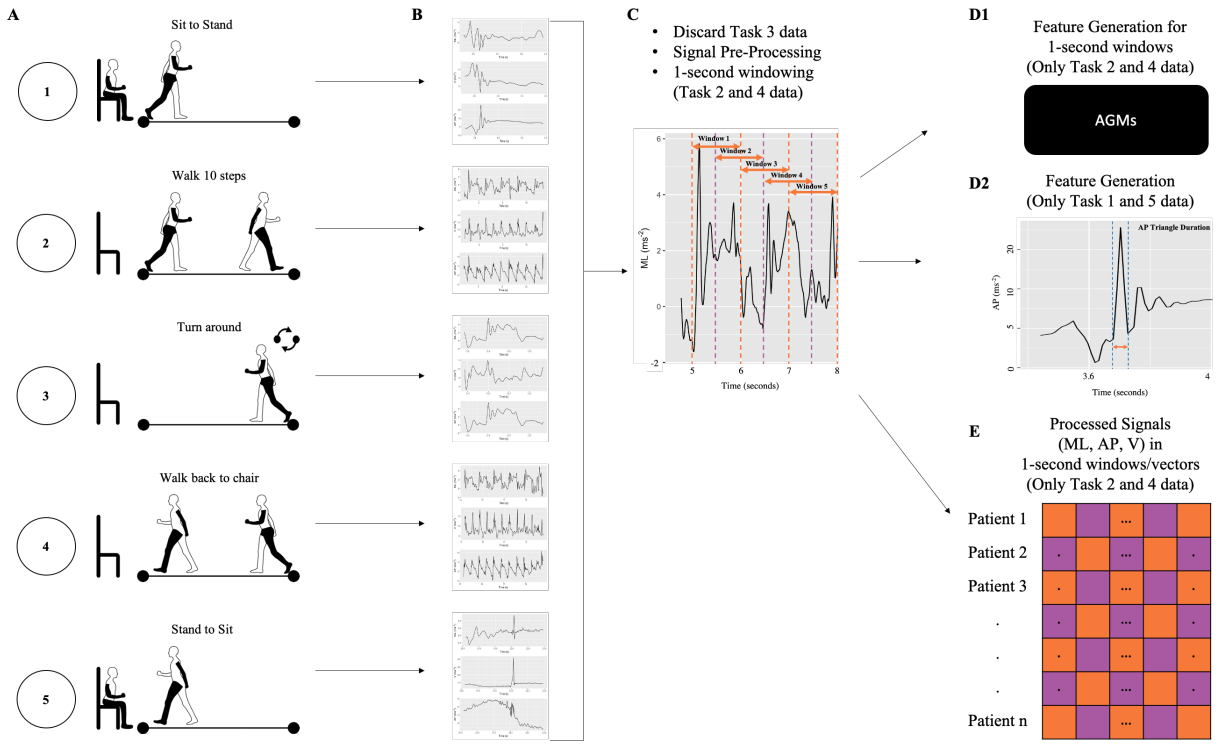
a timed GUG test (Figure 8.1-A). The GUG consisted of starting from a sitting position on a chair and then doing five tasks consecutively: (1) Stand up from the sitting position. (2) Walk ten steps forward. (3) Turn around 180 degrees. (4) Walk ten steps back to their original location. (5) Sit down on the chair (Figure 8.1-A and B). During this walking trial, we manually coded these tasks: 1, 2, 3, 4, and 5, respectively (Figure 8.1-A and -B).

When the participant completed the walking trial, we removed the smartphone from the belt. The accelerometer data from the smartphone was retrieved using the phyphox app (www.phyphox.org). The data from the smartphone was downloaded to a secure file and coded by an identification number.

Table 8.1: Demographic information and a selection of survey data from older adult patients in the emergency department.

Variables	**n = 134**
Age (mean (SD))	69.01 (8.11)
Identified Gender = Woman (%)	53 (39.6)
Fall(s) in the Last Year (%)	
Fallen in the Last Year	59 (44.0)
Have Not Fallen in the Last Year	75 (56.0)
Fall(s) 90 days after ED Visit (%)	
Fallen	14 (8.4)
Have Not Fallen	120 (72.3)
Fall Frequency in the Last Year (%) (n=59)	
Once	34 (58.6)
A few times	22 (37.9)
Monthly	1 (0.02)
Weekly	2 (3.4)
Fall Injuries in the Last Year = No (%)	32 (55.2)
Fall Worry in the Last Year = No (%)	39 (29.1)

Figure 8.1: A high-level overview of the data collection and signal pre-processing. (A) GUG tasks. (B) Raw acceleration signals for each GUG task. (C) Signal pre-processing for tasks 2 and 4, which includes an example of 1-second windowing with 50% overlap. Task 3 data was removed. (D1) AGM feature generation for each 1-second window, for tasks 2 and 4. (D2) AP triangle duration for tasks 1 and 5. (E) Processed acceleration signals for each 1-second window, for each patient.



Each accelerometer dataset consists of three streams of dependent variables (ML, V, and AP accelerations ($\frac{m}{s^2}$)) by the independent variable (time (s)) (Figure 8.1-B). Based on literature for classifying older adults' fall behavior based on wearables, we analyzed each participant's accelerometer data in the following way: the coded regions 1, 2, 4, and 5 were re-sampled to 100 Hz and filtered through a fifth-order, low pass Butterworth filter with a 12.5 Hz cut-off frequency [141] (Figure 8.1-C). A frequency response plot of the

Butterworth filter we used is depicted in the Appendix Figure 8.3. These pre-processing steps were crucial in reducing external noise sources from the accelerometer signal [110]. Subsequently, we segmented the coded regions 2 and 4 into windows of 1-second, each with a 50% overlap between two consecutive windows (Figure 8.1-C). Subsequently, we centered and scaled these 1-second windows. We used these 1-second windows as input vectors for the machine learning, based on previous literature using ranges of 0.1 - 10 sec windows for gait classification and prediction [234].

For each of the 1-second windows, we calculated a total of 26 features. This resulted in each of these 1-second intervals containing 26 elements in a single vector. These features include the descriptive statistical AGM features such as mean, maximum, standard deviation, the maximum difference (including maximum three-way difference), pair-wise correlation, and pair-wise covariance of the acceleration signals (ML, AP, V) [141] (Figure 8.1-C). We also included the harmonic ratio, which is the result of performing a fast Fourier transform of the acceleration signals (ML, AP, V); we generated the first twenty harmonics (Figure 8.1-C), and from these harmonics, we calculated the ratio of the sum of the amplitudes of the even harmonics (2,4,6, ..., n) to the sum of the amplitudes of the odd harmonics (1,3,5, ..., n) [141] (Figure 8.1-C). For coded regions 1 and 5, we computed the "AP triangle" (description given in [267]) (Figure 8.1-C). As shown in Rivolta et al. and Figure 8.1-C, the "AP triangle" is derived from the sit-to-stand and stand-to-sit activities where the AP signal has a local maxima signal amplitude peak, which looks like the titular triangle [331, 267]. The calculation for the "AP triangle" is to evaluate 1) the maximum slope of the triangle, 2) the minimum slope of the triangle, and 3) the acceleration peak of the triangle. This study uses the duration of the "AP triangle" as a feature. For task 3, we computed the step frequency by calculating the frequency peak from the power spectral density of the ML signal [267] (Figure 8.1-C). We also included the time to complete the GUG test both as a feature and as a possible interaction term.

For feature selection completed on the training set, we used the Wald test. The Wald test calculates the statistical significance of each coefficient in the model. If the test failed to reject the null (p-value > 0.05 used), those variables were removed from the model. The null hypothesis is that the coefficient of the variable in the model is significantly different

from zero.

8.2.2 Machine Learning Methods

Two parent datasets were used for testing and training the machine learning models: 1) 1-second windows of processed acceleration signals from the ML, V, and AP directions described above (Figure 8.1-C); and 2) vectors of AGMs which were calculated from the 1-second windows of the first parent dataset (Figure 8.1-D1, -D2, and -E). We implemented machine learning models to detect past falls separately and predict future falls (both binary outcomes). We used the leave-one-out testing model to capture how well other subjects' features and data can predict an individual's outcome. By combining all subjects' datasets, we trained each model on $n-1$ out of the n subjects, and then we tested the model on the remaining subject.

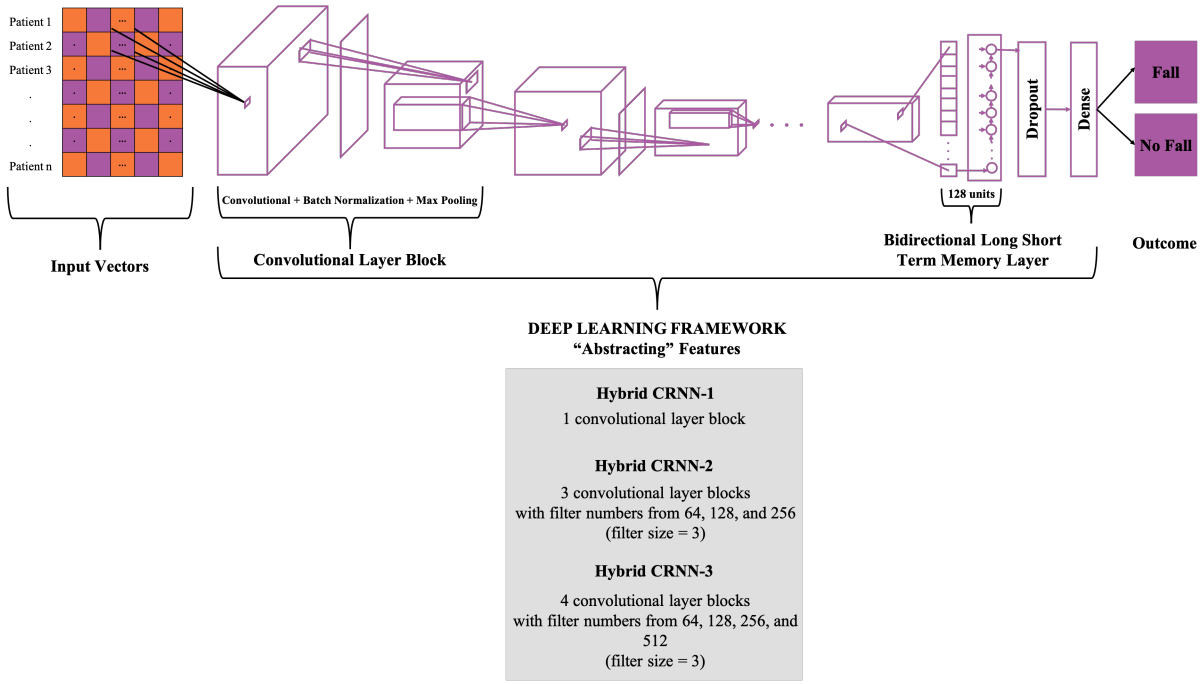
8.2.2.1 Deep Learning Models In our study, we implemented deep learning networks (neural networks, recurrent neural networks, CNNs, and hybrid convolutional recurrent neural networks) using the open-source packages of *tensorflow* and *keras* in R.

Our inputs are either in the form of acceleration data in 1-second vectors (overlapping 0.5 seconds), or the inputs are vectors consisting of acceleration gait measures as features from 1-second time intervals. Figure 8.2 shows an example of one of the input vectors from our training data. Each row of the input is a 1-second vector. The size of the tensor varies for each subject's inputs based on how large or small their task 2 and 4 data was. For example, subject 1's training tensor size was 665 by 2. In Table 8.2, the dimensions of the convolutional processes (for a CRNN) is shown as an example.

Neural network models work like how a brain or other sensory organ works. The analogy is like this: an input affects the brain, and that information is then directed through layers of neurons. These neurons are connected to other neurons, and information passes between them. A layer can be considered an output layer if a set of specific neurons receives enough information (or stimuli).

Analogies aside, the neural network model used in this analysis is a basic model with

Figure 8.2: Generalized, abstract architecture of the deep learning framework of the hybrid convolutional recurrent neural network model. Three configurations were formulated for this study.



one input layer, one hidden layer, and an output layer. Generally, neural networks are a compilation of mathematical operations. Each operation has an input and output (or tensors, which are in vector or matrix form). Each layer of the network has many tensors stored in it, and each layer is relational to the other layers via a weight system. The tensors in each layer are also called nodes; nodes are connected other nodes in another layer via weights. Each layer's output is determined by the activation function of the values in the layer multiplied by the weights.

Table 8.2: An example of the output shape of each of the layers and the parameter size for subject 1 (whose input layer was 665 by 2).

Layer	**(type)**	**Output Shape**	**No. of Parameters**
conv1d_2	(Conv1D)	(None, 1080, 24)	600
conv1d_1	(Conv1D)	(None, 1073, 48)	9264
batch_normalization_2	(BatchNormalizatio	(None, 1073, 48)	192
spatial_dropout1d	(SpatialDropout1D)	(None, 1073, 48)	0
conv1d	(Conv1D)	(None, 1066, 12)	4620
global_average_pooling1d	(GlobalAverageP	(None, 12)	0
batch_normalization_1	(BatchNormalizatio	(None, 12)	48
dropout_1	(Dropout)	(None, 12)	0
dense	(Dense)	(None, 48	624
batch_normalization	(BatchNormalization)	(None, 48	192
dropout	(Dropout)	(None, 48	0
dense_output	(Dense)	(None, 3)	147
Total			15,687
Trainable			15,471
Non-trainable			216

Recurrent neural networks differ from the general neural network models because they equally perform the same calculation for every input sequence element, or in other words, there are connections, which form a directed graph, between neurons with the same weights. Recurrent neural networks do not need to be fully connected. The outputs are dependent

on the prior calculation [110]. In our study, we use bidirectional Long Short Term Memory networks, a type of bidirectional recurrent neural network that uses a "gating mechanism to allow better modeling of long-term dependencies in the time-series data" [110]. Moreover, our inputs are the generated features or the 1-second acceleration windows, ten hidden layers, and two outputs. Weights are automatically hyper-tuned using the *caret* package in R.

CNNs is a type of fully fitted neural network where each neuron in one layer is connected to all the neurons in the next layer. The architectures of the convolutional recurrent neural network models used in this study consist of the following components: inputs, convolution layers, max-pooling layers, dropout layers, and output layers. The inputs are passed through a stack of convolutional layers. These convolutional layers have kernel sizes and filters. The kernel sizes we use are three by three, whereas the number of filters is increased with each convolutional layer. After each convolutional layer, we add a max-pooling layer with a size (2,2). Max-pooling is done to extract a more extensive set of features and, more specifically, select the most significant feature value from each set of feature vectors from the convolutional layer. The max-pooled result is concatenated to create a wide feature vector for each input vector. According to Giorgi et al., the previous three layers can be construed as a "convolutional layer block" and can be applied as often as needed [110]. Each wide feature vector from the max-pooling layer contains the most noticeable and pertinent extracted feature. After each convolutional layer, we added a batch normalization layer to expedite the training process. The batch normalization layer speeds up the training process as a result of improvements in parameter convergence during training. Then we use a bidirectional Long Short Term Memory layer followed by a dropout and dense layer. We primarily tested two hybrid convolutional recurrent neural networks: Hybrid convolutional recurrent neural network-1 and hybrid convolutional recurrent neural network-2 (Figure 8.2). A third hybrid convolutional recurrent neural network (hybrid convolutional recurrent neural network-3) tests whether more layers affect performance. Hybrid convolutional recurrent neural network-1 uses one convolutional layer, a batch normalization layer, a max-pooling layer, a bidirectional Long Short Term Memory layer with 256 units, a dropout layer, and a dense layer (Figure 8.2). Hybrid convolutional recurrent neural network-2 consists of several convolutional layers with filters from 64 to 256 (increasing by a factor of 2), separated by

batch normalization layers and max-pooling layers, then a bidirectional Long Short Term Memory layer with 128 units, dropout layer, and dense layer (Figure 8.2). Hybrid convolutional recurrent neural network-3 has several convolutional layers with filters from 64 to 512 (increasing by a factor of 2), separated by batch normalization layers and max-pooling layers, and then a bidirectional Long Short Term Memory layer with 128 units, dropout layer, and dense layer (Figure 8.2).

8.2.2.2 Baseline Algorithms We have chosen three standard or conventional machine learning classifiers to compare the performance of the deep learning algorithms: logistic regression, support vector machine, and random forest. All three approaches only use the generated AGM features as inputs (in other words, we did not perform these machine learning algorithms using just the acceleration signals from the 1-second acceleration windows). We tested these classifiers with all the features in the feature set and important features from the Wald test feature selection. We used the *caret* package in R to implement the other machine learning algorithms to compare against the deep learning models.

Logistic regression, a generalized linear model with a linear boundary, is a straightforward machine learning classifier because it is often used in the predictive analysis where there is a binary dependent variable (whether there was a fall or not). The model gives the relationship between this binary dependent variable (the odds of a fall vs. the odds of not a fall), and the predictor variables (the feature set), where β are the parameters and x represents the independent predictor variables (or features), as shown in the equation below (Equation 8.1). Odds are a measure of probability or likelihood of an event (here, it is a fall). In our study, odds are calculated as a ratio of the number of events that lead to a fall to the number of events that do not (Equation 8.1).

$$Log(odds) = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n \tag{8.1}$$

The support vector machine is a popular machine learning algorithm, and it operates by finding an optimal hyperplane that linearly separates the data points. Overall, support vector machines aim to maximize the training set's geometric margin and minimize the training error. The hyperplane equation splits the space into two areas, creating a decision

boundary that is often used to classify a binary outcome [276]. The support vector machine algorithm employs the Lagrange function, a linear kernel function, and a Lagrange multiplier in our study. The dataset is then mapped to a multidimensional space to become linearly separable [276].

The random forest algorithm is a bagging-based ensemble learning method that uses the Classification and Regression Tree algorithm to produce trees; the random forest algorithm splits nodes using the best split among a random subset of features [3]. Then the algorithm extends to several parallel trees. Generally, random forests are very robust and provide high accuracies in classification.

8.2.2.3 Validation Overall, validation techniques are used to evaluate the generalization of a trained classifier on a test set. We used a leave-one-out model for testing, with the binary outcome (fall vs. no fall) for each of the classification tasks (detecting past falls and predicting future falls). For each participant in the study, we trained on all the other participants. During the classifier's learning or training process, we ran 10-fold cross-validation on this training group to find the model's best parameters. Then, we tested the model on the participant that was left out of the training. The participant's predicted class was labeled. After doing this with each participant, we created a confusion matrix for all 134 participants. The confusion matrix evaluates the performance of a classifier or algorithm. The confusion matrix contrasts the algorithm's results from the predictions with the true classification labels; these result in the following sub-samples: true positives, true negatives, false positives, and false negatives [310]. We calculated and reported the true positive rate or sensitivity, true negative rate or specificity, and accuracy for our binary outcome [310].

Furthermore, we conducted two-way McNemar tests to compare the classifiers. McNemar tests check the discrepancies between two classifiers and checks if there is a significant difference between those counts [188]. The null hypothesis of the McNemar test is that the models will make errors in similar proportions [188].

8.2.3 Computing Environment

Computationally, we used the program, RStudio, which ran R version 3.5.2 and the following R packages: caret, signal, stats, and seewave [262]. The computer used to analyze the data had 2.2 GHz Intel Core i7 processor with 16 GB RAM and operated on a macOS Mojave operating system.

8.3 Results

The overall mean time to finish the GUG test was 26.55 seconds, with longer completion times among those who fell compared to no fall in the past year (27.12 vs. 26.12 sec; t-test p-value: 0.31). The mean time to finish the GUG test for those who fell within 90 days was longer than those who did not fall within 90 days (27.86 vs. 26.40 secs; t-test p-value: 0.38).

In the training phase of the hybrid CRNNs, we used a RMSprop optimizer. The training time for one iteration of hyperparameter tuning takes 66.5 minutes. The number of epochs were 200. Each layer consisted of several parameters that were hypertuned (e.g., 7,341,507 parameters for all the layers combined for one iteration).

In the testing phase, we evaluated and compared the performance of deep learning networks and standard machine learning models that used various input vectors during training. The input vectors used were 1-second windows and generated features from those 1-second windows. The overall mean number, over all 134 participants, of 1-second windows was 24 windows. The mean number of 1-second windows for those who fell in the past year was 24.90 windows. The mean number of 1-second windows for those who did not fall in the past year was 23.28 windows. The mean number of 1-second windows for those who fell within 90 days was 26.73 windows. The mean number of 1-second windows for those who did not fall within 90 days was 23.68 windows. The statistical spread of the 1-second windows was similar between the two groups (Figures 8.4 and 8.5) further indicating no significant difference of the GUG test timing between fallers and non-fallers for both before and after the ED visit.

For detecting past falls and future falls, after feature selection, the critical features (based on the Wald test) in the model were mean of ML signals, mean of AP signals, the standard deviation of V signals, the standard deviation of AP signals, the correlation between ML and V signals, the correlation between ML and AP signals, the correlation between V and AP signals, the covariance between V and AP signals, the covariance between ML and AP signals, the covariance between ML and V signals, the maximum value of the ML signal, AP duration during the coded 1 region, and AP duration during the coded 5 region (Table A1).

The sensitivity, specificity, and accuracy for our binary outcome are shown in (Table 8.3 and Table 8.4). Moreover, Table 8.5 and Table 8.6 show the results of the two-way McNemar tests between the performances of all the classifiers; if the p-value is less than 0.05, the null hypothesis is rejected, and both models have a different relative proportion of errors on the test set.

Table 8.3: Results from detecting past falls using generated features and acceleration signals using the leave-one-out model.

Models	Detecting Past Falls									
	Generated AGM Features (before feature selection/after feature selection)					Acceleration Signals				
	Sensitivity	Specificity	Accuracy	F1-score	No. Correctly Identified (n=134)	Sensitivity	Specificity	Accuracy	F1-score	No. Correctly Identified (n=134)
LR	0.30/0.30	0.30/0.30	0.09/0.07	0.02/0.02	12/10	-	-	-	-	-
SVM	0.34/0.17	0.34/0.17	0.08/0.04	0.02/0.02	11/5	-	-	-	-	-
RF	0.34/0.40	0.34/0.40	0.06/0.27	0.02/0.03	8/36	-	-	-	-	-
NN	0.30/0.46	0.30/0.40	0.09/0.30	0.02/0.03	12/40	0	0	0	0	0
RNN	0.60/0.59	0.71/0.71	0.69/0.69	0.60/0.60	93/92	0.30	0.30	0.09	0.02	12
CNN	0.75/0.75	0.76/0.76	0.75/0.75	0.60/0.60	100/100	0.51	0.40	0.49	0.45	66
Hybrid CRNN1	0.82/0.82	0.86/0.86	0.84/0.84	0.76/0.76	112/112	0.86	0.83	0.70	0.72	94
Hybrid CRNN2	0.82/0.82	0.86/0.86	0.84/0.84	0.76/0.76	112/112	0.96	0.93	0.92	0.86	123
Hybrid CRNN3	0.82/0.82	0.86/0.86	0.84/0.84	0.76/0.76	112/112	0.96	0.93	0.92	0.86	123

Table 8.4: Results for predicting future falls from using generated features and acceleration signals using the leave-one-out model. Note: LR=Logistic Regression. SVM=Support Vector Machine. RF=Random Forest. NN=neural network. RNN=recurrent neural network. CNN=convolutional neural network. CRNN: hybrid convolutional recurrent neural network.

Models	Predicting Future Falls									
	Generated AGM Features (before feature selection/after feature selection)					Acceleration Signals				
	Sensitivity	Specificity	Accuracy	F1-score	No. Correctly Identified (n=134)	Sensitivity	Specificity	Accuracy	F1-score	No. Correctly Identified (n=134)
LR	0.60/0.93	0.64/0.91	0.62/0.91	0.60/0.90	84/123	-	-	-	-	-
SVM	1.0/1.0	0.99/0.99	0.99/0.99	1.0/1.0	133/133	-	-	-	-	-
RF	0.96/0.97	0.96/0.97	0.90/0.93	0.96/0.97	121/125	-	-	-	-	-
NN	0.86/1.0	0.93/0.98	0.92/0.98	0.90/1.0	124/132	0.41	0.48	0.32	0.40	43
RNN	0.82/0.78	0.90/0.92	0.76/0.85	0.75/0.80	102/114	0.40	0.58	0.36	0.40	48
CNN	0.85/0.85	0.82/0.82	0.82/0.82	0.80/0.80	111/111	0.90	0.58	0.73	0.79	98
Hybrid CRNN1	0.90/0.90	0.93/0.93	0.92/0.92	0.89/0.90	123/123	0.96	0.88	0.88	0.93	118
Hybrid CRNN2	1.0/1.0	0.99/0.99	0.99/0.99	0.99/1.0	133/133	0.99	0.93	0.97	0.98	130
Hybrid CRNN3	1.0/1.0	0.99/0.99	0.99/0.99	0.99/1.0	133/133	0.99	0.93	0.97	0.98	130

Table 8.5: Two-way McNemar Tests of Model Significance done between the all models against all other models (for classifying past fall). Significant results (alpha level = 0.05) are bolded. Note: LR=Logistic Regression. SVM=Support Vector Machine. RF=Random Forest. NN=neural network. RNN=recurrent neural network. CNN=convolutional neural network. CRNN: hybrid convolutional recurrent neural network.

	McNemar Comparison (test statistic (p-value)) (before feature selection/after feature selection)								
Models	**LR**	**SVM**	**RF**	**NN**	**RNN**	**CNN**	**Hybrid CRNN1**	**Hybrid CRNN2**	**Hybrid CRNN3**
LR	-	0 (1) / 0 (0.062)	**0 (0.002) / 0 (0.008)**	0 (1) / 0 (0.125)	0 (0.5) / 0 (0.25)	0 (0.5) / 0 (0.5)	0 (1) / 0 (1)	0 (1) / 0 (1)	0 (1) / 0 (1)
SVM	-	-	**0 (0.004)** / 0 (0.125)	0 (1) / 0 (1)	**0 (0.004)** / 0 (1)	**0 (0.004)** / 0 (1)	0 (1) / 0 (1)	0 (1) / 0 (1)	0 (1) / 0 (1)
RF	-	-	-	**0 (0.031)** / 0 (0)	0 (0.125) / **0 (0.002)**	0 (0.125) / **0 (0.002)**	0 (1) / 0 (0.062)	0 (1) / 0 (0.062)	0 (1) / 0 (0.062)
NN	-	-	-	-	0 (0.5) / **0 (0.004)**	0 (0.5) / 0 (0)	**0 (0.001) / 0 (0.016)**	**0 (0.001) / 0 (0.016)**	**0 (0.001) / 0 (0.016)**
RNN	-	-	-	-	-	**0 (0.001) / 0 (0.008)**	**0 (0.001) / 0 (0.002)**	**0 (0.001) / 0 (0.002)**	**0 (0.001) / 0 (0.002)**
CNN	-	-	-	-	-	-	**0 (0.001) / 0 (0.001)**	**0 (0.001) / 0 (0.001)**	**0 (0.001) / 0 (0.001)**
Hybrid CRNN1	-	-	-	-	-	-	-	0 (1) / 0 (1)	0 (1) / 0 (1)
Hybrid CRNN2	-	-	-	-	-	-	-	-	0 (1) / 0 (1)
Hybrid CRNN3	-	-	-	-	-	-	-	-	-

Table 8.6: Two-way McNemar Tests of Model Significance done between the all models against all other models (for predicting future fall). Significant results (alpha level = 0.05) are bolded. Note: LR=Logistic Regression. SVM=Support Vector Machine. RF=Random Forest. NN=neural network. RNN=recurrent neural network. CNN=convolutional neural network. CRNN: hybrid convolutional recurrent neural network.

McNemar Comparison (test statistic (p-value)) (before feature selection/after feature selection)									
Models	**LR**	**SVM**	**RF**	**NN**	**RNN**	**CNN**	**Hybrid CRNN1**	**Hybrid CRNN2**	**Hybrid CRNN3**
LR	-	0 (1) / 0 (1)	0 (1) / **0 (0.002)**	0 (1) / 0 (1)	**0 (0.008)** / 0 (0.5)	0 (1) / 0 (0.5)	0 (1) / 0 (1)	0 (1) / 0 (1)	0 (1) / 0 (1)
SVM		-	**0 (0) / 0 (0.004)**	**0 (0.004)** / 0 (1)	0 (0) / 0 (0.5)	0 (0) / 0 (0.5)	**0 (0.001)** / 0 (1)	**0 (0.001)** / 0 (1)	**0 (0.001)** / 0 (1)
RF			-	0 (1) / **0 (0.031)**	0 (0) / 0 (0.125)	**0 (0.002)** / 0 (0.250)	0 (0.250) / 0 (1)	0 (0.250) / 0 (1)	0 (0.250) / 0 (1)
NN				-	0 (0) / 0 (0.5)	0 (0) / 0 (0.5)	0 (0.062) / 0 (1)	0 (0.062) / 0 (1)	0 (0.062) / 0 (1)
RNN					-	0 (1) / **0 (0.008)**	0 (1) / **0 (0.001)**	0 (1) / **0 (0.001)**	0 (1) / **0 (0.001)**
CNN						-	0 (1) / **0 (0.001)**	0 (1) / **0 (0.001)**	0 (1) / **0 (0.001)**
Hybrid CRNN1							-	0 (1) / 0 (1)	0 (1) / 0 (1)
Hybrid CRNN2								-	0 (1) / 0 (1)
Hybrid CRNN3									-

8.4 Discussion

This study evaluated several deep learning models and standard machine learning models to predict two binary outcomes: the incidence of past fall(s) and future fall(s) within 90 days of an emergency department visit for a sample of older community-dwelling adults. We used a smartphone device on the waist to measure acceleration signals in the ML, AP, and V directions while patients performed a timed GUG test. Our main finding was that the deep learning model we developed, a hybrid convolutional recurrent neural network model, had the highest accuracy for detecting past falls and predicting future falls.

The participants in our study were patients who came to the emergency department for non-fall-related reasons. This environment and setting introduce a bias in discovering fall-related behaviors. This sample represents realistic clinical constraints – the sample contains those who were not only emergency department patients but also those over the age of 60. This selection resulted in having a lower-than-ideal sample size, reducing the statistical power of our study. Ideally, a large number of training samples are required for deep learning, and

our sample size may have impacted our results (e.g., overfitting). While we don't have proof that we overcame overfitting, we tested on unseen data to prevent it. Furthermore, using convolutional layers on the data in our deep learning framework was another strategy to avoid overfitting [110].

Regression modeling and conventional machine learning methods rely on careful consideration of which AGM features to include in the model. For example, any insight from a patient's acceleration measurements may be confounded (for future falls) or deviate by/from their physical and mental condition during their emergency department visit. However, feature selection is a less critical methodological step with deep learning than with conventional machine learning techniques. Features are automatically selected from the acceleration signals within the neural network process. Our results, where the hybrid convolutional recurrent neural network-2 model performed with 92% and 97% accuracy for past falls and future falls, respectively, indicate that the deep learning model could filter through deviations from patient's typical health that the emergency department setting may have had on the accelerometer measurements. While, accuracies may be less descriptive and explanatory when there are highly imbalanced classes, we included F1-scores as well. Moreover, the hybrid convolutional recurrent neural network-3 model performed as well as the hybrid convolutional recurrent neural network-2 model. The hybrid convolutional recurrent neural network-2 model uses one less convolutional layer block than the hybrid convolutional recurrent neural network-3.

Despite having favorable results from conventional machine learning models with generated features, such as our support vector machine result from Table 8.4, using conventional machine learning models with AGMs may not be a sustainable tool for older adults who come into the emergency department, due to time constraints. Falls in the emergency department are common, but the wearable data from emergency department patients are often non-linear [8]. The deep learning models, particularly CNN models, will consistently undertake these non-linearities. Logistic regression and support vector machine (with a linear kernel) may not sustain good results under these conditions. Signal pre-processing steps were crucial to prepare our datasets for deep learning and for more discriminatory ability between models. Our McNemar tests indicate that the hybrid convolutional recurrent neural network models

performed just as well (if not better) as the other CNNs and RNNs.

It is essential to note some of the influential variables in our models that used generated features. Weiss et al. investigated whether an accelerometer could enhance their GUG test when evaluating Parkinson's disease [332]. They found that features derived from accelerometers were a useful measure that complemented both the GUG evaluation and other diagnostic criteria. Similarly, we could discern fall prognosis using only the accelerometer data from the GUG test's walking portions (coded regions 2 and 4). Using our rudimentary feature selection analysis in this dataset, the AGMs/features chosen were a mix of statistical AGMs and the AP triangle duration (Table A1). One of the most important features was the AP triangle duration during tasks 1 and 5 of the GUG test. Clinically, this finding makes sense because sitting up and down from the chair can also test balance, which is a crucial fall-risk behavior [198, 172, 235]. Another essential feature that we included in the model was the time to complete the GUG test. We included this feature because it is clinically used to determine the patient's mobility and the need for gait aid; longer GUG times indicate a higher risk for a future fall[198].

Our dataset had many more future non-fallers than fallers (Table 8.1). This sort of imbalance typically results in low performance from machine learning models. A statistical solution to this is to re-sample the data, either by under-sampling the majority class, over-sampling the minority class, or balanced-sampling, with both under-sampling and over-sampling [148]. However, these methods are controversial due to the lack of generalization of the machine learning experiments. In the computer vision field, where there are typically imbalanced classes, analysts often employ CNNs to tackle them. When we used CNNs and the hybrid convolutional recurrent neural network models, we achieved almost perfect classification accuracy (0.99 and 0.97 accuracies by using generated features and acceleration signals, respectively) [148].

Deploying the leave-one-out model was a tactical decision because we wanted to mimic clinical diagnosis conditions as much as possible. Typically, healthcare providers provide diagnoses based on their experience and evidence-based medical learning. However, with the innovative progress in biomedical healthcare, more data about each patient are being collected and stored [154]. By training our models on all the data from the patients in

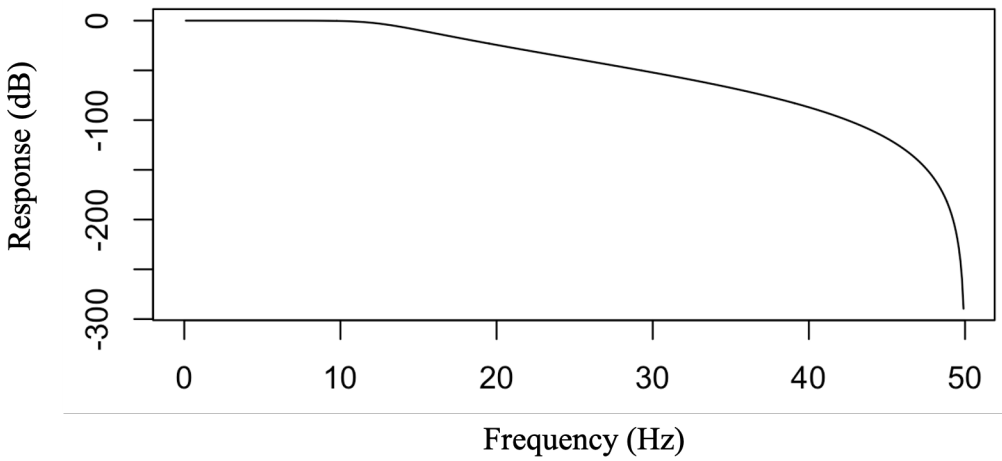
this population (except for the patient whom we are advising), we eliminate the need for adjusting for confounding (due to age, sex, or other confounding factors) and simulate an artificial environment that mimics the natural environment [187, 178]. In future applications, this type of leave-one-out modeling can further the fields of precision medicine and precision public health [178].

Findings from this study support further research using deep learning with acceleration signals to predict future falls in older adults. Our computing environment is very easy to set up: the computing environment did not require a GPU, and it is easily accessible. These results could be computed remotely and securely. Furthermore, standard machine learning models are using more traditional weighted predictors which most likely relate to current gait function (which relates closely to past falls) but deep learning models may identify features that presage these traditional impairments in gait that are yet to have influenced past falls.

Timed GUG tests and other clinical scales are useful as crude measures of fall risk. Future studies should seek to determine whether AGMs are more informative than simpler gait measures, such as speed and other observable metrics (other than inertial sensors or wearables) to assess fall-risk. A feasible and more data-driven approach using accelerometer-based signals during gait and balance tasks in clinical settings could improve real-time clinical decision-making. For example, if healthcare providers could efficiently conduct these tasks (as we did in this study) as part of routine care and programs that could provide real-time prognostic information like low/ high fall risk. These at-risk patients could be linked to prevention-based interventions. There may be other applications for data-driven gait and balance tasks for other diseases (e.g., Parkinsonism). Also, repeated testing over time in the same individual may reveal patterns of change not evident using a single time point for testing.

Moreover, with the population of older adults increases in the US, older adults' rates of ED visits continue to grow [64]. According to the American Geriatrics Society guidelines, ED care should include a multifactorial fall risk assessment that health professionals should perform for all older persons who live at home [250]. The American College of Emergency Physicians further endorses fall risk screening [28]. There are many fall risk assessment tools, few of which are practical to perform in the ED setting, and none have a good predictive

Figure 8.3: A frequency response plot of the Butterworth filter with a cut-off frequency at 12.5 Hz.



ability [204]. A feasible fall-risk screener such as the one we tested could help identify individuals who are at greater risk for falls and target them for physical therapy interventions focused on improving sit-to-stand transitions and gait stability.

8.5 Conclusion

We examined whether we could detect past and future falls using a deep learning approach applied to 3-axis accelerometer data collected during a brief gait task in the emergency department setting among older community-dwelling adults. With the leave-one-out model, the hybrid convolutional recurrent neural network deep learning approach achieved high accuracy for detecting past falls and predicting future falls. Looking to the future of clinical practice, the combination of such deep learning models with short gait and balance tasks and widely available smartphone sensors could transform the standard of care for older adults.

Figure 8.4: Boxplot of the number of 1-second windows in those who have not-fallen/fallen in the past year.

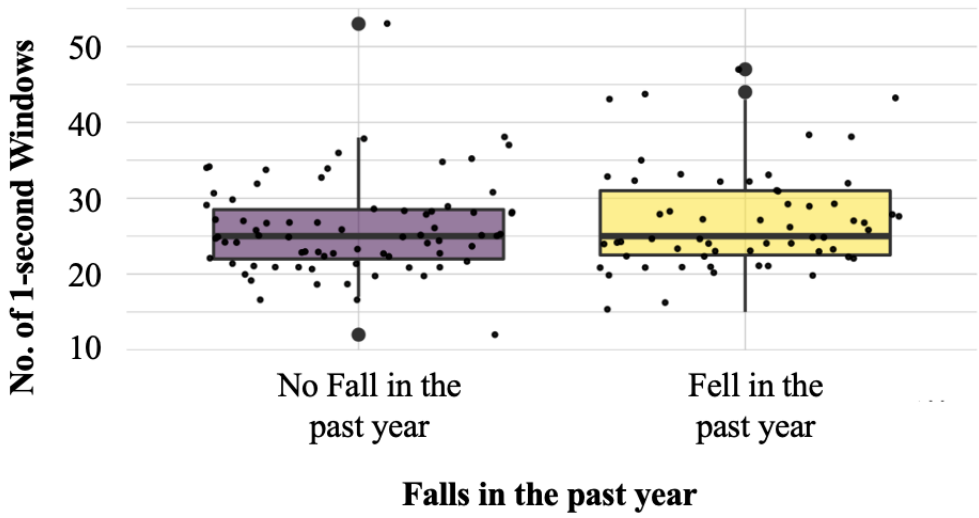


Figure 8.5: Boxplot of the number of 1-second windows in those who have not-fallen/fallen in the future 90 days.

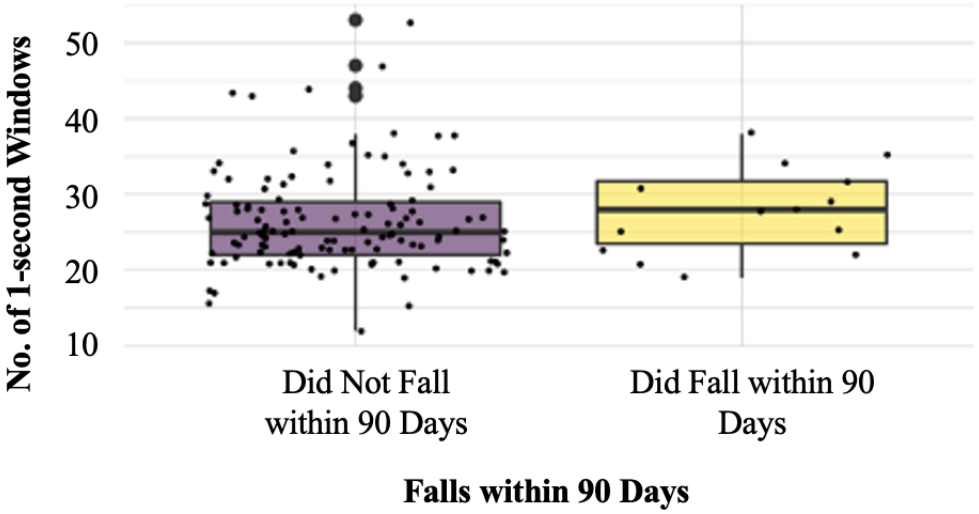


Table A1: Important features, after feature selection, for detecting past falls and future falls.

Important Features
Mean of the ML acceleration signal amplitudes
Mean of the AP acceleration signal amplitudes
Standard Deviation of the V acceleration signal amplitudes
Pair-wise correlation between the ML and V acceleration signal amplitudes
Pair-wise covariance between the V and AP acceleration signal amplitudes
Pair-wise covariance between the ML and AP acceleration signal amplitudes
Pair-wise covariance between the ML and V acceleration signal amplitudes
Maximum value of the ML acceleration signal amplitudes
AP triangle duration

9.0 Conclusion

The use of accelerometers and AGMs is increasing due to the ease of use and low cost. The ultimate goal is to develop screening measures for a walking-related physical-function decline. Also, acceleration-based models and AGMs could inform intervention strategy and monitor outcomes. However, currently, there is a disparity in the literature in reviewing the different mapping of AGMs to aspects of motor skill. In this dissertation, we characterized the three different modes of walking, defined seven motor skill aspects of walking, categorized five broad categories of AGMs, and discussed the typical AGMs used for the aspects of the motor skill of walking (Supplemental Figures in [78]). This dissertation elucidated how AGMs supplement simple measures and improve our understanding of how AGMs can be used to investigate locomotion.

In Chapter 4, we demonstrated that it is possible to create a non-linear, symbolic model of the human body, as a network of body sensors (or a body sensor network). Symbolic models that are individualized to a person and how they walk would have multiple applications to the clinical field. For example, these models can be used as a means of quantifying how close a patient's manifestation of walking is to an idealized model. If patient data does not fit these models well, it could be a sign of poor posture, injury, or some other underlying health problem. Furthermore, poor model fits can be used as a possible predictor of cognitive impairment or falls. These models could also be used to track therapy effectiveness; if the patient's data becomes closer to the idealized model throughout treatment, then there is a good indication that the therapy is working.

Personalized or precision medicine, or the concept of identifying and aiding individuals by their data, can significantly benefit from these individualized models, perhaps in combination with other types of models, because they could be useful in situations where strict monitoring would be valuable. For example, athletes may want to monitor their movements to check for abnormalities that may be an indication of an underlying problem such as degenerative bone disease. Deployed military personnel can also test stresses or injury with these types of models. Active monitoring could be done on elderly individuals at risk of falling and

suffering injury, as well. If these models begin detecting abnormalities and deviations from what is expected, it could warn the user to refocus as they are currently at risk of falling.

In Chapter 5, we found that the accelerometer signals, from six different body sites, used within two methods: the acceleration peak-finding method and the inverted pendulum method produced spatiotemporal estimates of walking. The innovation and the contribution of this work is that we performed similarity tests for all the spatiotemporal estimates from both methods compared to a treadmill - our results showed that the back accelerometer consistently was able to produce similar results. This improvement in using these algorithms is shown for slower walking. These results have the potential for improving clinical measures for patients and older adults. Further development and testing of models will allow application of these algorithms to a more generalizable population with varying gait speeds.

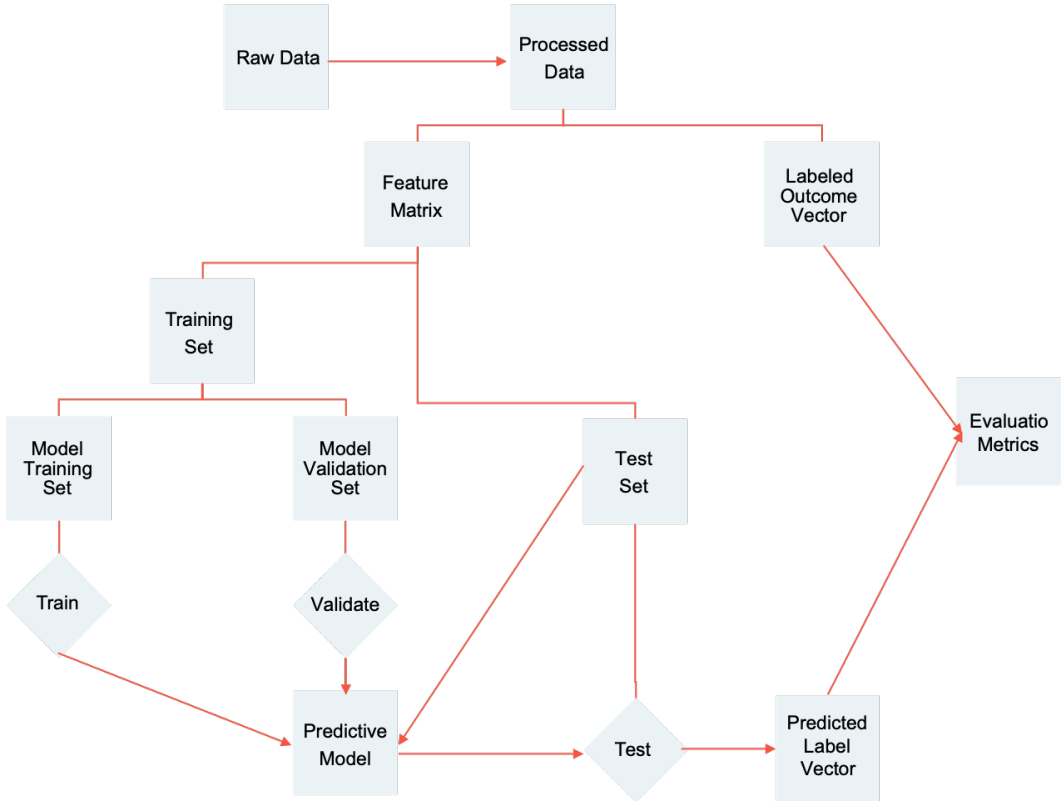
Typically, the spatiotemporal aspect of gait has been measured via the combination of many expensive machinery, such as inertial sensors, instrumented mats, force platforms, special insoles, and camera-tracking systems [246]. Cheaper methods have emerged, such as using the IMUs from accelerometers, gyroscopes, and magnetometers. However, as mentioned in [85], researchers have not calculate step width and step width variability from using a single tri-axial accelerometer alone. Spatiotemporal estimation is at the core of understanding various pathological and physiological states in older adults and, particularly, adults with medical conditions.

Our work in estimating step width and width variability first comes from characterizing the observable simple gait measures. Further work will consist of creating a three dimensional model that can estimate the foot angle during the step, from which we can determine step width.

AGM calculation and its use is at the heart of this dissertation. In Chapters 6 and 8, linking motor skills of walking to AGM metrics proved useful in quantifying declines due to aging and other neuromotor factors. In application, AGMs have been used to detect differences and changes in motor performance due to learning/expertise, or task and environment manipulations.

This process first starts with the raw ML, AP, and V accelerations from the accelerometer as input data. Signal pre-processing methods are applied on this input data. Stride and/or

Figure 9.1: An overview of the supervised machine learning pipeline.



window segmentation is performed to create samples/observations/rows for the feature set. Features are extracted as per the calculations required for each AGM. These features are ranked using the t-test; additionally, redundant features are eliminated if they are highly correlated. Features can also be transformed using PCA methodology.

Since AGMs are a promising component of motor skill research, machine learning with AGMs in the feature set may help older adults' quality of life and reduce the strain on healthcare. An overview of a typical flow of tasks for machine learning is shown in Figure 9.1.

In Chapter 7, we pointed out that even validated measures of aspects of the motor skill of walking (such as smoothness) may not be well-validated throughout the literature. In this case, the harmonic ratio, was a well-validated metric for measuring the smoothness of walking - we wanted to see if there were any comparable measures of smoothness of walking.

We used a concordance analysis technique to confirm whether other measures of smoothness used in the literature were concordant with the harmonic ratio. Only jerk-cost showed a level of high concordance. Future work includes the investigation of AGMs in all areas of the motor skill of walking, so that further work can be done in helping individuals whose walking motor skill is affected.

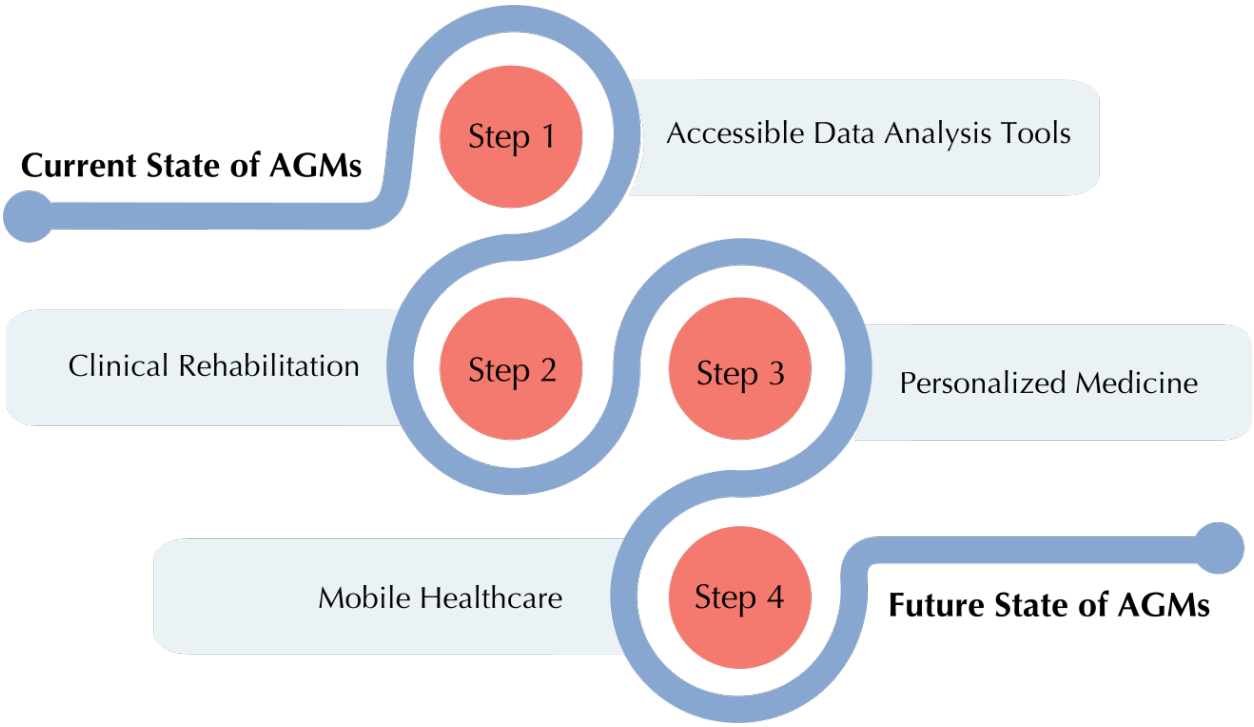
9.1 Possible Future Work with AGMs

Without reliable and accessible tools within an established signal pre-processing pipeline, the use of AGMs in research cannot be feasible. Acceleration signal pre-processing can be a time-consuming task and can get in the way of diagnosing or analyzing a clinical problem. The assessment of gait in the clinical space lacks maturity with the use of these signal pre-processing tasks.

Computing languages, packages, and toolboxes will come and go, but there will always be a constant need for technological tools that are more accessible to researchers of all levels. Some of the attributes any tool processing the acceleration signal to AGMs should have are the ability to visualize accelerations, packages that can filter out signal noise, and the ability to extract signal features into a data structure, that can later be used in statistical modeling. While MATLAB, Python, and the other current tools have all of these pieces, tools with greater ease of use and reduced requirements for programming could make these measures more available to a broader audience of researchers and clinicians.

In Figure 9.2, the future of this field and how gait accelerometry research can be ameliorated through the use of AGMs, not just in the clinical space but also in the hands of patients and consumers. For instance, AGMs combined with electronic health and medical records may be used to identify those with a high risk of falls [211]. Since wearables are increasingly reducing in size, they can be used as a means to provide digital medicine with a harmonious set of biomarkers (risk, diagnostic, monitoring, prognostic, etc.) [69].

Figure 9.2: Comparison between the current and future state of acceleration gait measure use in research.



10.0 Abbreviations

- **AGM**: Acceleration Gait Measures
- **ML**: Mediolateral
- **V**: Vertical
- **AP**: Anterior-posterior
- **IMU**: Inertial measurement unit
- **LDS**: Local dynamic stability
- **IQR**: Interquartile range
- **ANOVA**: Analysis of variance
- **DFT**: Discrete Fourier transform
- **FFT**: Fast Fourier transform
- **DWT**: Discrete wavelet transform
- **AIC**: Akaike information criterion
- **PCA**: Principal components analysis
- **PC**: Principal component
- **ICA**: Independent Component Analysis
- **SVM**: support vector machine
- **EA**: Evolutionary algorithms
- **GP**: Genetic programming
- **ANN**: Artificial neural networks
- **CNN**: Convolutional neural networks
- **TPR**: True positive rate
- **TNR**: True negative rate
- **FPR**: False-positive rate
- **FPR**: False-positive rate
- **PPV**: Positive predictive value
- **NPV**: Negative predictive value
- **AUC**: Area under the curve

- **ROC**: Receiver operating characteristic
- **MAE**: Mean absolute error
- **MAPE**: Mean absolute percentage error
- **PRIMA**: Program to Improve Mobility in Aging

Bibliography

[1] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[2] Mateo Aboy, Roberto Hornero, Daniel Abásolo, and Daniel Álvarez. Interpretation of the lempel-ziv complexity measure in the context of biomedical signal analysis. *IEEE Transactions on Biomedical Engineering*, 53(11):2282–2288, 2006.

[3] Koray Açıcı, Çağatay Berke Erdaş, Tunç Aşuroğlu, Münire Kılınç Toprak, Hamit Erdem, and Hasan Oğul. A random forest method to detect parkinson's disease via gait analysis. In *International conference on engineering applications of neural networks*, pages 609–619. Springer, 2017.

[4] Pia Addabbo, Mario Luca Bernardi, Filippo Biondi, Marta Cimitile, Carmine Clemente, and Danilo Orlando. Temporal convolutional neural networks for radar micro-doppler based gait recognition. *Sensors*, 21(2):381, 2021.

[5] Samira Ahmadi, Nariman Sepehri, Christine Wu, and Tony Szturm. Sample entropy of human gait center of pressure displacement: A systematic methodological analysis. *Entropy*, 20(8):579, 2018.

[6] Hirotugu Akaike. Factor analysis and aic. In *Selected papers of hirotugu akaike*, pages 371–386. Springer, 1987.

[7] Saad Albawi, Tareq Abed Mohammed, and Saad Al-Zawi. Understanding of a convolutional neural network. In *2017 International Conference on Engineering and Technology (ICET)*, pages 1–6. IEEE, 2017.

[8] Danette Alexander, Terry L Kinsley, and Christine Waszinski. Journey to a safe environment: fall prevention in an emergency department at a level i trauma center. *Journal of Emergency Nursing*, 39(4):346–352, 2013.

[9] Neil B Alexander and Allon Goldberg. Gait disorders: search for multiple causes. *Cleveland Clinic Journal of Medicine*, 72(7):586, 2005.

[10] Abdullah S Alharthi, Syed U Yunas, and Krikor B Ozanyan. Deep learning for monitoring of human gait: A review. *IEEE Sensors Journal*, 19(21):9575–9591, 2019.

[11] Diego Alvarez, Rafael C González, Antonio López, and Juan C Alvarez. Comparison of step length estimators from weareable accelerometer devices. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5964–5967. IEEE, 2006.

[12] Diego Alvarez, Rafael C González, Antonio López, and Juan C Alvarez. Comparison of step length estimators from wearable accelerometer devices. In *Encyclopedia of Healthcare Information Systems*, pages 244–250. IGI Global, 2008.

[13] Anne Felicia Ambrose, Geet Paul, and Jeffrey M Hausdorff. Risk factors for falls among older adults: a review of the literature. *Maturitas*, 75(1):51–61, 2013.

[14] Lorenza Angelini, Ilaria Carpinella, Davide Cattaneo, Maurizio Ferrarin, Elisa Gervasoni, Basil Sharrack, David Paling, Krishnan Padmakumari Sivaraman Nair, and Claudia Mazzà. Is a wearable sensor-based characterisation of gait robust enough to overcome differences between measurement protocols? a multi-centric pragmatic study in patients with multiple sclerosis. *Sensors*, 20(1):79, 2020.

[15] Cesare Angeloni, Patrick O Riley, and David E Krebs. Frequency content of whole body gait kinematic data. *IEEE Transactions on Rehabilitation Engineering*, 2(1):40–46, 1994.

[16] Janneke Annegarn, Martijn A Spruit, Hans HCM Savelberg, Paul JB Willems, Coby Van de Bool, Annemie MWJ Schols, Emiel FM Wouters, and Kenneth Meijer. Differences in walking pattern during 6-min walk test between patients with copd and healthy subjects. *PLoS ONE*, 7(5):e37329, 2012.

[17] Zachary Arnold, Danielle LaRose, and Emmanuel Agu. Smartphone inference of alcohol consumption levels from gait. In *2015 International Conference on Healthcare Informatics*, pages 417–426. IEEE, 2015.

[18] Siddharth Arora, Vinayak Venkataraman, Sean Donohue, Kevin M Biglan, Earl R Dorsey, and Max A Little. High accuracy discrimination of parkinson's disease participants from healthy controls using smartphones. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3641–3644. IEEE, 2014.

[19] Bernard Auvinet, Gilles Berrut, Claude Touzard, Laurent Moutel, Nadine Collet, Denis Chaleil, and Eric Barrey. Reference data for normal subjects obtained with an accelerometric device. *Gait & Posture*, 16(2):124–134, 2002.

[20] Bernard Auvinet, Richard Bileckot, Anne-Sophie Alix, Denis Chaleil, and Eric Barrey. Gait disorders in patients with fibromyalgia. *Joint Bone Spine*, 73(5):543–546, 2006.

[21] Damien Bachasson, Amélie Moraux, Gwenn Ollivier, Valérie Decostre, Isabelle Ledoux, Teresa Gidaro, Laurent Servais, Anthony Behin, Tanya Stojkovic, Luc J Hébert, et al. Relationship between muscle impairments, postural stability, and gait parameters assessed with lower-trunk accelerometry in myotonic dystrophy type 1. *Neuromuscular Disorders*, 26(7):428–435, 2016.

[22] Marc Bächlin, Jeffrey M Hausdorff, Daniel Roggen, Nir Giladi, Meir Plotnik, and Gerhard Tröster. Online detection of freezing of gait in parkinson's disease patients: A performance characterization. In *Proceedings of the Fourth International Conference on Body Area Networks*, page 11. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009.

[23] Suresh Balakrishnama and Aravind Ganapathiraju. Linear discriminant analysis-a brief tutorial. *Institute for Signal and Information Processing*, 18:1–8, 1998.

[24] Sivakumar Balasubramanian, Alejandro Melendez-Calderon, and Etienne Burdet. A robust and sensitive metric for quantifying movement smoothness. *IEEE transactions on biomedical engineering*, 59(8):2126–2136, 2011.

[25] Sivakumar Balasubramanian, Alejandro Melendez-Calderon, Agnes Roby-Brami, and Etienne Burdet. On the analysis of movement smoothness. *Journal of neuroengineering and rehabilitation*, 12(1):112, December 2015.

[26] Albert-László Barabási. Network medicine—from obesity to the "diseasome", 2007.

[27] Albert-László Barabási, Natali Gulbahce, and Joseph Loscalzo. Network medicine: a network-based approach to human disease. *Nature reviews genetics*, 12(1):56–68, 2011.

[28] Larry J Baraff, Richard Della Penna, Nicole Williams, and Arthur Sanders. Practice guideline for the ed management of falls in community-dwelling elderly persons. *Annals of emergency medicine*, 30(4):480–492, 1997.

[29] Ronny P Bartsch, Kang KL Liu, Amir Bashan, and Plamen Ch Ivanov. Network physiology: how organ systems dynamically interact. *PloS one*, 10(11):e0142143, 2015.

[30] Yoav Beck, Talia Herman, Marina Brozgol, Nir Giladi, Anat Mirelman, and Jeffrey M. Hausdorff. SPARC: a new approach to quantifying gait smoothness in patients with parkinson's disease. *Journal of NeuroEngineering and Rehabilitation*, 15(1):1–9, June 2018.

[31] JL Bellanca, KA Lowry, JM VanSwearingen, JS Brach, and MS Redfern. Harmonic ratios: a quantification of step to step symmetry. *Journal of Biomechanics*, 46(4):828–831, 2013.

[32] Chiraz BenAbdelkader, Ross Cutler, and Larry Davis. Stride and cadence as a biometric in automatic person identification and verification. In *Proceedings of Fifth IEEE international conference on automatic face gesture recognition*, pages 372–377. IEEE, 2002.

[33] S Binder. Injuries among older adults: the challenge of optimizing safety and minimizing unintended consequences. *Injury Prevention*, 8(suppl 4):iv2–iv4, 2002.

[34] J Martin Bland and DouglasG Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476):307–310, 1986.

[35] KayLynn Bland, Kristin Lowry, Alex Krajek, Taylor Woods, and Jessie VanSwearingen. Spatiotemporal variability underlying skill in curved-path walking. *Gait & Posture*, 67:137–141, 2019.

[36] Bastiaan R Bloem, Yvette AM Grimbergen, Monique Cramer, Mirjam Willemsen, and Aeilko H Zwinderman. Prospective assessment of falls in parkinson's disease. *Journal of Neurology*, 248(11):950–958, 2001.

[37] Richard W Bohannon, A Williams Andrews, and Michael W Thomas. Walking speed: reference values and correlates for older adults. *Journal of Orthopaedic & Sports Physical Therapy*, 24(2):86–90, 1996.

[38] Jaqueline M Bond and Meg Morris. Goal-directed secondary motor tasks: their effects on gait in subjects with parkinson disease. *Archives of Physical Medicine and Rehabilitation*, 81(1):110–116, 2000.

[39] Alberto G Bonomi, Annelies HC Goris, Bin Yin, and Klaas R Westerterp. Detection of type, duration, and intensity of physical activity using an accelerometer. *Medicine and Science in Sports and Exercise*, 41(9):1770–1777, 2009.

[40] Sheila de Melo Borges, Márcia Radanovic, and Orestes Vicente Forlenza. Correlation between functional mobility and cognitive performance in older adults with cognitive impairment. *Aging, Neuropsychology, and Cognition*, 25(1):23–32, 2018.

[41] Nikolaos V Boulgouris, Dimitrios Hatzinakos, and Konstantinos N Plataniotis. Gait recognition: a challenging signal processing technology for biometric identification. *IEEE Signal Processing Magazine*, 22(6):78–90, 2005.

[42] C VC Bouten, A AHJ Sauren, M Verduin, and JD Janssen. Effects of placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking. *Medical & Biological Engineering & Computing*, 35(1):50–56, 1997.

[43] Jennifer S Brach, Kristin Lowry, Subashan Perera, Victoria Hornyak, David Wert, Stephanie A Studenski, and Jessie M VanSwearingen. Improving motor control in walking: a randomized clinical trial in older adults with subclinical walking difficulty. *Archives of Physical Medicine and Rehabilitation*, 96(3):388–394, 2015.

[44] Jennifer S Brach, David McGurl, David Wert, Jessie M VanSwearingen, Subashan Perera, Rakie Cham, and Stephanie Studenski. Validation of a measure of smoothness of walking. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 66(1):136–141, 2010.

[45] Jennifer S Brach, David McGurl, David Wert, Jessie M VanSwearingen, Subashan Perera, Rakie Cham, and Stephanie Studenski. Validation of a measure of smoothness of walking. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 66(1):136–141, 2011.

[46] Jennifer S Brach, Jessie M Van Swearingen, Subashan Perera, David M Wert, and Stephanie Studenski. Motor learning versus standard walking exercise in older adults with subclinical gait dysfunction: a randomized clinical trial. *Journal of the American Geriatrics Society*, 61(11):1879–1886, 2013.

[47] Jennifer S Brach, Jessie M VanSwearingen, Alexandra Gil, Neelesh K Nadkarni, Andrea Kriska, Rakie Cham, and Subashan Perera. Program to improve mobility in aging (prima) study: Methods and rationale of a task-oriented motor learning exercise program. *Contemporary Clinical Trials*, 89:105912, 2020.

[48] Matthew AD Brodie, Hylton B Menz, and Stephen R Lord. Age-associated changes in head jerk while walking reveal altered dynamic stability in older people. *Experimental brain research*, 232(1):51–60, 2014.

[49] V Brooks. The neural basis of motor control. 1986.

[50] Sjoerd M Bruijn, Daan JJ Bregman, Onno G Meijer, Peter J Beek, and Jaap H van Dieën. Maximum lyapunov exponents as predictors of global gait stability: a modelling approach. *Medical Engineering & Physics*, 34(4):428–436, 2012.

[51] F Bugané, MG Benedetti, G Casadio, S Attala, F Biagi, Mario Manca, and Alberto Leardini. Estimation of spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal subjects by standard gait analysis. *Computer methods and programs in biomedicine*, 108(1):129–137, 2012.

[52] Benoit Caby, Suzanne Kieffer, Marie de Saint Hubert, Gerald Cremer, and Benoit Macq. Feature extraction and selection for objective gait analysis and fall risk assessment by accelerometry. *Biomedical engineering online*, 10(1):1–19, 2011.

[53] Tilman Calliess, Raphael Bocklage, Roman Karkosch, Michael Marschollek, Henning Windhagen, and Mareike Schulze. Clinical evaluation of a mobile sensor-based gait analysis method for outcome measurement after knee arthroplasty. *Sensors*, 14(9):15953–15964, 2014.

[54] Jorge Cancela, Matteo Pastorino, María Teresa Arredondo, Mario Pansera, Laura Pastor-Sanz, Federico Villagra, Maria A Pastor, and Ana P Gonzalez. Gait assessment in parkinson's disease patients through a network of wearable accelerometers in unsupervised environments. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2233–2236. IEEE, 2011.

[55] Carlotta Caramia, Ivan Bernabucci, Carmen D'Anna, Cristiano De Marchis, and Maurizio Schmid. Gait parameters are differently affected by concurrent smartphone-based activities with scaled levels of cognitive effort. *PloS one*, 12(10):e0185825, 2017.

[56] Carlotta Caramia, Cristiano De Marchis, and Maurizio Schmid. Optimizing the scale of a wavelet-based method for the detection of gait events from a waist-mounted accelerometer under different walking speeds. *Sensors*, 19(8):1869, 2019.

[57] Matteo Cesari, Stephen B Kritchevsky, Brenda WHJ Penninx, Barbara J Nicklas, Eleanor M Simonsick, Anne B Newman, Frances A Tylavsky, Jennifer S Brach,

Suzanne Satterfield, Douglas C Bauer, et al. Prognostic value of usual gait speed in well-functioning older people - results from the health, aging and body composition study. *Journal of the American Geriatrics Society*, 53(10):1675–1680, 2005.

[58] Tom Chau. A review of analytical techniques for gait data. part 2: neural network and wavelet methods. *Gait & Posture*, 13(2):102–120, 2001.

[59] Tom Chau, Scott Young, and Sue Redekop. Managing variability in the summary and comparison of gait data. *Journal of Neuroengineering and Rehabilitation*, 2(1):22, 2005.

[60] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4), 2015.

[61] Yize Chen, Marco Tulio Angulo, and Yang-Yu Liu. Revealing complex ecological dynamics via symbolic regression. *BioEssays*, 41(12):1900069, 2019.

[62] Yu M Chi and Gert Cauwenberghs. Wireless non-contact eeg/ecg electrodes for body sensor networks. In *2010 International Conference on Body Sensor Networks*, pages 297–301. IEEE, 2010.

[63] Francois Chollet and JJ Allaire. *Deep Learning with R*. Manning Publ., 2018.

[64] Shih-Chuan Chou, Olesya Baker, and Jeremiah D Schuur. Changes in emergency department care intensity from 2007–16: Analysis of the national hospital ambulatory medical care survey. *Western Journal of Emergency Medicine*, 21(2):209, 2020.

[65] Paul Cisek and John F Kalaska. Neural mechanisms for interacting with a world full of action choices. *Annual review of neuroscience*, 33:269–298, 2010.

[66] David J Clark. Automaticity of walking: functional significance, mechanisms, measurement and rehabilitation strategies. *Frontiers in Human Neuroscience*, 9:246, 2015.

[67] Ian Cleland, Basel Kikhia, Chris Nugent, Andrey Boytsov, Josef Hallberg, Kåre Synnes, Sally McClean, and Dewar Finlay. Optimal placement of accelerometers for the detection of everyday activities. *Sensors*, 13(7):9183–9200, 2013.

[68] Albert Cohen, Ingrid Daubechies, and Pierre Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1993.

[69] Andrea Coravos, Sean Khozin, and Kenneth D Mandl. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. *npj Digital Medicine*, 2(1):14, 2019.

[70] Grégoire Courtine and Marco Schieppati. Human walking along a curved path. i. body trajectory, segment orientation and the effect of vision. *European Journal of Neuroscience*, 18(1):177–190, 2003.

[71] Ronita L Cromwell and Roberta A Newton. Relationship between balance and gait stability in healthy older adults. *Journal of aging and physical activity*, 12(1):90–100, 2004.

[72] KM Culhane, M O'connor, D Lyons, and GM Lyons. Accelerometers in rehabilitation medicine for older adults. *Age and Ageing*, 34(6):556–560, 2005.

[73] David Cunado, Mark S Nixon, and John N Carter. Automatic extraction and description of human gait models for recognition purposes. *Computer vision and image understanding*, 90(1):1–41, 2003.

[74] Adele Cutler, D Richard Cutler, and John R Stevens. Random forests. In *Ensemble Machine Learning*, pages 157–175. Springer, 2012.

[75] Matthew D Czech, Dimitrios Psaltos, Hao Zhang, Tomasz Adamusiak, Monica Calicchio, Amey Kelekar, Andrew Messere, Koene RA Van Dijk, Vesper Ramos, Charmaine Demanuele, et al. Age and environment-related differences in gait in healthy adults using wearables. *NPJ Digital Medicine*, 3(1):1–9, 2020.

[76] Anthony Dalton, Hanan Khalil, Monica Busse, Anne Rosser, Robert van Deursen, and Gearóid ÓLaighin. Analysis of gait and balance through a single triaxial accelerometer in presymptomatic and symptomatic huntington's disease. *Gait & Posture*, 37(1):49–54, 2013.

[77] Burcu F Darst, Kristen C Malecki, and Corinne D Engelman. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. *BMC Genetics*, 19(1):1–6, 2018.

[78] Pritika Dasgupta, Jessie Van Swearingen, Alan Godfrey, Mark Redfern, Manuel Montero-Odasso, and Ervin Sejdić. Acceleration gait measures as proxies for motor skill of walking: a narrative review. *IEEE transactions on neural systems and rehabilitation engineering*, 2020.

[79] Pritika Dasgupta, Jessie VanSwearingen, and Ervin Sejdić. You can tell by the way i use my walk. predicting the presence of cognitive load with gait measurements. *BioMedical Engineering OnLine*, 17(1):122, 2018.

[80] Roy B Davis. Reflections on clinical gait analysis. *Journal of Electromyography and Kinesiology*, 7(4):251–257, 1997.

[81] Ihana Thaís Guerra de Oliveira Gondim, Caroline de Cássia Batista de Souza, Marco Aurélio Benedetti Rodrigues, Izaura Muniz Azevedo, Maria das Graças Wanderley de Sales, Otávio Gomes Lins, et al. Portable accelerometers for the evaluation of spatio-temporal gait parameters in people with parkinson's disease: an integrative review. *Archives of Gerontology and Geriatrics*, page 104097, 2020.

[82] Silvia Del Din, Alan Godfrey, and Lynn Rochester. Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and parkinson's disease: toward clinical and at home use. *IEEE Journal of Biomedical and Health Informatics*, 20(3):838–847, 2015.

[83] Silvia Del Din, Aodhán Hickey, Naomi Hurwitz, John C Mathers, Lynn Rochester, and Alan Godfrey. Measuring gait with an accelerometer-based wearable: influence of device location, testing protocol and age. *Physiological Measurement*, 37(10):1785, 2016.

[84] AR Den Otter, ACH Geurts, M De Haart, T Mulder, and Jaak Duysens. Step characteristics during obstacle avoidance in hemiplegic stroke. *Experimental Brain Research*, 161(2):180–192, 2005.

[85] Silvia Del Din, Alan Godfrey, Brook Galna, Sue Lord, and Lynn Rochester. Free-living gait characteristics in ageing and parkinson's disease: impact of environment and ambulatory bout length. *Journal of NeuroEngineering and Rehabilitation*, 13(1), May 2016.

[86] Jonathan B Dingwell and JoSeph P Cusumano. Nonlinear time series analysis of normal and pathological human walking. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 10(4):848–863, 2000.

[87] Milica D Djurić-Jovičić, Nenad S Jovičić, Saša M Radovanović, Iva D Stanković, Mirjana B Popović, and Vladimir S Kostić. Automatic identification and classification of freezing of gait episodes in parkinson's disease patients. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):685–694, 2014.

[88] Emer P Doheny, Barry R Greene, Timothy Foran, Clodagh Cunningham, Chie Wei Fan, and Rose Anne Kenny. Diurnal variations in the outcomes of instrumented gait and quiet standing balance assessments and their association with falls history. *Physiological Measurement*, 33(3):361, 2012.

[89] Takehiko Doi, Soichiro Hirata, Rei Ono, Kota Tsutsumimoto, Shogo Misu, and Hiroshi Ando. The harmonic ratio of trunk acceleration predicts falling among older people: results of a 1-year prospective study. *Journal of Neuroengineering and Rehabilitation*, 10(1):7, 2013.

[90] Takehiko Doi, Hyuma Makizako, Hiroyuki Shimada, Daisuke Yoshida, Kengo Ito, Takashi Kato, Hiroshi Ando, and Takao Suzuki. Brain atrophy and trunk stability during dual-task walking among older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 67(7):790–795, 2012.

[91] Takehiko Doi, Ryota Yamaguchi, Tsuyoshi Asai, Minoru Komatsu, Daisuke Makiura, Masanori Shimamura, Soichiro Hirata, Hiroshi Ando, and Masahiro Kurosaka. The effects of shoe fit on gait in community-dwelling older adults. *Gait & Posture*, 32(2):274–278, 2010.

[92] Lars Donath, Oliver Faude, Eric Lichtenstein, Corina Nüesch, and Annegret Mündermann. Validity and reliability of a portable gait analysis system for measuring spatiotemporal gait characteristics: comparison to an instrumented treadmill. *Journal of neuroengineering and rehabilitation*, 13(1):6, 2016.

[93] Amira El-Attar, Amira S Ashour, Nilanjan Dey, Hatem Abdelkader, Mostafa M Abd El-Naby, and R Simon Sherratt. Discrete wavelet transform-based freezing of gait detection in parkinson's disease. *Journal of Experimental & Theoretical Artificial Intelligence*, pages 1–17, 2018.

[94] Deborah D Espy, Feng Yang, Tanvi Bhatt, and Y-C Pai. Independent influence of gait speed and step length on stability and fall risk. *Gait & posture*, 32(3):378–382, 2010.

[95] Y Rakhshani Fatmehsari and F Bahrami. Lempel-ziv complexity criteria for nonlinear analysis of gait in patients with parkinson's disease. In *2011 18th Iranian Conference of Biomedical Engineering (ICBME)*, pages 137–141. IEEE, 2011.

[96] Patrik Fazio, Gino Granieri, Ilaria Casetta, Edward Cesnik, Sante Mazzacane, Pietro Caliandro, Francesco Pedrielli, and Enrico Granieri. Gait measures with a triaxial

accelerometer among patients with neurological impairment. *Neurological Sciences*, 34(4):435–440, 2013.

[97] Yuanyuan Feng, Christopher K Wong, Vandana Janeja, Ravi Kuber, and Helena M Mentis. Comparison of tri-axial accelerometers step-count accuracy in slow walking conditions. *Gait & Posture*, 53:11–16, 2017.

[98] Curtis S Florence, Gwen Bergen, Adam Atherly, Elizabeth Burns, Judy Stevens, and Cynthia Drake. Medical costs of fatal and nonfatal falls in older adults. *Journal of the American Geriatrics Society*, 66(4):693–698, 2018.

[99] Laetitia Fradet, Gyusung Lee, and Natalia Dounskaia. Origins of submovements in movements of elderly adults. *Journal of Neuroengineering and Rehabilitation*, 5(1):28, 2008.

[100] Linda P Fried and Jack M Guralnik. Disability in older adults: evidence regarding significance, etiology, and risk. *Journal of the American Geriatrics Society*, 45(1):92–100, 1997.

[101] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *Annals of Statistics*, 28(2):337–407, 2000.

[102] A Gabell and USL Nayak. The effect of age on variability in gait. *Journal of Gerontology*, 39(6):662–666, 1984.

[103] Davrondzhon Gafurov, Einar Snekkenes, and Patrick Bours. Gait authentication and identification using wearable accelerometer sensor. In *2007 IEEE Workshop on. Automatic Identification Advanced Technologies*, pages 220–225. IEEE, 2007.

[104] Howard Gage. Accelerographic analysis of human gait. *The American Society of Mechanical Engineers*, 64, 1964.

[105] James R Gage, Peter A Deluca, and Thomas S Renshaw. Gait analysis: principles and applications. *JBJS*, 77(10):1607–1623, 1995.

[106] Sander PM Ganzevles, Peter J Beek, Hein AM Daanen, Bart MA Coolen, and Martin J Truijens. Differences in swimming smoothness between elite and non-elite swimmers. *Sports biomechanics*, pages 1–14, 2019.

[107] Davide Giavarina. Understanding bland altman analysis. *Biochemia Medica*, 25(2):141–151, 2015.

[108] Sophie Gillain, Mohamed Boutaayamou, Charlotte Beaudart, Marie Demonceau, Olivier Bruyère, Jean-Yves Reginster, Gaëtan Garraux, and Jean Petermans. Assessing gait parameters with accelerometer-based methods to identify older adults at risk of falls: a systematic review. *European Geriatric Medicine*, pages 1–14, 2018.

[109] Sophie Gillain, Emmanuelle Warzee, Françoise Lekeu, Vinciane Wojtasik, Didier Maquet, J-L Croisier, Eric Salmon, and Jean Petermans. The value of instrumental gait analysis in elderly healthy, mci or alzheimer's disease subjects and a comparison with other clinical tests used in single and dual-task conditions. *Annals of Physical and Rehabilitation Medicine*, 52(6):453–474, 2009.

[110] Giacomo Giorgi, Fabio Martinelli, Andrea Saracino, and Mina Sheikhalishahi. Walking through the deep: Gait analysis for user authentication through deep learning. In *IFIP International Conference on ICT Systems Security and Privacy Protection*, pages 62–76. Springer, 2018.

[111] ACRMDOG Godfrey, Richard Conway, David Meagher, and Gearoid ÓLaighin. Direct measurement of human movement by accelerometry. *Medical Engineering & Physics*, 30(10):1364–1386, 2008.

[112] Alan Godfrey, Silvia Del Din, Gillian Barry, JC Mathers, and Lynn Rochester. Instrumenting gait with an accelerometer: a system and algorithm examination. *Medical Engineering & Physics*, 37(4):400–407, 2015.

[113] Rafael C González, Antonio M López, Javier Rodriguez-Uría, Diego Alvarez, and Juan C Alvarez. Real-time gait event detection for normal subjects from lower trunk accelerations. *Gait & Posture*, 31(3):322–325, 2010.

[114] Huan Gou, Tengda Shi, Lei Yan, and Jiang Xiao. Gait and posture analysis method based on genetic algorithm and support vector machines with acceleration data. *Journal of Robotics and Mechatronics*, 28(3):418–424, 2016.

[115] Darwin Gouwanda and SMN Arosha Senanayake. Identifying gait asymmetry using gyroscopes: A cross-correlation and normalized symmetry index approach. *Journal of Biomechanics*, 44(5):972–978, 2011.

[116] Simone Guglielmetti, Antonio Nardone, Alessandro Marco De Nunzio, Marco Godi, and Marco Schieppati. Walking along circular trajectories in parkinson's disease. *Movement Disorders*, 24(4):598–604, 2009.

[117] Piyush Gupta and Tim Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering*, 61(6):1780–1786, 2014.

[118] Jack M Guralnik, Luigi Ferrucci, Carl F Pieper, Suzanne G Leveille, Kyriakos S Markides, Glenn V Ostir, Stephanie Studenski, Lisa F Berkman, and Robert B Wallace. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 55(4):M221–M231, 2000.

[119] Jack M Guralnik, Linda P Fried, and Marcel E Salive. Disability as a public health outcome in the aging population. *Annual Review of Public Health*, 17(1):25–46, 1996.

[120] Yara K Haddad, Gwen Bergen, and Curtis Florence. Estimating the economic burden related to older adult falls by state. *Journal of Public Health Management and Practice: JPHMP*, 25(2):E17, 2019.

[121] Laura Hak, Han Houdijk, Frans Steenbrink, Agali Mert, Peter van der Wurff, Peter J Beek, and Jaap H van Dieën. Speeding up or slowing down?: Gait adaptations to preserve gait stability in response to balance perturbations. *Gait & posture*, 36(2):260–264, 2012.

[122] Laura Hak, Han Houdijk, Frans Steenbrink, Agali Mert, Peter van der Wurff, Peter J Beek, and Jaap H van Dieën. Stepping strategies for regulating gait adaptability and stability. *Journal of Biomechanics*, 46(5):905–911, 2013.

[123] Lutz Hamel. Model assessment with roc curves. In *Encyclopedia of Data Warehousing and Mining, Second Edition*, pages 1316–1323. IGI Global, 2009.

[124] Antonia Hartmann, Kurt Murer, Rob A de Bie, and Eling D de Bruin. Reproducibility of spatio-temporal gait parameters under different conditions in older adults using a trunk tri-axial accelerometer system. *Gait & Posture*, 30(3):351–355, 2009.

[125] Klaus Hauer, Sarah E Lamb, Ellen C Jorstad, Chris Todd, and Clemens Becker. Systematic review of definitions and methods of measuring falls in randomised controlled fall prevention trials. *Age and ageing*, 35(1):5–10, 2006.

[126] Jeffrey M Hausdorff. Gait variability: methods, modeling and meaning. *Journal of Neuroengineering and Rehabilitation*, 2(1):19, 2005.

[127] Jeffrey M Hausdorff. Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. *Human movement science*, 26(4):555–589, 2007.

[128] Jeffrey M Hausdorff, Dean A Rios, and Helen K Edelberg. Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Archives of Physical Medicine and Rehabilitation*, 82(8):1050–1056, 2001.

[129] Jeffrey M Hausdorff, Galit Yogev, Shmuel Springer, Ely S Simon, and Nir Giladi. Walking is more like catching than tapping: gait in the elderly as a complex cognitive task. *Experimental Brain Research*, 164(4):541–548, 2005.

[130] Sven Heinrich, Kilian Rapp, Ulrich Rissmann, Caroline Becker, and H-H König. Cost of falls in old age: a systematic review. *Osteoporosis international*, 21(6):891–902, 2010.

[131] Jorunn L Helbostad and Rolf Moe-Nilssen. The effect of gait speed on lateral balance control during walking in healthy elderly. *Gait & Posture*, 18(2):27–36, 2003.

[132] Marius Henriksen, Hans Lund, R Moe-Nilssen, H Bliddal, and B Danneskiod-Samsøe. Test–retest reliability of trunk accelerometric gait analysis. *Gait & Posture*, 19(3):288–297, 2004.

[133] Caroline Hodt-Billington, Jorunn L Helbostad, and Rolf Moe-Nilssen. Should trunk movement or footfall parameters quantify gait asymmetry in chronic stroke patients? *Gait & Posture*, 27(4):552–558, 2008.

[134] Cheriel J Hofstad, Harmen van der Linde, Bart Nienhuis, Vivian Weerdesteyn, Jacques Duysens, and Alexander C Geurts. High failure rates when avoiding obstacles during treadmill walking in patients with a transtibial amputation. *Archives of Physical Medicine and Rehabilitation*, 87(8):1115–1122, 2006.

[135] Neville Hogan. An organizing principle for a class of voluntary movements. *Journal of Neuroscience*, 4(11):2745–2754, 1984.

[136] Neville Hogan and Dagmar Sternad. On rhythmic and discrete movements: reflections, definitions and implications for motor control. *Experimental Brain Research*, 181(1):13–30, 2007.

[137] Neville Hogan and Dagmar Sternad. Sensitivity of smoothness measures to movement duration, amplitude, and arrests. *Journal of motor behavior*, 41(6):529–534, 2009.

[138] John H Hollman, Eric M McDade, and Ronald C Petersen. Normative spatiotemporal gait parameters in older adults. *Gait & posture*, 34(1):111–118, 2011.

[139] Fabian Horst, Sebastian Lapuschkin, Wojciech Samek, Klaus-Robert Müller, and Wolfgang I Schöllhorn. Explaining the unique nature of individual gait patterns with deep learning. *Scientific reports*, 9(1):1–13, 2019.

[140] Han Houdijk, Franka M Appelman, Judith M Van Velzen, H Lucas, WV Van Der, and CAM Van Bennekom. Validity of dynaport gaitmonitor for assessment of spatiotemporal parameters in amputee gait. *Journal of Rehabilitation Research and Development*, 45:5–11, 2008.

[141] Jennifer Howcroft, Jonathan Kofman, and Edward D Lemaire. Feature selection for elderly faller classification based on wearable sensors. *Journal of Neuroengineering and Rehabilitation*, 14(1):47, 2017.

[142] Jennifer Howcroft, Jonathan Kofman, Edward D Lemaire, and William E McIlroy. Analysis of dual-task elderly gait in fallers and non-fallers using wearable sensors. *Journal of Biomechanics*, 49(7):992–1001, 2016.

[143] Jennifer Howcroft, Edward D Lemaire, and Jonathan Kofman. Prospective elderly fall prediction by older-adult fall-risk modeling with feature selection. *Biomedical Signal Processing and Control*, 43:320–328, 2018.

[144] Jennifer D Howcroft, Edward D Lemaire, Jonathan Kofman, and William E McIlroy. Analysis of dual-task elderly gait using wearable plantar-pressure insoles and accelerometer. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5003–5006. IEEE, 2014.

[145] David Howell, Louis Osternig, and Li-Shan Chou. Monitoring recovery of gait balance control following concussion using an accelerometer. *Journal of Biomechanics*, 48(12):3364–3368, 2015.

[146] Alan Hreljac. Stride smoothness evaluation of runners and other athletes. *Gait & Posture*, 11(3):199–206, 2000.

[147] Alan Hreljac and Philip E Martin. The relationship between smoothness and economy during walking. *Biological Cybernetics*, 69(3):213–218, 1993.

[148] Chen Huang, Yining Li, Chen Change Loy, and Xiaoou Tang. Learning deep representation for imbalanced classification. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5375–5384, 2016.

[149] Ryan P Hubble, Peter A Silburn, Geraldine A Naughton, and Michael H Cole. Assessing stability in mild and moderate parkinson's disease: Can clinical measures provide insight? *Gait & posture*, 49:7–13, 2016.

[150] James Alexander Hughes, Joseph Alexander Brown, and Adil Mehmood Khan. Smartphone gait fingerprinting models via genetic programming. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 408–415. IEEE, 2016.

[151] James Alexander Hughes, Sheridan Houghten, and Joseph Alexander Brown. Models of parkinson's disease patient gait. *IEEE journal of biomedical and health informatics*, 24(11):3103–3110, 2019.

[152] B Huijben, KS Van Schooten, JH Van Dieën, and M Pijnappels. The effect of walking speed on quality of gait in older adults. *Gait & Posture*, 65:112–116, 2018.

[153] Jessie M Huisinga, Martina Mancini, Rebecca J St George, and Fay B Horak. Accelerometry reveals differences in gait variability between patients with multiple sclerosis and healthy controls. *Annals of Biomedical Engineering*, 41(8):1670–1679, 2013.

[154] Tim Hulsen, Saumya S Jamuar, Alan R Moody, Jason H Karnes, Orsolya Varga, Stine Hedensted, Roberto Spreafico, David A Hafler, and Eoin F McKinney. From big data to precision medicine. *Frontiers in Medicine*, 6:34, 2019.

[155] Espen AF Ihlen, Aner Weiss, Yoav Beck, Jorunn L Helbostad, and Jeffrey M Hausdorff. A comparison study of local dynamic stability measures of daily life walking

in older adult community-dwelling fallers and non-fallers. *Journal of Biomechanics*, 49(9):1498–1503, 2016.

[156] Trienke IJmker and Claudine JC Lamoth. Gait and cognition: the relationship between gait stability and variability with executive function in persons with and without dementia. *Gait & Posture*, 35(1):126–130, 2012.

[157] Keaton A Inkol, Andrew H Huntley, and Lori Ann Vallis. Repeated exposure to forward support-surface perturbation during overground walking alters upper-body kinematics and step parameters. *Journal of motor behavior*, 51(3):318–330, 2019.

[158] Marco Iosa, Augusto Fusco, Giovanni Morone, and Stefano Paolucci. Development and decline of upright gait stability. *Frontiers in Aging Neuroscience*, 6:14, 2014.

[159] Marco Iosa, Tiziana Marro, Stefano Paolucci, and Daniela Morelli. Stability and harmony of gait in children with cerebral palsy. *Research in Developmental Disabilities*, 33(1):129–135, 2012.

[160] Marco Iosa, Daniela Morelli, Enrica Nisi, Carlo Sorbara, Stefano Negrini, Paola Gentili, Stefano Paolucci, and Augusto Fusco. Assessment of upper body accelerations in young adults with intellectual disabilities while walking, running, and dual-task running. *Human Movement Science*, 34:187–195, 2014.

[161] Norio Ishigaki, Teiji Kimura, Yuki Usui, Kaoru Aoki, Nobuyo Narita, Masayuki Shimizu, Kazuo Hara, Nobuhide Ogihara, Koichi Nakamura, Hiroyuki Kato, et al. Analysis of pelvic movement in the elderly during walking using a posture monitoring system equipped with a triaxial accelerometer and a gyroscope. *Journal of Biomechanics*, 44(9):1788–1792, 2011.

[162] Mitsuaki Ishii and Hideaki Mashimo. Accelerometer based analysis of gait initiation failure in advanced juvenile parkinsonism: a single subject study. *Journal of Physical Therapy Science*, 28(11):3252–3256, 2016.

[163] Plamen Ch Ivanov, Kang KL Liu, and Ronny P Bartsch. Focus on the emerging new fields of network physiology and network medicine. *New journal of physics*, 18(10):100201, 2016.

[164] Delaram Jarchi, James Pope, Tracey KM Lee, Larisa Tamjidi, Amirhosein Mirzaei, and Saeid Sanei. A review on accelerometry-based gait analysis and emerging clinical applications. *IEEE Reviews in Biomedical Engineering*, 11:177–194, 2018.

[165] Kimberlee Jordan, John H Challis, and Karl M Newell. Walking speed influences on gait cycle variability. *Gait & Posture*, 26(1):128–134, 2007.

[166] MP Kadaba, HK Ramakrishnan, ME Wootten, J Gainey, G Gorton, and GVB Cochran. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *Journal of Orthopaedic Research*, 7(6):849–860, 1989.

[167] Hyun Gu Kang and Jonathan B Dingwell. Dynamic stability of superior vs. inferior segments during walking in young and older adults. *Gait & Posture*, 30(2):260–263, 2009.

[168] Snehal N Kathale and Supriya Solaskar. A method for identifying human by using gait cycle. In *Intelligent Communication Technologies and Virtual Mobile Networks*, pages 655–666. Springer, 2019.

[169] Justin J Kavanagh and Hylton B Menz. Accelerometry: a technique for quantifying movement patterns during walking. *Gait & Posture*, 28(1):1–15, 2008.

[170] Robert B Kerstein and J Radke. Computer-guided occlusal treatment improves smoothness, timing and velocity of gum chewing. *Adv Dent Technol Tech*, 1(2):11–21, 2019.

[171] Samina Khalid, Tehmina Khalil, and Shamila Nasreen. A survey of feature selection and feature extraction techniques in machine learning. In *2014 science and information conference*, pages 372–378. IEEE, 2014.

[172] H Khalil, R van Deursen, L Quinn, A Rosser, and M Busse. F18 clinical measurement of sit to stand performance in people with huntington's disease: reliability and validity for 30 seconds chair sit to stand test. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(Suppl 1):A28–A28, 2010.

[173] Hiromitsu Kobayashi, Wataru Kakihana, and Tasuku Kimura. Combined effects of age and gender on gait symmetry and regularity assessed by autocorrelation of trunk acceleration. *Journal of Neuroengineering and Rehabilitation*, 11(1):109, 2014.

[174] Dylan Kobsar, Chad Olson, Raman Paranjape, Thomas Hadjistavropoulos, and John M Barden. Evaluation of age-related differences in the stride-to-stride fluctuations, regularity and symmetry of gait using a waist-mounted tri-axial accelerometer. *Gait & Posture*, 39(1):553–557, 2014.

[175] Anton Kocheturov, Panos M Pardalos, and Athanasia Karakitsiou. Massive datasets and machine learning for computational biomedicine: trends and challenges. *Annals of Operations Research*, 276(1-2):5–34, 2019.

[176] Motonaga Kojima, Shuichi Obuchi, Osamu Henmi, and Noriaki Ikeda. Comparison of smoothness during gait between community dwelling elderly fallers and non-fallers using power spectrum entropy of acceleration time-series. *Journal of Physical Therapy Science*, 20(4):243–248, 2008.

[177] Motonaga Kojima, Shuichi Obuchi, Kousuke Mizuno, Osamu Henmi, and Noriaki Ikeda. Power spectrum entropy of acceleration time-series during movement as an indicator of smoothness of movement. *Journal of Physiological Anthropology*, 27(4):193–200, 2008.

[178] Michael R Kosorok and Eric B Laber. Precision medicine. *Annual review of statistics and its application*, 6:263–286, 2019.

[179] John R Koza. *Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems*, volume 34. Stanford University, Department of Computer Science Stanford, CA, 1990.

[180] John R Koza. Genetic programming as a means for programming computers by natural selection. *Statistics and computing*, 4(2):87–112, 1994.

[181] John R Koza and James P Rice. Automatic programming of robots using genetic programming. In *AAAI*, volume 92, pages 194–207. Citeseer, 1992.

[182] Joy P Ku, Jennifer L Hicks, Trevor Hastie, Jure Leskovec, Christopher Ré, and Scott L Delp. The mobilize center: an nih big data to knowledge center to advance human movement research and improve mobility. *Journal of the American Medical Informatics Association*, 22(6):1120–1125, 2015.

[183] Max Kuhn, Kjell Johnson, et al. *Applied predictive modeling*, volume 26. Springer, 2013.

[184] Stefan Kuhn, Björn Egert, Steffen Neumann, and Christoph Steinbeck. Building blocks for automated elucidation of metabolites: Machine learning methods for nmr prediction. *BMC Bioinformatics*, 9(1):1–19, 2008.

[185] Arthur D Kuo, J Maxwell Donelan, and Andy Ruina. Energetic consequences of walking like an inverted pendulum: step-to-step transitions. *Exercise and sport sciences reviews*, 33(2):88–97, 2005.

[186] Robert Kwiecien, Annette Kopp-Schneider, and Maria Blettner. Concordance analysis: part 16 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, 108(30):515, 2011.

[187] Demetrios N Kyriacou and Roger J Lewis. Confounding by indication in clinical research. *Jama*, 316(17):1818–1819, 2016.

[188] Peter A Lachenbruch. Mcnemar test. *Wiley StatsRef: Statistics Reference Online*, 2014.

[189] Daniël Lakens. Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4):355–362, 2017.

[190] CJC Lamoth, Peter J Beek, and Onno G Meijer. Pelvis–thorax coordination in the transverse plane during gait. *Gait & posture*, 16(2):101–114, 2002.

[191] Claudine J Lamoth, Floor J van Deudekom, Jos P van Campen, Bregje A Appels, Oscar J de Vries, and Mirjam Pijnappels. Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people. *Journal of Neuroengineering and Rehabilitation*, 8(1):2, 2011.

[192] Claudine JC Lamoth, Simone R Caljouw, and Klaas Postema. Active video gaming to improve balance in the elderly., 2011.

[193] BS Lay, WA Sparrow, KM Hughes, and NJ O'Dwyer. Practice effects on coordination and control, metabolic energy expenditure, and muscle activation. *Human movement science*, 21(5-6):807–830, 2002.

[194] Brandi S Row Lazzarini and Theodore J Kataras. Treadmill walking is not equivalent to overground walking for the study of walking smoothness and rhythmicity in older adults. *Gait & Posture*, 46:42–46, 2016.

[195] Song-Mi Lee, Sang Min Yoon, and Heeryon Cho. Human activity recognition from accelerometer data using convolutional neural network. In *2017 ieee international conference on big data and smart computing (bigcomp)*, pages 131–134. IEEE, 2017.

[196] Robert Lemoyne, Cristian Coroian, Timothy Mastroianni, and Warren Grundfest. Accelerometers for quantification of gait and movement disorders: a perspective review. *Journal of Mechanics in Medicine and Biology*, 8(02):137–152, 2008.

[197] Shuzhao Li, Andrei Todor, and Ruiyan Luo. Blood transcriptomics and metabolomics for personalized medicine. *Computational and structural biotechnology journal*, 14:1–7, 2016.

[198] Mau-Roung Lin, Hei-Fen Hwang, Ming-Hsia Hu, Hong-Dar Isaac Wu, Yi-Wei Wang, and Fu-Chao Huang. Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and tinetti balance measures in community-dwelling older people. *Journal of the American Geriatrics Society*, 52(8):1343–1348, 2004.

[199] FJ López-Sanromán, R Holmbak-Petersen, I Santiago, IA Gómez de Segura, and Eric Barrey. Gait analysis using 3d accelerometry in horses sedated with xylazine. *The Veterinary Journal*, 193(1):212–216, 2012.

[200] Sue Lord, Tracey Howe, Julia Greenland, Linda Simpson, and Lynn Rochester. Gait variability in older adults: a structured review of testing protocol and clinimetric properties. *Gait & Posture*, 34(4):443–450, 2011.

[201] Kristin A Lowry, Andrew J Carrel, Jessica M McIlrath, and Ann L Smiley-Oyen. Use of harmonic ratios to examine the effect of cueing strategies on gait stability in persons with parkinson's disease. *Archives of Physical Medicine and Rehabilitation*, 91(4):632–638, 2010.

[202] Kristin A Lowry, Ann L Smiley-Oyen, Andrew J Carrel, and John P Kerr. Walking stability using harmonic ratios in parkinson's disease. *Movement Disorders*, 24(2):261–267, 2009.

[203] Kristin A Lowry, Jessie M VanSwearingen, Subashan Perera, Stephanie A Studenski, and Jennifer S Brach. Walking smoothness is associated with self-reported function after accounting for gait speed. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 68(10):1286–1290, 2013.

[204] Michelle M Lusardi, Stacy Fritz, Addie Middleton, Leslie Allison, Mariana Wingood, Emma Phillips, Michelle Criss, Sangita Verma, Jackie Osborne, and Kevin K Chui. Determining risk of falls in community dwelling older adults: a systematic review and meta-analysis using posttest probability. *Journal of geriatric physical therapy (2001)*, 40(1):1, 2017.

[205] Trieu Phat Luu, KH Low, Xingda Qu, HB Lim, and KH Hoon. An individual-specific gait pattern prediction model based on generalized regression neural networks. *Gait & posture*, 39(1):443–448, 2014.

[206] Jamie Macbeth and Majid Sarrafzadeh. Shrinking symbolic regression over medical and physiological signals. In *2010 2nd International Conference on Signal Processing Systems*, volume 1, pages V1–461. IEEE, 2010.

[207] Walter Maetzler, Martina Mancini, Inga Liepelt-Scarfone, Katharina Müller, Clemens Becker, Rob C Van Lummel, Erik Ainsworth, Markus Hobert, Johannes Streffer, Daniela Berg, et al. Impaired trunk stability in individuals at high risk for parkinson's disease. *PLoS One*, 7(3):e32240, 2012.

[208] Brian E Maki. Gait changes in older adults: predictors of falls or indicators of fear? *Journal of the American Geriatrics Society*, 45(3):313–320, 1997.

[209] Arthur I Mallinon and Neil S Longridge. Increasing the usefulness of tandem walking evaluation. *Journal of Otolaryngology–Head & Neck Surgery*, 37(6), 2008.

[210] Martina Mancini, Fay B Horak, Cris Zampieri, Patricia Carlson-Kuhta, John G Nutt, and Lorenzo Chiari. Trunk accelerometry reveals postural instability in untreated parkinson's disease. *Parkinsonism & Related Disorders*, 17(7):557–562, 2011.

[211] Allison Marier, Lauren EW Olsho, William Rhodes, and William D Spector. Improving prediction of fall risk among nursing home residents using electronic medical records. *Journal of the American Medical Informatics Association*, 23(2):276–282, 2015.

[212] Eladio Martin. Novel method for stride length estimation with body area network accelerometers. In *2011 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems*, pages 79–82. IEEE, 2011.

[213] Hiromi Matsumoto, Hiroshi Hagino, Mari Osaki, Shinji Tanishima, Chika Tanimura, Akihiro Matsuura, and Tomoyuki Makabe. Gait variability analysed using an accelerometer is associated with locomotive syndrome among the general elderly population: The gaina study. *Journal of Orthopaedic Science*, 21(3):354–360, 2016.

[214] Eric T McAdams, Claudine Gehin, Norbert Noury, Carolina Ramon, Ronald Nocua, Bertrand Massot, Aurélien Oliveira, André Dittmar, Chris D Nugent, and Jim

McLaughlin. Biomedical sensors for ambient assisted living. In *Advances in Biomedical Sensing, Measurements, Instrumentation and Systems*, pages 240–262. Springer, 2010.

[215] Hylton B Menz, Stephen R Lord, and Richard C Fitzpatrick. Acceleration patterns of the head and pelvis when walking on level and irregular surfaces. *Gait & Posture*, 18(1):35–46, 2003.

[216] Hylton B Menz, Stephen R Lord, and Richard C Fitzpatrick. Age-related differences in walking stability. *Age and Ageing*, 32(2):137–142, 2003.

[217] Jairo H Migueles, Cristina Cadenas-Sanchez, Ulf Ekelund, Christine Delisle Nyström, Jose Mora-Gonzalez, Marie Löf, Idoia Labayen, Jonatan R Ruiz, and Francisco B Ortega. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. *Sports Medicine*, 47(9):1821–1845, 2017.

[218] Jeremy Miles. R squared, adjusted r squared. *Wiley StatsRef: Statistics Reference Online*, 2014.

[219] Alexandre Millecamps, Kristin A Lowry, Jennifer S Brach, Subashan Perera, Mark S Redfern, and Ervin Sejdić. Understanding the effects of pre-processing on extracted signal features from gait accelerometry signals. *Computers in Biology and Medicine*, 62:164–174, 2015.

[220] Kristin M Mills, Scott Sadler, Karen Peterson, and Lorrin Pang. An economic evaluation of preventing falls using a new exercise program in institutionalized elderly. *Journal of physical activity and health*, 15(6):397–402, 2018.

[221] I Minami, K Oogai, T Nemoto, T Nakamura, Y Igarashi, and N Wakabayashi. Measurement of jerk-cost using a triaxial piezoelectric accelerometer for the evaluation of jaw movement smoothness. *Journal of Oral Rehabilitation*, 37(8):590–595, 2010.

[222] Chihiro Mizuike, Shohei Ohgi, and Satoru Morita. Analysis of stroke patient walking dynamics using a tri-axial accelerometer. *Gait & Posture*, 30(1):60–64, 2009.

[223] R Moe-Nilssen. A new method for evaluating motor control in gait under real-life environmental conditions. part 2: Gait analysis. *Clinical Biomechanics*, 13(4-5):328–335, 1998.

[224] Rolf Moe-Nilssen. A new method for evaluating motor control in gait under real-life environmental conditions. part 1: The instrument. *Clinical biomechanics*, 13(4-5):320–327, 1998.

[225] Rolf Moe-Nilssen. A new method for evaluating motor control in gait under real-life environmental conditions. part 1: The instrument. *Clinical Biomechanics*, 13(4-5):320–327, 1998.

[226] Rolf Moe-Nilssen, Mona Kristin Aaslund, Caroline Hodt-Billington, and Jorunn L Helbostad. Gait variability measures may represent different constructs. *Gait & Posture*, 32(1):98–101, 2010.

[227] Rolf Moe-Nilssen and Jorunn L Helbostad. Estimation of gait cycle characteristics by trunk accelerometry. *Journal of Biomechanics*, 37(1):121–126, 2004.

[228] Rolf Moe-Nilssen and Jorunn L Helbostad. Interstride trunk acceleration variability but not step width variability can differentiate between fit and frail older adults. *Gait & Posture*, 21(2):164–170, 2005.

[229] JRW Morris. Accelerometry: A technique for the measurement of human body movements. *Journal of biomechanics*, 6(6):729–736, 1973.

[230] Meg E Morris, Frances Huxham, Jennifer McGinley, Karen Dodd, and Robert Iansek. The biomechanics and motor control of gait in parkinson disease. *Clinical Biomechanics*, 16(6):459–470, 2001.

[231] Niyousha Mortaza, Noor Azuan Abu Osman, and N Mehdikhani. Are the spatio-temporal parameters of gait capable of distinguishing a faller from a non-faller elderly. *Eur J Phys Rehabil Med*, 50(6):677–691, 2014.

[232] Lilian Genaro Motti Ader, Barry R Greene, Killian McManus, Niall Tubridy, and Brian Caulfield. Short bouts of gait data and body-worn inertial sensors can provide reliable measures of spatiotemporal gait parameters from bilateral gait data for persons with multiple sclerosis. *Biosensors*, 10(9):128, 2020.

[233] Malaak Nasser Moussa, Crystal D Vechlekar, Jonathan H Burdette, Matt R Steen, Christina E Hugenschmidt, and Paul J Laurienti. Changes in cognitive state alter human functional brain networks. *Frontiers in human neuroscience*, 5:83, 2011.

[234] Muhammad Muaaz and René Mayrhofer. Orientation independent cell phone based gait authentication. In *Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia*, pages 161–164, 2014.

[235] Marcella Mun-San Kwan, Sang-I Lin, Ching-Huey Chen, Jacqueline CT Close, and Stephen R Lord. Sensorimotor function, balance abilities and pain influence timed up and go performance in older community-living people. *Aging clinical and experimental research*, 23(3):196–201, 2011.

[236] Mario Munoz-Organero and Ramona Ruiz-Blazquez. Time-elastic generative model for acceleration time series in human activity recognition. *Sensors*, 17(2):319, 2017.

[237] United Nations. World population prospects 2019. 2019.

[238] Kirsten K Ness, James G Gurney, and Gillian H Ice. Screening, education, and associated behavioral responses to reduce risk for falls among people over age 65 years attending a community health fair. *Physical Therapy*, 83(7):631–637, 2003.

[239] Shu Nishiguchi, Minoru Yamada, Koutatsu Nagai, Shuhei Mori, Yuu Kajiwara, Takuya Sonoda, Kazuya Yoshimura, Hiroyuki Yoshitomi, Hiromu Ito, Kazuya Okamoto, et al. Reliability and validity of gait analysis by android-based smartphone. *Telemedicine and e-Health*, 18(4):292–296, 2012.

[240] Domen Novak, Peter Reberšek, Stefano Marco Maria De Rossi, Marco Donati, Janez Podobnik, Tadej Beravs, Tommaso Lenzi, Nicola Vitiello, Maria Chiara Carrozza, and Marko Munih. Automated detection of gait initiation and termination using wearable sensors. *Medical Engineering & Physics*, 35(12):1713–1720, 2013.

[241] Obinna Nwanna. *Validation of an Accelerometry Based Method of Human Gait Analysis*. PhD thesis, Cleveland State University, 2014. unpublished thesis.

[242] MN Nyan, FEH Tay, KHW Seah, and YY Sitoh. Classification of gait patterns in the time–frequency domain. *Journal of Biomechanics*, 39(14):2647–2656, 2006.

[243] Tommy Öberg, Alek Karsznia, and Kurt Öberg. Basic gait parameters: reference data for normal subjects, 10-79 years of age. *Journal of Rehabilitation Research and Development*, 30:210–210, 1993.

[244] Ciara M O'Connor, Susannah K Thorpe, Mark J O'Malley, and Christopher L Vaughan. Automatic detection of gait events using kinematic data. *Gait & Posture*, 25(3):469–474, 2007.

[245] Sooyol Ok, Kazuo Miyashita, and Kazunori Hase. Evolving bipedal locomotion with genetic programming-a preliminary report. In *Proceedings of the 2001 Congress on Evolutionary Computation (IEEE Cat. No. 01TH8546)*, volume 2, pages 1025–1032. IEEE, 2001.

[246] Laurent Oudre, Rémi Barrois-Müller, Thomas Moreau, Charles Truong, Aliénor Vienne-Jumeau, Damien Ricard, Nicolas Vayatis, and Pierre-Paul Vidal. Template-based step detection with inertial measurement units. *Sensors*, 18(11):4033, 2018.

[247] Kazuaki Oyake, Tomofumi Yamaguchi, Masafumi Sugasawa, Chihiro Oda, Shigeo Tanabe, Kunitsugu Kondo, Yohei Otaka, and Kimito Momose. Validity of gait asymmetry estimation by using an accelerometer in individuals with hemiparetic stroke. *Journal of Physical Therapy Science*, 29(2):307–311, 2017.

[248] Luca Palmerini, Sabato Mellone, Guido Avanzolini, Franco Valzania, and Lorenzo Chiari. Quantification of motor impairment in parkinson's disease using an instrumented timed up and go test. *IEEE transactions on neural systems and rehabilitation engineering*, 21(4):664–673, 2013.

[249] Luca Palmerini, Laura Rocchi, Sinziana Mazilu, Eran Gazit, Jeffrey M Hausdorff, and Lorenzo Chiari. Identification of characteristic motor patterns preceding freezing of gait in parkinson's disease using wearable sensors. *Frontiers in Neurology*, 8:394, 2017.

[250] American Geriatrics Society Panel on Prevention of Falls in Older Persons and British Geriatrics Society. Summary of the updated american geriatrics society/british geriatrics society clinical practice guideline for prevention of falls in older persons. *Journal of the American Geriatrics Society*, 59(1):148–157, 2011.

[251] Anisoara Paraschiv-Ionescu, Christopher J Newman, Lena Carcreff, Corinna N Gerber, Stephane Armand, and Kamiar Aminian. Locomotion and cadence detection using a single trunk-fixed accelerometer: validity for children with cerebral palsy in daily life-like conditions. *Journal of neuroengineering and rehabilitation*, 16(1):24, 2019.

[252] Ilaria Pasciuto, Elena Bergamini, Marco Iosa, Giuseppe Vannozzi, and Aurelio Cappozzo. Overcoming the limitations of the harmonic ratio for the reliable assessment of gait symmetry. *Journal of Biomechanics*, 53:84–89, 2017.

[253] Aftab E Patla, Allan Adkin, and Tonya Ballard. Online steering: coordination and control of body center of mass, head and body reorientation. *Experimental Brain Research*, 129(4):629–634, 1999.

[254] Kara K Patterson, William H Gage, Dina Brooks, Sandra E Black, and William E McIlroy. Evaluation of gait symmetry after stroke: a comparison of current methods and recommendations for standardization. *Gait & Posture*, 31(2):241–246, 2010.

[255] Massimiliano Pau, Serena Mandaresu, Giuseppina Pilloni, Micaela Porta, Giancarlo Coghe, Maria Giovanna Marrosu, and Eleonora Cocco. Smoothness of gait detects early alterations of walking in persons with multiple sclerosis without disability. *Gait & Posture*, 58:307–309, 2017.

[256] Roshni Mary Peter, Alex Joseph, and KR John. The cost of falls among the elderly in kattankulathur block, tamil nadu. *Indian Journal of Public Health Research & Development*, 10(9), 2019.

[257] Vasco Ponciano, Ivan Miguel Pires, Fernando Reinaldo Ribeiro, Gonçalo Marques, Nuno M Garcia, Nuno Pombo, Susanna Spinsante, and Eftim Zdravevski. Is the timed-up and go test feasible in mobile devices? a systematic review. *Electronics*, 9(3):528, 2020.

[258] Jan M Potter, Alan L Evans, and George Duncan. Gait speed and activities of daily living function in geriatric patients. *Archives of physical medicine and rehabilitation*, 76(11):997–999, 1995.

[259] Stephen J Preece, John Y Goulermas, Laurence PJ Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. Activity identification using body-mounted sensors: a review of classification techniques. *Physiological measurement*, 30(4):R1, 2009.

[260] Michiel Punt, Sjoerd M Bruijn, Kimberley S van Schooten, Mirjam Pijnappels, Ingrid G van de Port, Harriet Wittink, and Jaap H van Dieën. Characteristics of daily life gait in fall and non fall-prone stroke survivors and controls. *Journal of Neuroengineering and Rehabilitation*, 13(1):1–7, 2016.

[261] Lien Quach, Andrew M Galica, Richard N Jones, Elizabeth Procter-Gray, Brad Manor, Marian T Hannan, and Lewis A Lipsitz. The nonlinear relationship between gait speed and falls: the maintenance of balance, independent living, intellect, and zest in the elderly of boston study. *Journal of the American Geriatrics Society*, 59(6):1069–1073, 2011.

[262] R Core Team. R: A language and environment for statistical computing. 2020.

[263] Lisa Reider. Life expectancy and assessment of functional status in older adults. In *Acetabular Fractures in Older Patients*, pages 5–19. Springer, 2020.

[264] Fabienne Reynard, Philippe Vuadens, Olivier Deriaz, and Philippe Terrier. Could local dynamic stability serve as an early predictor of falls in patients with moderate neurological gait disorders? a reliability and comparison study in healthy individuals and in patients with paresis of the lower extremities. *PLoS ONE*, 9(6):e100550, 2014.

[265] Sietse M Rispens, Mirjam Pijnappels, Kimberley S van Schooten, Peter J Beek, Andreas Daffertshofer, and Jaap H van Dieën. Consistency of gait characteristics as determined from acceleration data collected at different trunk locations. *Gait & Posture*, 40(1):187–192, 2014.

[266] F Riva, MJP Toebes, MAGM Pijnappels, R Stagni, and JH Van Dieën. Estimating fall risk with inertial sensors using gait stability measures that do not require step detection. *Gait & Posture*, 38(2):170–174, 2013.

[267] Massimo W Rivolta, Md Aktaruzzaman, Giovanna Rizzo, Claudio L Lafortuna, Maurizio Ferrarin, Gabriele Bovi, Daniela R Bonardi, Andrea Caspani, and Roberto Sassi. Evaluation of the tinetti score and fall risk assessment via accelerometry-based movement analysis. *Artificial intelligence in medicine*, 95:38–47, 2019.

[268] David A Rosenbaum. *Human Motor Control*. Academic press, 2009.

[269] David A Rosenbaum, Katie M Chapman, Chase J Coelho, Lanyun Gong, and Breanna Erin Studenka. Choosing actions. *Frontiers in Psychology*, 4:273, 2013.

[270] Jan Rueterbories, Erika G Spaich, Birgit Larsen, and Ole K Andersen. Methods for gait event detection and analysis in ambulatory systems. *Medical Engineering & Physics*, 32(6):545–552, 2010.

[271] Heydar Sadeghi, Paul Allard, François Prince, and Hubert Labelle. Symmetry and limb dominance in able-bodied gait: a review. *Gait & Posture*, 12(1):34–45, 2000.

[272] Catherine M Said, Patricia A Goldie, Aftab E Patla, Elsie Culham, William A Sparrow, and Meg E Morris. Balance during obstacle crossing following stroke. *Gait & Posture*, 27(1):23–30, 2008.

[273] Arash Salarian, Fay B Horak, Cris Zampieri, Patricia Carlson-Kuhta, John G Nutt, and Kamiar Aminian. itug, a sensitive and reliable measure of mobility. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):303–310, 2010.

[274] Albert Samà, Carlos Pérez-López, J Romagosa, D Rodriguez-Martin, Andreu Català, Joan Cabestany, DA Perez-Martinez, and AleJandro Rodríguez-Molinero. Dyskinesia and motor state detection in parkinson's disease patients with a single movement sensor. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1194–1197. IEEE, 2012.

[275] Guto Leoni Santos, Patricia Takako Endo, Kayo Henrique de Carvalho Monteiro, Elisson da Silva Rocha, Ivanovitch Silva, and Theo Lynn. Accelerometer-based human fall detection using convolutional neural networks. *Sensors*, 19(7):1644, 2019.

[276] Suzana Petrović Savić, Branko Ristić, Nikola Prodanović, and Goran Devedžić. Gait classification using a support vector machine algorithm. In *2020 9th Mediterranean Conference on Embedded Computing (MECO)*, pages 1–4. IEEE, 2020.

[277] Giulia Schifino, Veronica Cimolin, Massimiliano Pau, Maira Jaqueline da Cunha, Bruno Leban, Micaela Porta, Manuela Galli, and Aline Souza Pagnussat. Functional electrical stimulation for foot drop in post-stroke people: Quantitative effects on step-to-step symmetry of gait using a wearable inertial sensor. *Sensors*, 21(3):921, 2021.

[278] Michael Schmidt and Hod Lipson. Comparison of tree and graph encodings as function of problem complexity. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pages 1674–1679. ACM, 2007.

[279] Michael D Schmidt and Hod Lipson. Coevolving fitness models for accelerating evolution and reducing evaluations. In *Genetic Programming Theory and Practice IV*, pages 113–130. Springer, 2007.

[280] Michael D Schmidt and Hod Lipson. Coevolution of fitness predictors. *IEEE Transactions on Evolutionary Computation*, 12(6):736–749, 2008.

[281] Michael D Schmidt, Ravishankar R Vallabhajosyula, Jerry W Jenkins, Jonathan E Hood, Abhishek S Soni, John P Wikswo, and Hod Lipson. Automated refinement and inference of analytical models for metabolic networks. *Physical Biology*, 8(5):055011, 2011.

[282] K Schneider and RF Zernicke. Jerk-cost modulations during the practice of rapid arm movements. *Biological Cybernetics*, 60(3):221–230, 1989.

[283] Roman Schniepp, Ken Möhwald, and Max Wuehr. Clinical and automated gait analysis in patients with vestibular, cerebellar, and functional gait disorders: perspectives and limitations. *Journal of Neurology*, 266(1):118–122, 2019.

[284] Wolfgang I Schöllhorn. Applications of artificial neural nets in clinical biomechanics. *Clinical Biomechanics*, 19(9):876–898, 2004.

[285] Ervin Sejdić, Igor Djurović, and Jin Jiang. Time–frequency feature representation using energy concentration: An overview of recent advances. *Digital Signal Processing*, 19(1):153–183, 2009.

[286] Ervin Sejdić, Kristin A Lowry, Jennica Bellanca, Subashan Perera, Mark S Redfern, and Jennifer S Brach. Extraction of stride events from gait accelerometry during treadmill walking. *IEEE journal of translational engineering in health and medicine*, 4:1–11, 2015.

[287] Ervin Sejdić, Kristin A Lowry, Jennica Bellanca, Subashan Perera, Mark S Redfern, and Jennifer S Brach. Extraction of stride events from gait accelerometry during treadmill walking. *IEEE Journal of Translational Engineering in Health and Medicine*, 4:1–11, 2016.

[288] Ervin Sejdić, Kristin A Lowry, Jennica Bellanca, Mark S Redfern, and Jennifer S Brach. A comprehensive assessment of gait accelerometry signals in time, frequency and time-frequency domains. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):603, 2014.

[289] Ervin Sejdić, Catriona M Steele, and Tom Chau. The effects of head movement on dual-axis cervical accelerometry signals. *BMC Research Notes*, 3(1):269, 2010.

[290] Masaki Sekine, Toshiyo Tamura, Masaki Yoshida, Yuki Suda, Yuichi Kimura, Hiroaki Miyoshi, Yoshifumi Kijima, Yuji Higashi, and Toshiro Fujimoto. A gait abnormality

measure based on root mean square of trunk acceleration. *Journal of Neuroengineering and Rehabilitation*, 10(1):118, 2013.

[291] R Senden, B Grimm, IC Heyligers, HHCM Savelberg, and K Meijer. Acceleration-based gait test for healthy subjects: reliability and reference data. *Gait & Posture*, 30(2):192–196, 2009.

[292] R Senden, K Meijer, IC Heyligers, HHCM Savelberg, and B Grimm. Importance of correcting for individual differences in the clinical diagnosis of gait disorders. *Physiotherapy*, 98(4):320–324, 2012.

[293] R Senden, HHCM Savelberg, B Grimm, IC Heyligers, and K Meijer. Accelerometry-based gait analysis, an additional objective approach to screen subjects at risk for falling. *Gait & Posture*, 36(2):296–300, 2012.

[294] Rachel Senden, Bernd Grimm, Kenneth Meijer, Hans Savelberg, and Ide C Heyligers. The importance to including objective functional outcomes in the clinical follow up of total knee arthroplasty patients. *The Knee*, 18(5):306–311, 2011.

[295] Rachel Senden, Ide C Heyligers, Kenneth Meijer, Hans Savelberg, and Bernd Grimm. Acceleration-based motion analysis as a tool for rehabilitation: exploration in simulated functional knee limited walking conditions. *American Journal of Physical Medicine & Rehabilitation*, 90(3):226–232, 2011.

[296] S Shin, D An, and W Yoo. Effect of visual acuity in older females on energy expenditure during obstacle navigation. *European Geriatric Medicine*, 7(2):105–110, 2016.

[297] Joana Silva, João Madureira, Cláudia Tonelo, Daniela Baltazar, Catarina Silva, Anabela Martins, Carlos Alcobia, and Inês Sousa. Comparing machine learning approaches for fall risk assessment. In *International Conference on Bio-inspired Systems and Signal Processing*, volume 5, pages 223–230. SCITEPRESS, 2017.

[298] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[299] Gary L Smidt, Jasbir S Arora, and Richard C Johnston. Accelerographic analysis of several types of walking. *American Journal of Physical Medicine & Rehabilitation*, 50(6):285–300, 1971.

[300] Jacob J Sosnoff, Myla D Goldman, and Robert W Motl. Real-life walking impairment in multiple sclerosis: preliminary comparison of four methods for processing accelerometry data. *Multiple Sclerosis Journal*, 16(7):868–877, 2010.

[301] Julien Stamatakis, Julien Crémers, Didier Maquet, Benoit Macq, and Gaëtan Garraux. Gait feature extraction in parkinson's disease using low-cost accelerometers. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 7900–7903. IEEE, 2011.

[302] Ljubisa Stankovic, Milos Dakovic, and Thayannathan Thayaparan. *Time-Frequency Signal Analysis with Applications*. Artech House, 2014.

[303] Naoto Takayanagi, Motoki Sudo, Yukari Yamashiro, Sangyoon Lee, Yoshiyuki Kobayashi, Yoshifumi Niki, and Hiroyuki Shimada. Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. *Scientific Reports*, 9(1):1–6, 2019.

[304] Ayumi Tanigawa, Saori Morino, Tomoki Aoyama, and Masaki Takahashi. Gait analysis of pregnant patients with lumbopelvic pain using inertial sensor. *Gait & Posture*, 65:176–181, 2018.

[305] Weijun Tao, Tao Liu, Rencheng Zheng, and Hutian Feng. Gait analysis using wearable sensors. *Sensors*, 12(2):2255–2283, 2012.

[306] Philippe Terrier and Olivier Dériaz. Kinematic variability, fractal dynamics and local dynamic stability of treadmill walking. *Journal of NeuroEngineering and Rehabilitation*, 8(1):12, February 2011.

[307] Philippe Terrier and Fabienne Reynard. Effect of age on the variability and stability of gait: a cross-sectional treadmill study in healthy individuals between 20 and 69 years of age. *Gait & Posture*, 41(1):170–174, 2015.

[308] Philip D Thompson. Gait disorders. *Neurology in Clinical Practice: Principles of Diagnosis and Management*, 1:323, 2004.

[309] Stefan Timmermans and Marc Berg. *The Gold Standard: The Challenge of Evidence-based Medicine and Standardization in Health Care*. Temple University Press, 2010.

[310] Kai Ming Ting. Confusion matrix. *Encyclopedia of Machine Learning and Data Mining*, 260, 2017.

[311] Yuki Tochigi, Neil A Segal, Tanawat Vaseenon, and Thomas D Brown. Entropy analysis of tri-axial leg acceleration signal waveforms for measurement of decrease of physiological variability in human gait. *Journal of Orthopaedic Research*, 30(6):897–904, 2012.

[312] Diana TroJaniello, Andrea Ravaschio, Jeffrey M Hausdorff, and Andrea Cereatti. Comparative assessment of different methods for the estimation of gait temporal parameters using a single inertial sensor: application to elderly, post-stroke, parkinson's disease and huntington's disease subjects. *Gait & Posture*, 42(3):310–316, 2015.

[313] Phuc Huu Truong, Jinwook Lee, Ae-Ran Kwon, and Gu-Min Jeong. Stride counting in human walking and walking distance estimation using insole sensors. *Sensors*, 16(6):823, 2016.

[314] John W Tukey. *Exploratory data analysis*, volume 2. Reading, Mass., 1977.

[315] Can Tunca, Gülüstü Salur, and Cem Ersoy. Deep learning for fall risk assessment with inertial sensors: utilizing domain knowledge in spatio-temporal gait parameters. *IEEE journal of biomedical and health informatics*, 24(7):1994–2005, 2019.

[316] Andrea Tura, Michele Raggi, Laura Rocchi, Andrea G Cutti, and Lorenzo Chiari. Gait symmetry and regularity in transfemoral amputees assessed by trunk accelerations. *Journal of NeuroEngineering and Rehabilitation*, 7(1):4, 2010.

[317] Wim Van Drongelen. *Signal Processing for Neuroscientists*. Academic press, 2018.

[318] Kimberley S van Schooten, Mirjam Pijnappels, Sietse M Rispens, Petra JM Elders, Paul Lips, and Jaap H van Dieën. Ambulatory fall-risk assessment: amount and quality of daily-life gait predict falls in older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 70(5):608–615, 2015.

[319] Kimberley S van Schooten, Sietse M Rispens, Mirjam Pijnappels, Andreas Daffertshofer, and Jaap H van Dieen. Assessing gait stability: the influence of state space reconstruction on inter-and intra-day reliability of local dynamic stability during over-ground walking. *Journal of Biomechanics*, 46(1):137–141, 2013.

[320] Jessie M VanSwearingen, Subashan Perera, Jennifer S Brach, David Wert, and Stephanie A Studenski. Impact of exercise to improve gait efficiency on activity and participation in older adults with mobility limitations: a randomized controlled trial. *Physical Therapy*, 91(12):1740–1751, 2011.

[321] Jessie M VanSwearingen and Stephanie A Studenski. Aging, motor skill, and the energy cost of walking: implications for the prevention and treatment of mobility decline in older persons. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 69(11):1429, 2014.

[322] Peter H Veltink, HansB J Bussmann, Wiebe De Vries, WimL J Martens, and Rob C Van Lummel. Detection of static and dynamic activities using uniaxial accelerometers. *IEEE Transactions on Rehabilitation Engineering*, 4(4):375–385, 1996.

[323] Joe Verghese, Roee Holtzer, Richard B Lipton, and Cuiling Wang. Quantitative gait markers and incident fall risk in older adults. *The Journals of Gerontology: Series A*, 64(8):896–901, 2009.

[324] Aliénor Vienne, Rémi P Barrois, Stéphane Buffat, Damien Ricard, and Pierre-Paul Vidal. Inertial sensors to assess gait quality in patients with neurological disorders: a systematic review of technical and analytical challenges. *Frontiers in psychology*, 8:817, 2017.

[325] James C Wall, Julie Devlin, Rhonda Khirchof, and Brian Lackey. Measurement of step widths and step lengths: a comparison of measurements made directly from a grid with those made from a video recording. *Journal of Orthopaedic & Sports Physical Therapy*, 30(7):410–417, 2000.

[326] Elizabeth Walshe. *Falling Head Over Heels: Investigating the higher-level cognitive and electrophysiological processes underlying gait control and falls in older adults and stroke survivors*. PhD thesis, National University of Ireland Maynooth, 2016.

[327] Liang Wang, Weiming Hu, and Tieniu Tan. A new attempt to gait-based human identification. In *Object recognition supported by user interaction for service robots*, volume 1, pages 115–118. IEEE, 2002.

[328] Ning Wang, Eliathamby Ambikairajah, Branko G Celler, and Nigel H Lovell. Accelerometry based classification of gait patterns using empirical mode decomposition. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 617–620. IEEE, 2008.

[329] Ning Wang, Eliathamby Ambikairajah, Stephen J Redmond, Branko G Celler, and Nigel H Lovell. Classification of walking patterns on inclined surfaces from accelerometry data. In *2009 16th International Conference on Digital Signal Processing*, pages 1–4. IEEE, 2009.

[330] Robert L Waters and Sara Mulroy. The energy expenditure of normal and pathologic gait. *Gait & Posture*, 9(3):207–231, 1999.

[331] Alexander Weiss, T Herman, M Plotnik, M Brozgol, N Giladi, and JM Hausdorff. An instrumented timed up and go: the added value of an accelerometer for identifying fall risk in idiopathic fallers. *Physiological measurement*, 32(12):2003, 2011.

[332] Aner Weiss, Talia Herman, Meir Plotnik, Marina Brozgol, Inbal Maidan, Nir Giladi, Tanya Gurevich, and Jeffrey M Hausdorff. Can an accelerometer enhance the utility of the timed up & go test when evaluating patients with parkinson's disease? *Medical engineering & physics*, 32(2):119–125, 2010.

[333] Aner Weiss, Sarvi Sharifi, Meir Plotnik, Jeroen PP van Vugt, Nir Giladi, and Jeffrey M Hausdorff. Toward automated, at-home assessment of mobility among patients with parkinson disease, using a body-worn accelerometer. *Neurorehabilitation and Neural Repair*, 25(9):810–818, 2011.

[334] Michael W Whittle. *Gait Analysis: An Introduction*. Butterworth-Heinemann, 2014.

[335] A Th M Willemsen, Jan A van Alste, and HBK Boom. Real-time gait assessment utilizing a new way of accelerometry. *Journal of Biomechanics*, 23(8):859–863, 1990.

[336] DA Winter, CD MacKinnon, GK Ruder, and C Wieman. An integrated emg/biomechanical model of upper body balance and posture during human gait. In *Progress in Brain Research*, volume 97, pages 359–367. Elsevier, 1993.

[337] David A Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[338] Alan Wolf, Jack B Swift, Harry L Swinney, and John A Vastano. Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–317, 1985.

[339] Raymond E Wright. Logistic regression. 1995.

[340] Jun Yabuki, Kazuto Yamaguchi, Takashi Ito, Kazunori Akizuki, and Yukari Ohashi. Relationship between correcting error and smoothness of movement in the process of motor learning. *Journal of Physical Therapy Science*, 32(10):691–697, 2020.

[341] H John Yack and Roseanne C Berger. Dynamic stability in the elderly: identifying a possible measure. *Journal of Gerontology*, 48(5):M225–M230, 1993.

[342] Ryota Yamaguchi, Soichiro Hirata, Takehiko Doi, Tsuyoshi Asai, Junichiro Inoue, Daisuke Makiura, Hiroshi Ando, Masahiro Kurosaka, and Yasushi Miura. The usefulness of a new gait symmetry parameter derived from lissajous figures of tri-axial acceleration signals of the trunk. *Journal of Physical Therapy Science*, 24(5):405–408, 2012.

[343] Che-Chang Yang and Yeh-Liang Hsu. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors*, 10(8):7772–7788, 2010.

[344] Che-Chang Yang, Yeh-Liang Hsu, Kao-Shang Shih, and Jun-Ming Lu. Real-time gait cycle parameter recognition using a wearable accelerometry system. *Sensors*, 11(8):7314–7326, 2011.

[345] Mitsuru Yoneyama, Yosuke Kurihara, Kajiro Watanabe, and Hiroshi Mitoma. Accelerometry-based gait analysis and its application to parkinson's disease assessment: Part 2: A new measure for quantifying walking behavior. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(6):999–1005, 2013.

[346] Mitsuru Yoneyama, Yosuke Kurihara, Kajiro Watanabe, and Hiroshi Mitoma. Accelerometry-based gait analysis and its application to parkinson's disease assessment: part 1: detection of stride event. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3):613–622, 2014.

[347] Mitsuru Yoneyama, Hiroshi Mitoma, Nobuo Sanjo, Maya Higuma, Hiroo Terashi, and Takanori Yokota. Ambulatory gait behavior in patients with dementia: a comparison with parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(8):817–826, 2016.

[348] Jang-Hee Yoo, Doosung Hwang, Ki-Young Moon, and Mark S Nixon. Automated human recognition by gait using neural network. In *2008 First Workshops on Image Processing Theory, Tools and Applications*, pages 1–6. IEEE, 2008.

[349] Jang-Hee Yoo, Mark S Nixon, and Chris J Harris. Extracting human gait signatures by body segment properties. In *Proceedings Fifth IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 35–39. IEEE, 2002.

[350] Agnes Zijlstra, Eling D de Bruin, Nienke Bruins, and Wiebren Zijlstra. The step length–frequency relationship in physically active community-dwelling older women. *European Journal of Applied Physiology*, 104(3):427, 2008.

[351] Wiebren Zijlstra. Assessment of spatio-temporal parameters during unconstrained walking. *European Journal of Applied Physiology*, 92(1-2):39–44, 2004.

[352] Wiebren Zijlstra and At L Hof. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait & Posture*, 18(2):1–10, 2003.

[353] Cong Zong, Mohamed Chetouani, and Adriana Tapus. Automatic gait characterization for a mobility assistance system. In *2010 11th International Conference on Control Automation Robotics & Vision*, pages 473–478. IEEE, 2010.

[354] Qin Zou, Yanling Wang, Qian Wang, Yi Zhao, and Qingquan Li. Deep learning-based gait recognition using smartphones in the wild. *IEEE Transactions on Information Forensics and Security*, 15:3197–3212, 2020.