1-1-2021

Personalized Stroke Rehabilitation

Bryant A. Seamon
Medical University of South Carolina

PERSONALIZED STROKE REHABILITATION

BY

Bryant A. Seamon

A dissertation submitted to the faculty of the Medical University of South Carolina in partial fulfillment of the requirements for the degree Doctor of Philosophy
in the College of Health Professions

DEDICATION

To God who was there every step of the way and has given me joy from the thrill of scientific discovery. It has been a precious gift to learn the secrets of creation and share them with others.

To Cameron Seamon, who has been by my side and supported my goals every day of our life together. Each day together is exciting and full. With you as my teammate, I am fiercely confident we can take on anything.

To Jaylen Seamon, who arrived in the middle of this journey, reminding me how to play and how fun life can be. I look forward to seeing the man you will become.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the contributions of several exceptional people. I want to take a moment to express my gratitude.

I would first like to highlight the support of my committee members during my graduate studies. I have shared several excitements and disappointments with them over my training and many of our conversations laid the framework for the work in this dissertation. Dr. Steve Kautz, my committee chair, has been an incredible mentor. From my first step on campus as an eager physical therapist, ready to change clinical practice, to today, he has offered an abundance of guidance, support, and advice. While Dr. Kautz did give me an exceptional amount of expertise in studying human movement, he also taught me how to approach challenges, think about questions, and how to convey my stories. I am so thankful that he saw a scientist in me before I ever could see it myself. Next, Dr. Mark Bowden has shaped my transition from the clinic to the laboratory and classroom by giving me a roadmap to success. Dr. Bowden was one of the first to instill in me a passion for rehabilitation during my undergraduate training and still serves as an inspiration for where I hope my career will lead. The way I think about integration of research from the lab to clinical practice and training is from his example. Lastly, Dr. Craig Velozo opened my eyes to the potential for measurement to radically change clinical practice. Much of the way I approach thinking about science can be credited to our many conversations about current projects, grants in progress, and banter on life.

In addition to my committee members, I would like to recognize Dr. Shraddha Srivastava, Heather Knight, and Dr. Viswanathan Ramakrishnan. Each of these individuals provided me with help and contributed to data collection, analysis, and statistical support. However, more importantly, they taught me the benefit of working on a team and how powerful diversity of thought is when thinking about how to approach a problem.

Next, I would like to acknowledge the research participants who donated their time and data for these studies to be possible. Their contributions can never be overlooked because without them, science would not be able to happen. I also want to thank my patients. My patients have always inspired and motivated my research. They provided the foundation for my curiosity, supplied many of my research ideas, and motivated me to better clinical practice.

Lastly, I have been extremely fortunate to have the support of funding during my training. I want to acknowledge the financial support from the Department of Veteran’s Affairs, the Foundation for Physical Therapy Research, the National Institutes of Health, and the Medical University of South Carolina.

Abstract of Dissertation Presented to the
Doctor of Philosophy Program in Health and Rehabilitation Science
Medical University of South Carolina
In Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

PERSONALIZED STROKE REHABILITATION

By

Bryant A. Seamon

Chairperson: Steven A. Kautz, PhD
Committee: Mark G. Bowden, PT, PhD
Craig A. Velozo, OTR/L, PhD

Body of Abstract

Stroke continues to be a leading cause of disability. Many individuals who suffer a stroke will receive specialized rehabilitation designed to maximize recovery and restore independence with daily activities. Yet, recovery is still highly variable with heterogeneity in treatment response that is still poorly understood. In recent years, the growing burden of stroke has pushed for moving away from a “one size fits all” approach with the concept of precision or personalized rehabilitation to generate evidence for clinicians to provide the right care, to the right patient, at the right time. The collection of research in this dissertation addresses key areas of personalized rehabilitation related to standardized outcome measurement, identification of biomarkers, and individualized response to treatment parameters in three parts. Part I illustrates how principals of item response theory can be used to inform a personalized measurement approach through clinical applications of Rasch analysis. Part II demonstrates how muscle coordination is linked with biomechanical variables of walking performance that can be used as potential biomarkers of recovery. Lastly, Part III shows how interindividual differences in treatment response could inform individualized prescription for transcranial direct stimulation for recovery of walking post-stroke as an example of how to approach personalized rehabilitation.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 2

CHAPTER 4

PART I

MANUSCRIPT #1

MANUSCRIPT #2

MANUSCRIPT #3

PART II

MANUSCRIPT #1

PART III

MANUSCRIPT #1

LIST OF TABLES

CHAPTER 4

ABBREVIATIONS

ABC	Activities-specific Balance Confidence Scale
ANPT	Academy of Neurologic Physical Therapy
AP	Anterior-posterior
AUC	Area Under the Curve
CFA	Confirmatory factor analysis
CI	Confidence interval
cMDC	Conditional Minimal Detectable Change
COBRE	Center for Biomedical Research Excellence
EFA	Exploratory factor analysis
EMG	Electromyography
FGA	Functional Gait Assessment
FM	Fugl- Meyer
GM	Gluteus medius
GRF	Ground reaction force
ICC	Intraclass correlation coefficient
IQR	Interquartile range
LE	Lower extremity
LEAPS	Locomotor-Experience Applied Post-Stroke
LH	Lateral hamstring
MDC	Minimal Detectable Change
MG	Medial gastrocnemius
MH	Medial hamstring
NIH	National Institutes of Health
NINDS	Neurological Disorders and Stroke
NNMF	Non-negative Matrix Factorization
PE	Parameter estimate
RF	Rectus femoris
SD	Standard deviation
SE	Standard error
SEM	Standard Error of Measurement
SO	Soleus
TA	Tibialis anterior
tDCS	Transcranial Direct Current Stimulation
TMS	Transcranial Magnetic Stimulation
VM	Vastus medialis

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION TO THE PROBLEM

Unfortunately, over two thirds of individuals who suffer a stroke will not make a full recovery (Wade, Langton-Hewer, Wood, Skilbeck, & Ismail, 1983). Stroke continues to be a leading cause of disability with 80% of stroke survivors have lasting sensorimotor deficits (Wade et al., 1983) and 50% not returning to community levels of ambulation (Perry, Garrett, Gronley, & Mulroy, 1995) in the United States. In the United States approximately 610,000 individuals will suffer a stroke for the first time next year (Benjamin et al., 2018) and many of them will receive specialized rehabilitation services designed to restore home and community independence (Winstein et al., 2016).

Yet, despite advances in stroke rehabilitation care, recovery is highly heterogeneous and interindividual variability in treatment response remains poorly understood (Bernhardt et al., 2017). Stroke is unique in this problem because of innate heterogeneity in the pathological presentation of the disease. Stroke rehabilitation research and clinical practice have often used a “one size fits all” approach because there is a lack of evidence that addresses interindividual variability (Cramer et al., 2017). Addressing interindividual variability should provide researchers and clinicians with evidence

for administering individuals the right treatment at the right time to improve outcomes, otherwise known as precision or personalized rehabilitation.

The emphasis on personalized rehabilitation has increased in response to the growing burden stroke has placed on the United States. In 2016, the National Institute of Neurological Disorders and Stroke (NINDS) formed the National Institutes of Health (NIH) StrokeNet to develop a strategic plan for advancing personalization of stroke rehabilitation. The group presented the need for characterizing interindividual differences with respect to treatment response as one of the key specific issues to address. They list potential sources of interindividual differences including variation in (1) neural injury and clinical presentation, (2) treatment response, and (3) time since initial injury in regard to natural recovery and neural plasticity. In response to these sources of variation, the group stressed the need for research to (1) advance the use of standardized outcome measures, (2) identify of biomarkers or phenotypes to best inform subgrouping individuals, and (3) evaluate individual or subgroup response to treatments (Cramer et al., 2017).

The overarching goal of this dissertation is to address each of these three needs to generate evidence for personalized stroke rehabilitation in research and clinical practice. The dissertation is broken into three parts, with each part addressing an identified research area and subsequent specific aims.

Part I addresses standardized outcome measurement. Part I examines how item-level psychometrics and clinical applications of item response theory can allow standardized outcome measures in stroke rehabilitation to be used to

inform personalized clinical decision making. Part II addresses the need to identify biomarkers. Part II will examine whether biomechanical variables of walking performance can serve as a biomarker for impairments to walking-specific motor control which is directly linked to neural damage. Part III will address individual treatment response and examining whether there is interindividual response to different prescriptions of non-invasive brain stimulation.

1.2 RESEARCH AIMS

Part I: Item Response Theory and Clinical Measurement

The primary goal for Part I of the dissertation is to analyze two common standardized measures in stroke rehabilitation using item response theory and present the benefits of individual-level measurement data. Specifically, Part I will use the Rasch model to derive item-level psychometrics and item difficulty hierarchies for the Activities-specific Balance Confidence Scale (ABC) and the Functional Gait Assessment (FGA). Part I will demonstrate the benefit of using item-level psychometrics for personalization of treatment planning (FGA ability map demonstration) and deriving patient specific change scores (ABC scale conditional minimal detectable change demonstration).

Aim 1: Evaluate item-level psychometrics for a patient-reported (the Activities-specific Balance Confidence Scale) and clinician-observed (the

Functional Gait Assessment) outcome measure commonly used in stroke rehabilitation.

Aim 2: Demonstrate how item-level psychometrics can be used to inform personalized treatment planning using an ability map with the Functional Gait Assessment.

Aim 3: Demonstrate how item-level psychometrics can be used to derive change scores using individual-level patient data with the Activities-specific Balance Confidence Scale.

Part II: Searching for Biomarkers to Inform Patient Subgroups

The primary goal for Part II of the dissertation is to evaluate the potential for biomechanical variables to serve as a biomarker for stroke rehabilitation of walking-specific motor control impairments. Part II will quantify the association between walking-specific motor control and biomechanical variables of walking performance in healthy individuals. Understanding the association in healthy individuals will help inform data reduction techniques to identify which biomechanical variables have the strongest ability to serve as a biomarker. This information can then be used to inform personalized treatments to address biomechanical deficits that impact underlying deficits to motor control caused by the stroke.

Aim 1: Quantify the association between walking-specific motor control and quantifiable biomechanical variables during steady-state self-selected walking for healthy individuals.

Part III: Evaluating Treatment Prescription

The primary goal for Part III of the dissertation is to demonstrate a framework for examining the effects of manipulating available parameters for an intervention on patient specific impairments. Part III will specifically examine the effects of manipulating tDCS electrode montage on walking-specific motor control during post-stroke walking.

Aim 1: Determine whether electrode montage type acutely influences walking performance.

CHAPTER 2

LITERATURE REVIEW

2.1 BACKGROUND

This chapter will review the scientific literature relative to each part of the dissertation including topics related to measurement theory, measurement of walking-specific motor control and mechanics, and transcranial direct current stimulation (tDCS) prescription and treatment effects. Knowledge gaps and research needs are identified throughout this chapter.

2.2 MEASUREMENT THEORY:

Classical Test Theory and Item Response Theory

Standardized outcome measures in rehabilitation are primarily patient-reported or clinician-observed tools. Traditionally, these types of measurement tools have been informed by classical test theory. Classical test theory is used to derive a model for test score into (1) the true score and (2) error (Albano, 2016). The relative simplicity of this model supports why it has been the primary choice for examining a tool's measurement properties and deriving psychometrics. However, the relative simplicity of the model is also its primary limitation; psychometrics are sample and test dependent (Albano, 2016). Sample and test dependency lack an overall ability estimate (a measure of the underlying

construct of interest) and score interpretations are not always generalizable across samples. Another limitation of classical test theory is the reliance on group-level data for generating error estimates. Group-level data assumes that error is the same across a measurement scale and limits the interpretability of individual-level data. The limitations of classical test theory have some distinct implications in research and clinical practice. The first, is that individual scores are difficult to interpret and often lack meaning for informing decision making (Haley & Fragala-Pinkham, 2006; Velozo & Woodbury, 2011). The second, is that group-level data error estimates may cause misinterpretation of individual-level measures, especially when assessing for change (Haley & Fragala-Pinkham, 2006; Kozlowski, Cella, Nitsch, Heinemann, & Rehabilitation, 2016; Paul W. Stratford et al., 1996). These implications are likely the reason that implementation of standardized outcome measures in physical therapy have been met with several barriers. For example, physical therapists have cited concerns about time (administration time is lengthy, requires more time than information is worth), lack of clinically relevant information (information is too subjective, items are not relevant to the patient), and limited interpretability (difficult to interpret, does not contribute to the plan of care) as barriers to measurement use (Jette, Halbert, Iverson, Miceli, & Shah, 2009; J. G. Stevens & Beurskens, 2010). These barriers may be addressed by another approach to examining test psychometrics, item response theory.

Item response theory places a focus on individual items in contrast to the overall score (Velozo & Woodbury, 2011). Item response theory fits statistical

models to response data for estimating parameters that represent the relative locations of persons and items on the continuum of a latent ability (Thissen & Wainer, 2001). These statistical models generate an item difficulty hierarchy that can be used to assess the clinical and theoretical validity of the instrument. Compared to classical test theory, item response theory attempts to measure the person's ability level relative to item difficulty. This approach generates an ability estimate that is sample and test independent (Velozo, Kielhofner, & Lai, 1999). Item response theory models also generate item level psychometrics and measure-level error estimates (Kozlowski et al., 2016). These two benefits of item response theory may be useful for providing meaning to individual-level scores and allow individual-level data to be the primary source of decision making. These benefits may also address many of the barriers listed by physical therapists preventing standardized measurement use by addressing the limitations of classical test theory.

The Rasch Model

The Rasch model is a one parameter logistic model and the simplest of item response theory models (Bond & Fox, 2015). The model is based on a probabilistic relationship between item difficulty and person ability. Persons should have a higher probability of passing items that are easier than their ability level and a lower probability of passing items that are harder. Persons should also have about a 50/50 chance of passing items that are at their difficulty level. The Rasch model examines the fit of individual's response data to generate an

item difficulty hierarchy against the distribution of individual's ability level on an interval logit scale. Fit statistics are used to understand how well items and persons fit the Rasch model and whether items overlap in their difficulty.

The Rasch model has two key assumptions that need to be evaluated. The first is that the scale is unidimensional, meaning that the scale is only measuring one construct. The second, is that the scale uses an ordered rating scale. Factor analysis methods can be used to test and explore for additional constructs (Reeve et al., 2007) and fit statistics from the Rasch analysis can be used to evaluate the rating scale structure (J.M. Linacre, 2003), as well as test the implications for misfitting persons or items (Wright & Linacre, 1994).

Applications of the Rasch Model and Connections with Personalized Rehabilitation

The ability of the Rasch model to generate an item difficulty hierarchy, interval measure scale, and measure-level error estimates has broad implications for how to approach standardized measurement in research and clinical use. We explore two applications, keyforms and conditional minimal detectable change, that have direct implications for personalized rehabilitation and may address many barriers limiting standardize measurement use.

Keyforms

Item difficulty hierarchies have provided a means for examining the validity of a scale against theoretical and clinical expectations, but they also give

understanding for how individuals may move along a construct's continuum (Haley & Fragala-Pinkham, 2006). Measures from the Rasch model correspond with this continuum and can give meaning to an individual's score. Typically, the relationship of item difficulty to ability level is visually presented using an item-person map, also called a Wright Map. Item-person maps have different varieties but generally present the distribution of person ability against item difficulty on the measure scale. These maps are ideal for examining how well the item difficulty hierarchy matches the underlying construct and understanding how much of the ability is captured by the overall scale. However, these maps do not provide information with respect to a specific individual. Linacre et. al. introduced the keyform to address this need (J. M. Linacre, 1997).

Initially, the keyform was developed to provide a way for clinicians or researchers to generate instantaneous measure values using the Rasch informed scale (Kielhofner, Dobria, Forsyth, & Basu, 2005; J. M. Linacre, 1997). Keyforms are generated from the Rasch analysis and show the item hierarchy on the left vertical axis with items progressing from easiest to hardest with the easiest item on the bottom. Items are clustered together by difficulty level and these clusters are separated by a blank horizontal space. Each row extending horizontally from an item contains each of the available categories on the rating scale. The x-axis shows the interval measurement scale. Raters can use keyforms by circling the persons score on each item and then drawing a vertical line "through" the bulk of the circles (J. M. Linacre, 1997). The point where the line intersects the x-axis is an estimate of the person's measure. Hence the idea

of spontaneous measurement. However, keyforms can also provide an ability map for understanding a person's measure with respect to their ability along the latent construct (Grattan, Velozo, Skidmore, Page, & Woodbury, 2019; Velozo & Woodbury, 2011; Woodbury et al., 2016).

The concept of an ability map was presented by Haley et. al. in 2006 (Haley & Fragala-Pinkham, 2006) and built upon by others (Grattan et al., 2019; Velozo & Woodbury, 2011; Woodbury et al., 2016). Keyforms show two key pieces of information about an individual: their measure, and their response pattern to items. A rater can look at a keyform and instantly see how an individual's measure corresponds to their ability (i.e., response pattern) with respect to the items' difficulties. This representation can be thought of as the ability map. In rehabilitation, this type of map is important for personalization because the ability map can show how someone should progress along the construct of interest's continuum with respect to their current ability. Thus, individual measure values have meaning with respect to a person's ability level and with respect to how they should progress over time. Items at the person's ability level reflect tasks that are representative of their current place on the continuum and could be thought of as the "just right challenge" for an individual. Similarly, items that are harder for an individual can be used to set goals for rehabilitation and recovery (Grattan et al., 2019; Velozo & Woodbury, 2011; Woodbury et al., 2016). This approach to measurement informs how to interpret measured values for personalizing rehabilitation decision making, treatment design, and care plans.

Conditional Minimal Detectable Change

Error estimates are important for understanding the amount of error associated with a person's score, but they also are critical for identifying change. Historically, error estimates from classical test theory were used to calculate a minimal detectable change (MDC) threshold. MDCs are often presented with a confidence interval and represent the amount of change necessary to exceed the measurement error of the scale (Riddle & Stratford, 2013; Paul W. Stratford et al., 1996; P. W. Stratford & Binkley, 1999). MDCs are used by researchers and clinicians for classifying individuals as having a detectable change for pre- and post-tests. However, there are some key limitations to the MDC approach.

One limitation is that MDCs rely on a single error estimates derived from classical test theory which is problematic because of the concerns presented earlier regarding classical test theory. First, MDCs are derived using a single error estimate that is assumed to be stable across all measures on a scale. However, statistically, we know that this isn't the case and that error changes across a measurement scale with more error being associated with extreme ends and less error at the midrange of a scale (J. M. Linacre, 2007). The idea that different measures have different amounts of error was recognized by Stratford et. al. when introducing the MDC concept for physical therapists but the single error estimate was preferred because it was less computationally intensive under classical test theory (Paul W. Stratford et al., 1996). Second, MDC's are sample and test dependent. Thus, MDC's have limited generalizability to persons with

characteristics that do not match the sample characteristics from which the MDC value was generated.

The next limitation is the violation of mathematical principles used to calculate MDC's when scales are ordinal. There has been criticism of the MDC approach because ordinal data summarized with mean and standard deviation statistics is not ideal from a pure mathematical standpoint (Anselmi, Vidotto, Bettinardi, & Bertolotti, 2015; Caronni, Picardi, Gilardone, & Corbo, 2021; Kahler, Rogausch, Brunner, & Himmel, 2008; S. S. Stevens, 1946). Since many rehabilitation standardized outcome measures that are patient-reported or clinician-observed often rely on ordinal scales many would argue that MDCs should not be used for these tools.

The Rasch model is able to generate measure-level error estimates that can be used to generate MDCs that are conditional to an individual's pre- and post-score (Kozlowski et al., 2016). The conditional minimal detectable change (cMDC) approach directly addresses all of the key limitations described for the MDC approach. First, cMDC thresholds are based on measure values from the Rasch informed interval scaling. This aspect of the cMDC approach addresse criticisms related to using ordinal data. Second, the cMDC approach accounts for error associated with the pre- and post-score. This accounts for differences in error that are present at different points on the scale. Accounting for error differences could prevent instances where people may have been misclassified as either having or not having changed. As a result, cMDCs provide the means for using individual-level data to understand error and identify change. Clinicians

and researchers alike can use cMDCs and individual-level data for a personalized approach to evaluating change.

2.3 MEASUREMENT OF WALKING-SPECIFIC MOTOR CONTROL AND MECHANICS

Walking is a complex activity requiring successful execution of biomechanical tasks that are controlled by well-coordinated muscle activity (Gottlieb, 1998). Normal walking is a cyclical movement pattern where one gait cycle is defined as the period when the reference limb moves from heel strike to the next heel strike. A normal gait cycle consists of three observable phases: (1) double limb support, (2) swing phase, and (3) single-leg stance phase. These phases can be further divided into the following 6 phases or bins of the gait cycle: (1) initial contact/loading response (initial double support), (2) first half of single-leg stance, (3) second half of single-leg stance (4) second double support, (5) first half of swing, (6) second half of swing (D. J. Clark, Ting, Zajac, Neptune, & Kautz, 2010). During each phase of the gait cycle, there are kinematic and kinetic changes to reflect the accomplishment of key biomechanical tasks; body support, forward propulsion and leg swing (Zajac, Neptune, & Kautz, 2003). Appropriate kinematic and kinetic changes and successful biomechanical task completion are all the result of appropriately activated and timed muscle activity.

Measuring Walking-Specific Motor Control

There have been several approaches used to measure motor control during walking. One prominent approach is to quantify the coordination of spatial and temporal components of measurable gait mechanics (Krasovsky & Levin, 2010; Plotnik, Giladi, & Hausdorff, 2008). However, these approaches do have some limitations including how to determine an acceptable level of variability (e.g., cyclograms) (Field-Fote & Tepavac, 2002; Krasovsky & Levin, 2010), distinguishing between spatial and temporal consistency (e.g., Discrete Relative Phase Index) (Krasovsky & Levin, 2010), and differentiating symmetry and variability when accounting for interlimb similarities (e.g., Phase Coordination Index) (Plotnik et al., 2008).

Another approach to measuring motor control is to use surface EMG. Surface EMG signals provide a non-invasive means of observing gross interactions between the nervous system and motor output because EMG is reflective of the common neural drive a muscle receives. Since walking requires the contributions of many muscles, it is difficult to quantify comprehensive measures of the multi-channel EMG and studies examining motor control during walking were largely qualitative (Knutsson & Richards, 1979). Recently, the use of statistical factorization methods have provided a way to reduce EMG data from multiple muscles into a small set of modules, also referred to as modes or synergies, for quantifying walking-specific motor control (D. J. Clark et al., 2010; Ivanenko, Cappellini, Dominici, Poppele, & Lacquaniti, 2005; Ivanenko, Poppele, & Lacquaniti, 2004). Factorization provides a composition of muscle activity (relative muscle patterns) and temporal representation (pattern activation and

timing) for each module. Thus, modules are a representation of coordinated muscle activity during walking that have a link with underlying neural activity (Bryant A. Seamon, Richard R. Neptune, & Steven A. Kautz, 2018). What is currently missing from module-based analysis when measuring coordination is their association with biomechanical task success. Simulation and observational studies have begun to establish these associations, alleviating some of this concern (Allen, Kautz, & Neptune, 2013; Bowden, Clark, & Kautz, 2010; D. J. Clark et al., 2010; Bryant A. Seamon et al., 2018)

Modules

Modules have been used to measure coordination by quantifying the complexity of coordination (low, medium, high/normal) (D. J. Clark et al., 2010) and individual muscle pattern weightings and activation curves have been used to evaluate the quality of coordination when accomplishing biomechanical tasks of walking (Routson, Clark, Bowden, Kautz, & Neptune, 2013). Initially, there were concerns that statistical factorization methods (e.x. principal component analysis, non-negative matrix factorization) may produce findings that are not physiologically relevant because they are simply mathematical formulas. However, modules have been relatively consistent across studies that examine healthy walking (Cappellini, Ivanenko, Poppele, & Lacquaniti, 2006; D. J. Clark et al., 2010; d'Avella, 2016; Ivanenko et al., 2005; Ivanenko et al., 2004; Tresch, Cheung, & d'Avella, 2006). Findings from these studies have supported that non-negative matrix factorization and a common set of eight muscles can produce 4 modules that can account for over 90% of the variation in observed EMG data.

The composition of muscle activity and activation of each module during the gait cycle theoretically aligns with biomechanical tasks required for walking and computer simulations have supported this association (D. J. Clark et al., 2010; McGowan, Neptune, Clark, & Kautz, 2010; Neptune, Clark, & Kautz, 2009).

The Early Stance module (Module 1) is primarily made up of muscle activity from knee extensors (rectus femoris and vastus medialis) and gluteus medius. This module contributes to the biomechanical functions of body support and backward propulsion (braking). The Late Stance module (Module 2) is primarily gastrocnemius and soleus activity and contributes to body support and forward propulsion (acceleration). The Double Burst module (Module 3) contains two peaks of muscle activity primarily in rectus femoris and tibialis anterior. One peak contributes to foot clearance during early swing and the second to control of the foot at heel strike. The Swing to Stance module (Module 4), primarily composed of medial and lateral hamstring activity, contributes to leg declaration at the end of swing phase before accelerating the leg in early stance to accomplish forward progression. Figure 1 below shows the relative muscle contributions to each module and their relative activation profile over the gait cycle.

Figure 1: Modules During Healthy Walking

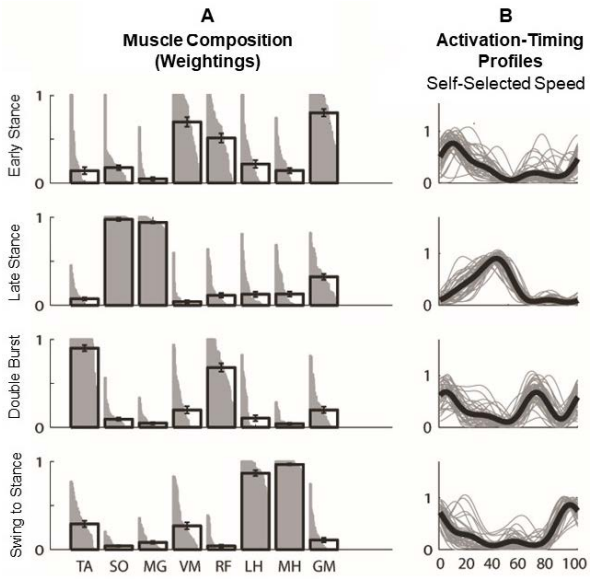


Figure 1: Muscle compositions (A) and activation-timing profiles (B) for the 4 modules seen in healthy walking at self-selected speeds (n=20). **A.** Group averages for muscle weightings are presented with black boxes and standard error bars. If a muscle was fully represented in a module across all individuals then the box would be completely gray. **B.** Individual (thin gray) and group (black) curves are pictured for each module. The activation timing profiles show the magnitude of the module's activation over the gait cycle (averaged to 100 points). Modified from Clark et. al., 2010 (D. J. Clark et al., 2010).

Muscle Abbreviations: TA, tibialis anterior; SO, soleus; MG, medial gastrocnemius; VM, vastus medialis; RF, rectus femoris; LH and MH, lateral and medial hamstring; GM, gluteus medius.

Modules have also been useful for understanding changes in walking-specific motor control for persons post-stroke (B. A. Seamon, R. R. Neptune, & S. A. Kautz, 2018). Persons post-stroke typically walk with fewer modules on the paretic leg. These modules appear to be merged versions of the healthy modules described above and correlate with poorer walking performance (Allen et al., 2013; Barroso et al., 2017; Bowden et al., 2010; D. J. Clark et al., 2010; Coscia et al., 2015; Gizzi, Nielsen, Felici, Ivanenko, & Farina, 2011; Hashiguchi et al., 2016; Kautz, Bowden, Clark, & Neptune, 2011; Routson et al., 2013; Routson, Kautz, & Neptune, 2014). This finding supports the conventional theory of motor recovery post-stroke that proposes stroke severity is associated with how well an individual can move out of mass flexion and extension muscle coordination patterns often seen with merged or fewer modules. As individuals recover, they progressively gain the ability to move away from these mass activation patterns and gain more independent control of their movements (D. J. Clark et al., 2010). Under this framework, more modules equate to more independent control of muscle coordination patterns (Ferrante et al., 2016; Hashiguchi et al., 2016). This link between modules and theory is important because clinicians and researchers have used this theory of motor recovery to hypothesize that an individual's motor control level is the cause of altered walking mechanics or biomechanical deficits commonly seeing in post-stroke walking (Knutsson & Richards, 1979; Ting et al., 2015). Modules appear to provide a way to quantify coordination that is representative of the underlying neuropathology caused from a stroke and may

serve as a biomarker to identify interindividual differences from a mechanistic viewpoint.

The benefit of a mechanistic biomarker in stroke rehabilitation for walking-specific motor control recovery

A biomarker for walking-specific motor control recovery would have several explicit benefits for stroke rehabilitation. Stroke rehabilitation has been plagued by a reliance on measures that lack a singular mechanistic link with underlying pathology making it difficult for clinicians and researchers to identify interindividual differences and impairments for treatment targets (Cramer et al., 2017). This is particularly true for walking where persons can improve performance on activity measures (e.x., gait speed) by addressing a variety of impairments such as strength, balance, and walking-specific motor control (Bowden, Embry, & Gregory, 2011). A biomarker for walking-specific motor control recovery could alleviate these problems and provide clinicians with a framework for providing personalized rehabilitation that targets patient specific impairments (L. Awad, Reisman, & Binder-Macleod, 2019; L. N. Awad, Lewek, Kesar, Franz, & Bowden, 2020; Bowden et al., 2012; Bowden, Hannold, Nair, Fuller, & Behrman, 2008).

The primary challenge for using modules as a biomarker in stroke rehabilitation is the measure's limited potential for clinical use. Module analyses require a sophisticated lab-based gait-analysis, time, and expertise to conduct. All these factors make it highly unlikely that modules would regularly be

measured in the clinic. One alternative is to further explore the association between modules and biomechanics to identify variables that could serve as a surrogate measure.

Biomechanical variables have been explored as potential biomarkers in stroke rehabilitation for walking because of their ability to differentiate individuals post-stroke (L. N. Awad et al., 2020; Louis N. Awad, Reisman, Pohlig, & Binder-Macleod, 2016; Balasubramanian, Bowden, Neptune, & Kautz, 2007; Bowden, Balasubramanian, Neptune, & Kautz, 2006; Mulroy, Gronley, Weiss, Newsam, & Perry, 2003) and their relative potential for clinical translation given technological advancements in wearable sensors and movement analysis. Modules provide a unique way to approach biomechanical variable selection and testing because of the potential to link biomechanical variables to underlying neural pathology. Simulation studies support that there is causality between modules and biomechanical functions in health adults (Neptune et al., 2009). But this still needs to be explored in *in vivo* gait studies for both healthy individuals and stroke survivors to determine the suitability of potential surrogate measures. More research is needed to understand which biomechanical variables specifically link with each module observed in healthy persons to inform baseline biomarker expression. Furthermore, testing is needed to understand how these biomarkers present in person's post stroke relative to their individual nervous system damage and walking ability.

2.4 PRESCRIPTION AND TREATMENT EFFECTS OF TRANSCRANIAL DIRECT CURRENT STIMULATION

Transcranial Direct Current Stimulation and Stroke Rehabilitation

Transcranial direct current stimulation (tDCS) is a type of non-invasive brain stimulation that is a relatively easy to administer, low cost, and safe way to modulate cortical activity (Schlaug, Renga, & Nair, 2008; Woods et al., 2016). The potential for modulating the amount of excitability in the cortex has prompted tDCS to be investigated as a potential adjunct therapeutic in stroke rehabilitation where interhemispheric imbalances are commonly present (Hummel & Cohen, 2005; Nowak, Grefkes, Ameli, & Fink, 2009). Several studies have already found positive effects of tDCS on stroke rehabilitation including areas of aphasia (Fridriksson et al., 2019; Fridriksson et al., 2018), upper (Boggio et al., 2007; Fregni et al., 2005; Hummel & Cohen, 2005; Kim et al., 2010) and lower extremity (Nowak et al., 2009; Reis & Fritsch, 2011) function, and walking (Madhavan & Stinear, 2010; Madhavan, Weber, & Stinear, 2011). However, these promising findings primarily consist of small effects and bring into question the realistic therapeutic utility of tDCS, especially with respect to walking (Lefaucheur et al., 2017; Li, Fan, Yang, He, & Li, 2018; Vaz et al., 2019).

Addressing tDCS Limitations in Stroke Rehabilitation

The high degree of heterogeneity and interindividual differences that is typical of stroke pathology has become a leading proposed mechanism for the limited effects of tDCS (Vaz et al., 2019). Understanding what the interindividual

differences are and their impact on treatment effects of tDCS will be important for moving this body of work forward. One way to approach this need is by placing a greater emphasis on study designs that test the effect of varying tDCS parameters for informing prescription (Fridriksson et al., 2019; Lefaucheur et al., 2017; Lefebvre & Liew, 2017; Vaz et al., 2019). There are several parameters that can be investigated. These include parameters related to treatment frequency, duration, and length, and parameters related to tDCS deliver including current strength and electrode montage.

The Case for Examining Electrode Montage

A strong case can be made for examining the effects of different electrode montages with respect to prescription because various pad placements can induce different modulatory effects. In general, an increase in cortical excitability can be achieved by placing an anode pad over the target brain region (i.e., excitatory stimulus) and a decrease in cortical excitability (i.e., inhibitory stimulus) can be achieved with a cathode pad placed over the target region. The ability to target specific brain locations with excitatory or inhibitory stimuli indicates the potential for addressing interindividual variations within stroke because prescription of electrode montage can be informed by neuropathology. For example, pad placement could be targeted to excite the lesioned hemisphere, inhibit the contralesional hemisphere, or both. Using neuropathology to inform electrode montage has shown positive effects in studies examining tDCS effects

on upper extremity and hand function (Boggio et al., 2007; Fregni et al., 2005; Hummel & Cohen, 2005; Kim et al., 2010).

Current Evidence on Electrode Montage Prescription and Walking

Walking is a unique activity because it is a bilateral, cyclical movement that engages both cortical hemispheres. Unilateral stroke lesions can be disruptive to coordination between hemispheres and an underlying impairment for poor walking performance (Li et al., 2018). Prescribing electrode montage to correct this imbalance could, in theory, improve walking performance. Studies examining the effects of tDCS prescribed to treat interhemispheric imbalances have shown improvements in bilateral upper extremity tasks, but this has not been explored for walking. It is known that tDCS has shown potential for augmenting rehabilitation treatment for walking ability in persons post-stroke (Lefaucheur et al., 2017; Li et al., 2018; Madhavan & Stinear, 2010; Madhavan et al., 2011; Vaz et al., 2019). Studies specifically examining the effects of electrode montage have only tested a dual electrode (Ojardias et al., 2019; Tahtis, Kaski, & Seemungal, 2014; van Asseldonk & Boonstra, 2016). These studies found acute improvements in walking function (Timed Up and Go (Tahtis et al., 2014) and 6-minute walk distance (Ojardias et al., 2019)) and paretic power (van Asseldonk & Boonstra, 2016). However, to date there are no studies that compare the differences between an excitatory (anode over target M1 – usually ipsilesional), inhibitory (cathode over target M1 – usually contralesional), and dual montage (usually anode over ipsilesional M1 and cathode over contralesional M1). Studies

are needed that test the immediate effects of each montage compared to sham stimulation to inform the best prescription for improving walking performance. Additionally, there are also no studies to date that examine the acute effects of tDCS with respect to montage on a comprehensive assessment of walking performance including functional ability, biomechanics, and walking-specific motor control. Studies evaluating treatment prescription need to take comprehensive assessments to understand the effects of tDCS on contributors of walking performance such as biomechanical or neuromotor impairments. This approach is necessary for understanding interindividual differences that underpin variations in treatment response which is critical for personalization of prescription in stroke rehabilitation.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

This chapter is divided into three parts. Each part presents methodology specific to one of the three research areas.

3.2 RESEARCH AREA 1: Item Response Theory and Clinical Measurement

Aim 1: Evaluate item-level psychometrics for a patient-reported (the Activities-specific Balance Confidence Scale) and clinician-observed (the Functional Gait Assessment) outcome measure commonly used in stroke rehabilitation.

Aim 2: Demonstrate how item-level psychometrics can be used to inform personalized treatment planning using an ability map with the Functional Gait Assessment.

Aim 3: Demonstrate how item-level psychometrics can be used to derive change scores using individual-level patient data with the Activities-specific Balance Confidence Scale.

3.2.1 Study Design

We used a cross-sectional research design to perform tests of unidimensionality and a Rasch analysis for the Activities-specific Balance Confidence (ABC) scale and the Functional Gait Assessment (FGA) to answer Aim 1. Item-level psychometrics generated from the FGA were used to construct an ability map for informing personalized treatment planning to address Aim 2. Item-level psychometrics from the ABC scale were used to generate change scores based on individual-level data to address Aim 3.

All methodology for Part I was considered secondary research. All data was free of individual identifying information and did not require Institutional Review Board approval under the revised Federal Policy for the Protection of Human Subjects (Revised Common Rule)(HHS.gov, 2017)

3.2.2 Dataset Descriptions

Data for analyses with the ABC scale were obtained from the NIH's National Institute of Neurologic Disease and Stroke's Archived Clinical Research Dataset. The dataset included ABC scale response data for 406 individuals approximately 2-months post-stroke who were participants in the Locomotor-Experience Applied Post-Stroke (LEAPS) trial (P. Duncan, 2012; Pamela W. Duncan et al., 2011). Individuals who participated in the LEAPS trial were over the age of 18, could walk at least 10-feet with maximum assistance of one person, walk slower than 0.8 m/s, and were living in the community at the time of

enrollment (P. W. Duncan et al., 2007). The LEAPS trial was a phase 3, multisite, randomized control trial and the full inclusion and exclusion criteria can be found in Duncan et. al., 2007 (P. W. Duncan et al., 2007). All response data for the ABC scale were from the LEAPS baseline assessment and there were no missing data points.

Data for the analyses with the FGA were provided by a shared database maintained by the NIH Center for Biomedical Research Excellence (COBRE) in Stroke Recovery at the Medical University of South Carolina. The database includes data for research participants who participated in studies held at the COBRE in Stroke Recovery and consented to having their information archived for future research studies. The shared database is approved by the Institutional Review Board at the Medical University of South Carolina. The database provided a dataset with FGA response data for 101 individuals post-stroke. The date for FGA data collection was used to link demographic variables and calculate participant's ages and time since stroke. The FGA data collection date was also used to link to lower-extremity Fugl-Meyer scores and overground self-selected walking speed. Trained physical therapy staff conducted and scored the FGA for all individuals. Individuals completed the FGA without an assistive device or orthotics. However, an aircast was allowed when individuals had severe paretic ankle instability to prevent injury.

3.2.3 Rasch Analysis and Tests of Unidimensionality

Rasch analysis for both scales was completed in Winsteps [version 3.93.1; John Lincare/Winsteps.com, Beaverton, OR, USA]. Tests for unidimensionality and local dependence were completed with Mplus [version 7.4; Muthén & Muthén, Los Angeles, CA, USA] for the ABC scale and in R [version 3.6.1**;** R Foundation for Statistical Computing; Vienna, Austria] (Team, 2019) with the following packages: *'lavaan'* (Rosseel, 2012), *'psyc'* (Revelle, 2020), and *'polycor' (Fox, 2019)* for the FGA.

Rating-Scale Structure

Rating-scale structure was evaluated with Linacre's 3 category rating-scale criteria: (1) a minimum of 10 observations per rating-scale category, (2) rating-scale category average measures advance monotonically (i.e., demonstrate increasing item difficulty with increasing category value) and (3) outfit mean-squares are less than 2 (J.M. Linacre, 2003).

Category probability curves were visually inspected to confirm distinct peaks for each category of the rating-scale. Distinct peaks for category probability curves indicate each rating-scale category is the most probable response for a specific portion of the measurement scale (Bond & Fox, 2015).

If the rating-scale structure did not meet the designated criteria or demonstrate distinct peaks for category probability curves, then we applied modifications to the rating-scale to best fit the established criteria for the Rasch model.

Unidimensionality

We tested the ABC scale and FGA for unidimensionality after making modifications to the rating-scale structure if they were required. A one-factor confirmatory factor analysis was used to evaluate unidimensionality. Recommendations from Reeve et. al., 2007 were used to evaluate the model fit: (1) comparative fit index >0.95, (2) root mean square error approximation <0.06, (3) Tucker-Lewis Index >0.95, and (4) standardized root mean residuals <0.08 (Reeve et al., 2007). If the one-factor model did not meet the recommendations, we performed an exploratory factor analysis to test for additional factors. We compared model fit from the exploratory factor analysis to the same recommendations used in the confirmatory factor analysis. In addition, we weighed accepting additional factors using: (1) the eigenvalue ratio between the first and second factors (greater than 4 indicates sufficient unidimensionality), (2) scree plot visualization, and clinical interpretation.

Local Dependence

The Rasch model assumes that items have local independence. Local independence of items means that there are no significant associations among item responses when controlling for the measure's dominant factor.(Reeve et al., 2007) We evaluated for local independence of items using a residual correlation matrix from the one-factor confirmatory factor analysis. If residual correlations were >0.2 or <-0.2 between two items we considered them to be locally dependent. (Reeve et al., 2007)

Item and Person Fit

The fit of items and persons to the Rasch model were evaluated with fit statistics. Items or individuals were considered misfitting when outfit statistics had mean-square standardized residuals greater than or equal to 1.4 and standardized z-scores greater than or equal to 2 (J.M. Linacre, 2003; Wright & Linacre, 1994).

Item Difficulty Hierarchy

We evaluated the FGA's theoretical construct validity with the item difficulty hierarchy generated by the Rasch model. We also used the item difficulty hierarchy to evaluate the extent that items overlapped. Items were considered to overlap if the measure estimates for any two items were within 2 standard errors.

Person-Item Match

We evaluated for floor and ceiling effects. We considered the scale to have a floor and/or ceiling effect if more than 15% of individuals scored the worst or best possible measure value (Lim et al., 2015).

Separation Index

The person separation index was used to quantify how well each scale differentiated people into statistically distinct strata. The following formula was

used to calculate the number of strata from the person separation index (Wright & Masters, 2002):

$$\text{Strata} = [4*(\text{person separation index})+1]/3$$

3.2.4 Ability Map Development

We developed an ability map for the FGA from a keyform that is generated by Winsteps during the Rasch analysis to address Aim 2. A keyform is a tool that can be used to score persons on individual items and then quickly generate the person's measure value. The keyform presents the items in difficulty hierarchy on the right vertical axis and the measure values on the horizontal axis. The items are clustered with items of similar difficulty and clusters are separated by a horizontal blank space. Each row for the items contains the available rating-scale categories. The rating-scale categories for each FGA item correspond to the original score system: 0=severe impairment, 1=moderate impairment, 2=mild impairment, and 3 = normal.

A rater can use a keyform by circling the score for each item and then drawing a vertical line through the "bulk" of the circles to x-axis. The line intersects with the x-axis to provide an approximation of the person's measure value on the Rasch model. Measure values are typically generated on the logit scale. The logit scale anchors the mean to 0 and standard deviation to 1. Logit scales can be transformed by selecting a new anchor for the mean and rescaling the measures. We chose to convert the logit scale to a scale from 0-30 to mimic

the traditional FGA scale and enhance the clinical interpretability of measure values.

We presented three patient scenarios to demonstrate how a keyform can be used as an ability map to generate personalized treatment planning from individual patient responses to items on the FGA. We separated our sample into tertials based on their measure values to represent high, moderate, and low ability levels. We randomly selected an individual from each tertial and evaluated their fit statistics to ensure they did not misfit the Rasch model. Each individual's response data was used to complete a keyform score sheet and generate an estimate of their measure value. We discuss how the keyform can be interpreted as an ability map by clinicians for personalized treatment planning.

3.2.5 Change Threshold Determinations

We calculated conditional change thresholds, also known as conditional minimal detectable change (cMDC), using item-level psychometrics generated from our Rasch analysis of the ABC scale to address Aim 3. We also calculated a traditional change threshold, also known as minimal detectable change (MDC), using group-level data for the ABC scale from our dataset. We used theoretical patient examples to demonstrate the differences between cMDCs and an MDC for classifying patients as having a detectable change.

Calculating cMDC Thresholds

Rasch analysis generates an interval scale with measure values in logits. Each logit has an associated standard of error (SE). We converted the logit scale back to a 100-point scale (0-100) by anchoring the mean and scaling measure values with the same approach described for the FGA analysis. We applied this transformation to help improve the clinical interpretability of measure values on the interval scale. The transformation also helps improve the interpretability of comparisons between the cMDC and MDC, which relies on the traditional ABC scale scoring system.

We used the SE generated with each measure value to calculate a cMDC threshold with a 95% confidence interval ($cMDC_{95}$) for every possible pair of pre- and post-score combinations using the following formula (Kozlowski et al., 2016):

$$cMDC_{95} = ([SE_{pre\text{-}score}\text{-}SE_{post\text{-}score}]/2)*1.96*\sqrt{(2)}$$

Calculating MDC Thresholds

We calculated the standard error of measurement (SEM) which represents group-level error data for the ABC scale with original scores (i.e., prior to Rasch analysis) in the following formula (Wyrwich, 2004):

$$\text{Standard Error of Measurement (SEM)} = \text{Standard Deviation} *\sqrt{(1\text{-Cronbach Alpha})}$$

Next, we calculated one MDC threshold with a 95% confidence interval (MDC_{95}) with the following formula (P. W. Stratford, Binkley, Riddle, & Guyatt, 1998):

$$MDC_{95} = 1.96*\sqrt{(2)}*(SEM)$$

Comparing cMDC and MDC Thresholds

We plotted all possible $cMDC_{95}$ thresholds associated with each person measure on the cMDC scale to visualize the effect of individual-level standard error on the threshold for detecting change across the ABC scale. We also plotted the MDC_{95} as a horizontal line across the scale to illustrate differences in the approach for detecting change. Finally, we extracted three of the $cMDC_{95}$ curves to show three theoretical patient scenarios to examine if there are differences in how change is detected between approaches. We selected a low (25), moderate (50), and high (75) initial score for the patient scenarios to capture the differences between a $cMDC_{95}$ and MDC_{95} approach across three representative ability levels.

3.3 RESEARCH AREA 2: Searching for Biomarkers to Inform Patient Subgroups

Aim 1: Quantify the association between walking-specific motor control and quantifiable biomechanical variables during steady-state self-selected walking for healthy individuals.

3.3.1 Study Design

We used a cross-sectional research design with health, community-dwelling individuals. Participants completed a laboratory gait analysis consisting of walking on an instrumented treadmill for three, 30-second bouts of self-selected treadmill walking. Study procedures were approved by the Institutional Review Board at the Medical University of South Carolina and conformed to the Declaration of Helsinki. All participants signed a written informed consent form prior to enrollment.

3.3.2 Participants

We enrolled 20 healthy adults between 40 to 85 years of age. Participants were excluded if they had any neurologic diseases or significant orthopedic impairments in their legs (e.g., pain, amputation, severe osteoarthritis, etc.) that would limit walking ability or alter motor control.

3.3.3 Experimental Procedures

All participants walked for three, 30-second bouts of self-selected walking on an instrumented treadmill (Bertec; Columbus, OH). Kinetic, kinematic, and electromyographic (EMG) data were collected during each trial. We sampled ground reaction forces at 2000 Hz to derive kinetic variables. Ground reaction

force data was filtered with a 4th order Savitzky-Golay filter and resampled at 100 Hz. We collected whole-body 3D kinematic measures using a 16-camera motion capture system (PhaseSpace, Inc.; San Leandro, CA). Motion capture data was sampled at 120 Hz, filtered with a 4^{th} order Savitzky-Golay filter, and resampled at 100 Hz. EMG data were collected from surface bipolar pre-amplified electrodes (Motion Lab Systems; Baton Rouge, LN, USA) and sampled at 2000 Hz. We collected EMG data from eight muscles bilaterally: tibialis anterior, soleus, medial gastrocnemius, vastus medialis, rectus femoris, medial hamstrings, lateral hamstrings, and gluteus medius. EMG data was high-pass filtered at 40 Hz, demeaned, rectified, low-pass filtered at 4 Hz using a zero-lag forth order Butterworth filter, and resampled to 100 Hz. EMG amplitude for each muscle was averaged within each region of the gait cycle (or bin): (1) initial contact/loading response (initial double support), (2) first half of single-leg stance, (3) second half of single-leg stance (4) second double support, (5) first half of swing, (6) second half of swing) for each step. The bin with the highest average was used to normalize EMG amplitude across all trials for an individual. Lastly, data for each step was time normalized to the gait cycle (0-100%).

Data Analysis

BIOMECHANICAL VARIABLES: We derived biomechanical variables from kinetic and kinematic data using custom MATLAB software (Mathworks, Natick, MA) to represent the biomechanical subtasks of walking. Each biomechanical

variable was calculated per bin per step for each leg. Variables of interest included:

(1) Changes in ground reaction force or impulse (the time integral of the ground reaction force) in the anterior-posterior (AP) and vertical directions. We calculated net impulse (area under the curve; AUC) from the ground reaction force components.

(2) Changes in anterior-posterior leg angle (sagittal plane of motion). Leg angle was calculated by taking the angle between a vertical line from the center of mass of the pelvis to the ground and a line from the center of mass of the pelvis to the center of mass of the foot. The change in anterior-posterior leg angle was represented by the net AUC of the leg angle measurement over each bin of the gait cycle.

(3) Changes in leg length. Leg length was calculated as the distance from the center of mass of the pelvis to the center of mass of the foot. Leg length was normalized for each participant to the distance between their pelvis center of mass and foot center of mass during static standing. The change in leg length was represented by the positive AUC of the leg length over each bin of the gait cycle.

EMG VARIABLES: We quantified muscle coordination using a module analysis of EMG data with Non-negative Matrix Factorization (NNMF). EMG from each muscle for per leg was combined into an *m* x *t* matrix (EMG_O), where *m* is the number of muscles (8) and *t* is the time base (*t* = number of gait cycles x 101). We then applied an NNMF algorithm to the *m* x *t* matrix for each participant. NNMF creates two matrices that define a pre-selected number of modules (*n*). We pre-selected the number of modules to be 4 based on previous research in healthy individuals from our lab (D. J. Clark et al., 2010). One matrix from NNMF is an *m* x *n* matrix with the relative weighting of each muscle within each module (*W* matrix). The second is an *n* x *t* matrix indicating the activation timing profile of each module for each step in the trial (*H* matrix). A key assumption of NNMF is that muscle weightings are fixed across all steps and muscles can belong to more than one module. NNMF runs an iterative optimization procedure to minimize the error between EMG_O and a reconstructed EMG signal (EMG_R) created by the *m* x *n* and *n* x *t* matrices. We quantified module activation by taking the AUC under each module's *H* matrix curve within each bin of the gait cycle per step.

Statistical Analysis

We fit linear mixed models with random coefficients (PROC GLIMMIX) to quantify the association between module and biomechanical variables at the group-level while controlling for variability at the individual-level. We identified the participant variable, participant by leg (i.e., right and left) interaction, and

participant by step interaction as random effects. We examined the following comparisons: Module 1 with anterior-posterior and vertical ground reaction forces in bins 1 and 2; Module 2 with anterior-posterior and vertical ground reaction forces, and trailing leg angle during bins 3 and 4; Module 3 with anterior-posterior ground reaction forces during bin 4, and leg angle and leg length during bins 5 and 6; and Module 4 with anterior-posterior and vertical ground reaction forces during bin 1, and leg angle in bin 6. We quantified the magnitude of association with parameter estimates and we used a Bonferroni correction for the number of tests (tests = 18) for an alpha level of 0.003. All analyses were done in SAS version 9.4 (SAS Inc., Cary, NC).

3.4 RESEARCH AREA 3: Evaluating Treatment Prescription

Aim 1: Determine whether electrode montage type acutely influences walking performance.

3.4.1 Study Design

We used a double-blind, randomized, cross-over experimental design. Participants completed a clinical battery of assessments including the lower-extremity Fugl-Meyer, Berg Balance Test, and Dynamic Gait Index at enrollment. Participants then completed three single session tDCS experimental conditions

where electrode montage was varied and one session of sham simulation. We block randomized sessions to control for order effects and separated sessions by a minimum 48-hour period for washout. Primary gait outcome measures were collected pre- and post-stimulation at each experimental session. EMG outcome measures were collected pre-stimulation on the participants first session and post-stimulation for all experimental sessions.

All study procedures were conducted by a team of licensed physical therapists and associated study staff. This study was approved by the Institutional Review Board at the Medical University of South Carolina and conformed to the Declaration of Helsinki. All participants signed written informed consent.

3.4.2 Participants

We enrolled 18 individuals with chronic stroke. One individual dropped out and one did not meet inclusion criteria. The remaining 16 individuals completed all study procedures and were included in the analysis. Inclusion criteria for the study were: 1) age 18 to 85 years old; 2) at least six-months post-stroke; 3) residual lower extremity paresis (Fugl-Meyer Lower Extremity motor score <34); 4) ability to walk independently at least 10 feet; 5) self-selected 10-meter gait speed < 0.8 m/s (at time of consent); and 6) provision of informed consent. Individuals were excluded if they met any of the following criteria: 1) significant musculoskeletal problems limiting hip and knee extension or ankle plantarflexion

to neutral joint positions; 2) self-reported history of unstable cardiovascular disease or severe osteoporosis, or 3) pregnancy.

3.4.3 Experimental Procedures

Participants received tDCS from an EMPI unit (Chattanooga; Hixson, TN). Sponges were cut to a 1.75 cm^2 size and prepped with a 0.9% saline solution to deliver a current density of 0.1 mA/cm^2 (Thair, Holloway, Newport, & Smith, 2017). Stimulation was ramped up to 2 mA and continued for 20 minutes to generate a dose rate of 40 mA/min. Participants were informed that they would feel a slight tingling sensation at the onset of stimulation that would resolve in approximately 60 seconds. Participants completed 3 single session tDCS experimental conditions where electrode montage was manipulated: 1) excitatory, 2) inhibitory, and 3) dual and one session of sham stimulation. In the excitatory condition the anode pad was placed over the target ipsilesional leg M1 area with the reference pad on the ipsilateral shoulder. In the inhibitory condition the cathode pad was placed over the target contralesional leg M1 area with the reference pad on the ipsilateral shoulder. In the dual condition we used two EMPI units to deliver simultaneous anodal stimulation over the ipsilesional M1 and cathodal stimulation over the contralesional M1. The reference pads for each EMPI unit were placed on the respective ipsilateral shoulder. The dual condition allowed for the excitatory and inhibitory montages to be applied simultaneously to target both leg M1 areas. This type of pad placement (active pads with a cortical placement, and reference pads on the ipsilateral shoulder) has been shown to

have a deeper current penetration depth in modeling studies (Noetscher, Yanamadala, Makarov, & Pascual-Leone, 2014). This pad placement also was shown to have a greater effect on excitability (Tatemoto, Yamaguchi, Otaka, Kondo, & Tanaka, 2013) and neuromotor output (Angius, Pageaux, Hopker, Marcora, & Mauger, 2016) in lower extremity areas of M1 compared to other types of pad placement. In addition to the experimental conditions, participants completed one session with sham stimulation. Sham stimulation was achieved by turning on the EMPI unit to apply 30 seconds of stimulation before an unblinded investigator turned the units off (Tanaka et al., 2011). During the sham stimulation, participants would feel the initial tingling sensation that would resolve in about 60 seconds like the experimental conditions and be unable to distinguish if the unit had been turned off or not. Blinding for participants and staff was maintained by using the same set up as described in the dual condition across all sessions. We estimated the location of M1 in each hemisphere by placing the pad 1cm lateral to the vertex and 1cm posterior to a hypothetical line between the tragi of each ear while the participant was in a forward seated position. This created a 2 cm gap between cephalic pad placements.

Each session (3 experimental conditions and sham) began with participants receiving the first 5 minutes of stimulation seated before walking the remaining 15 minutes on a treadmill while stimulation continued. Participants walked at their fastest-comfortable speed on the treadmill to mimic an adequate training stimulus used for rehabilitation (Lamontagne & Fung, 2004; Sullivan, Knowlton, & Dobkin, 2002). Several measures were taken to maintain patient

safety during testing. Participants were attached to a ceiling harness system to prevent falls and walking was paused every 5 minutes to assess vital signs. Walking was immediately discontinued if required for participant safety. Minimal physical assistance or verbal cues were provided to help participants avoid tripping during the 15-minute walking period with stimulation. No physical assistance or verbal cues were provided during pre- and post-stimulation data collections.

Data Analysis

Data collection for outcome measures were completed pre- and immediately post-stimulation for each experimental session. Participants walked over a 24-foot GaitRite (CIR Systems, Inc.; Franklin, NJ) for one trial at their self-selected walking speed and three trials at their fastest-comfortable walking speed during pre- and post-testing to determine overground walking speeds. Next, participants walked for three, 30-second trials at self-selected and again at fastest-comfortable walking speeds on a split belt instrumented treadmill (Bertec; Columbus, OH). Treadmill speeds were selected by participants independent from overground speeds and the two conditions did not have to match. Paretic step ratio and paretic propulsion were calculated from ground reaction force data sampled at 1000 Hz during treadmill walking using methods previously described by our lab (Bowden et al., 2006). We calculated paretic propulsion by dividing the positive anterior impulse of the paretic leg by the anterior impulse of both legs combined (Bowden et al., 2006). We calculated paretic step ratio from the

percentage of stride length performed by the paretic step (Balasubramanian et al., 2007). We expressed paretic step ratio and paretic propulsion as the absolute value of deviation from symmetry.

We collected EMG data during self-selected treadmill walking to examine muscle coordination changes in response to tDCS stimulation. We quantified muscle coordination by extracting modules from EMG data using a non-negative matrix factorization (NNMF) algorithm (Lee & Seung, 1999; Ting & Macpherson, 2005). We recorded surface EMG at 2000 Hz with bipolar pre-amplified electrodes (Motion Lab Systems; Baton Rouge, LN, USA) from eight leg muscles bilaterally: tibialis anterior, soleus, medial gastrocnemius, vastus medialis, rectus femoris, medial hamstrings, lateral hamstrings and gluteus medius (Hermens, Freriks, Disselhorst-Klug, & Rau, 2000). EMG signals were pre-processed before using the NNMF algorithm to select the number of modules for the paretic leg. EMG signal processing and module selection criteria can be found in Clark et. al., 2010 (D.J. Clark, Subramanian, Neptune, & SA, 2008). Pre-stimulation number of modules was determined for each individual at their initial session. Post-stimulation number of modules was determined immediately following tDCS or sham stimulation during self-selected treadmill walking trials.

Statistical Analysis

All statistical analyses were conducted using SAS version 9.4 (SAS Institute Inc. Cary, NC). We ran all analyses using change scores for gait speed, paretic step ratio and paretic propulsion at self-selected and fastest comfortable

speeds. We used a one-way ANOVA (or Kruskal-Wallis Test for non-parametric data) to examine the main effect of active stimulation across all three montages compared to sham for each variable of walking performance. Post-hoc testing with multiple t-tests (or Wilcoxon Rank-Sum tests) were performed to compare the effects of each electrode montage when indicated. We conducted a secondary, exploratory analysis to examine the effect of tDCS to sham accounting for an individual's best response to the three electrode montages. We identified everyone's best response for each variable with respect to electrode montage. We pooled these values to generate a group average response to stimulation and compared it to sham stimulation using t-tests (or Wilcoxon Rank-Sum tests). We accepted a higher alpha-value of 0.1 for this analysis to inform hypothesis generation. Changes in module number are reported observationally to generate hypotheses for future studies. We performed Spearman's correlation between changes in module number and other self-selected walking outcome variables to explore associations between muscle coordination and walking performance.

CHAPTER 4

MANUSCRIPTS

4.1 INTRODUCTION

Chapter 4 is divided into three parts. Each part addresses one of the research questions and contains manuscripts that were written to present the findings.

PART 1

Part 1 contains the following manuscripts:

1. Rasch Analysis of the Activities-Specific Balance Confidence Scale in Individuals Post-Stroke
2. Item-level Psychometrics for the Functional Gait Assessment in Persons with Stroke
3. Revisiting the concept of minimal detectable change for patient-reported outcome measures

PART 2

Part 2 contains the following manuscript:

1. Associations between biomechanical variables of walking performance and muscle coordination during self-selected steady-state walking

PART 3

Part 3 contains the following manuscript:

1. tDCS electrode montages may differentially impact variables of walking performance in individuals post-stroke: a preliminary study

4.2 PART I MANUSCRIPTS

Rasch Analysis of the Activities-Specific Balance Confidence Scale in Individuals Post-Stroke

Functional Gait Assessment Item-level Psychometrics for Measurement of Walking Balance Ability in Persons with Chronic Stroke

Revisiting the concept of minimal detectable change for patient-reported outcome measures

The following manuscript was accepted for publication in the *Archives of Rehabilitation Research and Clinical Translation* in December of 2019. The text below is from the final draft accepted prior to publication. A copy of the published version is open access and can be found at:
https://doi.org/10.1016/j.arrct.2019.100028

Title Page

(1) Running Head: Rasch analysis of the ABC Scale

(2) Title: Rasch Analysis of the Activities-Specific Balance Confidence Scale in Individuals Post-Stroke

(3) Full Name, Highest Degrees and Affiliations of the Authorship team:

Bryant A. Seamon PT, DPT, NCS, CSCS[1,2], Steven A. Kautz, PhD[1,2], Craig A. Velozo, PhD, OTR/L[2,3]

[1] Ralph H. Johnson VA Medical Center, Charleston, SC, USA

[2] Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[3] Division of Occupational Therapy, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

(5) This material has accepted as a poster presentation for the 2019 annual American Congress of Rehabilitation Medicine conference. The full manuscript has not been submitted for publication with any other journal.

(6) Funding Sources, Conflicts of Interest, Acknowledgements:

Partial funding for this project was provided by the VA Office of Research and Development (ORD), with additional support from the VA/ORD Rehabilitation R&D Service (1I01RX001935), support from the National Institutes of Health (NIH P20 GM109040) and the Promotion of Doctoral Studies Level 1 Scholarship from the Foundation for Physical Therapy Research. Data for the study was provided by the NIH National Institute of Neurological Disorders and Stroke from the Locomotor Experience Applied Post-stroke (LEAPS) trial (R01 NS050506). Any opinions expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Veteran Affairs or the NIH.

The authors have no conflicts of interest to report.

Acknowledgements

We would like to acknowledge the LEAPS investigator team [Principal Investigator: Pamela Duncan, PT, PhD, FAPTA, FAHA] for data collection and archival. We also want to acknowledge the National Institute of Neurological Disorders and Stroke for funding the LEAPS trial (R01 NS050506).

Authors Note

Data from the LEAPS trial can be obtained by contacting the National Institute of Neurological Disorders and Stroke at www.ninds@nih.org

(7) The name and address for the corresponding author:

Bryant A. Seamon, PT, DPT, NCS, CSCS

77 President Street, MSC 700

Charleston, SC 29425

843-310-3608

seamon@musc.edu

(8) Clinical Trial Number:

NCT00243919

Abstract

Objective: To examine the psychometric properties of the Activities-specific Balance Confidence (ABC) scale using Rasch analysis for individuals post-stroke.

Design: Retrospective cohort.

Setting: Data was extracted from the Locomotor Experience Applied Post-Stroke (LEAPS) phase three, multisite, randomized controlled clinical trial.

Participants: 406 community-dwelling, ambulatory, older-adults (mean age 61.97 years SD12.76; 45.07% female) approximately 2-months post-stroke.

Intervention: None.

Main Outcome Measures: We examined unidimensionality, local dependence, rating-scale structure, item and person fit, person-item match, and separation index of the ABC scale.

Results: Confirmatory and exploratory factor analysis showed the ABC scale was adequately unidimensional and three item pairs had local dependence. A collapsed 5-category rating-scale was superior to the 101-category scale. The hardest item was “walking outside on an icy sidewalk”, the easiest item was

“getting into or out of a car”, and no items misfit. The ABC scale had high person reliability (0.93), despite 10.5% of individuals misfitting the expected response pattern. Mean ability level of the sample was slightly lower (-0.56 logits) than the mean item difficulty indicating that the ABC scale adequately matched our sample’s balance confidence. The ABC scale did not have a floor or ceiling effect and separated individuals into 5 statistically distinct strata (separation index=3.71).

Conclusions: The Rasch model supports the use of the ABC scale to measure balance confidence in individuals’ post-stroke. The consistency of our results with previous Rasch analyses on the ABC scale demonstrates the instrument responds similarly across multiple populations; community-dwelling older-adults, outpatient orthopaedic physical therapy, stroke, Parkinson’s disease, and lower-limb amputation. Recommendations include collapsing the rating scale and developing a computerized-adaptive test version of the scale to enhance clinical utility.

Key Words: Stroke, Psychometrics, Outcome Assessment, Postural Balance

Abbreviations: ABC scale; Activities-specific Balance Confidence scale, ANPT; Academy of Neurologic Physical Therapy, CFA; confirmatory factor analysis, EFA; exploratory factor analysis

INTRODUCTION

Individuals post-stroke are at high risk for devastating consequences from falls including increased health care utilization and fracture rate[1, 2] with approximately 3 of 4 individuals falling during the first 6 months back at home and up to one quarter experiencing recurrent falls.[3, 4] Since falls are associated with fear of falling and balance confidence[5], measurement of an individual's confidence in their balance is an important component of clinical practice for physical therapists in stroke rehabilitation. The Activities-specific Balance Confidence (ABC) scale was designed to measure balance confidence and takes approximately 20 minutes to complete.[6] Individuals rate their confidence that they "will not lose their balance or become unsteady" when performing each daily task (item) on the scale from 0% (low confidence) to 100% (high confidence). A total score is calculated by averaging scores from all 16 items.

The ABC scale is widely used in stroke rehabilitation[7] and has psychometric evidence to support its use for quantifying balance confidence in stroke survivors.[8-13] Total scores on the ABC scale have concurrent validity with measures used to assess activity and participation domains of the International Classification of Functioning, Disability and Health including the Berg Balance Scale[9-11], walking speed[9, 10, 12, 13], Timed-up and Go[10, 12, 13], six-minute walk test[10, 12], Barthel Index[10], Lower Extremity Fugl-Meyer Assessment[11], five-time sit-to-stand test[11], modified Rivermead Mobility Index[13] and physical function scale of the SF-36[10, 12]. The ABC scale also has strong internal consistency[13], strong test-retest reliability for the scale's total score (intraclass correlation coefficient

[ICC]=0.85[9], 0.82[12]) and moderate to strong test-retest reliability at the item level (ICC range 0.53-0.93)[9]. The majority of individuals post-stroke score between 20-80% suggesting there is not a floor or ceiling effect and standard error of measurement has been reported between 6.81[9] and 5.05[10]. Cut-off values for distinguishing between individuals with a history of multiple falls and no falls after suffering a stroke has been reported as 81.1[8] and 63.75[14], where lower confidence is associated with more falls.

However, no studies have examined measurement characteristics of the ABC scale for stoke survivors using item response theory psychometric methods, like Rasch analysis. Rasch analysis takes advantage of probabilistic mathematical modeling to examine a measurement tool's ability to quantify abstract constructs in a meaningful way. This is accomplished by assuming the probability of successfully passing an item is dependent on the relationship between a person's ability and item difficulty. Results from a Rasch model orders a measure so scores can be interpreted linearly with set interval distances.[15] Often, Rasch analysis identifies items that overlap in measurement properties and can be used to develop short forms or computerized-adaptive tests to reduce the time required to administer an instrument.[16]

Previous studies examined the ABC scale with Rasch analysis in different populations; Arnadottir et. al. (community-dwelling older adults)[17], Sakakibara et. al. (lower-limb amputees)[18], Franchignoni et. al. (Parkinson's disease)[19], and Wang et. al. (outpatient physical therapy)[20]. These studies found similar psychometrics for the ABC scale indicating that Rasch methods may support the

comparison and use of the scale across patient populations, which is recommended for physical therapists in neurologic physical therapy practice by the Academy of Neurologic Physical Therapy (ANPT).[21]

Therefore, the purpose of this study is to examine the ABC scale using Rasch analysis for individuals post-stroke. We hypothesize the ABC scale will fit the Rasch model for these individuals similarly to other populations. Results from this analysis will provide support for comparison and use of the ABC scale across populations as recommended by the ANPT.

METHODS

Data Source

This study is a secondary analysis of data from 406 individuals post-stroke who participated in the Locomotor Experience Applied Post-Stroke (LEAPS) phase three, multisite, randomized controlled clinical trial.[22] Institutional Review Board approval of this secondary analysis was not required because data was free of identifiers. Included individuals had a stroke and (1) were greater than 18 years old, (2) were able to ambulate a minimum of 10 feet with maximum 1-person assist, (3) had a self-selected walking speed less than 0.8 m/s, and (4) were living in the community.[23] Individuals were excluded who had (1) additional neurologic pathology and co-morbidities, (2) severe pain, amputation or orthopaedic conditions limiting ambulation, or (3) severe cardiovascular comorbidities that would prevent participation in high intensity exercise.[23] Demographic data for the trial was collected during an enrollment window (within

30 days of diagnosis).[23] We analyzed ABC scale data collected at approximately 2 months post-stroke (Baseline Assessment). Summary demographic data was analyzed with SAS version 9.4[a] and presented in Table 1.

(Insert Table 1)

Rasch Analysis

Rasch analysis of the ABC scale was done with Winsteps version 3.93.1[b]. Tests of unidimensionality and local dependence were performed in Mplus version 7.4[c].

Rating-Scale Structure

Appropriateness of the rating scale structure was determined based on Linacre's 3 rating-scale criteria[24]; (1) each rating-scale category has a minimum of 10 observations, (2) average measures within each category advance monotonically (i.e., demonstrate increasing observed item difficulty with increasing category value) and (3) outfit mean-squares are less than 2. Category probability curves were examined to see if categories of the rating-scale had distinct peaks (indicating each category of the rating-scale is the most probable response for a given portion of the measure).[15] The rating-scale was collapsed for further analysis if it did not meet designated criteria.

Unidimensionality

An assumption of the Rasch model is that the measure is unidimensional. A confirmatory factor analysis (CFA) with one factor was performed on a random sample from our data (n=203) to assess unidimensionality. Model fit from the CFA was evaluated against recommendations from Reeve et. al., 2007:[25] (1) comparative fit index >0.95, (2) root mean square error approximation <0.06, (3) Tucker-Lewis Index >0.95, and (4) standardized root mean residuals <0.08. If model fit was poor, an exploratory factor analysis (EFA) was done on another random sample (n=203) to determine additional factors. Additional factors from EFA were evaluated on model fit (Reeve's recommendations[25]), eigenvalue ratio (greater than 4 indicates sufficient unidimensionality), visualization of the scree plot, and clinical interpretation.

Local Dependence

Local independence assumes no significant associations among items responses when controlling for the dominant factor of the measure.[25] A residual correlation matrix from the CFA was used to identify dependent item pairs. Residual item correlations >0.2 or <-0.2 were considered locally dependent.[25]

Item and Person Fit

Items or individuals were classified as misfitting if fit statistics had mean-square standardized residuals greater than or equal to 1.4 and standardized z-scores greater than or equal to 2.[24, 26]

Item Difficulty Hierarchy

Item difficulty was used to evaluate theoretical construct validity of the ABC scale. The Rasch model assigns item difficulty and person ability measures to a logit scale. Items that are easier or persons with lower ability are assigned lower values and items with higher difficulty or persons with high ability are assigned higher values. Item measure estimates were used to determine if items overlapped. Items were considered overlapping if the item's measure estimate was within 2 standard errors of another item.

Person-Item Match

Observation of the person-item map was used to evaluate for floor (within error of worst possible outcome [raw score 0/100]) and ceiling effects (within error of best possible outcome [raw score 100/100]). We considered the ABC scale to have a floor or ceiling effect if greater than 15% of individuals scored the worst or best possible outcomes.[27]

Separation Index

The person separation index was used to evaluate the ABC scale's ability to differentiate people into statistically distinct strata. The number of strata were determined from the following formula[28]:

$$\text{Strata} = [4*(\text{person separation index})+1]/3$$

RESULTS

Rating-Scale Structure

We initially analyzed the data using a 101-category (0-100%) rating-scale. Only a few rating-scale categories had more than 10 observations. No outfit mean-squares values exceeded 2.0, however, there were disordered rating-scale estimates. Collectively, this demonstrates that rating-scale categories were under-used, and the ABC scale was not adequately fitting the Rasch model. Therefore, we tested a collapsed 5-category rating-scale based on previous publications (0-9%, 10-30%, 31-60%, 61-90%, 91-100%).[18, 20] The new rating-scale had more than 10 observations per category, demonstrated appropriate rating-scale estimates of increasing ability level (rating-scale categories advanced monotonically), and no category exceeded the outfit mean square threshold of 2. The rating-scale structure results are presented in Table 2.

(Insert Table 2)

Unidimensionality

CFA using the collapsed rating-scale returned the following fit statistics; (1) comparative fit index 0.95 (>0.95 indicates good fit), (2) root mean square error approximation 0.15 (<0.06 indicates good fit), (3) Tucker-Lewis Index 0.95 (>0.95 indicates good fit), and (4) standardized root mean residuals 0.09 (<0.08 indicates good fit).

EFA returned two factors with eigenvalues >1.0 and the following fit statistics for a two factor model; (1) comparative fit index 0.96 (>0.95 indicates

good fit), (2) root mean square error approximation 0.14 (<0.06 indicates good fit), (3) Tucker-Lewis Index 0.95 (>0.95 indicates good fit), and (4) standardized root mean residuals 0.05 (<0.08 indicates good fit). Eigenvalues for the first two factors were 10.17 and 1.41, respectfully. The second factor included the following items; “stand on your tiptoes and reach for something above your head”, “stand on a chair and reach for something”, “sweep the floor”, “walk in a crowded mall where people rapidly walk past you”, “are bumped into by people as you walk through the mall”, “step onto or off an escalator while you are holding onto a railing”, “step onto or off an escalator while holding onto parcels such that you cannot hold onto the railing”, “walk outside on icy sidewalks”. The ratio of eigenvalues for the first and second factor is 7.21 and visual interpretation of the scree plot favors accepting only one factor. Although all criteria were not met, the results of the factor analysis support that the ABC scale adequately meets the assumption of unidimensionality.

Local Dependence

Three item pairs were found to have local dependence; (1) “walk outside on icy sidewalks”–“walk outside the house to a car parked in the driveway” (r=-0.23), (2) “walk outside the house to a car parked in the driveway”–“step onto or off an escalator while holding onto parcels such that you cannot hold onto the railing” (r=-0.25), and (3) “walk outside on icy sidewalks”–“reach for a small can off a shelf at eye level” (r=-0.21).

Item and Person Fit

No items misfit the Rasch model. The findings related to item fit are reported in Table 3. Forty-three individuals (10.6%) responses did not fit with the Rasch model (mean-square standardized residuals greater than or equal to 1.4 and standardized z-scores greater than or equal to 2 of fit statistics.[24, 26]). We found nearly identical results from the Rasch analysis when misfitting persons were removed. Therefore, we are reporting findings for the whole sample because (1) individuals included in the sample are largely representative of community-dwelling stroke survivors and (2) the ABC scale is designed and advocated to be broadly applicable for this patient population. The ABC had high person reliability (0.93) and Cronbach's alpha (0.95).

Item Difficulty and Person-Item Match

The results of the item difficulty analysis are presented in Table 3 and visually displayed using a Person-Item map in Figure 1. The hardest item was "walking outside on icy sidewalks" while the easiest item was "getting into or out of a car".

(Insert Table 3)

The Person-Item map in Figure 1 shows the distribution of (1) people based on ability (left: low ability, bottom; high ability, top) and (2) item difficulty (right: easy, bottom; hard, top). The range of the distribution was 11 logits with

the mean ability of our sample (-0.58 SE 0.38) below the mean difficulty of the items (anchored at 0) indicating the model adequately matched our participants' confidence. We did not have a ceiling effect (no individuals had a maximum score) and observed negligible floor effects (4 individuals had a minimum score [0.9%]). In addition to the overlapping items in Figure 1, we found occurrences where item measures were within 2 standard errors of another item indicating several items have overlapping difficulty. Item measures and standard errors are presented in Table 3.

(Insert Figure 1)

Separation Index

The ABC scale differentiated individuals in our sample into 5.28 statistically distinct strata (separation index= 3.71).

DISCUSSION

This is the first study to examine ABC scale psychometrics using Rasch analysis for individuals post-stroke. We found the Rasch model strongly supports the use of the ABC scale to measure balance confidence in these individuals. Like previous studies[18, 20], we found the scale fit the Rasch model better using a collapsed 5-category rating-scale. Collapsing the rating-scale corrected its disorder and prevented items misfit. We found the ABC scale was adequately unidimensional to meet the necessary assumption for Rasch analysis. Although

EFA suggested a second factor, we do not feel that it is warranted provided (1) the improvement in model fit is minimal, (2) the eigenvalue ratio for the first and second factors greatly exceeds the recommended value of 4, and (3) visualization of the scree plot confirms one dominant factor. Also, there is no clinical rationale to support items grouped in the second factor except for the fact they are the more difficult items on the scale. Therefore, there is not enough evidence that unidimensionality is violated, which is consistent with previous publications.[17-20] Our analysis of local dependence identified three item pairs with residual correlations greater than 0.2 magnitude. One recommendation is to remove items with high dependence when performing Rasch analysis because item dependence can be a threat to unidimensionality.[25] However, there is a discrepancy as to what magnitude of association constitutes removing items.[19, 25] Therefore, we reported results with all items under the caveat that effects of local dependence should be closely evaluated when translating the ABC scale into a computerized-adaptive test because item pairs with local dependence may need to be identified as “enemies”.[25]

We found other similarities and only minor differences in the ABC scale’s psychometrics for individuals post-stroke compared to other populations. We report a separation index (3.71) for the ABC scale indicating that the scale separated our sample into 5.28 statistically distinct strata based on balance confidence, similar to previous publications (5.2[20], 4[17]). Other congruent psychometrics include high Cronbach’s alphas (0.94[18], 0.95[19], 0.93[20], 0.95 in our study) and no floor or ceiling effects.[17, 18, 20] The range of the scale in our model

was 11 logits (-6 to 5), which was comparable to other publications.[17-20] The three items; “walk outside on an icy sidewalk”, “standing on a chair and reaching for something”, and “step onto or off an escalator while holding onto parcels such that you cannot hold onto the railing” are consistently (with the exception of order) the three most difficult items.[17-20]. However, there is more variability in item difficulty for easy items. We found the item “getting into or out of a car” to be the easiest item, which is comparable with two publications where this item was the second easiest[18, 20]. Yet, in other publications this item was considered moderately difficult and fell close to the center of the scale.[17, 19]

A component of validating measurement scales derived from Rasch analysis is to determine whether the item hierarchy is consistent with clinical and theoretical expectations.[19] The item hierarchy in Table 3 and Figure 1 show item difficulty progress from discrete stable tasks (i.e, reaching and transferring) to stable walking to walking tasks in conditions of increasing instability. Thus, item hierarchy is consistent with clinical and theoretical expectations for individuals post-stroke and more broadly, individuals with balance impairments.

Therefore, we can conclude that the ABC scale responds similarly for individuals post-stroke and other populations including; community-dwelling older-adults, outpatient orthopaedic physical therapy participants, individuals with lower-limb amputation and individuals with Parkinson’s disease. As a result, clinicians or researchers interested in measuring balance confidence for these clinical populations do not need to develop diagnosis specific versions of the instrument and can compare scores between patient groups. The ABC scale’s

ability to respond similarly across a variety of patient populations supports recent recommendations by the ANPT for the scale to be included in a core set of outcome measures in the rehabilitation of adults with neurologic diagnoses[21] and allows one to hypothesize that the instrument may be “diagnosis free”.

IMPLICATIONS FOR FUTURE RESEARCH AND PRACTICE

We present two recommendations for future research and practice to facilitate the clinical adoption of the ABC scale. First, we recommend implementing a 5-category rating scale (“no confidence” [0], “low confidence” [1], “moderate confidence” [2], “high confidence” [3], “complete confidence” [4])[18]. Second, we recommend reducing the number of items by creating short-forms and computerized-adaptive tests of the ABC scale based on the Rasch model. Although, three short-forms exist for the ABC scale,[29-31] they were not developed from a Rasch model and should be approached with caution as they may have diminished measurement characteristics relative to the full scale.[19, 32]
In general, these recommendations should facilitate clinical adoption of the ABC scale by reducing test administration time, a commonly cited barrier to outcome measurement use by practitioners.[33, 34]

STUDY LIMITATIONS

There are some limitations with this research. One limitation is that the authors were not in control of the data-collection procedure, which is typical of archival data secondary analyses. Selection bias associated with selection of

acute care facilities for urgent stroke care may exist potentially limiting our findings' generalizability. Generalizability of our findings may also be limited by the inclusion criteria required for individuals to participate in the LEAPS trial. Specifically, participants in the trial were community dwelling and able to ambulate indicating that our findings may not extend to more functionally limited individuals.

CONCLUSIONS

Consistent with calls to use the ABC scale across neurologic diseases in adult populations,[21] Rasch analysis supports the use of the ABC scale for measuring balance confidence in individuals post-stroke. The ABC scale's psychometrics are largely enhanced with a 5-category rating scale. We recommend using the ABC scale to quantify balance confidence in these individuals based on absent floor and ceiling effects and the scale's ability to distinguish 5 strata of individuals. Collapsing the ABC's rating scale and developing a computerized-adaptive test will enhance measurement capability and efficiency for clinicians and researchers working in stroke rehabilitation.

Table 1: Participant Demographics

n=406		
Demographic Characteristics		**Totals**
ABC Score **(2 months post-stroke)**		45.06% (23.88%)
Age		61.97 (12.76)
Sex		
	Male	54.93%
	Female	45.07%
Race		
	American Indian	1.23%
	Asian	13.3%
	Black or African American	22.17%
	White	57.64%
	Native Hawaiian	4.68%
	More than 1 race	0.74%
	Unknown	0.25%
	Hispanic or Latino	15.52%
Stroke Type		
	Ischemic	80.05%
	Hemorrhagic	18.72%
	Uncertain	1.23%
Stroke Location		
	Right Hemisphere	48.03%
	Left Hemisphere	35.22%
	Brainstem	62%
	Bilateral Hemispheres	6%

Continuous variables are presented in mean (standard deviation)
Categorical variables are presented as a percentage

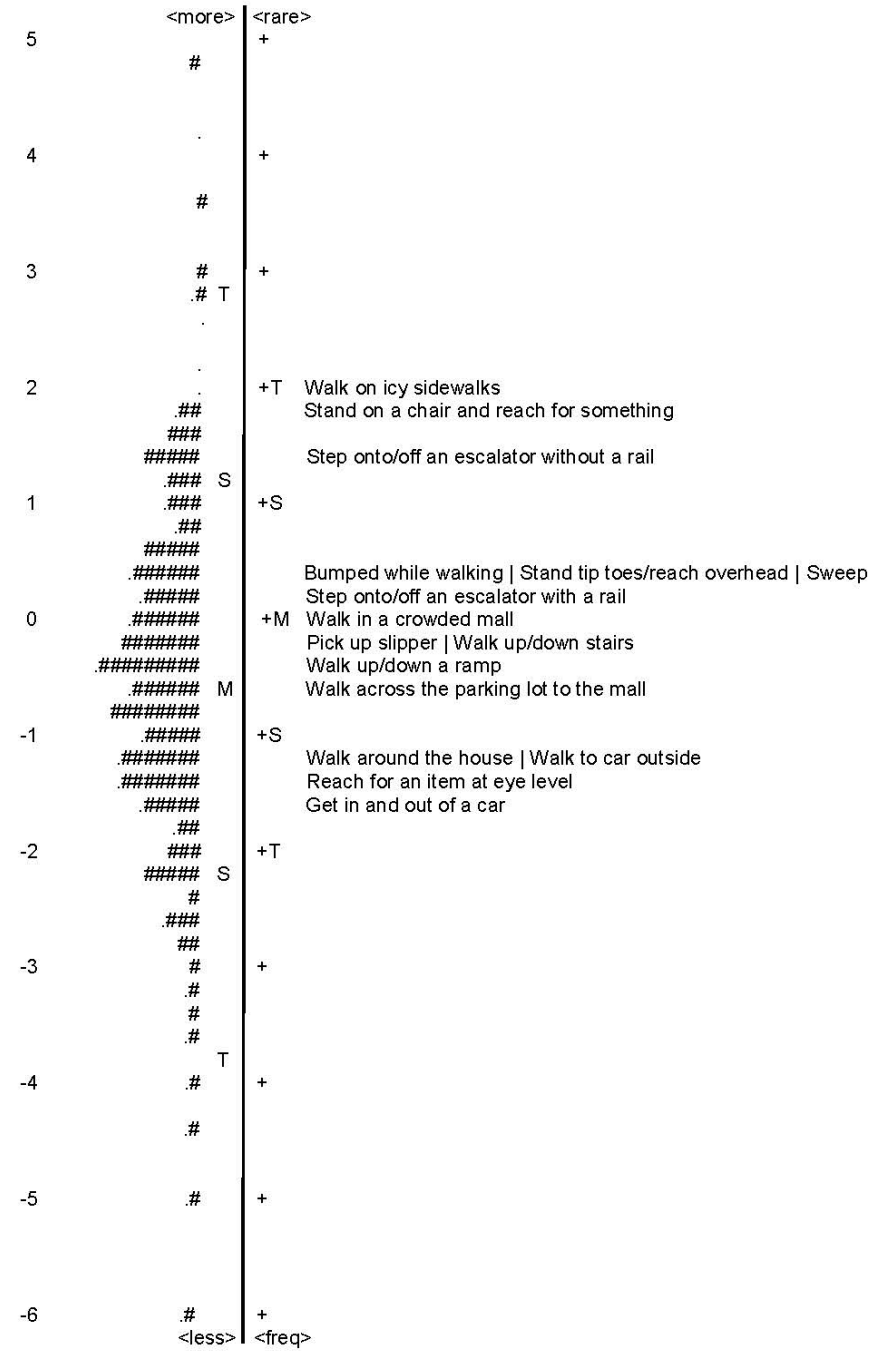
Table 2. Rating-Scale Structure

Score	Observed Average	Infit Mean-square	Outfit Mean-square	Frequency Counts (%)
0 ***(no confidence)***	-2.72	1.06	1.07	1521 (23%)
1 ***(low confidence)***	-1.32	0.90	0.84	1352 (21%)
2 ***(moderate confidence)***	-0.20	0.85	0.78	1594 (25%)
3 ***(high confidence)***	0.95	0.91	0.92	1400 (22%)
4 ***(complete confidence)***	2.25	1.40	1.33	629 (10%)

Table 3. Item Measure Order

Item	Item Number	Measure	Model Standard Error	Infit Mean-square	Infit z-score	Outfit Mean-square	Outfit z-score	Point Measure Correlation
“walk outside on icy sidewalks”	16	1.98	0.08	0.96	-0.4	0.94	-0.4	0.67
“stand on chair and reach for something”	6	1.79	0.08	1.29	3.4	1.10	0.8	0.65
“step onto or off an escalator while holding onto parcels such that you cannot hold onto the railing”	15	1.45	0.07	1.00	0.0	0.93	-0.6	0.71
“are bumped into by people as you walk through the mall”	13	0.49	0.07	0.73	-4.2	0.69	-4.1	0.80
“stand on your tip toes and reach for something above your head”	5	0.43	0.07	1.04	0.6	1.02	0.3	0.74
“sweep the floor”	7	0.36	0.07	1.35	4.5	1.31	3.4	0.71
“step onto or off an escalator while you are holding onto a railing”	14	0.11	0.07	1.05	0.7	0.98	-0.2	0.76
“walk in a crowded mall where people rapidly walk past you”	12	-0.04	0.07	0.65	-5.7	0.62	-5.6	0.82
“walk up or down stairs”	2	-0.11	0.07	1.03	0.5	1.02	0.2	0.75
“bend over and pick up a slipper from the front of a closet floor”	3	-0.27	0.07	1.01	0.1	0.97	-0.4	0.76
“walk up or down a ramp”	11	-0.44	0.07	0.81	-2.9	0.76	-3.4	0.80
“walk across a parking lot to the mall”	10	-0.51	0.07	0.88	-1.8	0.82	-2.4	0.80
“walk outside the house to a car parked in the driveway”	8	-1.11	0.07	0.88	-1.7	0.84	-2.1	0.80
“walk around the house”	1	-1.22	0.07	1.03	0.4	1.02	0.3	0.73
“reach for a small can off a shelf at eye level”	4	-1.35	0.07	1.23	3.0	1.26	3.0	0.73
“get into or out of a car”	9	-1.56	0.07	1.21	2.8	1.14	1.7	0.71

Figure 1. Person-Item Map



M = mean
S = 1 standard of deviation from the mean
T = tertile
# = 3 individuals
. = 1 individual

Item abbreviations: Bumped while walking (Are bumped into as you walk in a mall), Stand tip toes/reach overhead (Stand on tip toes and reach overhead), Pick up a slipper (Pick up a slipper from the floor), Walk to car outside (Walk outside to a car in the driveway)

SUPPLIERS

a. Winsteps version 3.93; John Lincare, Beaverton, OR: Winsteps.com
b. SAS version 9.4; SAS Institute Inc. 100 SAS Campus Drive Cary, NC 27513
c. Mplus version 7.4; Muthén & Muthén 3463 Stoner Avenue Los Angeles, CA 90066

REFERENCES

1. Pouwels S, Lalmohamed A, Leufkens B, de Boer A, Cooper C, van Staa T et al. Risk of hip/femur fracture after stroke: a population-based case-control study. Stroke 2009;40(10):3281-5.

2. Tilson JK, Wu SS, Cen SY, Feng Q, Rose DR, Behrman AL et al. Characterizing and identifying risk for falls in the LEAPS study: a randomized clinical trial of interventions to improve walking poststroke. Stroke 2012;43(2):446-52.

3. Forster A, Young J. Incidence and consequences of falls due to stroke: a systematic inquiry. BMJ (Clinical research ed) 1995;311(6997):83-6.

4. Walsh ME, Sorensen J, Galvin R, Williams DJP, Harbison JA, Murphy S et al. First year post-stroke healthcare costs and fall-status among those discharged to the community. European Stroke Journal 2018;3(3):254-62.

5. Friedman SM, Munoz B, West SK, Rubin GS, Fried LP. Falls and fear of falling: which comes first? A longitudinal prediction model suggests strategies for primary and secondary prevention. Journal of the American Geriatrics Society 2002;50(8):1329-35.

6. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) Scale. The journals of gerontology Series A, Biological sciences and medical sciences 1995;50a(1):M28-34.

7. Sullivan JE, Crowner BE, Kluding PM, Nichols D, Rose DK, Yoshida R et al. Outcome Measures for Individuals With Stroke: Process and Recommendations From the American Physical Therapy Association Neurology Section Task Force. Physical therapy 2013;93(10):1383-96.

8. Beninato M, Portney LG, Sullivan PE. Using the International Classification of Functioning, Disability and Health as a framework to examine the association between falls and clinical assessment tools in people with stroke. Physical therapy 2009;89(8):816-25.

9. Botner EM, Miller WC, Eng JJ. Measurement properties of the Activities-specific Balance Confidence Scale among individuals with stroke. Disability and rehabilitation 2005;27(4):156-63.

10. Salbach NM, Mayo NE, Robichaud-Ekstrand S, Hanley JA, Richards CL, Wood-Dauphinee S. Balance self-efficacy and its relevance to physical function and perceived health status after stroke. Archives of physical medicine and rehabilitation 2006;87(3):364-70.

11. An S, Lee Y, Lee D, Cho KH, Lee G, Park DS. Discriminative and predictive validity of the short-form activities-specific balance confidence scale for predicting fall of stroke survivors. J Phys Ther Sci 2017;29(4):716-21.

12. Forsberg A, Nilsagård Y. Validity and Reliability of the Swedish Version of the Activities-specific Balance Confidence Scale in People with Chronic Stroke. Physiotherapy Canada Physiotherapie Canada 2013;65(2):141-7.

13. Ylva N, Anette F. Psychometric properties of the Activities-Specific Balance Confidence Scale in persons 0-14 days and 3 months post stroke. Disability and rehabilitation 2012;34(14):1186-91.

14. Park EY, Lee YJ, Choi YI. The sensitivity and specificity of the Falls Efficacy Scale and the Activities-specific Balance Confidence Scale for hemiplegic stroke patients. J Phys Ther Sci 2018;30(6):741-3.

15. Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. 3rd ed. New York and London: Routledge; 2015.

16. Porter I, Goncalves-Bradley D, Ricci-Cabello I, Gibbons C, Gangannagaripalli J, Fitzpatrick R et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. Journal of comparative effectiveness research 2016;5(5):507-19.

17. Arnadottir SA, Lundin-Olsson L, Gunnarsdottir ED, Fisher AG. Application of Rasch analysis to examine psychometric aspects of the activities-specific balance confidence scale when used in a new cultural context. Archives of physical medicine and rehabilitation 2010;91(1):156-63.

18. Sakakibara BM, Miller WC, Backman CL. Rasch analyses of the Activities-specific Balance Confidence Scale with individuals 50 years and older with lower-

limb amputations. Archives of physical medicine and rehabilitation 2011;92(8):1257-63.

19. Franchignoni F, Giordano A, Ronconi G, Rabini A, Ferriero G. Rasch validation of the Activities-specific Balance Confidence Scale and its short versions in patients with Parkinson's disease. Journal of rehabilitation medicine 2014;46(6):532-9.

20. Wang YC, Sindhu B, Lehman L, Li X, Yen SC, Kapellusch J. Rasch Analysis of the Activities-Specific Balance Confidence Scale in Older Adults Seeking Outpatient Rehabilitation Services. The Journal of orthopaedic and sports physical therapy 2018;48(7):574-83.

21. Moore JL, Potter K, Blankshain K, Kaplan SL, O'Dwyer LC, Sullivan JE. A Core Set of Outcome Measures for Adults With Neurologic Conditions Undergoing Rehabilitation: A CLINICAL PRACTICE GUIDELINE. Journal of Neurologic Physical Therapy 2018;42(3):174-220.

22. Duncan PW, Sullivan KJ, Behrman AL, Azen SP, Wu SS, Nadeau SE et al. Body-Weight–Supported Treadmill Rehabilitation after Stroke. New England Journal of Medicine 2011;364(21):2026-36.

23. Duncan PW, Sullivan KJ, Behrman AL, Azen SP, Wu SS, Nadeau SE et al. Protocol for the Locomotor Experience Applied Post-stroke (LEAPS) trial: a randomized controlled trial. BMC neurology 2007;7:39.

24. Linacre JM. Rasch power analysis: size vs. significance: standardized chi-square fit statistic. Rasch Measurement Transactions 2003;17:918.

25. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JA et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Medical care 2007;45(5 Suppl 1):S22-31.

26. Wright BD, Linacre JM. Reasonable mean-square fit values. Rasch Measurement Transactions 1994;8(3):370.

27. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. BMJ Open 2015;5(7):e007765.

28. Wright BD, Masters GN. Number of Person or Item Strata: (4*Separation + 1)/3. Rasch Measurement Transactions 2002;16(3):888.

29. Lohnes CA, Earhart GM. External validation of abbreviated versions of the activities-specific balance confidence scale in Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 2010;25(4):485-9.

30. Oude Nijhuis LB, Arends S, Borm GF, Visser JE, Bloem BR. Balance confidence in Parkinson's disease. Movement disorders : official journal of the Movement Disorder Society 2007;22(16):2450-1.

31. Peretz C, Herman T, Hausdorff JM, Giladi N. Assessing fear of falling: Can a short version of the Activities-specific Balance Confidence scale be useful? Movement disorders : official journal of the Movement Disorder Society 2006;21(12):2101-5.

32. Wolfe EW, Smith EV, Jr. Instrument development tools and activities for measure validation using Rasch models: part II--validation activities. Journal of applied measurement 2007;8(2):204-34.

33. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. Physical therapy 2009;89(2):125-35.

34. Stevens JG, Beurskens AJ. Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. Physical therapy 2010;90(6):953-61.

Title Page

Item-level Psychometrics for the Functional Gait Assessment in Persons with Stroke

Bryant A. Seamon PT, DPT[1,2], Steven A. Kautz, PhD[1,2], Mark G. Bowden, PT, PhD[1,2,3], Jesse Dean, PhD[1,2,3], Chris Gregory, PT, PhD[1,2,3], Richard R. Neptune, PhD[4], Craig A. Velozo, PhD, OTR/L[2,5]

[1] Ralph H. Johnson VA Medical Center, Charleston, SC, USA

[2] Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[3] Division of Physical Therapy, Department of Health Professions, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[4] Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, USA

[5] Division of Occupational Therapy, Department of Health Professions, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

The name and address for the corresponding author:

Bryant A. Seamon, PT, DPT

77 President Street, MSC 700

Charleston, SC 29425

321-394-5351

seamon@musc.edu

Abstract

Objective: Item-level psychometrics of the Functional Gait Assessment (FGA), a measure of walking balance ability, have not been determined for persons with chronic stroke and made available for clinical use. The objective of this study was to evaluate the FGA's unidimensionality, report item-level psychometrics and the item hierarchy, and present an FGA ability map for clinical use to identify a person's measure and inform clinical decision-making.

Methods: We used retrospective response data from an NIH-funded center's shared research database containing 101 ambulatory persons with chronic stroke. Factor analysis was used to evaluate unidimensionality and item local dependence. Rasch analysis was used to examine rating-scale structure, item and person fit, item hierarchy and separation index of the FGA.

Results: Confirmatory and exploratory factor analyses confirmed the FGA's unidimensionality and showed that none of the items had local dependence. The category rating-scale advanced monotonically and met published criterion. The item hierarchy was like that for community-dwelling older adults with mean ability level of the sample slightly above the mean item difficulty (0.28 logits, 0.63 standard error). The FGA had high patient reliability (0.90) despite 3 items and 9.9% of the sample misfitting. The FGA did not have a floor or ceiling effect and was able to separate people into 4 strata.

Conclusions: Factor and Rasch analyses support the use of the FGA for measuring walking balance ability in ambulatory persons with chronic stroke. An FGA ability map can provide an instantaneous interval measure score for patients while informing personalized treatment design and goal setting. Results from this paper support clinical practice recommendations to use the FGA in outpatient stroke rehabilitation and should address barriers to clinical implementation.

Impact Statement: Item-level psychometrics and an FGA ability map provide clinicians with an understanding of the item hierarchy and a way to inform personalized treatment planning and goal setting. This should enhance clinical utility by improving patient specific interpretation and facilitate adoption of clinical practice guidelines in stroke rehabilitation.

Keywords: stroke, gait, balance, measurement – applied

Introduction

Measurement of walking ability is a necessary component of stroke rehabilitation. The Functional Gait Assessment (FGA) is a standardized clinical scale that was developed to measure a person's ability to maintain balance while walking.[1] The FGA consists of 10 items that measure an individuals' performance across a span of daily walking tasks that vary in difficulty. Items are scored using a 4-category rating-scale by a clinician observing an individual perform the tasks. An overall FGA score is calculated by summing the items' scores.

Conventional psychometric properties of the FGA have been reported for persons with chronic stroke.[2, 3] The FGA has excellent test/retest (ICC=0.95)[3], interrater (ICC=0.94)[2] and intrarater (ICC=0.97)[2] reliability and has strong correlations (r>0.7) with measures of mobility (i.e., Barthel Index, Rivermead Mobility Index), walking (ie. gait speed, Functional Ambulatory Category), and balance ability (i.e., Berg Balance Scale).[2] Similar psychometrics have also been reported for other patient populations including community-dwelling older adults[4], persons with vestibular disorders[1], and Parkinson's disease[5, 6]. This similarity across a clinical diagnoses commonly seen in outpatient neurologic rehabilitation resulted in the American Physical Therapy and Academy of Neurologic Physical Therapy's joint Clinical Practice Guideline recommendation for the FGA to be used as part of a core set of outcome measures.[7] Yet, there are still several barriers cited by clinicians for not using standardized measures. Several barriers reference clinical utility of measures including time (administration time is lengthy, requires more time than information is worth), lack of clinical relevancy

(information is too subjective, items are not relevant to the patient), and limited interpretability (difficult to interpret, does not contribute to the plan of care).[8, 9]

However, rather than disparage physical therapists for not using standardized measures, efforts should be focused on improving the usefulness of standardized measures for practicing clinicians, especially since individual test scores appear to have little value to the clinical reasoning process.[10, 11] Rasch analysis is one way to approach this problem since it generates item-level psychometrics and item difficulty hierarchies that can inform clinicians about an individual patient's ability level and how they will perform across items of various challenges on the assessment. The Rasch model uses a probabilistic relationship between a person's ability level and the item difficulty, with persons having a high probability of successfully completing items that are easier than their ability level and low probability of success on items that are harder than their ability.[12] A persons' ability level is determined when person ability matches item difficulty (when the patient has 50 percent probability of passing an item at a particular rating). This offers several benefits for clinicians including improved measurement efficiency and score interpretability by identifying where a patient's ability falls along a continuum.

Clinicians can take advantage of this improved clinical utility by using a keyform. Keyforms were first introduced by Linacre[13] as a Rasch informed score sheet for generating "instantaneous" measurement values on an interval scale from a person's performance or response to items. In addition, keyforms visually depicts a patient's response pattern in relation to the item difficulty hierarchy.

This allows a keyform to be used as an ”ability map” for physical therapists to identify ideal tasks for treatment targets in both short- and long-term care plans.[10, 14-17] Physical therapists will be able to quickly glance at a completed ability map to see what tasks the patient has mastery over, what tasks the patient is making progress towards mastery, and what tasks the patient demonstrates a poor ability to accomplish. Over time as a patient progresses through rehabilitation, the ability map should show that the patient is gaining mastery over more difficult items on the hierarchy.

A previous Rasch analysis of the FGA reported item-level psychometrics for community-dwelling older adults.[18] They showed the FGA was unidimensional, had an ordered rating-scale structure, and had a clinically valid item difficulty hierarchy. However, the primary findings were intended to demonstrate the FGA removed previous DGI ceiling effects. Since there is professional organizational support for the use of the FGA in adult neurological populations there is need to examine the item-level psychometrics in common neuropathology, such as stroke. This is needed because the performance-based rating-scale used in the FGA could result in a different item hierarchy for individuals with stroke due to lasting gait and balance impairments that are not commonly found community dwelling older-adults. Additionally, development of an FGA ability map may support the scale’s ability to inform clinical decision making with respect to treatment design. Thus, the purpose of this study was to examine the item-level psychometrics and item difficulty hierarchy of the FGA and to create an ability map for clinical use. We provide an illustration for how

ability maps can be used for informing clinical decision making with regards to treatment planning and lend interpretability to FGA scores.

Methods

Data Source

We used a database that included research participants who participated in studies at the NIH Center for Biomedical Research Excellence in Stroke Recovery at the Medical University of South Carolina. The shared collective database contains demographic information and research records for individuals who consent to having their information archived for future use and is approved by the Medical University of South Carolina's Institutional Review Board.

Data for 101 individuals post-stroke was available for analysis. The date of FGA data collection was used to calculate age and time since stroke (months), and to link lower extremity Fugl-Meyer scores and overgound self-selected walking speed. Participants completed the FGA without an assistive device or orthotics. An aircast was permitted for severe ankle instability on the paretic leg to prevent injury. Trained physical therapists and research staff oversaw and scored all participants. Demographic data was analyzed with SAS [version 9.4; SAS Institute, Cary, NC, USA].

Rasch Analysis

Rasch analysis of the FGA was completed using Winsteps [version 3.93.1; John Lincare/Winsteps.com, Beaverton, OR, USA]. Tests of unidimensionality

and local dependence were performed in R [version 3.6.1**;** R Foundation for Statistical Computing; Vienna, Austria][19] with the following packages: *'lavaan'*[20], *'psyc'*[21], and *'polycor'*[22].

Rating-Scale Structure

The rating scale structure was evaluated against Linacre's three essential rating-scale criteria[23]; (1) a minimum of 10 observations per rating-scale category, (2) rating-scale category average measures advance monotonically (i.e., demonstrate increasing item difficulty with increasing category value) and (3) outfit mean-squares are less than 2. If the rating-scale did not meet designated criteria we would apply modifications to best fit the criteria before continuing the analysis.

Unidimensionality

The Rasch model assumes that the measure is unidimensional. We performed a one-factor confirmatory factor analysis to test unidimensionality. Model fits was assessed against recommendations from Reeve et. al., 2007:[24] (1) comparative fit index >0.95, (2) root mean square error approximation <0.06, (3) Tucker-Lewis Index >0.95, and (4) standardized root mean residuals <0.08. If a one-factor model did not meet all the recommendations, we performed an exploratory factor analysis to determine if additional factors could be identified. We evaluated results from the exploratory factor analysis using Reeve's recommendations[24] for model fit, eigenvalue ratio between the first and second

factors (greater than 4 indicates sufficient unidimensionality), and clinical interpretation.

Local Dependence

Local independence of items assumes that there are no significant associations among item responses when controlling for the measure's dominant factor.[24] We used a residual correlation matrix from a one-factor confirmatory factor analysis to test for local independence and identify dependent item pairs. If residual item correlations were >0.2 or <-0.2 we considered those items to be locally dependent.[24]

Item and Person Fit

We evaluated the fit of items and persons to the Rasch model. Items or individuals were labeled as misfitting when fit statistics for outfit had mean-square standardized residuals greater than or equal to 1.4 and standardized z-scores greater than or equal to 2.[23, 25]

Item Difficulty Hierarchy

We used the item difficulty hierarchy evaluate the FGA's theoretical construct validity and to test the extend that item's overlapped. We considered items to be overlapping if measure estimates for any two items were within 2 standard errors.

Person-Item Match

The person-item map was used to evaluate for floor and ceiling effects. We required that greater than 15% of individuals scored the worst or best possible outcomes to consider the FGA to have a floor and/or ceiling effect.[26]

Separation Index

We used the person separation index to quantify how well the FGA can differentiate people into statistically distinct strata. The following formula was used to calculate the number of strata[27]:

$$\text{Strata} = [4*(\text{person separation index})+1]/3$$

Ability Map Development

A keyform score sheet was generated using the results from the Rasch analysis and is presented in Figure 1. The keyform shows the item hierarchy on the right vertical axis. Items progress from hardest to easiest with the hardest item at the top of the list. Items that are clustered together to show that these items have overlapping or similar difficulty. The clusters are separated by a blank line. The x-axis shows the interval-level measurement scale. The scale has been converted from logits to the original scoring to enhance clinical interpretation. The row for each item contains the possible category-rating scores: 0=severe impairment, 1=moderate impairment, 2=mild impairment, and 3 = normal. Clinicians can print the keyform score sheet and circle the rating for each item.

The clinicians can then draw a vertical line through the bulk of the circles to identify the patient's score on an interval scale.[13] We separated our sample into tertials to represent high, moderate, and low ability. We used a randomly selected individual from each tertial to demonstrate how the ability maps can be used in the clinic. Each representative individual's fit statistics were evaluated to ensure that they fit the Rasch model.

(INSERT FIGURE 1)

Role of the Funding Source

The funding sources for this research played no role in the design, conduct or reporting of this study and findings.

Results

The average age for our cohort was 58.6 years old (SD 12.6). We had 44 females (43.6%) and 42 individuals with left hemiparesis (58.4%). Mean lower extremity Fugl-Meyer score was 25.2 (SD 0.29) and mean overground walking speed was 0.76 m/s (SD 0.29). The average raw score for the FGA was 15.7 (SD 6.1). Descriptive statistics for our sample are presented in Table 1.

(INSERT TABLE 1 HERE)

Rating-Scale Structure

The FGA's current 4 category rating-scale fit the Rasch model well and satisfied each of Linacre's criteria[23]. Each category on the rating-scale had more than 10 observations with outfit mean square values less than 2.0 and average category measures advanced monotonically. Rating-scale structure results are presented in Table 2.

(INSERT TABLE 2)

Unidimensionality

A one-factor confirmatory factor analysis had the following fit statistics; (1) comparative fit index 0.98 (>0.95 indicates good fit), (2) root mean square error approximation 0.11 (<0.06 indicates good fit), (3) Tucker-Lewis Index 0.98 (>0.95 indicates good fit), and (4) standardized root mean residuals 0.07 (<0.08 indicates good fit). Since our sample is below 250 participants, the root mean square error approximation may not be an appropriate criterion for assessing model fit since it is sensitive to sample size.[28] To verify our findings, we ran an exploratory factor analysis to assess for other potential factors. The exploratory factor analysis recommended one factor and returned only one eigenvalue above 1. The eigenvalue ratio between the first and second factors was 8.9 (eigenvalues: factor 1 = 6.85, factor 2 = 0.77) with values greater than 4 supporting unidimensionality. Overall, the results from the exploratory factor analysis did not support exploring a second factor and we concluded that the FGA meets the unidimensionality assumption for Rasch analysis.

Local Dependence

No item pairs were found to have local dependence.

Item and Person Fit

The FGA has high person reliability (0.90) and a Chronbach's alpha value of 0.92). Three items; 1) gait with narrow base of support, 2) gait with eyes closed, and 3) gait and pivot turn and ten individuals (9.9%) responses met the criteria for misfitting the Rasch model (outfit mean-square standardized residuals greater than or equal to 1.4 and standardized z-scores greater than or equal to 2.[23, 25]). We removed missing fitting persons in serial order (most misfitting to least misfitting) reanalyzing the data to test the effect of the unexpected individual responses on overall fit of the items. Removing misfitting persons did not improve the fit of the 3 misfitting items against our criteria. Item fit statistics are presented in Table 3.

Item Difficulty and Person-Item Match

The item difficulty analysis is presented alongside item fit statistics in Table 3. The FGA's hardest item was "gait with narrow base of support" while the easiest item was "gait and pivot turn". The full hierarchy is presented in Figure 2 (Person-Item map).

(INSERT TABLE 3)

The Person-Item map shows the distribution of (1) people based on ability (left: low ability, bottom; high ability, top) against item difficulty (right: easy, bottom; hard, top) for each item rating-scale score (0, 1, 2, and 3). The range of ability levels on the FGA was 20 logits. The FGA did not have a ceiling or floor effect because no individuals had a minimum score and only 2 (2%) had a maximum score. Our sample's mean ability level was 0.28 logits (0.63 standard error) which is slightly above the mean difficulty level of the items (anchored at 0 logits). This indicates that the Rasch model adequately matched the walking balance ability of our participants.

Table 3 presents item measures and standard errors. Several items overlap in their coverage of the FGA's scale. This can be seen on the keyform score sheet (Figure 1) by relatively comparing the coverage of an item (i.e., range of categories on the rating scale) along the measurement scale to other items. This effect is also visible on the Person-Item Map (Figure 2) when determining the expanse of coverage for an item. For example, the item Gait Level Surface covers the expanse of -4 to 5 logits depending on the rating score.

(INSERT FIGURE 2)

Separation Index

The FGA differentiated our sample into 4.31 distinct strata with a separation index of 3.01.

Ability Maps

Figure 3 shows the completed FGA ability maps for our three representative individuals with low (3a), moderate (3b), and high (3c) walking balance ability. Actual scores for each item are circled. The solid vertical line traveling through the "bulk" of the items represents the patients score on an interval scale and the dashed vertical lines show two standard errors around this score.[13] Visualizing the response pattern to the item difficulty hierarchy provides an additional understanding of the persons score relative to their ability. In Figure 3a., which represents a person of low ability (score of 10; SE 1.17), this person shows mild impairment (i.e., 2) with the easiest item followed by moderate impairment (i.e., 1) on the next several items and eventually receiving severe impairment (i.e., 0) ratings for the hardest items. In Figure 3b., which represents a person of moderate ability (score of 15; SE 1.12), this person had a normal (i.e., 3) rating for the easiest item, primarily mild impairment (i.e., 2) ratings for the next cluster of items, moderate impairment (i.e., 1) for the next 4 items, followed by severe impairment (i.e., 0) for the most difficult item. In Figure 3b., which represents a person of high ability (score of 22; SE 1.30), this person obtains a normal (i.e., 3) rating for the first 5 items, begins to fluctuate between mild impairment (i.e., 2) and normal (i.e., 3) on the next three more difficult items, and then receives mild (i.e., 2) and moderate (i.e., 1) impairment for the two most difficult items. A dashed box around items presents "treatment targets". These items reflect challenges that are near the patient's ability level, identified as

ratings to the left of the black vertical line. These items reflect the next set of tasks the patient should gain mastery over as their ability level improves and the black line shifts to the right.

(INSERT FIGURE 3)

Discussion

This study evaluated the unidimensionality and measurement properties of the FGA using factor and Rasch analyses in persons post-stroke. Our results from this study provide support for the use of FGA in outpatient settings for assessing ambulatory individuals with chronic stroke as recommended by clinical practice guidelines. Factor analyses supported the unidimensionality of the FGA, implying the FGA is only measuring one construct, in this case, walking balance ability. Rasch analysis findings demonstrate that the FGA's rating scale structure, item difficulty hierarchy, and no floor or ceiling effect was sufficient for measuring a wide degree of walking balance ability on an interval scale in our sample, ambulatory persons with chronic stroke representative of individuals receiving outpatient physical therapy for walking balance related deficits. To our knowledge, this is the first study examining the measurement properties of the FGA for this patient population and the first to present an ability map for clinical use.

Item hierarchy, rating-scale structure, model fit

Ideally a measurement scale has sufficient items to assess the breath of difficulty for a specific construct and those items follow a difficulty hierarchy consistent with clinical and theoretical expectations.[29] The FGA's item difficulty hierarchy was previously analyzed using Rasch analysis in a sample of community-dwelling older adults[18] and had the benefit of being compared against the Dynamic Gait Index (DGI),[30-32] the FGA's predecessor.[1] Historically, the DGI had a logical order to the item hierarchy with easier items requiring less postural adjustments to perform the tasks and increasing item difficulty by adding additional task demands. The FGA's hierarchy is similar to the DGI, however, there is an exception with the item Gait Level Surface appearing at the middle of the hierarchy.[18] This may appear surprising because walking on a level surface without perturbations should theoretically be the easiest item. Possibilities for this finding could be related to the rating-scale structure or performance criteria. Beninato et. al. hypothesized that the reordering of item difficulty observed between the FGA and DGI was linked with the performance criterion used to assign a person a category on the rating-scale rather than poor rating-scale structure.[18] When we examine the category descriptions used as criterion for scoring there are several instances across items where individuals may have more difficulty reaching higher scores because of lasting sensorimotor deficits that are present with stroke pathology. Persons post-stroke are likely to have a difficult time obtaining higher than a 2 on the item Gait Level Surface because a 3 requires that the patient walk faster than 0.85 m/s without gait deviations or an assistive device. This contrasts with the item Gait and Pivot Turn, the easiest

item on our hierarchy, where ratings are not dependent on speed or gait deviations. This could explain some variation in item difficulty order between the community-dwelling older adult cohort previously studied, however, except for the easiest item difficulty was not starkly different. This may suggest the potential for the FGA to be “diagnosis free” and measurements could be interpreted similarly across diagnostic groups.

We found three items and approximately 10% of our sample misfit the Rasch model. It is expected data will depart from the model to an extent, with the question becoming ‘How much is tolerable?’.[25] A few misfitting people are unlikely to have a concerning impact on item-level psychometrics while a misfitting item may point to more important problems with unidimensionality, test administration, scoring, or definition of an item.[24, 25] Unfortunately, there is not a standardized approach to address misfit.[24] We report the fit statistics using the whole sample, rather than remove misfitting persons since; 1) removing misfitting persons did not improve the model fit and subsequent interpretation of measurement information from the FGA and 2) 10% or less individuals misfitting the model reflects the high degree of heterogeneity in this patient group. We left the misfitting items in our analysis for three reasons: 1) the FGA is unidimensional and there was no local dependence among items, 2) it has been suggested that mean-square values for items based on clinical observation can be higher than patient-reported scales (up to 1.7; meaning no FGA items misfit)[25], and 3) further evaluation of the items and their scoring criteria does not suggest they are measuring a different construct.[24]

Ability maps and their implications for clinical practice

The Rasch methodology provides an innovative way to approach clinical use of standardized measures with ability maps. Ability maps provides a way for clinicians to use Rasch informed interval scaling while visualizing the relationship between the patient's measure and ability.[10, 11, 13] Visualizing this relationship gives meaning and interpretability to an individual patient's score because clinicians can see items the patient has mastery over, are gaining mastery, and have no mastery. We identified items where the patient was rated just below their ability level to represent treatment targets. These items can be thought of as a "just right challenge" because should not be too easy or too difficult for the patient to complete while still challenging their ability level. These items can inform personalized short-term goals because they represent the next attainable ability level the patient should gain mastery over in rehabilitation (i.e., when their score shifts right with improvement).[10, 14-16] Ability maps can also be used over time to track progress and demonstrate a patient is gaining mastery over new functional tasks with increasing FGA scores.

Implications for future research

The degree of overlap between items (Figure 2) demonstrates the potential for the FGA to become a performance-based computerized adaptive test (CAT). An FGA CAT would have two key benefits for clinicians; 1) it would reduce test burden by requiring fewer items for measurement and 2) reduce

administration time by providing clinicians with immediate knowledge of the patient's walking balance ability level.[33-36]

Another important consideration that is identified by our results is the ability of the FGA to separate people into 4 distinct strata. Future work should explore the characteristics of the strata and test the ability of the FGA to separate persons into readily identifiable phenotypes using functional staging[37, 38] that can be used to inform personalized care.[39]

Finally, pragmatic clinical studies are needed to determine if the added information provided by the FGA ability map contributes to clinician goal setting and treatment planning compared to scores generated from the traditional FGA. These studies should also attempt to understand whether ability map help address many of the clinician-reported barriers to using standardized measures in practice.[16]

Limitations

There are several limitations to this study. First, we relied on retrospective data of higher functioning individuals with chronic stroke associated with one research site which can limit the generalizability of research findings. Second, measures and treatment target zones derived from an ability map are estimates and have sources of uncertainty. First, ratings are probabilistic in nature, not absolute, so they can fluctuate due to rater, patient, and environment factors. Next, the difficulty hierarchy of the items suggest that items are of increasing challenge but do not reflect the "measurement" distance between the items. [10, 13]

Lastly, there is added error in measuring ability levels at the extremes of the scale. Rasch measures corresponding to zero and perfect raw scores have infinite standard errors.[40] Infinite error is impractical to deal with in clinical practice so scales are slightly adjusted on the ability map when converting from the logit scale back to raw scoring. This is why the scale on the ability map is from 2 to 29[41] Despite these limitations, the ability map is able to provide a rapid estimation of a patient's ability and provides more interpretability to scores. (e.g., the participants in Figure 3 a-b show clear differences in their scoring patterns).[10]

Conclusions

In conclusion, Rasch analysis supports the use of the FGA for measuring walking balance ability in persons with chronic stroke. The FGA ability map provides clinicians with a new way to use and interpret patient responses on the FGA for clinical reasoning related to treatment design and goal setting that was not previously available. The FGA ability map should aid in addressing barriers to clinical use and promote the implementation of clinical practice guidelines for stroke rehabilitation.

Acknowledgements

None

Funding

Partial funding for this project was provided by the VA Office of Research and Development (ORD), with additional support from the VA/ORD Rehabilitation R&D Service (1I01RX001935, 1IK6RX003075, 1I01 RX002256-01A1, 1I01 RX000844, N2665-R), support from the National Institutes of Health (NIH P20 GM109040, 1 R21 HD083964-01A1) and the Promotion of Doctoral Studies Level I Scholarship from the Foundation for Physical Therapy Research. Data for the study was provided by The NIH Center for Biomedical Research Excellence in Stroke Recovery supported by NIH P20 GM109040.

Any opinions expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Veteran Affairs or the NIH.

Clinical Trial Registration:

Some of the data used in this study was collected from clinical trials registered at ClinicalTrials.gov (NCT03561246, NCT01970592, and NCT02964039).

Disclosures/Presentations

The authors report no conflicts of interest.

This material has not been previously presented at any conferences or meetings. The full manuscript has not been submitted for publication with any other journal.

References

1. Wrisley DM, Marchetti GF, Kuharsky DK, Whitney SL. Reliability, internal consistency, and validity of data obtained with the functional gait assessment. *Physical therapy.* Oct 2004;84(10):906-918.

2. Thieme H, Ritschel C, Zange C. Reliability and validity of the functional gait assessment (German version) in subacute stroke patients. *Archives of physical medicine and rehabilitation.* Sep 2009;90(9):1565-1570.

3. Lin JH, Hsu MJ, Hsu HW, Wu HC, Hsieh CL. Psychometric comparisons of 3 functional ambulation measures for patients with stroke. *Stroke.* Sep 2010;41(9):2021-2025.

4. Wrisley DM, Kumar NA. Functional gait assessment: concurrent, discriminative, and predictive validity in community-dwelling older adults. *Physical therapy.* May 2010;90(5):761-773.

5. Leddy AL, Crowner BE, Earhart GM. Functional gait assessment and balance evaluation system test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall. *Physical therapy.* Jan 2011;91(1):102-113.

6. Yang Y, Wang Y, Zhou Y, Chen C, Xing D, Wang C. Validity of the Functional Gait Assessment in patients with Parkinson disease: construct, concurrent, and predictive validity. *Physical therapy.* Mar 2014;94(3):392-400.

7. Moore JL, Potter K, Blankshain K, Kaplan SL, O'Dwyer LC, Sullivan JE. A Core Set of Outcome Measures for Adults With Neurologic Conditions

Undergoing Rehabilitation: A CLINICAL PRACTICE GUIDELINE. *Journal of Neurologic Physical Therapy.* 2018;42(3):174-220.

8. Stevens JG, Beurskens AJ. Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. *Physical therapy.* Jun 2010;90(6):953-961.

9. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Physical therapy.* Feb 2009;89(2):125-135.

10. Velozo CA, Woodbury ML. Translating measurement findings into rehabilitation practice: an example using Fugl-Meyer Assessment-Upper Extremity with patients following stroke. *Journal of rehabilitation research and development.* 2011;48(10):1211-1222.

11. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Physical therapy.* May 2006;86(5):735-743.

12. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 3rd ed. New York and London: Routledge; 2015.

13. Linacre JM. Instantaneous measurement and diagnosis. *Physical Medicine and Rehabilitation.* 1997;11(2):315-324.

14. Haley SM, Ni P, Ludlow LH, Fragala-Pinkham MA. Measurement Precision and Efficiency of Multidimensional Computer Adaptive Testing of Physical Functioning Using the Pediatric Evaluation of Disability Inventory.

Archives of physical medicine and rehabilitation. 2006/09/01/ 2006;87(9):1223-1229.

15. Grattan ES, Velozo CA, Skidmore ER, Page SJ, Woodbury ML. Interpreting Action Research Arm Test Assessment Scores to Plan Treatment. *OTJR: Occupation, Participation and Health.* 2019;39(1):64-73.

16. Woodbury ML, Anderson K, Finetto C, et al. Matching Task Difficulty to Patient Ability During Task Practice Improves Upper Extremity Motor Skill After Stroke: A Proof-of-Concept Study. *Archives of physical medicine and rehabilitation.* Nov 2016;97(11):1863-1871.

17. Hong I, Lim Y, Han H, Hay CC, Woo HS. Application of the Korean Version of the Modified Barthel Index: Development of a keyform for use in Clinical Practice. *Hong Kong journal of occupational therapy : HKJOT.* Jun 2017;29(1):39-46.

18. Beninato M, Ludlow LH. The Functional Gait Assessment in Older Adults: Validation Through Rasch Modeling. *Physical therapy.* Apr 2016;96(4):456-468.

19. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2019: http://www.R-project.org/

20. Yves Rosseel (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. URL http://www.jstatsoft.org/v48/i02/.

21. Revelle, W. (2020) psych: Procedures for Personality and Psychological Research, NorthwesternUniversity, Evanston, Illinois, USA, https://CRAN.Rproject.org/package=psych, Version = 2.0.9

22. John Fox (2019). polycor: Polychoric and Polyserial Correlations. R package version 0.7-10. https://CRAN.R-project.org/package=polycor

23. Linacre JM. Rasch power analysis: size vs. significance: standardized chi-square fit statistic. *Rasch Measurement Transactions.* 2003;17:918.

24. Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Medical care.* May 2007;45(5 Suppl 1):S22-31.

25. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measurement Transactions.* 1994;8(3):370.

26. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open.* Jul 27 2015;5(7):e007765.

27. Wright BD, Masters GN. Number of Person or Item Strata: (4*Separation + 1)/3. *Rasch Measurement Transactions.* 2002;16(3):888.

28. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal.* 1999/01/01 1999;6(1):1-55.

29. Franchignoni F, Giordano A, Ronconi G, Rabini A, Ferriero G. Rasch validation of the Activities-specific Balance Confidence Scale and its short

versions in patients with Parkinson's disease. *Journal of rehabilitation medicine.* Jun 2014;46(6):532-539.

30. Dye DC, Eakman AM, Bolton KM. Assessing the validity of the dynamic gait index in a balance disorders clinic: an application of Rasch analysis. *Physical therapy.* Jun 2013;93(6):809-818.

31. Chiu YP, Fritz SL, Light KE, Velozo CA. Use of item response analysis to investigate measurement properties and clinical validity of data for the dynamic gait index. *Physical therapy.* Jun 2006;86(6):778-787.

32. Marchetti GF, Whitney SL. Construction and validation of the 4-item dynamic gait index. *Physical therapy.* Dec 2006;86(12):1651-1660.

33. Wainer H. *Computerized adaptive testing: A primer, 2nd ed*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2000.

34. Porter I, Goncalves-Bradley D, Ricci-Cabello I, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *Journal of comparative effectiveness research.* Aug 2016;5(5):507-519.

35. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? *Health services research.* 2005;40(5 Pt 2):1694-1711.

36. Seamon BA, Kautz SA, Velozo CA. Measurement Precision and Efficiency of Computerized Adaptive Testing for the Activities-specific Balance Confidence Scale in People With Stroke. *Physical therapy.* Jan 22 2021.

37. Wang YC, Hart DL, Stratford PW, Mioduski JE. Clinical interpretation of computerized adaptive test-generated outcome measures in patients with knee impairments. *Archives of physical medicine and rehabilitation.* Aug 2009;90(8):1340-1348.

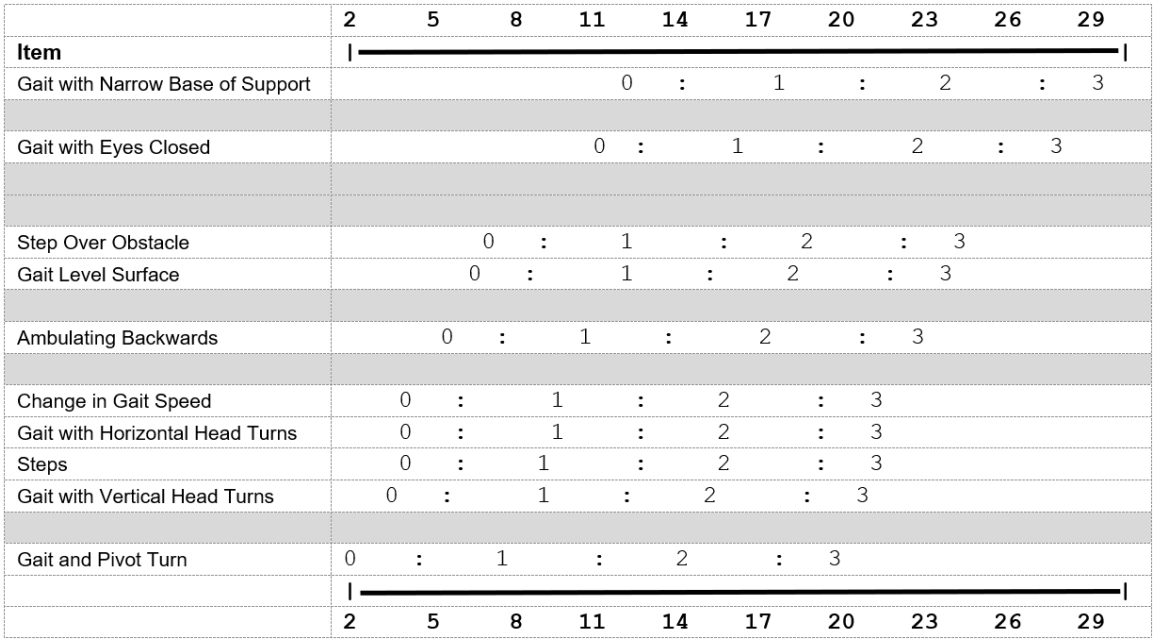
38. Wang YC, Sindhu B, Lehman L, Li X, Yen SC, Kapellusch J. Rasch Analysis of the Activities-Specific Balance Confidence Scale in Older Adults Seeking Outpatient Rehabilitation Services. *The Journal of orthopaedic and sports physical therapy.* Jul 2018;48(7):574-583.

39. Boyd LA, Hayward KS, Ward NS, et al. Biomarkers of stroke recovery: Consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. *International Journal of Stroke.* 2017;12(5):480-493.

40. Linacre JM. Standard Errors and Reliabilities: Rasch and Raw Score. *Rasch Measurement Transactions.* 2007;20(4).

41. Kozlowski AJ, Cella D, Nitsch KP, Heinemann AWDoPM, Rehabilitation FSoMNUCIL. Evaluating Individual Change With the Quality of Life in Neurological Disorders (Neuro-QoL) Short Forms. *Archives of physical medicine and rehabilitation.* 2016;97(4):650-654.

Figure 1. FGA Keyform Score Sheet



A keyform score sheet can be used to create an FGA ability map to determine a patient's measurement value and informing clinical decision-making based on that patient's performance on each item of the FGA. The FGA keyform score sheet is informed by the Rasch model and uses the interval scale scoring and item difficulty hierarchy. The top and bottom axes represent the interval score scale for the FGA. The logit scale produced by the Rasch model was converted back to the raw scores for ease of interpretation. The rows of the FGA keyform score sheet contains individual items with their category-rating scores. The items go from easiest to hardest with the easiest item at the bottom. Items are clustered by similar difficulty and each cluster is separated by a gray space. The rating-scale categories remain the same from the original FGA instructions (ie. 0=severe impairment, 1=moderate impairment, 2=mild impairment, 3=normal). A clinician can complete an ability map by circling the rating corresponding to a

patient’s performance for each item. The interval measure can be determined by drawing a vertical line through the “bulk” of the circles to the scale on the x-axis.

Table 1. Participant Demographic Information

Total Number of Participants = 101				
Age (years)	58.6 (12.6) [23-85]			
Male (n=57)	56.4%			
Female (n=44)	43.6%			
Right Hemiparesis (n=59)	41.6%			
Left Hemiparesis (n=42)	58.4%			
Lower Extremity Fugl-Meyer	25.2 (5.5) [14-34]			
Overground Self-Selected Walking Speed (m/s)	0.76 (0.29) [0.11-1.4]			
Functional Gait Assessment	15.7 (6.1) [6-30]			
	Rating-scale Categories			
Item-Level Information	**0**	**1**	**2**	**3**
(1) Gait Level Surface	1	62	27	11
(2) Change in Gait Speed	3	34	41	23
(3) Gait with Horizontal Head Turns	2	30	50	19
(4) Gait with Vertical Head Turns	1	30	48	22
(5) Gait and Pivot Turn	0	18	52	31
(6) Step Over Obstacle	23	34	24	20
(7) Gait with Narrow Base of Support	60	19	14	8
(8) Gait with Eyes Closed	38	41	19	3
(9) Ambulating Backwards	3	44	38	16
(10) Steps	0	29	57	15

Continuous variables are presented as mean (standard deviation) [range]

Categorical variables are presented as a percentage

Overground walking speeds were collected without use of an assistive device or orthotic unless an aircast was used to prevent ankle injury.

Frequency counts are provided for each category of the rating-scale per individual item on the FGA

Table 2. FGA Rating-Scale Structure

Rating-Scale Score	Observed Measure Average	Infit Mean-square	Outfit Mean-square	Frequency Count (%)
0	-3.53	.98	.98	131 (13%)
1	-1.17	.84	.81	341 (34%)
2	1.32	.95	1.18	370 (37%)
3	3.30	1.21	1.19	168 (17%)

Table 3. Item Measure Order

Item	Item Number	Measure Estimate	Standard Error	Infit Mean-square (z-score)	Outfit Mean-square (z-score)	Point Measure Correlation
Gait with Narrow Base of Support*	7	3.17	0.21	1.78 (4.4)	1.52 (2.4)	0.75
Gait with Eyes Closed*	8	2.48	0.20	1.60 (3.8)	1.54 (3.1)	0.64
Step Over Obstacle†	6	0.55	0.18	1.10 (0.7)	1.10 (0.8)	0.86
Gait Level Surface†	1	0.31	0.18	0.63 (-3.1)	0.64 (-2.9)	0.80
Ambulating Backwards	9	-0.34	0.19	0.73 (-2.1)	0.74 (-2.0)	0.79
Change in Gait Speed†	2	-0.93	0.19	0.83 (-1.3)	0.82 (-1.3)	0.78
Gait with Horizontal Head Turns†	3	-1.00	0.19	0.83 (-1.3)	0.82 (-1.3)	0.74
Steps†	10	-1.03	0.19	0.78 (-1.7)	0.92 (-.5)	0.71
Gait with Vertical Head Turns†	4	-1.18	0.19	0.75 (-1.9)	0.74 (-2.0)	0.77
Gait and Pivot Turn*	5	-2.04	0.20	0.98 (-0.1)	1.52 (2.7)	0.65

* Indicates a misfitting item

† Indicates an overlapping item

Items are presented in difficulty order to show their hierarchy with the hardest item first followed by items in decreasing level of difficulty

Figure 2. Person-Item Map

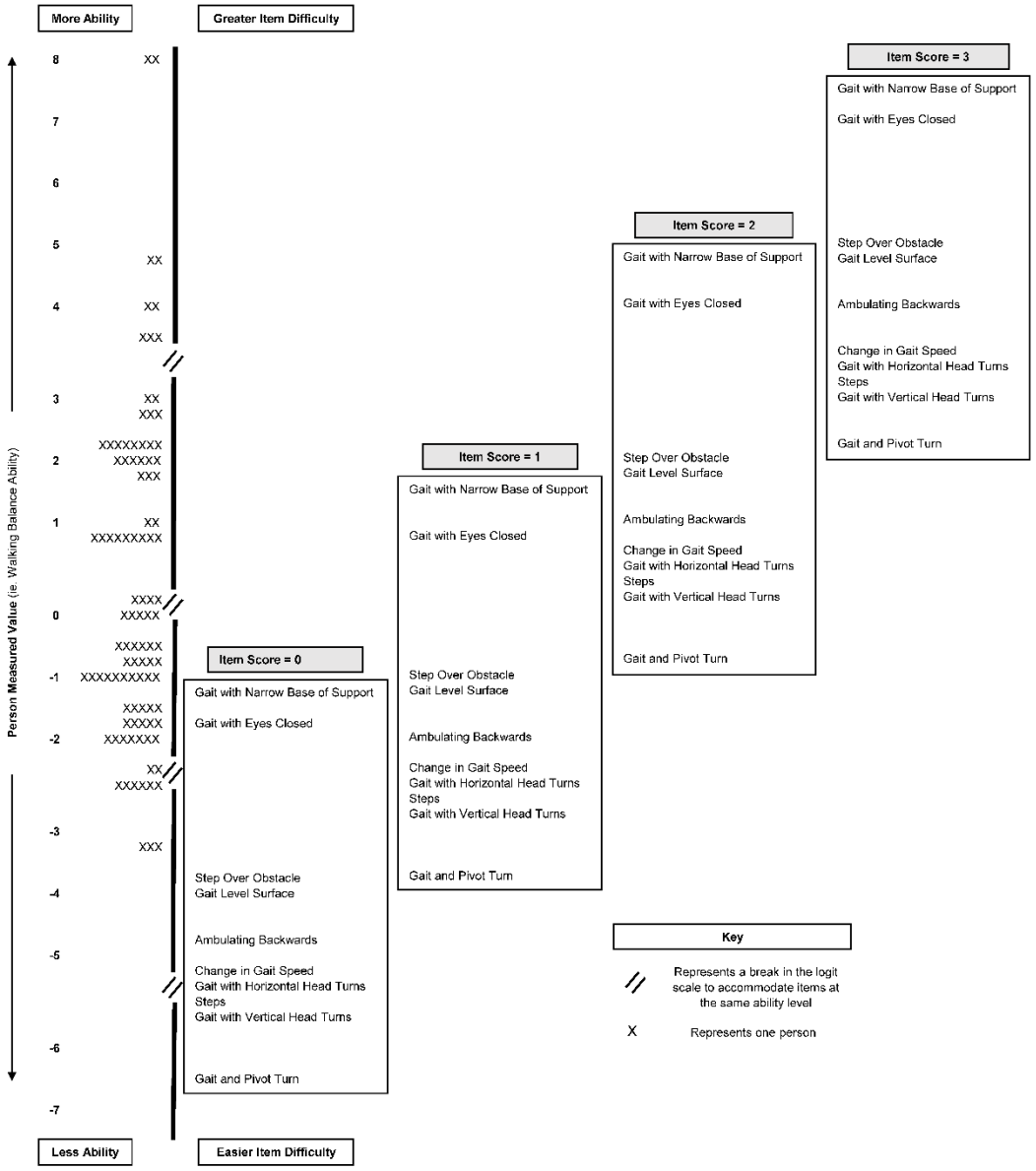


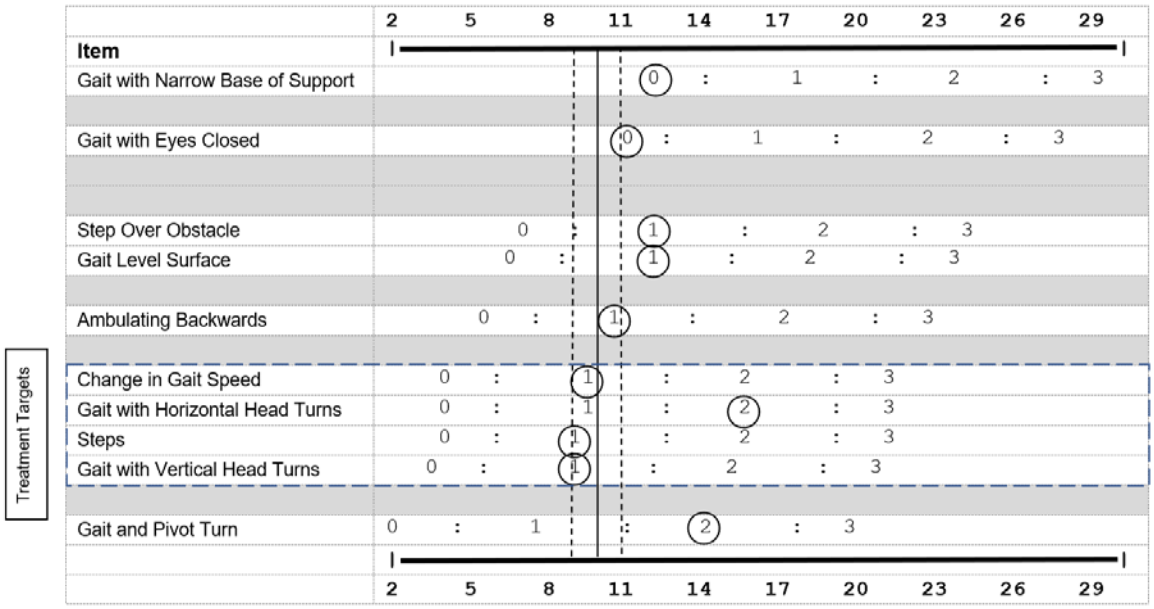
Figure 2 presents the distribution of ability levels (ie. walking balance ability) for our sample to the left of the black line. To the right of the black line, the items are presented in a hierarchy from most difficult (top) to least difficult (bottom). The hierarchy is repeated based on the highest probability of person receiving a

specified score for an item (0, 1, 2, or 3) corresponding to their ability level. This part of the figure can be used to predict a person's score for each item based on their ability level by extending a horizontal line across the figure at a given ability level. Persons will be most likely to receive the highest available score for an item that is at or below their ability level.

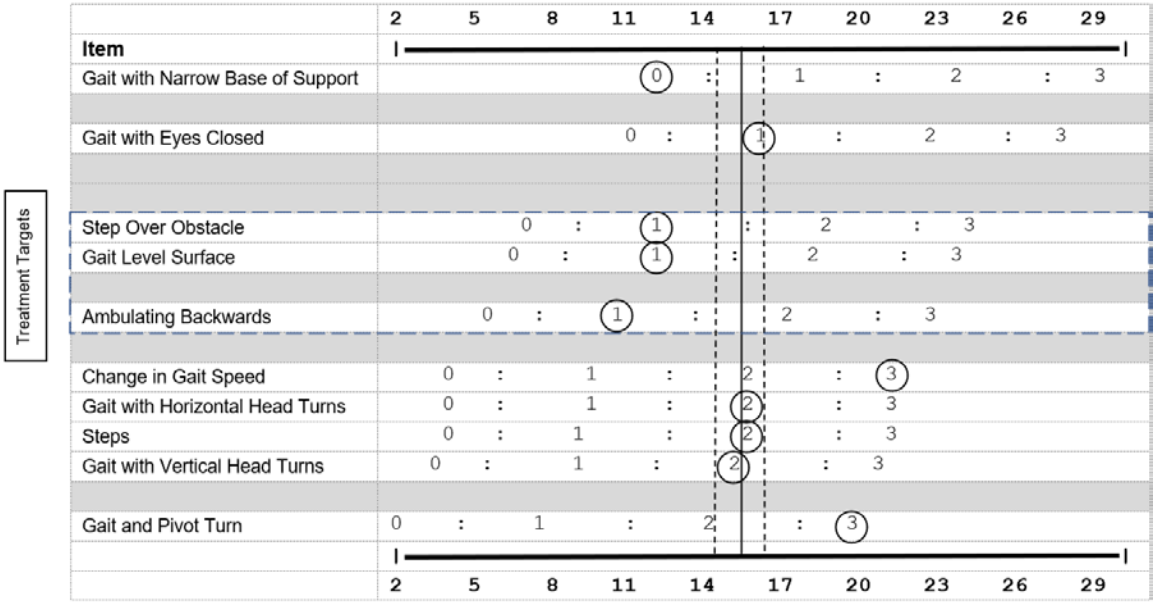
For example, a person with an ability level of 0 logits (approximately the mean ability level) would have the highest probability of scoring a 0 on the items Gait with Narrow Base of Support and Gait with Eyes Closed, a 2 on the item Gait Pivot and Turn, a 1 on all remaining items. Similarly, a person with an ability level of 4 would be most likely to score a 1 on the item Gait with Narrow Base of Support, 2 on Gait with Eyes Closed, Step Over Obstacle, Gait Level Surface and 3 on Ambulating Backwards and all remaining items.

Figure 3. Examples of FGA ability maps for clinical measurement and decision-making

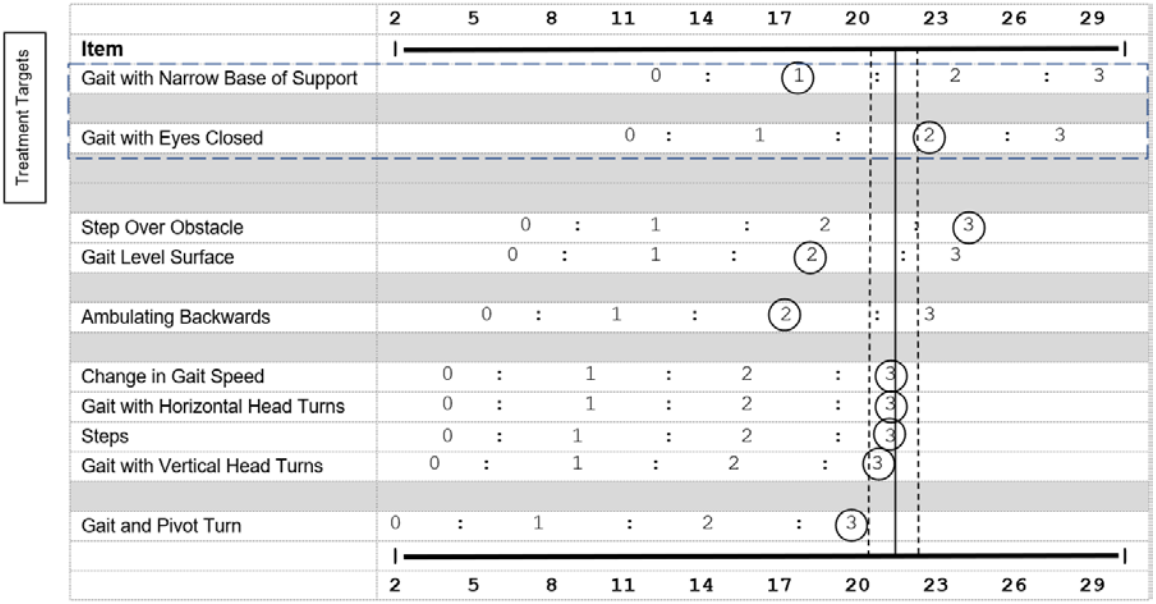
3a. Person with Low Ability [Person Measure = 10; standard error = 1.17]



3b. Person with Moderate Ability [Person measure = 15 ; standard error = 1.12]



3c. Person with High Ability [Person measure = 22; standard error = 1.30]



Completed FGA ability maps are presented for a representative individual with low (3a.), moderate (3b), and high (3c) walking balance ability. The persons true ability level (determined by the Rasch model) is presented by a solid vertical line. The standard error associated with the person's ability level is presented as two dashed vertical lines. Circles are placed around the individual's rating-scale score for each item. The rating-scale from 0-3 matches the original rating-scale used with the FGA. Clinicians can obtain the person's measure by drawing the vertical lines through the "bulk" of the circles. Items that represent treatment targets are represented by a short-term goal planning box. These items are reflective of challenges to walking balance ability that are not too easy or difficult for the patient and are ideal treatment targets. The treatment target range should shift up as patient's progress in rehabilitation and gain mastery over skills below their ability level (i.e., left of the black line).

Title Page

Revisiting the concept of minimal detectable change for patient-reported outcome measures

Bryant A. Seamon PT, DPT[1,2], Steven A. Kautz, PhD[1,2], Mark G. Bowden[1,2,3], Craig A. Velozo, PhD, OTR/L[2,4]

[1] Ralph H. Johnson VA Medical Center, Charleston, SC, USA

[2] Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[3] Division of Physical Therapy, Department of Rehabilitation Sciences, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[4] Division of Occupational Therapy, Department of Rehabilitation Sciences, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

The name and address for the corresponding author:

Bryant A. Seamon, PT, DPT

77 President Street, MSC 700

Charleston, SC 29425

843-310-3608

seamon@musc.edu

Abstract

Interpreting change is a requisite component of clinical decision making for physical therapists. Physical therapists often interpret change using Minimal Detectable Change (MDC) threshold values. Current MDC threshold formulas are informed by Classical Test Theory and calculated with group-level error data. A Classical Test Theory approach assumes measurement error is the same across the entirety of the measure's scale and confines the MDC to the sample characteristics of the study. An item response theory informed approach accounts for variability in error by converting ordinal scales into interval measures based on latent trait ability that have their own associated error estimates. Error estimates at the measure-level can be used to determine a conditional minimal detectable change (cMDC) threshold for individual patients based on their unique pre- and post-score combination. cMDC thresholds can provide clinicians with a means for using individual score data to interpret change scores. cMDCs provide a personalized approach that should lower the threshold for change compared to the MDC and enhance precision of care decisions by preventing misclassification of patients. The purpose of this perspective paper is to present how principles of item response theory and findings from a Rasch analysis can address MDC thresholds limitations for informing clinical practice. We demonstrate how a conditional minimal detectable change (cMDC) threshold can be generated from item-level psychometrics derived from the Rasch model using the patient-reported Activities-specific Balance Scale (ABC) commonly used in stroke rehabilitation. We also illustrate how the cMDC compares to the

MDC when accounting for changes in measurement error across a scale. With theoretical patient examples, we highlight how reliance on the MDC can result in misclassification of patient change and cMDCs can help prevent this from occurring. This personalized approach for interpreting change can be used by physical therapists to enhance precision of care decisions.

Keywords: stroke, balance, measurement – applied

Introduction

Precision rehabilitation requires that physical therapists accurately interpret a patient's change over time to make appropriate clinical decisions such as whether an intervention is efficacious. Standardized measurement tools can be used to quantify a patient's change by calculating the difference between a pre-intervention and post-intervention scores. However, the clinical meaning of change scores is not intuitively apparent for physical therapists.[1, 2] Physical therapists need to be able to understand the responsiveness of the measure to interpret whether the change scores are reflective of detectable change. Responsiveness of a standardized measure is commonly quantified with a minimal detectable change (MDC) threshold. MDC thresholds represent the amount of change needed to exceed measurement error for a specific measure and are calculated with an accompanying confidence interval.[3-5] For example, an MDC threshold with a 95% confidence interval (MDC_{95}) equal to 15 points means that 95% of individuals who are unchanged will have a random fluctuation in their scores on pre- and post-tests up to 15.[6] Clinically, this means a patient would need to change greater than 15 points to have a detectable change that exceeds measurement error, regardless of their initial and final scores. The ability to have a threshold for establishing "real" change on a measurement tool holds a great deal of clinical benefit because physical therapists can easily identify whether patients benefit or not from treatment and inform care decisions.

However, there are downsides to using MDC thresholds stemming from underlying assumptions that limit the clinical interpretation and application for

physical therapists.[1, 4] A primary concern with MDC thresholds is the reliance on group-level error data to calculate the value. Group-level data is heavily dependent on the characteristics of the sample used to quantify measurement error.[7] This means that MDC thresholds are sample specific, limiting the validity of the MDC in situations where an individual patient is not entirely reflective of the sample's characteristics used to derive the change threshold. Group-level data, also, assumes that error is consistent across the whole measurement scale.[7] Measures in rehabilitation rarely have consistent error across the whole scale as the precision is typically greater at midrange scores and less at the extremes.[8] This means that detecting change is dependent on the error specifically associated with an individual patient's pre- and post-scores. Determining MDC values from pre- and post-score combinations should be preferable because it may improve (i.e., lower) a measurement tool's threshold for responsiveness, especially when patients have scores near the midrange of the scale.[4, 9] More precise thresholds could help prevent physical therapists from making inappropriate care decisions that result from misclassifying change in patients. This method also promotes a more personalized approach to rehabilitation by allowing physical therapists to interpret change from individual-level data.

The downsides to MDC thresholds were somewhat apparent when the concept was initially introduced into physical therapy practice by Stratford and colleagues.[4] Stratford and colleagues hypothesized that a more precise approach for determining change thresholds could come from individualizing MDC values to the error associated with a person's specific initial and final score

combinations, also known as a conditional minimal detectable change (cMDC).[4] This conditional approach to calculating an MDC threshold should address concerns with changes in error across a measurement scale and allow physical therapists to use individual-level data to make personalized decisions regarding change for patients. The idea of using cMDC thresholds did not initially take hold largely because of the classical test theory approach to measurement and error. Typically, classical test theory has been used in measurement to explain the patient's overall score into two parts: 1) the true score and 2) error. This approach gives an understanding of the construct for the whole instrument (total score) but relies on group-level data error making it difficult to calculate cMDC thresholds.[7] While there are now mathematical formulas for deriving cMDC thresholds under classical test theory approach, the calculations are intensive and resulting values can be hard to interpret when scales are ordinal.[4, 9]

One way to overcome challenges with classical test theory is to use an item response theory model to generate cMDC thresholds.[9] Item response theory models, like the Rasch model, examine the individual item fit to the measurement model by examining the probabilistic relationship between a person's ability and the difficulty of the item.[10] For example, persons should have a high probability of doing well on items less difficult than their ability levels, a low probability of doing well on items that are more difficult than their ability level, and a 50% probability of doing well on items that are at their ability level. When measures fit the Rasch model, an item difficulty hierarchy (i.e., easy to hard item difficulty ordering) can be generated with a linear, interval measurement scale.[10] This scale is

independent of the sample and allows a level of precision, standard error, to be generated for every person measure available.[8] The standard errors can be used to quickly generate cMDC thresholds using a simple formula for every pre- and post-score combination.[9]

The purpose of this paper is to demonstrate how cMDC thresholds can be derived from Rasch informed measure-level error and to illustrate the benefits of individualized detectable change thresholds using a common patient-reported standardized outcome measure in stroke rehabilitation, the Activities-specific Balance Confidence (ABC) scale. We will demonstrate how changes in error across a scale comparatively affect cMDC and traditional MDC threshold calculations. We will also use theoretical patient examples to illustrate how reliance on a single threshold for detecting change can result in misclassification and inappropriate clinical decision making compared to a personalized conditional threshold, which we argue will enhance care precision.

Methods

Previous Rasch Analysis Findings

We previously performed Rasch analysis of the ABC scale using response data from persons post-stroke in Winsteps [version 3.93.1; John Lincare/Winsteps.com, Beaverton, OR, USA].[11] Data for the ABC scale was taken from the Locomotor Experience Applied Post-Stroke (LEAPS) trial and included 406 individuals approximately 2-months post-stroke.[12]

Results of the Rasch analysis for the ABC scale showed that the original rating-scale structure did not adequately fit the Rasch model because many of the categories in the rating-scale (i.e., 0-100%) were underused. We addressed this issue by collapsing the rating scale to 5 categories: no confidence (0-9%), low confidence (10-30%), medium confidence (31-60%), high confidence (61-90%), full confidence (91-100%). Collapsing the rating-scale improved the fit of the ABC scale to the Rasch model and generated item difficulty hierarchies and item-level psychometrics that were similar to those found in other studied populations.[11, 13-16]

Rasch analysis generates an interval scale using logits for person measures that quantify a person's ability level, in this case balance confidence, and each person measure has a standard error (SE). The logit scale is anchored with a mean value of 0 and standard deviation of 1, which allows for negative and decimal point values. The logit scale may confuse clinicians because the current scoring for the ABC scale does not provide negative values or decimals. Logits can be converted to different values by selecting a new anchor for the mean and re-scaling the person measures.[17] We converted persons measures and SE from logits to a 100-point scale to mimic current scoring for the ABC scale and to help with clinical interpretation.

Calculating MDC Thresholds

Group-level error data for the ABC scale was calculated using original scores (i.e., prior to Rasch analysis) with the following formula[18]:

$$\text{Standard Error of Measurement (SEM)} = \text{Standard Deviation} * \sqrt{(1-\text{Cronbach Alpha})}$$

The Standard Error of Measurement (SEM) was used to calculate an MDC threshold with a 95% confidence interval (MDC_{95}) with the following formula[19]:

$$MDC_{95} = 1.96 * \sqrt{(2)} * (SEM)$$

Calculating cMDC Thresholds

We calculated cMDC thresholds with a 95% confidence interval ($cMDC_{95}$) for every possible pair of pre- and post-score combinations using the following formula[9]:

$$cMDC_{95} = ([SE_{pre\text{-}score} - SE_{post\text{-}score}]/2) * 1.96 * \sqrt{(2)}$$

Comparing MDC and cMDC Thresholds

We plotted all the $cMDC_{95}$ thresholds associated with each patient measure on the ABC scale to observe the effect of individual-level standard error on detectable change across the measurement scale. In addition, we plotted the MDC_{95} threshold to illustrate the differences between using a fixed and conditional threshold for detecting change. We extracted three $cMDC_{95}$ threshold curves to show theoretical patient examples with an initial high (75), moderate

(50), and low (25) measure and plotted the $cMDC_{95}$ thresholds against change scores to show when there are mismatches between $cMDC_{95}$ and MDC_{95} for determining patient change. We use these theoretical examples to show the value of using $cMDC_{95}$ thresholds to detect change in individual patients.

Results

Figure 1 shows a plot of all possible $cMDC_{95}$ thresholds for each initial and final measure combination compared to the MDC_{95} threshold. The x-axis represents the patient's final ABC measure. The y-axis represents the change threshold value required for detectable change. Each line on the plot contains the $cMDC_{95}$ values associated with an initial measure value. A vertical line can be draw up from the final measure value on the x-axis to where it intersects the line associated with a person's initial measure. The point of intersection on the y-axis the associated $cMDC_{95}$. The dash line, in Figure 1, represents the MDC_{95} threshold. Our MDC_{95} was 14.72, which is similar to those reported for the ABC scale[20, 21] in previous studies with persons post-stroke. The MDC_{95} threshold dash line is fixed across the whole scale because it is not dependent on the patient's initial or final measures.

The $cMDC_{95}$ changes across the scale because it is a function of the SE associated with the initial and final measures. In the plot, each line has a fixed initial measure and associated SE while the final measure and associated SE is changing. The different heights for each line are reflective of the amount of SE that is associated with the fixed initial measure value. Lines at the top of the plot

represent high or low extreme initial measures (i.e., 0 or 100). The lines begin to move down towards the bottom of the plot as the initial measure value is closer to mid-range of the scale because there is less SE associated with these scores.

(INSERT FIGURE 1 HERE)

Figure 2 contains three theoretical patient examples. Figure 2A. is for a patient with a moderate initial measure value of 50, 2B. is for a patient with a low initial measure value of 25, and 2C. is for a patient with a high initial measure value of 75. Each of the three plots have the change score required for detectable change on the y-axis and all possible change scores for that patient. The blue dash line in each plot is the $cMDC_{95}$ curve and the black dash line is the MDC_{95} value. Detectable change based on the $cMDC_{95}$ and MDC_{95} value can be determined by comparing the x-axis value and the associated y-axis value for each line ($cMDC_{95}$ and MDC_{95}). The patient has achieved a detectable change when the x-axis value exceeds the y-axis value. Each plot also has a Misclassification Zone. The Misclassification Zone represents a range of change scores where there is a mismatch between the $cMDC_{95}$ and MDC_{95} thresholds.

For example, in Figure 2A., the patient has initial measure of 50 and the x-axis represents change scores from 50 to 100 (max value on the scale). Change scores from 0 to 6 (i.e., final measure values of 50-56) would not be considered detectable change by the $cMDC_{95}$ or MDC_{95}, change scores from 7 to 14 would be considered detectable change by the $cMDC_{95}$ but not the MDC_{95} and change

scores from 15 to 50 would be considered detectable change by both the $cMDC_{95}$ or MDC_{95}. The range of change scores from 7 to 14 where there is a mismatch in change detection by the $cMDC_{95}$ or MDC_{95} creates the Zone of Misclassification. Figure 2B. has the same pattern as 2A. The initial range of change scores 0 to 8 is not considered detectable change by the $cMDC_{95}$ or MDC_{95}. Change scores between 9 and 14 are considered detectable change by the $cMDC_{95}$, not the MDC_{95}, and creates a Zone of Misclassification. Change scores greater than 15 are then considered detectable change under both thresholds. Figure 2C still displays a Zone of Misclassification, but the associated alignment between $cMDC_{95}$ and MDC_{95} thresholds is different. For this patient, with a high initial measure value of 75, the change scores from 0 to 14 would not be considered detectable change. However, change scores from 15 to 25 would be considered detectable change with the MDC_{95} and not the $cMDC_{95}$. In this case, the relationship between the $cMDC_{95}$ and MDC_{95} is reversed for the Zone of Misclassification.

(INSERT FIGURE 2 HERE)

Discussion

The purpose of this paper was to demonstrate that cMDC thresholds can be derived from a Rasch informed version of the ABC scale and to examine the differences between an MDC and cMDC approach for detecting change using theoretical patient examples. Our results show that cMDC thresholds can easily

be derived when patient measure-level data is available for a standardized outcome measure. We were able to demonstrate how cMDC thresholds are a function of the change in error across a scale with more error associated at the extreme scores following a “u-shaped” trajectory.[8] This feature allows cMDC’s to generally reduce the threshold for detecting change in comparison to an MDC threshold – particularly for persons with measures near the mid-range of the scale. Our theoretical patient examples show that there are ranges of change scores that fall into a Misclassification Zone where detectable change under the MDC and cMDC are not in agreement. These examples highlight the benefits of using individual-level data compared to group-level data for making interpretations about unique patients. The examples also demonstrate how patients are at risk of being misclassified as either changing or not changing when using group-level data informed thresholds for detectable change (i.e., MDCs).

Misclassification of patients based on the MDC approach has wide reaching implications for clinical decision making.[9] In a broad sense, the most concerning misinterpretation of change would be classifying a patient as having plateaued or not responded to physical therapy treatment and ending an episode of care when the patient had in fact made a measurable change. However, misclassification of change also has implications for clinical decision making related to treatment prescription. Physical therapists regularly make decisions to manipulate variables related to treatment prescription to either progress patients or regress them when a prescription has detrimental effects. Misinterpretation of

change, positive or negative, could result in clinicians inappropriately altering the volume, frequency, intensity, or mode of treatment. Regardless of whether these inappropriate decisions have serious adverse events (i.e., over prescribing treatment), treatment precision is lost, and individual patient outcomes suffer.

In addition to clinical applications, the use of cMDC vs MDC thresholds has implications for research. Often MDC thresholds are used to determine sample size, determine treatment efficacy, calculate measures used to inform clinical decisions like number needed to treat, and serve as anchors for identifying minimally important differences or minimally important clinical differences.[9, 22-25] Specifically, using an MDC threshold will likely cause samples sizes to be overestimated, which may cause trials to not appear feasible because of the required number of participants. In addition, MDC thresholds may cause beneficial (or detrimental) treatment effects to be missed, especially when the primary outcome is dichotomized based on whether a participant had a detectable change. These implications are especially relevant in rehabilitation studies where patient-reported or clinician-observed measures, instead of laboratory measures, are the primary outcome. Researchers should consider using item-response theory informed measures and cMDC thresholds to prevent overpowering studies or inappropriately concluding that a treatment has no effect.

Limitations and Future Directions

There are still limitations to the cMDC approach despite the benefits of using individual-level data. The primary limitation is that cMDC thresholds must be more easily attainable to be commonly used by researchers and physical therapists.[26] There are a couple of ways to address this barrier. First, standardized measures that use ordinal scales need to be informed by item response theory methodology (e.g., Rasch analysis) to generate interval-level scaling and measure-level precision estimates (i.e., standard error). This should quell concerns about using ordinal level data for calculating MDC thresholds.[26-29] Second, measure-level precision should be reported in item response theory papers generating item-level psychometrics for generating cMDCs. This would allow clinicians to be able to generate cMDC's and provide researchers to with access to error information for study design and statistical analysis planning. Another limitation for adopting a cMDC approach is the fact each standardized measure will have a change threshold for every possible score combination. It is unlikely to expect clinicians to use cMDCs if they are not easily accessible in real time, especially when the alternative is committing to memory a few MDC thresholds that are relevant to common measures they use and patient populations they treat. To address this concern, we recommend creating web-based calculators or apps that can integrate with electronic medical records to rapidly generate an individual's cMDC threshold in real time for clinical decision making. These products also have potential to be integrated with electronic versions of standardized measures, such as a computerized adaptive test, to provide a real-time indicator to clinicians when patients have a detectable

change. Future studies should also aim to examine whether measure-level precision estimates are unique to specific populations. It is feasible that some standardized measurement tools may be “diagnosis-free” [11], but more research in this area is necessary before applying cMDC’s across patient populations. Also, while we have shown how cMDC’s work for a patient-reported outcome measure, the Misclassification Zone may not be as large or display a similar result for measures that already have appropriate rating-scale structure prior to performing a Rasch analysis. We recommend that future studies examining the differences between MDC and cMDC thresholds investigate the effect of rating-scale structure on error.

Lastly, clinicians should be aware that cMDC thresholds allow for the detection of change beyond measurement error for individual patients but do not attempt to inform whether that change was meaningful to the patient. There are several proposed methods for assigning meaningfulness to change scores.[1, 24, 25] However, many of these approaches utilize group-level MDC calculations and do not allow for individual interpretation. We recommend that clinicians consider the individual patient goals in relation to their overall function with the cMDC to make clinical decisions regarding the “meaningfulness” of change until more empirical methodology is presented to promote a more personalized rehabilitation approach.

Conclusions

In conclusion, cMDC thresholds can be easily generated from Rasch informed measure-level precision estimates. These cMDC thresholds have explicit benefits over the traditional MDC approach including the use of individual-level data to inform change detection. The cMDC approach allows clinicians to interpret change relative to their individual patients and greatly reducing the opportunity for misclassifying change or no change. This is especially true where clinicians are at risk of underreporting change for measures in the middle of a scale and overreporting change at the extreme high end using a traditional MDC. Clinicians can use cMDCs to provide a personalized view of interpreting change that should enhance clinical decision making and precision rehabilitation.

Role of the Funding Source

The funding sources for this research played no role in the design, conduct or reporting of this study and findings.

Acknowledgements

We would like to acknowledge the LEAPS investigator team [Principal Investigator: Pamela Duncan, PT, PhD, FAPTA, FAHA] for data collection and archival. We also want to acknowledge the National Institute of Neurological Disorders and Stroke for funding the LEAPS trial (R01 NS050506).

Authors Note

Data from the LEAPS trial can be obtained by contacting the National Institute of Neurological Disorders and Stroke at www.ninds@nih.org

Funding

Partial funding for this project was provided by the VA Office of Research and Development (ORD), with additional support from the VA/ORD Rehabilitation R&D Service (1I01RX001935), support from the National Institutes of Health (NIH P20 GM109040,) and the Promotion of Doctoral Studies Level I Scholarship from the Foundation for Physical Therapy Research. Data for the study was provided by the NIH National Institute of Neurological Disorders and Stroke from the Locomotor Experience Applied Post-stroke (LEAPS) trial (R01 NS050506).

Any opinions expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Veteran Affairs or the NIH.

Clinical Trial Registration:

Some of the data used in this study was collected from clinical trials registered at ClinicalTrials.gov (NCT00243919).

Disclosures/Presentations

The authors report no conflicts of interest.

Partial findings from this manuscript were presented at the Medical University of South Carolina's Research Day in 2020. The full manuscript has not been submitted for publication with any other journal.

References

1. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. *Physical therapy.* May 2006;86(5):735-743.
2. Velozo CA, Woodbury ML. Translating measurement findings into rehabilitation practice: an example using Fugl-Meyer Assessment-Upper Extremity with patients following stroke. *Journal of rehabilitation research and development.* 2011;48(10):1211-1222.
3. Riddle DL, Stratford PW. Is This Change Real? : Interpreting Patient Outcomes in Physical Therapy. Philadelphia: F.A. Davis Company; 2013.
4. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the Minimum Level of Detectable Change for the Roland-Morris Questionnaire. *Physical therapy.* 1996;76(4):359-365.
5. Stratford PW, Binkley JM. Applying the results of self-report measures to individual patients: an example using the Roland-Morris Questionnaire. *The Journal of orthopaedic and sports physical therapy.* Apr 1999;29(4):232-239.
6. Stratford PW, Riddle DL. When minimal detectable change exceeds a diagnostic test-based threshold change value for an outcome measure: resolving the conflict. *Physical therapy.* Oct 2012;92(10):1338-1347.
7. Albano AD. Reviewing Classical Test Theory. *Introduction to Educational and Psychological Measurement Course Notes*: University of Nebraska-Lincoln; 2016.

8. Linacre JM. Standard Errors and Reliabilities: Rasch and Raw Score. *Rasch Measurement Transactions.* 2007;20(4).

9. Kozlowski AJ, Cella D, Nitsch KP, Heinemann AWDoPM, Rehabilitation FSoMNUCIL. Evaluating Individual Change With the Quality of Life in Neurological Disorders (Neuro-QoL) Short Forms. *Archives of physical medicine and rehabilitation.* 2016;97(4):650-654.

10. Bond TG, Fox CM. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 3rd ed. New York and London: Routledge; 2015.

11. Seamon BA, Kautz SA, Velozo CA. Rasch Analysis of the Activities-Specific Balance Confidence Scale in Individuals Poststroke. *Archives of Rehabilitation Research and Clinical Translation.* 2019/12/01/ 2019;1(3):100028.

12. Duncan PW, Sullivan KJ, Behrman AL, et al. Protocol for the Locomotor Experience Applied Post-stroke (LEAPS) trial: a randomized controlled trial. *BMC neurology.* Nov 8 2007;7:39.

13. Arnadottir SA, Lundin-Olsson L, Gunnarsdottir ED, Fisher AG. Application of Rasch analysis to examine psychometric aspects of the activities-specific balance confidence scale when used in a new cultural context. *Archives of physical medicine and rehabilitation.* Jan 2010;91(1):156-163.

14. Franchignoni F, Giordano A, Ronconi G, Rabini A, Ferriero G. Rasch validation of the Activities-specific Balance Confidence Scale and its short

versions in patients with Parkinson's disease. *Journal of rehabilitation medicine.* Jun 2014;46(6):532-539.

15. Sakakibara BM, Miller WC, Backman CL. Rasch analyses of the Activities-specific Balance Confidence Scale with individuals 50 years and older with lower-limb amputations. *Archives of physical medicine and rehabilitation.* Aug 2011;92(8):1257-1263.
16. Wang YC, Sindhu B, Lehman L, Li X, Yen SC, Kapellusch J. Rasch Analysis of the Activities-Specific Balance Confidence Scale in Older Adults Seeking Outpatient Rehabilitation Services. *The Journal of orthopaedic and sports physical therapy.* Jul 2018;48(7):574-583.
17. Linacre JM. A user's guide to Winsteps Rasch-model computer programs. *Help for Winsteps Rasch Measurement and Rasch Analysis Software*2009: http://www.winsteps.com/a/winsteps.pdf. Accessed 2019.
18. Wyrwich KW. Minimal important difference thresholds and the standard error of measurement: is there a connection? *Journal of biopharmaceutical statistics.* Feb 2004;14(1):97-110.
19. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. *Physical therapy.* Nov 1998;78(11):1186-1196.
20. Botner EM, Miller WC, Eng JJ. Measurement properties of the Activities-specific Balance Confidence Scale among individuals with stroke. *Disability and rehabilitation.* Feb 18 2005;27(4):156-163.

21. Salbach NM, Mayo NE, Robichaud-Ekstrand S, Hanley JA, Richards CL, Wood-Dauphinee S. Balance self-efficacy and its relevance to physical function and perceived health status after stroke. *Archives of physical medicine and rehabilitation.* Mar 2006;87(3):364-370.

22. van de Graaf VA, Noorduyn JCA, Willigenburg NW, et al. Effect of Early Surgery vs Physical Therapy on Knee Function Among Patients With Nonobstructive Meniscal Tears: The ESCAPE Randomized Clinical Trial. *Jama.* 2018;320(13):1328-1337.

23. Froud R, Eldridge S, Lall R, Underwood M. Estimating the number needed to treat from continuous outcomes in randomised controlled trials: methodological challenges and worked example using data from the UK Back Pain Exercise and Manipulation (BEAM) trial. *BMC Medical Research Methodology.* 2009/06/11 2009;9(1):35.

24. Malec JF, Ketchum JM. A Standard Method for Determining the Minimal Clinically Important Difference for Rehabilitation Measures. *Archives of physical medicine and rehabilitation.* 2020/06/01/ 2020;101(6):1090-1094.

25. Collins JP. Measures of Clinical Meaningfulness and Important Differences. *Physical therapy.* 2019;99(11):1574-1579.

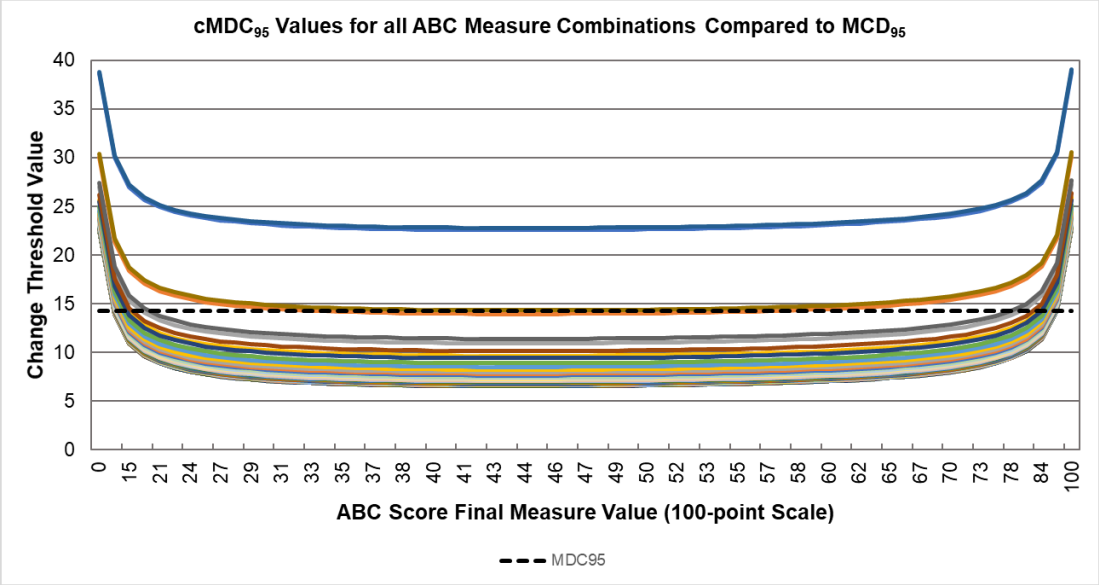
26. Caronni A, Picardi M, Gilardone G, Corbo M. The McNemar Change Index worked better than the Minimal Detectable Change in demonstrating the change at a single subject level. *Journal of clinical epidemiology.* 2021/03/01/ 2021;131:79-88.

27. Stevens SS. On the theory of scales of measurement. *Science (New York, N.Y.).* Jun 7 1946;103(2684):677-680.

28. Kahler E, Rogausch A, Brunner E, Himmel W. A parametric analysis of ordinal quality-of-life data can lead to erroneous results. *Journal of clinical epidemiology.* May 2008;61(5):475-480.

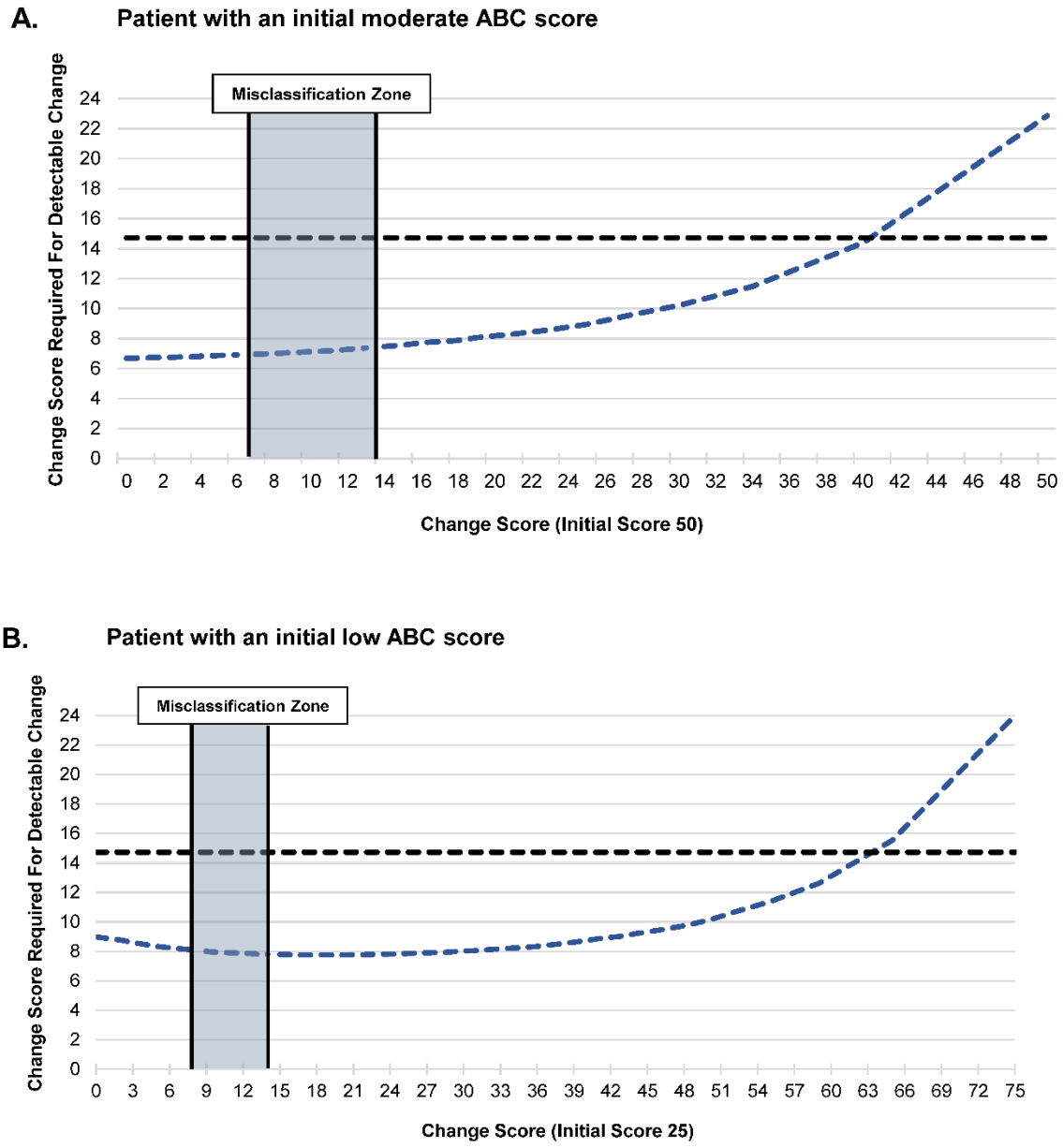
29. Anselmi P, Vidotto G, Bettinardi O, Bertolotti G. Measurement of change in health status with Rasch models. *Health and quality of life outcomes.* Feb 7 2015;13:16.

Figure 1. cMDC$_{95}$ and MDC$_{95}$ thresholds for the ABC scale



Each line contains all possible cMDC$_{95}$ thresholds associated with a fixed initial score. The dash line represents the MDC$_{95}$ threshold. The dash line representing the MDC$_{95}$ threshold is fixed because it is not dependent on a specific initial and final measure combination. The x-axis represents the patient's final measure value, and the y-axis reflects the minimal detectable change threshold value associated with the combination of initial and final measures. A specific cMDC$_{95}$ threshold can be identified selecting a final measure value and drawing a vertical line from the measure's location on the x-axis to a line that represents a fixed initial measure of interest. The corresponding y-axis value where the lines intersect is the cMDC$_{95}$ threshold for that measure pairing. The different heights of the curves are related to the amount of standard error associated with the fixed initial score. For example, the higher a line is on the plot, the closer the initial measure is to an extreme end of the scale and the lower a line is on the plot the closer the initial measure is to the middle of the scale.

Figure 2. Comparing $cMDC_{95}$ and MDC_{95} thresholds with Patient Examples



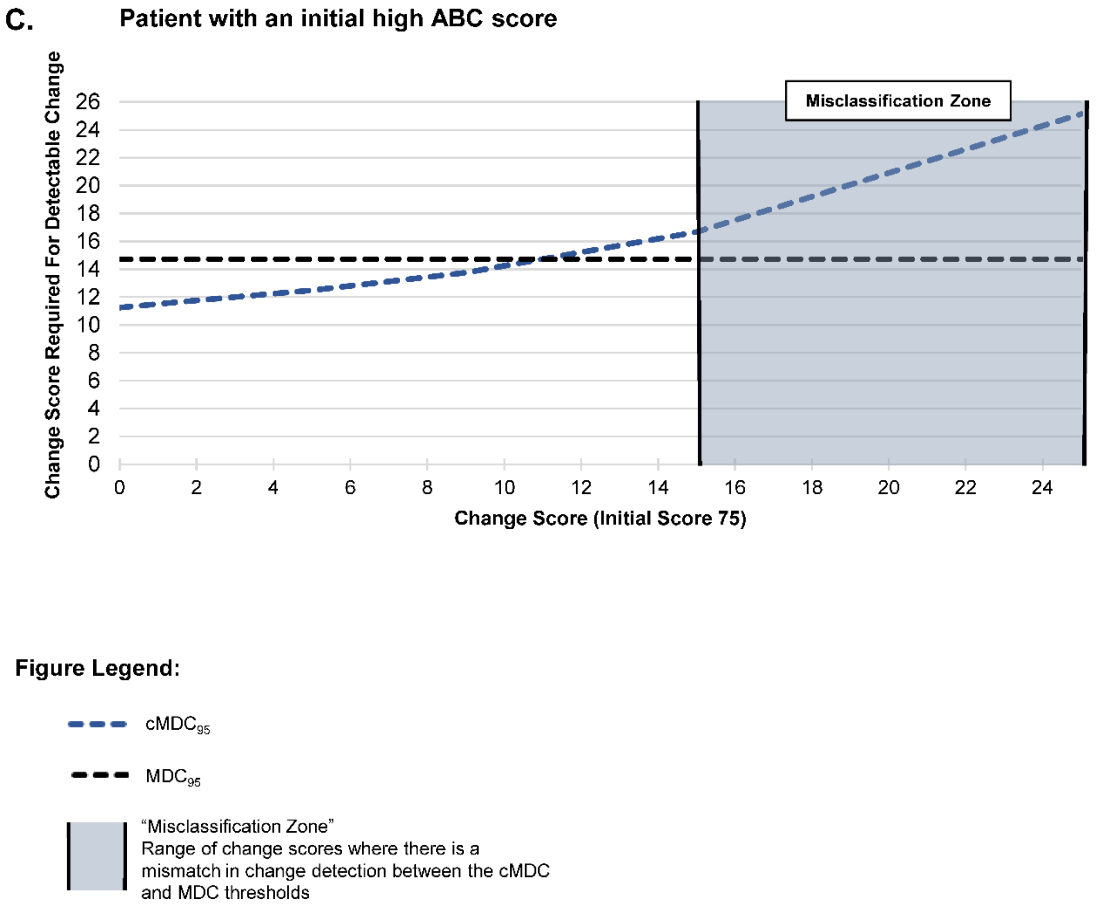


Figure 2 presents three theoretical patient examples to demonstrate how misclassification of change occurs when using the MDC_{95} compared to a $cMDC_{95}$. The x-axis represents the change scores based on the person's initial measure value on the ABC. The y-axis represents the change scores required for detectable change. The blue dash line represents the $cMDC_{95}$ curve associated with the initial measure value for the theoretical patient and the black dash line represents the MDC_{95}. The change threshold can be identified by drawing a vertical line from the x-axis of the patient's change score to where it intersects with the blue dash line. The associated y-value with the intersection is the

required threshold for detectable change. The change score magnitude can be compared to the threshold to see if the patient has a detectable change. These figures compare whether the patient had detectable change under the $cMDC_{95}$ and MDC_{95} approach. The Misclassification Zone is noting the range of scores where the $cMDC_{95}$ and MDC_{95} approach do not agree. In A and B the $cMDC_{95}$ identified a change but the MDC_{95} did not in the Misclassification Zone. In C the $cMDC_{95}$ did not identify change and the MDC_{95} did in the Misclassification Zone. Outside of the Misclassification Zone the $cMDC_{95}$ and MDC_{95} were in agreement.

4.3 PART II MANUSCRIPT

Associations between biomechanical variables of walking performance and muscle coordination during self-selected steady-state walking

Title Page

Associations between biomechanical variables of walking performance and muscle coordination during self-selected steady-state walking

Bryant A. Seamon PT, DPT[a,b], Shraddha Srivastava, PhD[a,b], Viswanathan Ramakrishnan, PhD[c], Mark G. Bowden, PT, PhD[a,b,d], Richard R. Neptune, PhD[e], Steven A. Kautz, PhD[a,b]

[a] Ralph H. Johnson VA Medical Center, Charleston, SC, USA

[b] Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[c] Department of Public Health Sciences, College of Medicine, Medical University of South Carolina, Charleston, SC, USA

[d] Division of Physical Therapy, Department of Health Professions, College of Health Professions, Medical University of South Carolina, Charleston, SC, USA

[e] Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX, USA

The name and address for the corresponding author:

Bryant A. Seamon, PT, DPT

77 President Street, MSC 700

Charleston, SC 29425

843-310-3608

seamon@musc.edu

Funding

Partial funding for this project was provided by the VA Office of Research and Development (ORD), with additional support from the VA/ORD Rehabilitation R&D Service (1I01RX001935), support from the National Institutes of Health (NIH P20 GM109040,) and the Promotion of Doctoral Studies Level I Scholarship from the Foundation for Physical Therapy Research.

Any opinions expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Veteran Affairs or the NIH.

Abstract

Background

Factorization of EMG data into modules has been used to quantify muscle coordination during gait. However, no one has examined the association between modules and observable biomechanical variables during walking outside of computer simulation models.

Research Question

Are modules associated with observable measures of biomechanical functions during steady-state walking?

Methods

We used a cross-sectional design with twenty, healthy individuals who completed 3, 30-second trials of steady-state walking at self-selected speeds on an instrumented treadmill with motion capture. EMG was collected from 8 lower extremity muscles bilaterally. Non-negative matrix factorization was used to extract 4 modules. Changes in biomechanical variables were quantified using the area under the curve for each bin of the gait cycle per step. Variables included anterior-posterior and vertical ground reaction force, leg angle, and leg length. Generalized mixed linear models with random coefficients for person by leg and person by step interactions were used to quantify the association between changes in module activation with biomechanical measures of walking. Comparisons within bins between modules and biomechanical variables were

selected *a priori* based on hypothesized relationships suggested from previous computer simulation work. Bonferroni corrections were applied for evaluating statistical significance.

Results

We found significant positive associations between Module 1 and vertical ground reaction force during early stance; Module 2 and AP and vertical ground reaction forces during late stance; Module 3 and changes in leg angle and leg length during swing phase, and decreased AP ground reaction forces in early stance; Module 4 and increased vertical ground reaction force in early stance, and decreased leg angle in late swing phase.

Significance

Our findings demonstrate that there are physiologically expected associations between biomechanical variables and modules during steady-state walking. Results from this study can be used for comparisons with pathological conditions impacting walking performance and informing biomarkers of walking recovery.

Keywords: Gait, biomechanics, electromyography, coordination, modules

Introduction

Walking is a particularly complex task that requires a high degree of well-coordinated muscle activity.[1] Recently, growing evidence has supported that the complex coordination of muscle activity during walking can be quantified using a module-based analysis.[2-5] Module-based analyses apply factorization algorithms to surface electromyographic (EMG) data collected from multiple muscles to identify modules that contain co-excited muscles with their activation profile during the gait cycle. Four modules are typically identified in healthy walking that have robust consistency in composition and timing of activation during the gait cycle across healthy individuals.[2, 4] Module 1, active in early stance, primarily contains activity from gluteus medius, rectus femoris, and vastus medialis. Module 2, active in late stance, consists of soleus and medial gastrocnemius activity. Module 3, active in early stance and swing, consists of tibialis anterior and rectus femoris. Module 4, active in late stance and early swing, consists of medial and lateral hamstring activity. However, factorization algorithms used to generate modules, are primarily a means of data reduction and may not necessarily produce factors, or modules, that are physiologically relevant.[6]

Neptune and colleagues[6] addressed this concern using a muscle-actuated forward dynamic computer simulation model with module data from Clark et. al..[2] They were able to show that the typical 4 modules were sufficient for producing coordinated waking. Additionally, they identified the contributions of each module to key biomechanical functions during walking. Module 1 provided

body support and deceleration during early stance, Module 2 provided body support and forward propulsion in late stance, while Module 3 and 4 provided leg acceleration in early swing and deceleration in late swing to prepare for foot contact.

No one to date has demonstrated the association between observable walking biomechanics and modules during healthy walking, outside of simulation. Therefore, while simulation has shown the biomechanical functions the modules should perform, no one has yet linked module performance with quantified biomechanical performance during walking. The purpose of this study was to quantify the associations between modules and biomechanical variables of walking that reflect the functions described in Neptune et al., 2009 during steady state walking at self-selected speeds in healthy adults. Specifically, we compared the association between module activation with changes in anterior-posterior (i.e., propulsion) and vertical (i.e., body support) ground reaction forces, leg angle (i.e. stance to swing transition), and length (i.e., leg clearance). Specifically, we hypothesized that Module 1 would be associated with changes in anterior-posterior and vertical ground reaction forces during early stance; Module 2 would be associated with changes in anterior-posterior and vertical ground reaction forces, and trailing leg angle during late stance; Module 3 would be associated with changes in anterior-posterior ground reaction forces during late stance, and leg angle and leg length during swing phase; and Module 4 would be associated with changes in anterior-posterior and vertical ground reaction forces during early stance, and leg angle in late swing phase.

Methods

Participants:

Participants for this study included 20 healthy older adults between the ages of 40-85. Participants were excluded if they had any neurologic disease or significant orthopedic impairments in the lower limb that would limit walking performance. All participants signed a written informed consent form approved by the Institutional Review Board at the Medical University of South Carolina and conformed to the Declaration of Helsinki.

Procedures:

Participants completed three, 30 second, walking trials at self-selected walking speed on a split belt instrumented treadmill (Bertec; Columbus, OH). Ground reaction forces were sampled at 2000 Hz. Force data was then filtered with a 4th order Savitzky-Golay filter and resampled at 100 Hz. Whole-body 3D kinematic data was recorded by a 16-camera motion capture system (PhaseSpace, Inc.; San Leandro, CA) and sampled at 120 Hz, filtered with a 4th order Savitzky-Golay filter, and resampled at 100 Hz. Bipolar pre-amplified electrodes (Motion Lab Systems; Baton Rouge, LN, USA) were used to collect surface EMG from eight muscles bilaterally: tibialis anterior, soleus, medial gastrocnemius, vastus medialis, rectus femoris, medial hamstrings, lateral hamstrings, and gluteus medius. Surface EMG was sampled at 2000 Hz, high-pass filtered at 40 Hz, demeaned, rectified, and low-pass filtered at 4 Hz using a

zero-lag forth order Butterworth filter and resampled to 100 Hz. EMG amplitude for each muscle was averaged within each region of the gait cycle (or bin): (1) initial contact/loading response (initial double support), (2) first half of single-leg stance, (3) second half of single-leg stance (4) second double support, (5) first half of swing, (6) second half of swing) for each step. The bin with the highest average was used to normalize EMG amplitude across all trials for an individual. Data for each step was time normalized to the gait cycle (0-100%)

Data Analysis:

NON-NEGATIVE MATRIX FACTORIZATION (NNMF): Quantifying muscle coordination was done using a module analysis with NMMF. EMG from each muscle for per leg was combined into an *m* x *t* matrix (EMG_O), where *m* is the number of muscles (8) and *t* is the time base (*t* = number of gait cycles x 101). An NNMF algorithm is applied to the *m* x *t* matrix for each person. NNMF creates two matrices that define a pre-selected number of modules (*n*), in this case *n* = 4. One is an *m* x *n* matrix indicating the relative weighting of each muscle within each module, also known as the *W* matrix. The second is an *n* x *t* matrix indicating the activation timing profile of each module for each step in the trial, also known at the *H* matrix. NNMF assumes that the muscle weightings are fixed across all steps and allows for muscles to be belong to more than one module. NNMF runs an iterative optimization procedure to minimize the error by adjusting the muscle weightings (*W* matrix) and activation profiles (*H* matrix). The *H* matrix

for each module per person was divided into separate steps and an area under the curve was calculated for each bin of the gait cycle per step.

BIOMECHANICAL VARIABLES FOR WALKING: Biomechanical variables were quantified from force and motion capture data using custom software in MATLAB (Mathworks, Natick, MA). Specific kinematic and kinetic variables of interest were selected to represent the biomechanical subtasks of walking. These variables included:

(1) Changes in ground reaction force or impulse (the time integral of the ground reaction force) in the anterior-posterior (AP), and vertical directions. Impulse was calculated by taking the net area under the curve (AUC) for each component (AP, vertical) of the ground reaction force in each bin of the gait cycle.

(2) Changes in leg angle in the anterior-posterior direction. Leg angle was calculated by taking the angle from a vertical line between the ground and center of mass of the pelvis and a line between the center of mass of the pelvis to the center of mass of the foot. The changes in leg angle were represented by the net AUC in each bin for each step.

(3) Changes in leg length. Leg length was calculated as the distance from the center of mass of the pelvis to the center of mass of the foot. We normalized leg length to for each participant to the distance between their pelvis center of mass

and foot center of mass during static standing. Changes in leg length were represented by the AUC in each bin for each step.

Statistical Analysis:

All statistical analyses were completed in SAS version 9.4 (SAS Inc., Cary, NC). PROC GLIMMIX was used to fit linear mixed models with random coefficients. The participant variable, participant by leg (i.e., right and left) interaction, and participant by step interaction were included as random effects in the model. The random coefficients model allowed us to examine the association between module and biomechanical variables while controlling for variability at the individual and group levels. Magnitude of association was quantified using parameter estimates. We used a Bonferroni correction for the number of tests (tests = 18) resulting in an alpha level of 0.003.

Results

Our sample consisted of 20 healthy adults with a mean age of 58 years. Twelve individuals were female, 17 were Caucasian, and 3 Black. Demographic data is presented in Table 1.

(*INSERT TABLE 1 HERE)*

Module Weighting and Activation Profiles

Module weighting and activation profiles are presented in Figure 1. The four modules' weightings were like previously published muscle compositions and displayed similar timing curves, with the exception of increased rectus femoris activity in Module 1 compared to Module 3.[2, 6] Visually, there was consistency across participants and between legs for module weightings and activation profiles. Module 1 was primarily active in early stance and consisted mainly of muscle activity from the gluteus medius, rectus femoris, and vastus medialis. Module 2 was primarily active in late stance and consisted mainly of muscle activity from the medial gastrocnemius and soleus. Module 3 was primarily active in swing phase and early stance, consisting mainly of muscle activity from the tibialis anterior and rectus femoris. Module 4 was primarily active in late swing phase and early stance, consisting mainly of muscle activity from the medial and lateral hamstrings. Module weighting and activation profiles are presented in Figure 1.

(INSERT FIGURE 1 HERE)

Module Associations with Biomechanical Variables

Average curves for each biomechanical variable during the gait cycle are presented in Figure 2. Each plot contains the group mean (black line), 1 (light gray) and 2 (dark gray) standard deviation ranges. Individual averages across steps for each leg (right and left) are shown as gray curves. Descriptive statistics for each biomechanical variable mean AUC and standard deviation are

presented by bin in Table 3. All parameter estimates and p-values from the linear mixed models are presented in Table 4 by module.

(INSERT FIGURE 2 HERE)

(INSERT TABLE 2 HERE)

MODULE 1: Module 1 was found to have a positive association with increases in vertical ground reaction forces in bin 1 (parameter estimate (PE)=68.38 standard error (6.05); $p<0.0001$) and bin 2 (PE=39.56 (7.66); $p<0.0001$). Module 1 also had a positive association with increases in AP ground reaction force in bin 2 (PE=5.49 (1.54); $p<0.0001$) and no significant association with bin 1.

MODULE 2: Module 2 was found to have a positive association with cumulative increases in vertical and AP ground reaction forces across bins 3 and 4, with the greatest magnitude in bin 4 (Vertical – PE=286.39 (8.12), $p<0.0001$; AP – PE=38.17 (1.48), $p<0.0001$). Additionally, Module 2 was associated with cumulative increased leg extension (i.e., increased negative AUC for leg angle) across bins 3 (PE=-2.35 (0.16); $p<0.0001$) and 4 (PE=-8.43 ()0.32); $p<0.0001$).

MODULE 3: Module 3 was associated with decreases in AP ground reaction force in bin 1 4 (PE=-6.0 (1.44); $p<0.0001$); a decrease in leg angle, or leg extension, in bin 5 (PE=-1.45 (0.28); $p<0.0001$); and an increase in leg length in

bin 6 (PE=0.03 (0.01); p<0.0017). Module 3 did not have any significant associations with AP leg angle in Bin 6 or leg length in Bin 5.

MODULE 4: Module 4 was positively associated with increases in vertical ground reaction force during bin 1 (PE=37.48 (4.43); p<0.0001) and with decreases in leg angle during bin 6 (PE=-2.46 (0.39); p<0.0001). We did not find any significant associations between Module 4 and AP ground reaction force in bin 1.

(INSERT TABLE 3 HERE)

Discussion

Our results demonstrate that there is a strong association with modular organization of muscle coordination and biomechanical variables of walking performance at self-selected speeds. We found several associations that supported our proposed hypotheses which were informed by previous computer simulation work[6] and are consistent with the theory that neural organization drives biomechanical output during movement.[3, 4, 6-10] The 4 modules we identified for each individual were consistent across participants and previous analyses using similar methodology, with the exception of increased rectus femoris activity in Module 1.[2, 11]

We hypothesized that Module 1 would be associated with changes in AP and vertical ground reaction forces in bins 1 and 2 because simulation showed Module 1 contributed to body support and braking during early stance phase.[2,

6] We observed that Module 1 was primarily associated with increases in vertical ground reaction force in bins 1 and 2. This association is supported by the primary contributions of vastus medialis and rectus femoris in Module 1. Several simulations studies have shown that vastus medialis contributes to body support through vertical ground reaction force[6, 12-14] and that rectus femoris contributes to body support with vastus medialis when co-active in early stance.[14, 15] However, contrary to our hypothesis, we did not see an association with AP ground reaction forces during bin 1 where one may expect to see a negative AP impulse (i.e., braking or backward propulsion) to correspond with deceleration and vastus medialis activity. [6, 12-14] Our observation of a positive association with increase AP ground reaction force in bin 2 could potentially be explained by gluteus medius activity since this muscle can contribute to forward propulsion, in addition to body support, during early stance.[12, 13]

We hypothesized Module 2 would be associated with changes in AP and vertical ground reaction forces in bins 3 and 4 because Module 2 has been shown to contribute to body support and forward propulsion during stance phase in simulation.[2, 6] Additionally, we hypothesized that Module 2 would be associated with increased leg extension which has been associated with propulsion in terminal stance during walking.[16, 17] Consistent with our hypotheses, we observed that Module 2 was primarily associated with a positive increase in vertical and anterior-posterior ground reaction forces across bins 3 and 4. We also observed that Module 2 was also strongly associated with leg

angle movements into extension across mid to late stance, bins 3 and 4. These findings are consistent with previous studies linking trailing leg angle and forward propulsion. [16, 17] Our findings suggest that muscle activity is coordinated to place the foot in an optimal position to maximize propulsion forces with respect to vertical forces generated by the ankle plantarflexors. We also found that Module 2 was associated with an increase in foot velocity during early swing in the anterior-posterior direction.

We hypothesized that Module 3 would be associated with changes in leg angle and leg length during swing phase because it is thought to coordinate with Module 4 to accomplish swing phase.[2, 6] We found that Module 3 was associated with leg angle (increased extension) in bin 5 and increases in leg length in bin 6, but we did not find any significant association for leg angle in bin 6 or leg length in bin 5, partially proving our hypothesis. We also hypothesized that Module 3 would be associated with decreases in AP ground reaction force (i.e., propulsion) during bin 4. We confirmed this hypothesis which supports previous findings that increased tibialis anterior activity in late stance can counteract propulsion by the ankle plantarflexors.[18]

Lastly, we hypothesized that Module 4 would be associated with leg angle in late swing phase (capturing leg deceleration) and increases in AP and vertical ground reaction forces during early stance.[2, 6] Our findings partially confirmed our hypotheses because we observed associations between Module 4 and leg angle in late swing and with vertical ground reaction forces in early stance. The negative association we saw in late swing with leg angle would suggest that

Module 4 is contributing to a decrease in leading limb angle in preparation for heel strike and stance phase.

Limitations

There are several limitations to this study. We selected biomechanical variables based on theorized linkages between the biomechanical functions of body support, propulsion, and leg swing. Our results suggest that there may have been other variables with stronger associations to these biomechanical functions, especially with respect to leg swing and propulsion; for example, using pelvic and foot center of mass acceleration to capture changes in velocity at the trunk and leg segments. Another limitation was our constraint to the sagittal plane of movement. Future studies should examine associations of modules with key biomechanical variables in the frontal plane.[6] Also, we constrained our quantification of biomechanical variables to the net area under the curve for each bin. It is reasonable to consider exploring associations with positive and negative AUCs especially for variables that may have primarily negative or positive values during a single bin. Additionally, our large sample size left us overpowered. Although generalized mixed linear models are very robust[19] and able to control for repeated measures (i.e., between legs and steps), we were at risk for identifying significant findings that were not meaningful[20]. We took steps to address this, including hypothesizing associations *a priori* and applying a more conservative correction for the number of statistical tests. Lastly, our study design is not able to prove causation between modular activation and changes in

biomechanical variables, although the simulation work has shown that the changes are consistent with module contributions to the dynamics of walking. Prospective study designs are needed to confirm our findings are causal.

Future Directions

Our findings provided more evidence to strengthen the theory that modules share a mechanistic link with observable biomechanical variables that are measurable during walking. This seems to suggest that modules may have a biomechanical signature that is expressed which could potentially be used as a biomarker for walking-specific motor control recovery in persons with pathological gait. Future studies should examine whether observed changes in module activity have similar associations with changes in biomechanical variables in pathological walking. This would be especially helpful for individuals with stroke where modules may reflect the loss of movement fractionation common in this patient population[2, 21-29] and can change with interventions.[11, 21, 30, 31] A biomechanical biomarker for modules could enable clinicians to have a surrogate measure for the measurement of muscle coordination during walking and inform patient specific interventions.[32]

Role of the Funding Source: Funding sources did not play a role in study design; the collection, analysis and/or interpretation of data; writing of the manuscript; or decision to submit for publication.

Declarations of Interest: None

References

[1] Gottlieb GL. Muscle activation patterns during two types of voluntary single-joint movement. J Neurophysiol. 1998;80:1860-7.

[2] Clark DJ, Ting LH, Zajac FE, Neptune RR, Kautz SA. Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. J Neurophysiol. 2010;103:844-57.

[3] d'Avella A. Modularity for motor control and motor learning. Advances in Experimental Medicine and Biology2016. p. 3-19.

[4] Ivanenko YP, Cappellini G, Dominici N, Poppele RE, Lacquaniti F. Coordination of locomotion with voluntary movements in humans. J Neurosci. 2005;25:7238-53.

[5] Cappellini G, Ivanenko YP, Poppele RE, Lacquaniti F. Motor patterns in human walking and running. J Neurophysiol. 2006;95:3426-37.

[6] Neptune RR, Clark DJ, Kautz SA. Modular control of human walking: a simulation study. J Biomech. 2009;42:1282-7.

[7] d'Avella A, Saltiel P, Bizzi E. Combinations of muscle synergies in the construction of a natural motor behavior. Nature neuroscience. 2003;6:300-8.

[8] Ting LH, Macpherson JM. A limited set of muscle synergies for force control during a postural task. J Neurophysiol. 2005;93:609-13.

[9] Ivanenko YP, Poppele RE, Lacquaniti F. Five basic muscle activation patterns account for muscle activity during human locomotion. J Physiol. 2004;556:267-82.

[10] McGowan CP, Neptune RR, Clark DJ, Kautz SA. Modular control of human walking: Adaptations to altered mechanical demands. J Biomech. 2010;43:412-9.

[11] Ferrante S, Chia Bejarano N, Ambrosini E, Nardone A, Turcato AM, Monticone M, et al. A Personalized Multi-Channel FES Controller Based on Muscle Synergies to Support Gait Rehabilitation after Stroke. Front Neurosci. 2016;10:425.

[12] Anderson FC, Pandy MG. Individual muscle contributions to support in normal walking. Gait Posture. 2003;17:159-69.

[13] Liu MQ, Anderson FC, Pandy MG, Delp SL. Muscles that support the body also modulate forward progression during walking. J Biomech. 2006;39:2623-30.

[14] Neptune RR, Zajac FE, Kautz SA. Muscle force redistributes segmental power for body progression during walking. Gait Posture. 2004;19:194-205.

[15] Zajac FE, Neptune RR, Kautz SA. Biomechanics and muscle coordination of human walking: part II: lessons from dynamical simulations and clinical implications. Gait Posture. 2003;17:1-17.

[16] Bowden MG, Balasubramanian CK, Neptune RR, Kautz SA. Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking. Stroke. 2006;37:872-6.

[17] Awad LN, Binder-Macleod SA, Pohlig RT, Reisman DS. Paretic Propulsion and Trailing Limb Angle Are Key Determinants of Long-Distance Walking Function After Stroke. Neurorehabil Neural Repair. 2015;29:499-508.

[18] Turns LJ, Neptune RR, Kautz SA. Relationships between muscle activity and anteroposterior ground reaction forces in hemiparetic walking. Archives of physical medicine and rehabilitation. 2007;88:1127-35.

[19] Schielzeth H, Dingemanse NJ, Nakagawa S, Westneat DF, Allegue H, Teplitsky C, et al. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods in Ecology and Evolution. 2020;11:1141-52.

[20] Sullivan GM, Feinn R. Using Effect Size-or Why the P Value Is Not Enough. J Grad Med Educ. 2012;4:279-82.

[21] Seamon BA, Neptune RR, Kautz SA. Using a Module-Based Analysis Framework for Investigating Muscle Coordination during Walking in Individuals Poststroke: A Literature Review and Synthesis. Appl Bionics Biomech. 2018;2018:3795754.

[22] Allen JL, Kautz SA, Neptune RR. The influence of merged muscle excitation modules on post-stroke hemiparetic walking performance. Clin Biomech (Bristol, Avon). 2013;28:697-704.

[23] Allen JL, Kautz SA, Neptune RR. Step length asymmetry is representative of compensatory mechanisms used in post-stroke hemiparetic walking. Gait Posture. 2011;33:538-43.

[24] Bowden MG, Clark DJ, Kautz SA. Evaluation of abnormal synergy patterns poststroke: relationship of the Fugl-Meyer Assessment to hemiparetic locomotion. Neurorehabil Neural Repair. 2010;24:328-37.

[25] L.G. Brough SAK, M.G. Bowden, C.M. Gregory, R.R. Neptune. Merged Plantarflexor Muscle Activity is Predictive of Poor Walking Performance in Post-Stroke Hemiparetic Subjects. J Biomech. 2018.

[26] Routson RL, Kautz SA, Neptune RR. Modular organization across changing task demands in healthy and poststroke gait. Physiol Rep. 2014;2.

[27] Gizzi L, Nielsen JF, Felici F, Ivanenko YP, Farina D. Impulses of activation but not motor modules are preserved in the locomotion of subacute stroke patients. J Neurophysiol. 2011;106:202-10.

[28] Barroso FO, Torricelli D, Molina-Rueda F, Alguacil-Diego IM, Cano-de-la-Cuerda R, Santos C, et al. Combining muscle synergies and biomechanical analysis to assess gait in stroke patients. J Biomech. 2017;63:98-103.

[29] Kautz SA, Bowden MG, Clark DJ, Neptune RR. Comparison of motor control deficits during treadmill and overground walking poststroke. Neurorehabil Neural Repair. 2011;25:756-65.

[30] Routson RL, Clark DJ, Bowden MG, Kautz SA, Neptune RR. The influence of locomotor rehabilitation on module quality and post-stroke hemiparetic walking performance. Gait Posture. 2013;38:511-7.

[31] Hashiguchi Y, Ohata K, Kitatani R, Yamakami N, Sakuma K, Osako S, et al. Merging and Fractionation of Muscle Synergy Indicate the Recovery Process in Patients with Hemiplegia: The First Study of Patients after Subacute Stroke. Neural Plast. 2016;2016:5282957.

[32] Cramer SC, Wolf SL, Adams HP, Jr., Chen D, Dromerick AW, Dunning K, et al. Stroke Recovery and Rehabilitation Research: Issues, Opportunities, and the National Institutes of Health StrokeNet. Stroke. 2017;48:813-9.

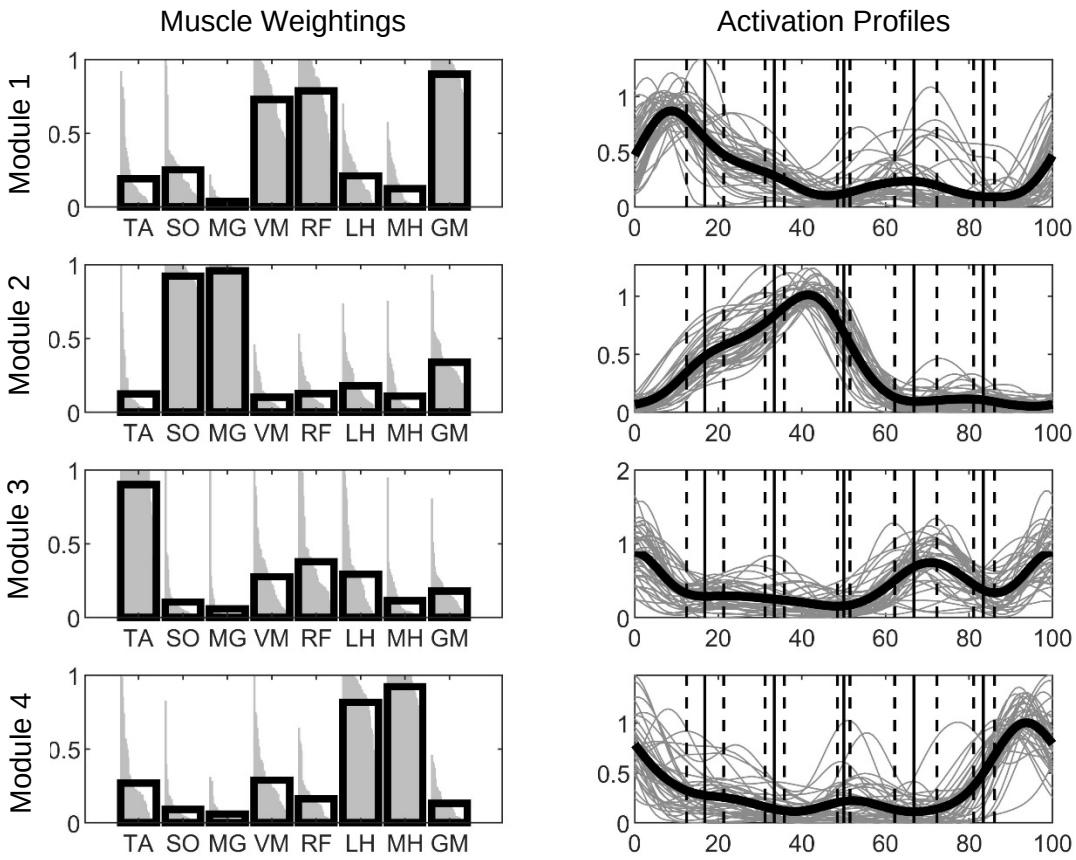
Table 1. Participant Demographics

n = 20	
Age (years)	58.2 (10.1) [40-79]
Sex	
Male	8 (40%)
Female	12 (60%)
Race	
White	17 (85%)
Back	3 (15%)

Continuous variables are presented as mean (standard deviation) [range]

Categorical variables are presented as count (frequency - percentage)

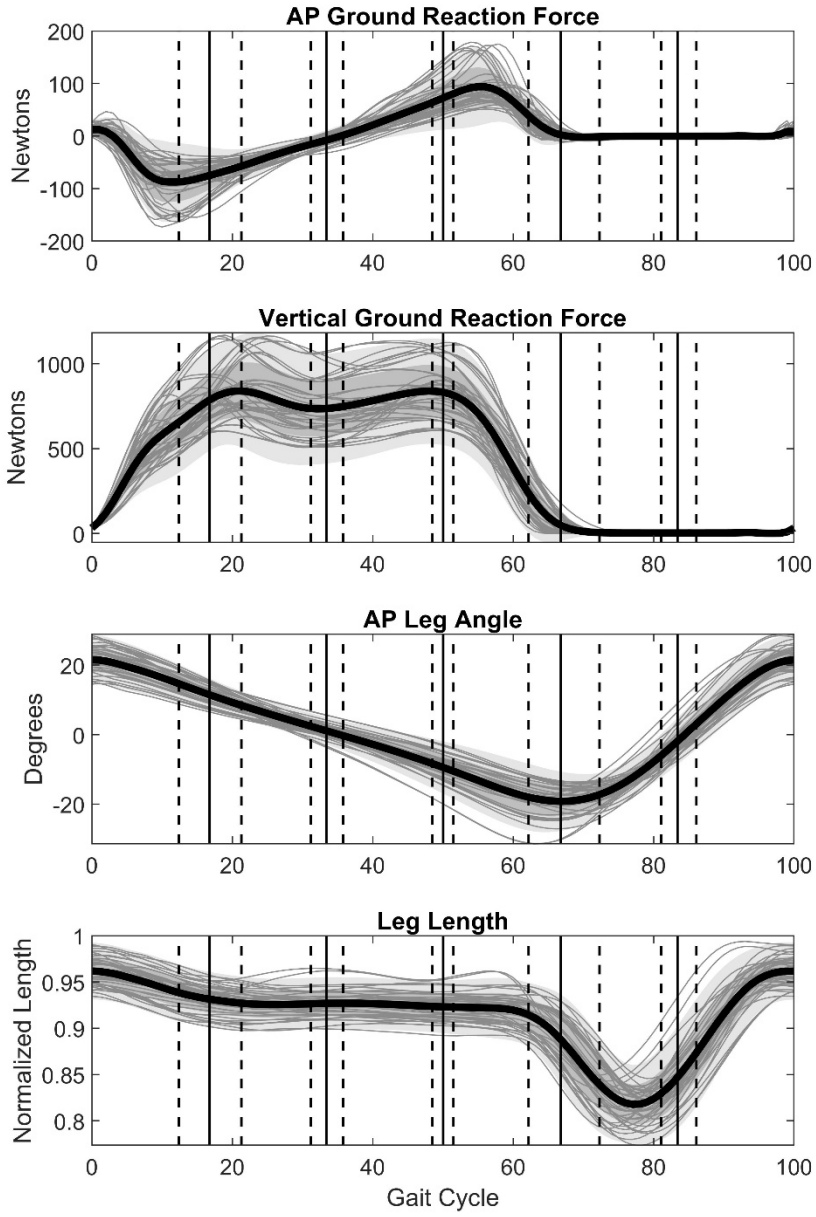
Figure 1. Muscle weightings and activation profiles for each module



Muscle weighting plots show the strength of representation for each muscle within a module. Gray bars are used to show the muscle weightings for each leg per participant per module. The black bar represents the group mean. Activation profiles represent the activation of the module over the gait cycle. Thin gray lines represent the profiles for each leg per participant averaged across all steps. The black line represents the group mean. Solid vertical lines represent the group mean for the start and stop of each bin. Dash vertical lines are presented to the right and left of the mean to show the range of values observed from our sample (i.e., average per person per leg) for the start and stop of each bin.

TA – tibialis anterior; SO – soleus; MG – medial gastroc; VM – vastus medialis; RF – rectus femoris; LH – lateral hamstring; MH – medial hamstring; GM – gluteus medius

Figure 2. Average curves for biomechanical variables of interest across the gait cycle



Biomechanical plots show the change for a variable across the gait cycle. Gray lines represent the average curve for each leg per participant. The black line represents the group mean. A dark shaded region surrounding the black line represents 1 standard of deviation and lighter shaded region represents 2

standards of deviation. Solid vertical lines represent the group mean for the start and stop of each bin. Dash vertical lines are presented to the right and left of the mean to show the range of values observed from our sample (i.e., average per person per leg) for the start and stop of each bin.

Table 2. Descriptive Statistics for Biomechanical Variable Areas Under the Curve

	Bin 1	**Bin 2**	**Bin 3**	**Bin 4**	**Bin 5**	**Bin 6**
AP Ground Reaction Force	-9.66 (3.26)	-9.42 (3.98)	5.77 (3.31)	13.21 (3.75)	.	.
Vertical Ground Reaction Force	94.18 (34.06)	158.02 (31.08)	157.60 (31.96)	96.95 (28.81)	.	.
AP Leg Angle	3.44 (0.78)	1.31 (0.48)	-0.80 (0.58)	-2.97 (0.69)	-2.78 (0.91)	2.45 (0.91)
Leg Length	0.19 (0.05)	0.19 (0.02)	0.19 (0.02)	0.18 (0.04)	0.17 (0.02)	0.18 (0.02)

Values are presented as mean (standard deviation)

Table 3. Associations between Biomechanical Variables and Modules

A. Module 1

	Bin 1	**Bin 2**
AP Ground Reaction Force	1.68 (0.81) p=0.039	5.49 (1.54) p<0.0001*
Vertical Ground Reaction Force	68.38 (6.05) p<0.0001*	39.56 (7.66) p<0.0001*

B. Module 2

	Bin 3	**Bin 4**
AP Ground Reaction Force	16.11 (1.13) p<0.0001*	38.17 (1.48) p<0.0001*
Vertical Ground Reaction Force	223.97 (6.96) p<0.0001*	286.39 (8.12) p<0.0001*
AP Leg Angle	-2.35 (0.16) p<0.0001*	-8.43 (0.32) p<0.0001*

C. Module 3

	Bin 4	**Bin 5**	**Bin 6**
AP Ground Reaction Force	-6.0 (1.44) p<0.0001*	.	.
AP Leg Angle	.	-1.45 (0.28) p<0.0001*	0.73 (0.38) p=0.053
Leg Length	.	0.01 (0.01) p=0.52	0.03 (0.01) p=0.0017*

D. Module 4

	Bin 1	**Bin 6**
AP Ground Reaction Force	0.95 (0.59) p=0.106	.
Vertical Ground Reaction Force	37.48 (4.43) p<0.0001*	.
AP Leg Angle	.	-2.46 (0.39) p<0.0001*

Values presented are parameter estimates with p-values

* indicates significant finding, $p<0.003$

4.3 PART III MANUSCRIPT

tDCS electrode montages may differentially impact variables of walking performance in individuals post-stroke: a preliminary study

The following manuscript was accepted for publication in the *Journal of Clinical Neurophysiology* in March of 2021. The text below is from the final draft accepted prior to publication.

tDCS electrode montages may differentially impact variables of walking performance in individuals post-stroke: a preliminary study

Bryant A. Seamon, PT, DPT [a, b], Mark G. Bowden, PT, PhD [a, b, c *], John H. Kindred, PhD [a, c], Aaron E. Embry, PT, DPT [a, b, c], Steven A. Kautz, PhD [a, b, c]

[a] Ralph H. Johnson VA Medical Center, 109 Bee St, Charleston, SC 29401, USA

[b] Department of Health Sciences and Research, College of Health Professions, Medical University of South Carolina, 77 President Street, Charleston, SC 29425, USA

[c] Division of Physical Therapy, College of Health Professions, Medical University of South Carolina, 151-B Rutledge Avenue, Charleston, SC 29425, USA

*** Corresponding Author:**

Dr. Mark G. Bowden

151-B Rutledge Avenue, MSC 962

Charleston, SC 29425

843-792-5036

bowdenm@musc.edu

Running Title: tDCS and post-stroke walking performance

Conflict of Interest:

The authors declare no competing financial interests.

Funding:

Funding for this project was provided by the American Heart Association (12IRG9430057), the VA Office of Research and Development (ORD), with additional support from the VA/ORD Rehabilitation R&D Service (1IK6RX003075 and 1I01RX001935) and from the National Institutes of Health (NIH P20 GM109040 and NIH P2C HD086844). This publication was partially supported by a Promotion of Doctoral Studies Level I Scholarship from the Foundation for Physical Therapy Research.

Any opinions expressed in this publication are those of the authors and do not necessarily reflect the view of the U.S. Department of Veteran Affairs or the NIH.

Abstract:

Background:

Transcranial direct current stimulation (tDCS) has mixed effects on walking performance in individuals post-stroke. This is likely the result of variations in tDCS electrode montages and individualized responses. The purpose of this study was to quantify the effects of a single session of tDCS using various electrode montages on post-stroke walking performance.

Methods:

Individuals with chronic stroke (n=16) participated in a double-blind, randomized cross-over study with sham stimulation and three tDCS electrode montages. Gait speed, paretic step ratio and paretic propulsion were assessed pre- and post-stimulation at self-selected and fastest comfortable speeds. Changes in muscle activation patterns with self-selected walking were quantified by the number of modules derived from non-negative matrix factorization of EMG signals for hypothesis generation.

Results:

There was no significant effect of active stimulation montages compared to sham. Comparisons between each participant's best response to tDCS and sham show personalized tDCS may have a positive effect on fastest comfortable overground gait speed (p=0.084), paretic step ratio (p=0.095) and paretic propulsion (p=0.090), and self-selected paretic step ratio (p=0.012). Participants

with 2 or 3 modules at baseline increased module number in response to the all experimental montages and sham, but responses were highly variable.

Conclusions:

A single session of tDCS may affect clinical and biomechanical walking performance, but effects appear to be dependent on individual response variability to different electrode montages. Our findings are consistent with responses to various tDCS electrode montages being the result of underlying neuropathology and we recommend examining how individual factors affect responses to tDCS.

Key Words: (6 allowed)

Brain stimulation, electromyography, stroke, walking, biomechanics, rehabilitation

Individuals post-stroke commonly report impaired walking performance that is associated with the severity of central nervous system damage.(1) Transcranial direct current stimulation (tDCS) holds promise as a potential therapeutic adjuvant capable of modifying or modulating the central nervous system and may be able to augment standard rehabilitation strategies for recovery of walking function.(2, 3) tDCS is a non-invasive brain stimulation technique that uses a small electrical current to modulate cortical activity and is simple to administer, low-cost, and low-risk.(4, 5) tDCS neuromodulation has been shown to up or down regulate cortical excitability and effect interhemispheric imbalances that can result from a stroke(6, 7) including deeper structures like the leg area of the motor cortex in healthy individuals(2, 8, 9) and persons post-stroke.(3) Regulating cortical excitability or interhemispheric imbalance has been hypothesized as a mechanism for improving motor function(6, 10) (primarily in the upper extremity(7, 11-13)) and aphasia(14, 15). Treating interhemispheric imbalance may have a larger influence on walking performance given the bilateral task requires coordination between hemispheres(16) although this model has not been conclusively translated to the lower extremity.

Despite the promise of tDCS acting on impaired cortical activity to improve motor function the effects are often small bringing into question the therapeutic utility of tDCS in clinical practice especially for walking rehabilitation (i.e. gait speed and endurance).(16-18) One hypothesized reason for this is the limited evidence to guide personalization of tDCS prescription to combat the inherent

heterogeneity of the pathology.(18) Few experiments examine the effect of varying available parameters such as; treatment frequency, duration and length, current strength, and electrode montage.(14, 17-19)

Electrode montage is of specific therapeutic interest in post-stroke rehabilitation for walking because electrode placement informed by neuroanatomical and physiological pathology has had a positive effect on clinical measures of upper extremity and hand function.(7, 11-13) Improvements are seen with either excitatory (anode applied over ipsilesional M1)(7) or inhibitory stimulation (cathode is applied over contralesional M1)(11) when compared to sham.(12, 13) Studies examining the effects of dual montages (anode over ipsilesional M1 and cathode over contralesional M1 simultaneously) had positive findings on measures of walking function (decreased Timed Up and Go times(20), increased 6-minute walk test distance(21), and increased paretic power(22)) in response to a single session of tDCS. Yet, it remains unknown how dual montages compare to single anode or cathode configurations against sham stimulation on walking performance.(17)

The purpose of this study was to quantify the effects of a single treatment of tDCS delivered during treadmill walking on gait speed, paretic step ratio and paretic propulsion in individuals post-stroke and to compare three electrode montages to sham stimulation: 1.) Excitatory (anode over ipsilesional M1), 2.) Inhibitory (cathode over contralesional M1, and 3.) Dual (an anode over ipsilesional M1, and a cathode over contralesional M1). We hypothesized the dual electrode montage would have the largest effect on measures of walking

performance as walking is a bilateral, coordinated activity resulting from restored interhemispheric balance. In an exploratory analysis, we examined the effect of tDCS stimulation and electrode montage on muscle activation patterns in individuals post-stroke during treadmill walking.

Materials and Methods:

Participants: Eighteen individuals with chronic stroke were enrolled and sixteen completed all study procedures. One participant dropped-out and one did not meet the inclusion criteria. Demographic data are presented in Table 1.

(INSERT TABLE 1 HERE)

Inclusion criteria were: 1) age 18 to 85 years old; 2) at least six-months post-stroke; 3) residual lower extremity paresis (Fugl-Meyer Lower Extremity motor score <34); 4) ability to walk independently at least 10 feet; 5) self-selected 10-meter gait speed < 0.8 m/s (at time of consent); and 6) provision of informed consent. Participants were excluded for: 1) significant musculoskeletal problems limiting hip and knee extension or ankle plantarflexion to neutral joint positions; 2) self-reported history of unstable cardiovascular disease or severe osteoporosis, or 3) pregnancy. Screening, testing and tDCS interventions were completed by a team of licensed physical therapists and associated study staff. All participants signed a written informed consent form approved by the Institutional Review

Board at the Medical University of South Carolina and conformed to the Declaration of Helsinki.

Experimental Procedure: We used a double-blind, randomized, cross-over experimental design. A timeline of procedures is presented in Figure 1. Participants were screened and completed a clinical assessment, which included the lower extremity motor portion of the Fugl-Meyer(23), Berg Balance Test(24), and Dynamic Gait Index(25). Participants then completed three single sessions of tDCS and one sham stimulation session. Sessions were blocked randomized to control for order effects and separated by a minimum 48-hour washout period. Participants completed pre- and post-stimulation testing for each session.

(INSERT FIGURE 1 HERE)

Transcranial direct current stimulation: tDCS was delivered using an EMPI unit (Chattanooga; Hixson, TN) and 1.75 cm2 sponges prepped with 0.9% saline solution. This created a current density of 0.1 mA/cm2 consistent with recommendations.(26) We informed participants they may feel a slight tingling sensation that should subside within approximately 60 seconds. Stimulation was ramped up to 2mA x at a dose rate of 40mA/min for a total of 20 minutes. Experimental conditions administered tDCS with one of the following electrode montages illustrated in Figure 2: 1) excitatory (anode over target ipsilesional leg M1 area), 2) inhibitory (cathode over target contralesional leg M1 area) or 3) dual

(both excitatory and inhibitory montages applied simultaneously to target both leg M1 areas). Reference pads for the excitatory and inhibitory montages were placed on the ipsilateral shoulder. Modeling work has shown that this type of extracephalic pad placement creates a more focal concentration of current under the electrode, increasing penetration depth.(27) Studies with healthy individuals suggest that extracephalic pad placement has a greater effect on cortical excitability (9) and neuromotor output (28) compared to cephalic placement for deeper M1 areas of the leg. The dual montage used two EMPI units to deliver simultaneous active anodal stimulation over ipsilesional M1 and cathodal stimulation over contralesional M1. Reference pads for both units were placed on the respective ipsilateral shoulder. In each montage, M1 location was determined in a forward seated position approximately 1cm lateral to the vertex and 1cm posterior to a hypothetical line between the tragi creating a 2cm gap between the cephalic pads in the dual montage configuration.

(INSERT FIGURE 2 HERE)

Stimulation parameters were set prior to each session by an unblinded investigator and participants were fitted with two EMPI units using pad placement described in the dual experimental setup to maintain participant and staff blinding to active stimulation parameters. Sham stimulation was done by turning on the EMPI units to apply 30 seconds of stimulation before manually turning the units off by an unblinded investigator (per published guidelines).(29) Participants

received tDCS or sham stimulation for 5 minutes in a seated position and then continued receiving stimulation while walking for 15-minutes on a treadmill. Participants walked at their fastest comfortable speed on the treadmill to provide an adequate training stimulus. Faster walking speeds are commonly used in rehabilitation programs having been shown to have immediate and long-term effects on walking performance for persons post-stroke.(30, 31) Walking was paused every 5-minutes to assess cardiovascular response to exercise (i.e. blood pressure, heart rate, and activity tolerance). Walking immediately resumed unless continuation was contraindicated for safety. A ceiling harness system without body-weight support was used to prevent falls or injury. Verbal cues to alter gait pattern were not provided. Minimal physical assistance was provided to prevent tripping or interruptions in walking and not given during data collection trials.

Data analysis: GaitRite (CIR Systems, Inc.; Franklin, NJ) data was used to calculate self-selected and fastest comfortable overground gait speeds. Participants walked over a 24-foot GaitRite for one trial at their self-selected speed and three trials at their fastest comfortable speed during pre- and post-testing. Participants also walked for three, 30-second trials on a split belt instrumented treadmill (Bertec; Columbus, OH) at self-selected and fastest comfortable speeds, which did not have to match overground speeds. Ground reaction force (GRF) data was sampled at 1000 Hz to derive paretic step ratio and paretic propulsion using methods previously described by our lab.(32)

Paretic propulsion was calculated by dividing the positive anterior impulse of the paretic leg by the anterior impulse of both legs combined.(32) Paretic step ratio was calculated from the percentage of stride length performed by the paretic step.(33) Paretic step ratio and paretic propulsion were expressed as the absolute value of deviation from symmetry (0.5).

Muscle coordination patterns were quantified for each participant during self-selected treadmill walking by extracting modules using a non-negative matrix factorization (NNMF) algorithm.(34, 35) Surface EMG was recorded at 2000 Hz with bipolar pre-amplified electrodes (Motion Lab Systems; Baton Rouge, LN, USA) at the following eight muscle locations bilaterally: tibialis anterior, soleus, medial gastrocnemius, vastus medialis, rectus femoris, medial hamstrings, lateral hamstrings and gluteus medius.(36) Specific post-processing of EMG signals and selection of modules for the paretic leg can be found in Clark et. al., 2010.(37)

Statistical analyses were conducted using SAS version 9.4 (SAS Institute Inc. Cary, NC) on change scores for gait speed, paretic step ratio and paretic propulsion at self-selected and fastest comfortable speeds. A one-way ANOVA (or Kruskal-Wallis Test for non-parametric data) examined the main effect of active stimulation across all three montages compared to sham for each variable of walking performance. Post-hoc testing with multiple t-tests (or Wilcoxon Rank-Sum tests) were performed to compare the effects of each electrode montage. Corrections for multiple comparisons were not performed for this preliminary

study. Changes in module number are reported observationally to generate hypotheses for future studies.

Results:

Walking performance: No significant main effect of active stimulation montage was observed for overground walking speed at self-selected (ANOVA; F=0.44, p=0.723, df=3) or fastest comfortable speed (Kruskal-Wallis; X^2=2.419, p=0.490, df=3). Additionally, no significant effect was found for paretic step ratio at self-selected (Kruskal-Wallis; X^2=3.013, p=0.389, df=3) or fastest comfortable speeds (Kruskal-Wallis; X^2=1.357, p=0.716, df=3) or paretic propulsion at self-selected (ANOVA; F=0.31, p=0.819, df=3) or fastest comfortable speeds (Kruskal-Wallis; X^2=0.749, p=0.862, df=3). Group descriptive statistics for each variable are presented in Table 2.

(INSERT TABLE 2 HERE)

Visual inspection of the data showed a high degree of overall response variability to each electrode montage for each measured variable. Table 3 shows each participant's best response to an electrode montage for each variable.

(INSERT TABLE 3 HERE)

To examine the possibility participants may exhibit preferential responses to specific montages, we compared each participant's best response to stimulation with sham for each variable using t-tests (or Wilcoxon Rank-Sum test for non-parametric data). We also accepted a higher false positive rate of 10% (alpha=0.1) for generating exploratory hypotheses. We observed a significant difference and improved fastest comfortable overground gait speed (mean difference=0.06 m/s, 95% CI [-0.008 – 0.12], t-test p=0.084) but not self-selected (mean difference=0.05, 95% CI [-0.036 – 0.137], t-test p=0.242). There was a significant effect for improvement in paretic step ratio (median=0.017, IQR [0.023], Wilcoxon; p=0.012) and paretic propulsion (mean difference=0.035, 95% CI [-0.006 – 0.077], t-test p=0.090) at self-selected speeds and paretic step ratio (mean difference=0.01, 95% CI [-0.002 – 0.02], t-test p=0.095) at fastest comfortable speeds. No differences were found for paretic propulsion (median=0.002, IQR [0.068], Wilcoxon; p=0.645) at fastest comfortable speeds.

(INSERT TABLE 4 HERE)

Muscle activation patterns: In our sample, 6 participants used 2 modules, 6 participants used 3 modules and 4 participants used 4 modules for self-selected comfortable treadmill walking at baseline. The change in module number for the participants in response to tDCS with each electrode montage and sham are presented in Table 5.

(INSERT TABLE 5 HERE)

Only one of the 6 participants with 2 modules did not change their module number after tDCS or sham stimulation. Four of the remaining 5 participants increased module number with the excitatory, inhibitory and/or dual electrode montage and 4 of the 6 participants increased module number with sham stimulation. One 2 module participant increased module number with all electrode montages and sham stimulation.

One of the 6 participants with 3 modules did not change their module number after tDCS or sham stimulation. Two participants did not change their module number in response to any tDCS condition but did to sham stimulation. One of these individuals improved module number and the other decreased. One participant increased module number after the inhibitory montage, one in response to the excitatory montage and dual montage and one in response to the inhibitory and dual montage. However, all three of these participants also increased module number in response to sham stimulation.

Although module number cannot increase from 4, we observed two instances where module number decreased in those with 4 modules at baseline. One participant reduced module number to 3 in response to the dual montage and sham stimulation, the other reduced module number to 3 in response to the inhibitory montage.

A small positive association was found between a change in module number and change in paretic propulsion symmetry ($r=0.29$; $p=0.0251$) across all

conditions; including sham stimulation. However, this association was not present within experimental and sham conditions separately. There were no associations between changes in module number and changes in gait speed, paretic step ratio or paretic propulsion at self-selected walking speeds when examining all tDCS conditions or within each experimental montage group.

Discussion:

Our aim was to compare the immediate effects of three tDCS electrode montages and sham stimulation on post-stroke walking performance. We used a double blind, placebo controlled, randomized cross-over design to evaluate changes in gait speed, paretic step ratio and paretic propulsion. We found no group main effects for any of the electrode montages compared to sham stimulation on walking performance immediately following one session of tDCS, which was inconsistent with our hypothesis.

Our lack of a single session effect on post-stroke walking performance is comparable with findings from other experiments.(20-22) The immediate effect of tDCS with an excitatory or dual electrode montage has not had a significant effect on walking performance with the exception of the 6-minute walk and Timed-up and Go tests, compared to sham.(20-22) The authors hypothesized the large variation in participant response to tDCS likely caused the negative findings(20-22) and this heterogeneity continues to be a key challenge in post-stroke tDCS neuromodulation research.(18) In an exploratory attempt to address variation in our sample, we tested the main effect of tDCS by comparing each

participant's best response of the electrode montages to sham stimulation. This is based on the assumption that individuals may respond to different montages based on unknown characteristics likely arising from the variety of motor network impairments result from lesion location, size and cortical reorganization.(22, 38, 39) A higher false positive rate, alpha=0.1, was accepted to generate hypotheses for future research. We found tDCS stimulation had a positive effect on fastest comfortable gait speed, paretic step ratio during self-selected and fastest comfortable speeds, and paretic propulsion at self-selected speeds. We also observed the "best montage" often varied for each measure of walking performance (Table 3) within individuals. The lack of a specific pattern lends support to the idea that "one-size does not fit all" in tDCS prescription.(40) Investigators should consider that different electrode montages may impact different features of walking performance on an individual level based on gait speed, clinical features like stroke chronicity(40), or presence, type and degree of interhemispheric imbalance. It is important to note we did not investigate the reproducibility and robustness of our observed effects (i.e., would the electrode configuration that shows the best result be the same under a second test). Nor did we have a large enough sample to examine whether clinical performance or baseline walking performance could predict tDCS response. Previous investigations have demonstrated the reproducibility and benefit of personalized tDCS electrode montage in language rehabilitation.(41, 42) Our results suggest this effect should also be tested in walking rehabilitation and whether markers of clinical or baseline walking performance predict participant response.

Investigators should screen individuals for inter-hemispheric imbalances using TMS in addition to prioritizing assessment of neurophysiological effects from tDCS to establish associations between stimulation and neural pathophysiological changes.

Our exploratory comparison of the effect of tDCS on muscle activation patterns offers one potential window into the mechanisms by which different tDCS electrode configurations may influence walking. Many participants with more severe impairment (ie, 2 or 3 modules) were able to move away from mass flexion and extension muscle activation patterns in response to tDCS. We hypothesize that tDCS can modulate the cortex to enhance voluntary muscle activity during walking for some individuals. This is supported by evidence that tDCS can increase force production in lower extremity muscles and helps explain the association we found between paretic propulsion and improved muscle activation patterns.(22, 29, 43) The individual variability we observed could be explained by the fact that walking is also influenced by subcortical structures. Thus, individuals with more severe cortical impairments may have a greater response to tDCS. However, like our findings related to the best response condition, we cannot be certain that we have captured a true effect since we did not test the reproducibility of our findings and we have a very small sample for each module number at baseline limiting our statistical power. We recommend that more research examine the effects of tDCS on muscle activation patterns since they can provide a mechanistic understanding for biomechanical changes in task performance.(44)

Finally, there are a few methodological choices in our design that may have impacted the results. First, we are unable to know whether current was shunted during the dual montage experimental condition. It is possible that current may have crossed between the two cephalic pads creating a different stimulation environment than hypothesized. Second, our washout period of 48 hours may not have been sufficiently long enough. There is precedence for a 48-hour washout in post-stroke tDCS literature(3, 11) but no formal investigation into the optimal length of time has been done and recent recommendations call for a minimum 1-week period(26). Lastly, the robust response to sham for many subjects suggests that twenty minutes of walking may provide a neuromodulating effect as potent as a single tDCS session and appears to have been an active ingredient in our experiment. Ojardias et. al.(21) saw a similar response to walking on walking performance in individuals post-stroke after examining a single session of tDCS and there is recent evidence to support that moderate intensity aerobic activity can increase neurophysiological markers of corticospinal excitability.(45-47) An intriguing question is whether the increase in corticospinal excitability that accompanies tDCS has a similar mechanism to the increase with walking practice, and whether the effects are additive when the two stimuli are combined. Future research should be designed to further investigate the effects of walking practice on excitability, with and without concomitant tDCS.

In summary, our observation that individuals may have an optimal tDCS electrode montage to elicit improvements in walking performance is perhaps the most important finding of this study. The possibility that individuals post-stroke

likely need personalized stimulation parameters has important implications for hypothesis generation and future tDCS studies attempting to optimize tDCS prescription. It is imperative for investigators to employ research methods to best understand how electrode placement will impact walking performance considering clinical presentation corticomotor response, neuroanatomy and tractography to address heterogeneity in participant response to tDCS.

REFERENCES:

1. Mayo NE, Wood-Dauphinee S, Cote R, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. Archives of physical medicine and rehabilitation. 2002 Aug;83(8):1035-42.

2. Madhavan S, Stinear JW. Focal and bi-directional modulation of lower limb motor cortex using anodal transcranial direct current stimulation. Brain stimulation. 2010 Jan;3(1):42.

3. Madhavan S, Weber KA, 2nd, Stinear JW. Non-invasive brain stimulation enhances fine motor control of the hemiparetic ankle: implications for rehabilitation. Exp Brain Res. 2011 Mar;209(1):9-17.

4. Schlaug G, Renga V, Nair D. Transcranial direct current stimulation in stroke recovery. Arch Neurol. 2008 Dec;65(12):1571-6.

5. Woods AJ, Antal A, Bikson M, Boggio PS, Brunoni AR, Celnik P, et al. A technical guide to tDCS, and related non-invasive brain stimulation tools. Clin Neurophysiol. 2016 Feb;127(2):1031-48.

6. Nowak DA, Grefkes C, Ameli M, Fink GR. Interhemispheric competition after stroke: brain stimulation to enhance recovery of function of the affected hand. Neurorehabil Neural Repair. 2009 Sep;23(7):641-56.

7. Hummel F, Cohen LG. Improvement of motor function with noninvasive cortical stimulation in a patient with chronic stroke. Neurorehabilitation and neural repair. 2005 Mar;19(1):14-9.

8. Jeffery DT, Norton JA, Roy FD, Gorassini MA. Effects of transcranial direct current stimulation on the excitability of the leg motor cortex. Experimental Brain Research. 2007 2007/09/01;182(2):281-7.

9. Tatemoto T, Yamaguchi T, Otaka Y, Kondo K, Tanaka S, editors. Anodal Transcranial Direct Current Stimulation over the Lower Limb Motor Cortex Increases the Cortical Excitability with Extracephalic Reference Electrodes. 2013; Berlin, Heidelberg: Springer Berlin Heidelberg.

10. Reis J, Fritsch B. Modulation of motor performance and motor learning by transcranial direct current stimulation. Current opinion in neurology. 2011 Dec;24(6):590-6.

11. Fregni F, Boggio PS, Mansur CG, Wagner T, Ferreira MJ, Lima MC, et al. Transcranial direct current stimulation of the unaffected hemisphere in stroke patients. Neuroreport. 2005 Sep 28;16(14):1551-5.

12. Boggio PS, Nunes A, Rigonatti SP, Nitsche MA, Pascual-Leone A, Fregni F. Repeated sessions of noninvasive brain DC stimulation is associated with motor function improvement in stroke patients. Restorative neurology and neuroscience. 2007;25(2):123-9.

13. Kim DY, Lim JY, Kang EK, You DS, Oh MK, Oh BM, et al. Effect of transcranial direct current stimulation on motor recovery in patients with subacute stroke. American journal of physical medicine & rehabilitation / Association of Academic Physiatrists. 2010 Nov;89(11):879-86.

14. Fridriksson J, Basilakos A, Stark BC, Rorden C, Elm J, Gottfried M, et al. Transcranial direct current stimulation to treat aphasia: Longitudinal analysis of a randomized controlled trial. Brain stimulation. 2019 Jan - Feb;12(1):190-1.

15. Fridriksson J, Rorden C, Elm J, Sen S, George MS, Bonilha L. Transcranial Direct Current Stimulation vs Sham Stimulation to Treat Aphasia After Stroke: A Randomized Clinical Trial. JAMA neurology. 2018 Dec 1;75(12):1470-6.

16. Li Y, Fan J, Yang J, He C, Li S. Effects of transcranial direct current stimulation on walking ability after stroke: A systematic review and meta-analysis. Restorative neurology and neuroscience. 2018;36(1):59-71.

17. Lefaucheur J-P, Antal A, Ayache SS, Benninger DH, Brunelin J, Cogiamanian F, et al. Evidence-based guidelines on the therapeutic use of transcranial direct current stimulation (tDCS). Clinical Neurophysiology. 2017 2017/01/01/;128(1):56-92.

18. Vaz PG, Salazar A, Stein C, Marchese RR, Lukrafka JL, Plentz RDM, et al. Noninvasive brain stimulation combined with other therapies improves gait speed after stroke: a systematic review and meta-analysis. Top Stroke Rehabil. 2019 Apr;26(3):201-13.

19. Lefebvre S, Liew SL. Anatomical Parameters of tDCS to Modulate the Motor System after Stroke: A Review. Frontiers in neurology. 2017;8:29.

20. Tahtis V, Kaski D, Seemungal BM. The effect of single session bi-cephalic transcranial direct current stimulation on gait performance in sub-acute stroke: A pilot study. Restorative neurology and neuroscience. 2014;32(4):527-32.

21. Ojardias E, Aze OD, Luneau D, Mednieks J, Condemine A, Rimaud D, et al. The Effects of Anodal Transcranial Direct Current Stimulation on the Walking Performance of Chronic Hemiplegic Patients. Neuromodulation : journal of the International Neuromodulation Society. 2019 May 23.

22. van Asseldonk EH, Boonstra TA. Transcranial Direct Current Stimulation of the Leg Motor Cortex Enhances Coordinated Motor Output During Walking With a Large Inter-Individual Variability. Brain stimulation. 2016 Mar-Apr;9(2):182-90.

23. Fugl-Meyer A, Jaasko L, Leyman I, Olsson S, Seglind S. The post-stroke hemiplegic patient: a method of evaluation of physical performance. Scandinavian journal of rehabilitation medicine. 1975;7:13-31.

24. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Canadian journal of public health Revue canadienne de sante publique. 1992 Jul-Aug;83 Suppl 2:S7-11.

25. Shumway-Cook A, Woollacott M. Motor Control: Theory and Practical Applications. Philadelphia: Lippincott Williams and Wilkins; 2001.

26. Thair H, Holloway AL, Newport R, Smith AD. Transcranial Direct Current Stimulation (tDCS): A Beginner's Guide for Design and Implementation. Frontiers in neuroscience. 2017;11:641.

27. Noetscher GM, Yanamadala J, Makarov SN, Pascual-Leone A. Comparison of cephalic and extracephalic montages for transcranial direct current stimulation--a numerical study. IEEE transactions on bio-medical engineering. 2014 Sep;61(9):2488-98.

28. Angius L, Pageaux B, Hopker J, Marcora SM, Mauger AR. Transcranial direct current stimulation improves isometric time to exhaustion of the knee extensors. Neuroscience. 2016 Dec 17;339:363-75.

29. Tanaka S, Takeda K, Otaka Y, Kita K, Osu R, Honda M, et al. Single session of transcranial direct current stimulation transiently increases knee extensor force in patients with hemiparetic stroke. Neurorehabilitation and neural repair. 2011 Jul-Aug;25(6):565-9.

30. Lamontagne A, Fung J. Faster is better: implications for speed-intensive gait training after stroke. Stroke. 2004 Nov;35(11):2543-8.

31. Sullivan KJ, Knowlton BJ, Dobkin BH. Step training with body weight support: effect of treadmill speed and practice paradigms on poststroke locomotor recovery. Archives of physical medicine and rehabilitation. 2002 May;83(5):683-91.

32. Bowden MG, Balasubramanian CK, Neptune RR, Kautz SA. Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking. Stroke. 2006 Mar;37(3):872-6.

33. Balasubramanian CK, Bowden MG, Neptune RR, Kautz SA. Relationship between step length asymmetry and walking performance in subjects with chronic hemiparesis. Archives of physical medicine and rehabilitation. 2007 Jan;88(1):43-9.

34. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature. 1999 Oct 21;401(6755):788-91.

35. Ting LH, Macpherson JM. A limited set of muscle synergies for force control during a postural task. J Neurophysiol. 2005 Jan;93(1):609-13.

36. Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. Development of recommendations for SEMG sensors and sensor placement procedures. J Electromyogr Kinesiol. 2000 Oct;10(5):361-74.

37. Clark DJ, Subramanian S, Neptune RR, SA K, editors. Fewer basic activation patterns account for lower extremity EMG during walking in adults post-stroke compared to healthy controls. Neural Control of Movement Annual Meeting; 2008; Naples, FL.

38. Jones TA. Motor compensation and its effects on neural reorganization after stroke. Nature reviews Neuroscience. 2017 May;18(5):267-80.

39. Lefebvre S, Dricot L, Laloux P, Gradkowski W, Desfontaines P, Evrard F, et al. Neural substrates underlying stimulation-enhanced motor skill learning after stroke. Brain : a journal of neurology. 2015 Jan;138(Pt 1):149-63.

40. McCambridge AB, Stinear JW, Byblow WD. Revisiting interhemispheric imbalance in chronic stroke: A tDCS study. Clin Neurophysiol. 2018 Jan;129(1):42-50.

41. Shah-Basak PP, Norise C, Garcia G, Torres J, Faseyitan O, Hamilton RH. Individualized treatment with transcranial direct current stimulation in patients with chronic non-fluent aphasia due to stroke. Front Hum Neurosci. 2015;9:201.

42. Sebastian R, Tsapkini K, Tippett DC. Transcranial direct current stimulation in post stroke aphasia and primary progressive aphasia: Current

knowledge and future clinical applications. NeuroRehabilitation. 2016;39(1):141-52.

43. Sohn MK, Jee SJ, Kim YW. Effect of transcranial direct current stimulation on postural stability and lower extremity strength in hemiplegic stroke patients. Annals of rehabilitation medicine. 2013 Dec;37(6):759-65.

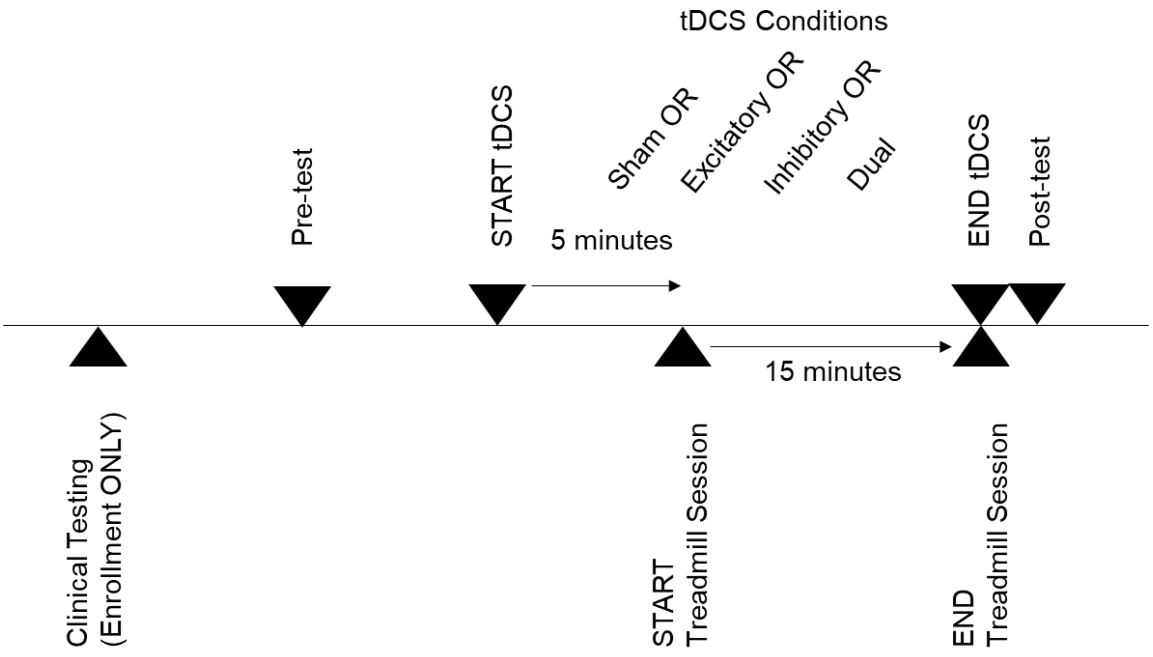
44. Bowden MG, Clark DJ, Kautz SA. Evaluation of abnormal synergy patterns poststroke: relationship of the Fugl-Meyer Assessment to hemiparetic locomotion. Neurorehabil Neural Repair. 2010 May;24(4):328-37.

45. Garnier YM, Lepers R, Stapley PJ, Papaxanthis C, Paizis C. Changes in cortico-spinal excitability following uphill versus downhill treadmill exercise. Behavioural brain research. 2017 Jan 15;317:242-50.

46. Bonnard M, Camus M, Coyle T, Pailhous J. Task-induced modulation of motor evoked potentials in upper-leg muscles during human gait: a TMS study. The European journal of neuroscience. 2002 Dec;16(11):2225-30.

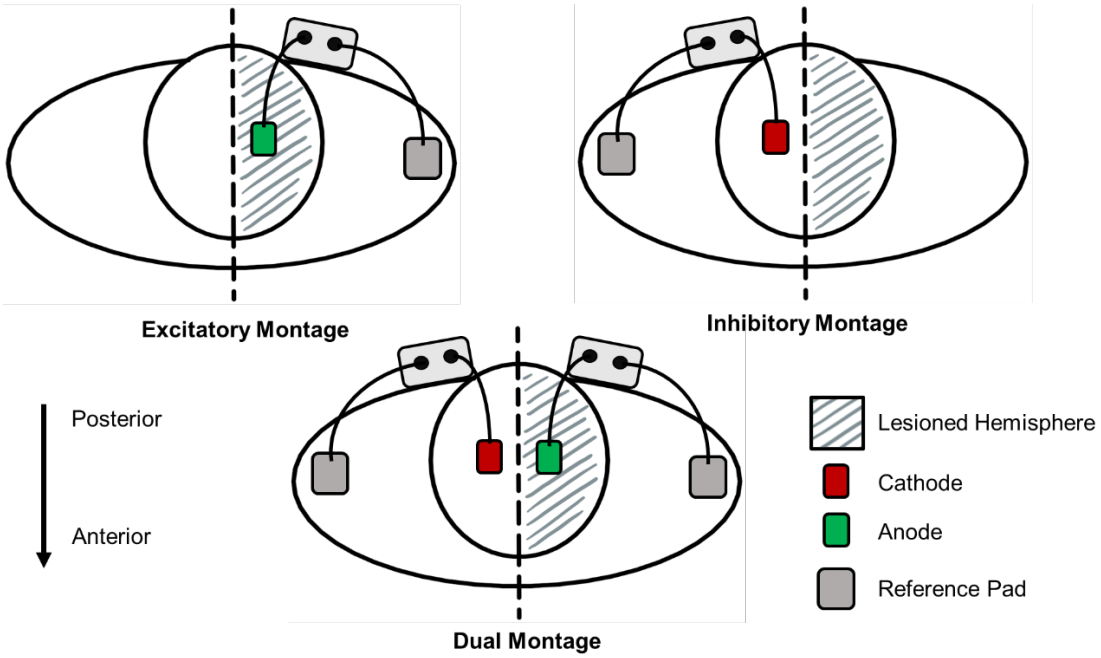
47. Yamazaki Y, Sato D, Yamashiro K, Nakano S, Onishi H, Maruyama A. Acute Low-Intensity Aerobic Exercise Modulates Intracortical Inhibitory and Excitatory Circuits in an Exercised and a Non-exercised Muscle in the Primary Motor Cortex. Frontiers in physiology. 2019;10:1361.

Figure 1: Timeline of Experimental Procedures



Participants completed clinical testing during enrollment. For each of the four experimental sessions (3 active and one sham), participants completed pre-testing followed by active tDCS or sham stimulation with treadmill walking and concluded with post-testing.

Figure 2: Diagram of experimental tDCS montages



Excitatory Montage: Anode pad was placed over the target ipsilesional M1 leg area, reference pad was placed on the ipsilateral shoulder

Inhibitory Montage: Cathode pad was placed over the target contralesional M1 leg area, reference pad was placed on the ipsilateral shoulder.

Dual Montage: Combination of the Excitatory and Inhibitory montages using 2 tDCS units with one delivering the excitatory (anode placed over the target ipsilesional M1 leg area) and the other inhibitory (cathode pad was placed over the target contralesional M1 leg area) currents.

In each montage the cephalic pad was placed 1cm lateral to the vertex and 1 cm posterior to an imaginary line between the tragi. This created a 2cm gap between cephalic electrodes in the Dual montage.

Table 1: Participant Demographics

	n	Descriptive Statistics
Age (years) mean (SD) [range]	16	59 (11.5) [30-77]
Sex (male) frequency (%)	16	11/16 (68.8%)
Hemiparetic side (right) frequency (%)	16	9/16 (56.3%)
Chronicity (months) mean (SD) [range]	16	54.56 (76.5) [10-325]
FM-Total LE mean (SD) [range]	16	23.94 (5.78) [13-32]
FM-Synergy mean (SD) [range]	16	15.56 (4.15) [8-21]
Dynamic Gait Index mean (SD) [range]	15	16 (4.55) [7-22]
Berg Balance Scale mean (SD) [range]	11	46.91 (8.94) [25-55]
Gait Speed (self-selected) (m/s) mean (SD) [range]	16	0.82 (0.33) [0.23-1.43]
Paretic Step Ratio mean (SD) [range]	16	0.51(0.06) [0.37-0.63]
Modules mean (SD) [range]	16	2.94 (0.85) [2-4]

Table 2: Descriptive statistics for change scores as a result of each tDCS electrode montage and sham stimulation

		tDCS Experimental Condition			
		Sham	**Excitatory**	**Inhibitory**	**Dual**
Self-Selected	**Gait Speed m/s**	0.086 (0.132)	0.043 (0.128)	0.068 (0.092)	0.085 (0.109)
	Paretic Step Ratio	0.005 (0.012)	0.009 (0.022)	0.003 (0.032) ǂ	0.002 (0.032)
	Paretic Propulsion	0.015 (0.061)	0.006 (0.069)	0.027 (0.043)	0.023 (0.074)
Fastest Comfortable	**Gait Speed m/s**	0.059 (0.051)	0.042 (0.134)	0.020 (0.081)	0.036 (0.100) ǂ
	Paretic Step Ratio	-0.003 (0.012)	-0.000 (0.017)	-0.005 (0.019)	-0.020 (0.0.28) ǂ
	Paretic Propulsion	0.012 (0.056)	-0.003 (0.041) ǂ	-0.007 (0.048) ǂ	0.005 (0.040)

Statistics are presented as mean (standard deviation)

ǂ Indicates median (interquartile range)

Gait speed was calculated from overground walking trials.

Paretic Step Ratio and Paretic Propulsion were calculated from treadmill walking trials.

Table 3: Individual variability in response to tDCS stimulation based on electrode montage

	Self-Selected Speed			Fastest Comfortable Speed		
Participant Number	**Gait Speed**	**Paretic Step Ratio**	**Paretic Propulsion**	**Gait Speed**	**Paretic Step Ratio**	**Paretic Propulsion**
1	2	1	3	1	2	2
2	2	3	1	2	1	3
3	2	1	2	1	1	3
4	3	1	2	3	2	1
6	3	2	1	3	3	1
7	3	3	3	1	1	1
8	2	2	2	2	2	1
9	3	3	2	3	3	3
10	3	3	2	3	1	1
11	3	1	×	1	3	×
12	3	2	2	1	1	3
13	1	1	3	2	3	1
15	1	1	1	3	1	3
16	3	2	1	2	1	1
17	1	1	2	1	2	2
18	3	1	3	1	1	2

Numeric codes for each variable indicate the montage (electrode placement and stimulation parameters) that elicited the best response for each participant.

1 = Excitatory montage (color = light orange)

2 = Inhibitory montage (color = light blue)

3 = Dual montage (color = white)

× Missing data due to poor GRF quality during data collection.

Table 4: Descriptive statistics for change scores comparing the effect of tDCS compared to sham stimulation using a participant's best response to each of the three electrode montages

		Sham	Pooled Best Response to tDCS
Self-Selected	**Gait Speed m/s**	0.086 (0.132)	0.136 (0.106)
	Paretic Step Ratio	0.005 (0.019) ǂ	0.017 (0.023) ǂ **
	Paretic Propulsion	0.015 (0.061)	0.051 (0.049) *
Fastest Comfortable	**Gait Speed m/s**	0.059 (0.051)	0.114 (0.111) *
	Paretic Step Ratio	-0.003 (0.012)	0.007 (0.019) *
	Paretic Propulsion	0.002 (0.068) ǂ	0.026 (0.053) ǂ

** Indicates statistical significance at $p<0.05$.

* Indicates statistical significance at $p<0.10$.

Statistics are presented as mean (standard deviation)

ǂ Indicates median (interquartile range)

Pooled Best Response was created by pooling each participant's best response to any of the tDCS electrode montages for a given variable.

Gait speed was calculated from overground walking trials.

Paretic Step Ratio and Paretic Propulsion were calculated from treadmill walking trials.

Table 5: Individual variability in muscle activation pattern response to tDCS stimulation

		Change in Module Number			
Participant Number	**Baseline Module Number**	**Excitatory**	**Inhibitory**	**Dual**	**Sham**
1	**2**	0	1	2	2
7	**2**	1	1	0	2
11	**2**	0	0	0	0
12	**2**	0	2	1	2
16	**2**	1	0	1	0
17	**2**	1	1	1	1
3	**3**	0	1	0	1
4	**3**	1	0	1	1
8	**3**	0	0	0	1
10	**3**	0	0	0	-1
15	**3**	0	1	1	1
2	**4**	0	0	0	0
6	**4**	0	0	0	×
9	**4**	0	×	-1	-1
13	**4**	0	0	0	0
18	**4**	0	-1	0	0

× Missing data due to poor EMG quality during data collection.

CHAPTER 5

CONCLUSION

5.1 CONCLUSION

Personalized rehabilitation holds the promise of improved health outcomes because it can account for individual differences between people (Denny & Collins, 2021). Addressing individual differences has been a prominent barrier for advances in stroke rehabilitation that has typically been relegated to a “one size fits all” approach because of the lack of evidence-based tools for personalized rehabilitation (Cramer et al., 2017). As stroke rehabilitation embraces personalized approaches clinicians and researchers should use and generate evidence that addresses individual variability to improve outcomes by providing the right treatment to the right person at the right time. The work in this dissertation provides a body of evidence for personalized stroke rehabilitation by addressing three research needs identified by StrokeNet for understanding individual variability.

Part I was designed to show the benefit of standardized outcome measurement on personalized stroke rehabilitation. Our work demonstrated how applications of item-psychometrics can give meaningfulness to standard outcome measurement scores. We used an ability map for the FGA to show how clinicians

can use information about item difficulty hierarchies and patient response patterns to inform treatment planning and plan of care development for an individual patient. We also used item-level psychometrics generated by a Rasch analysis to demonstrate how error is different for different portions of a measurement scale and how that information can be used to evaluate change more precisely at the individual-level.

Part II was an initial step towards identifying a biomarker for walking-specific motor control recovery. Biomarkers need to be linked with underlying mechanisms for pathophysiology. We showed how modules, a measure of muscle coordination (i.e., walking-specific motor control) during walking, is strongly associated with observable biomechanical measures that are related to key functional tasks for successful walking. Our findings suggests that modules may have an expressed biomechanical signature of that could be used as a biomarker for walking-specific motor control recovery. This is important for stroke rehabilitation because modules share a strong association with neural damage and differences in motor recovery among individuals.

Part III demonstrated that more emphasis should be placed on evaluating how manipulation of treatment parameters will influence individual treatment response. We showed that the type of tDCS montage may to be a source of the high individual variability in treatment response to non-invasive brain stimulation. Our findings illustrate the importance of examining treatment parameters to provide the best evidence for individualized prescription.

5.2 FUTURE DIRECTIONS

The collection of studies in this dissertation have provided a “road map” for future studies in the areas of standardized measurement, biomarker development, and evaluating variability in treatment response.

Many of the applications from item response theory have yet to be tested in a pragmatic way. We have demonstrated how ability maps can inform clinical decision making to elevate personalized rehabilitation, but future studies need to examine whether using ability maps have an impact on clinical outcomes. Future pragmatic trials could be designed to evaluate stroke rehabilitation outcomes between traditional practice and a cohort using ability maps. Alongside this study, investigators could survey physical therapists to determine whether ability maps addressed many of the previously reported barriers to standardized outcome assessment that we theorized the maps would. Based on our work in Part I, we make the argument that cMDCs will improve clinical care by preventing misclassification of patients’ change status. We also argue that cMDCs would reduce sample size needs and prevent Type II errors in research studies. However, more work is needed to determine whether both would be the case. We recommend using simulation studies to examine the differences between the MDC and cMDC on sample size and Type II error rates, as well as retrospectively examine change data from large trials to see what proportion of individuals would have been misclassified under the cMDC approach.

Biomarkers for stroke recovery hold a great deal of promise. However, it has been difficult to identify biomarkers for functional tasks, like walking. Our findings in Part II take an important step towards identifying a biomarker for stroke recovery of walking-specific motor control using modules and biomechanical variables. Our findings provided more evidence to strengthen the theory that modules share a mechanistic link with observable, measurable biomechanical variables during healthy walking. However, we limited ourselves to sagittal plane mechanics and found that anterior-posterior leg angle and leg length only partially captured swing mechanics. Future work should attempt to expand on the measures we used in our design to create a more comprehensive biomechanical signature for each module. Specific examples include expanding the analyses to include frontal plane biomechanical variables and measures of leg velocity or acceleration to better capture swing phase mechanics. Biomechanical variables have always held interest as a potential marker for stroke walking performance because deviations in movement patterns can often be observed and used as treatment targets in rehabilitation. However, there is not a well understood framework for linking biomechanics back to the level of stroke pathology that would help differentiate individuals for more personalized treatment. Modules are a potential way to solve this problem because they capture individual differences in motor recovery linked with mechanisms of underlying neural pathology. Once a comprehensive biomechanical signature is established for each module, future work that can explore module and biomechanical variable associations in stroke pathology. Specifically, studies can

examine how changes in modules post-stroke are related to changes in biomechanical signatures and whether altered signatures can be predictive of modules or walking-specific motor control. Lastly, future prospective studies can examine whether treatments designed to treat abnormal biomechanical signatures result in changes to modules.

Finally, there is still more work needed in the field of intervention prescription for stroke rehabilitation. Individual variability in treatment response continues to impede progress for tDCS to be used regularly in clinical practice. Our findings demonstrate that individual variability could be a result of limited knowledge about how manipulation of treatment parameters influence outcomes. Future studies should examine the effects of individual parameters on outcomes and attempt to replicate those effects by re-treating or testing individuals based on findings from a test period. For example, with respect to electrode montage in tDCS, future studies would have a second arm of the study where individuals received their “optimal” prescription and evaluate whether their “optimal” prescription reproduced their previous results. Secondly, our results point to the importance for intervention prescription to have a theoretical mechanistic link to pathology that can be tested. This is important because a common framework for relating individual stroke neuropathology to recommended interventions is still a bit disparate. We theorized pad placement montages in our design based on predicted neural damage, however, our study would have been strengthened by testing our underlying hypotheses for each montage. Measures of neural physiology and brain imaging could have been used to test whether our

proposed neuropathology was present and whether the tDCS prescription has an impact on the underlying neuropathology. Another potential direction is to examine the effect of tDCS on impairment measures that reflect changes in individual neuropathology. This can be particularly useful for clinicians because impairment measures may reveal a phenotypic expression of neuropathology that serves to identify treatment targets for a specific individual. Our study examined the effects of tDCS on modules which is an impairment measure for walking-specific motor control. Our results suggested that tDCS may have an acute effect on modules and that treatment response may be dependent on individual variability in walking-specific motor control impairment. However, the robustness of our findings needs to be investigated. Still, modules may serve as an impairment measure that can be used to understand variability between individuals and inform treatment prescription. Future studies should keep this conceptual model of approaching study design for interventions to strengthen our ability to differential between individuals and personalize treatment prescription.

REFERENCES

Albano, A. D. (2016). Reviewing Classical Test Theory. In Introduction to Educational and Psychological Measurement Course Notes: University of Nebraska-Lincoln.

Allen, J. L., Kautz, S. A., & Neptune, R. R. (2013). The influence of merged muscle excitation modules on post-stroke hemiparetic walking performance. *Clin Biomech (Bristol, Avon), 28*(6), 697-704. doi:10.1016/j.clinbiomech.2013.06.003

Angius, L., Pageaux, B., Hopker, J., Marcora, S. M., & Mauger, A. R. (2016). Transcranial direct current stimulation improves isometric time to exhaustion of the knee extensors. *Neuroscience, 339*, 363-375. doi:10.1016/j.neuroscience.2016.10.028

Anselmi, P., Vidotto, G., Bettinardi, O., & Bertolotti, G. (2015). Measurement of change in health status with Rasch models. *Health Qual Life Outcomes, 13*, 16. doi:10.1186/s12955-014-0197-x

Awad, L., Reisman, D., & Binder-Macleod, S. (2019). Distance-Induced Changes in Walking Speed After Stroke: Relationship to Community Walking Activity. *J Neurol Phys Ther, 43*(4), 220-223. doi:10.1097/npt.0000000000000293

Awad, L. N., Lewek, M. D., Kesar, T. M., Franz, J. R., & Bowden, M. G. (2020). These legs were made for propulsion: advancing the diagnosis and

treatment of post-stroke propulsion deficits. *J Neuroeng Rehabil, 17*(1), 139. doi:10.1186/s12984-020-00747-6

Awad, L. N., Reisman, D. S., Pohlig, R. T., & Binder-Macleod, S. A. (2016). Identifying candidates for targeted gait rehabilitation after stroke: better prediction through biomechanics-informed characterization. *Journal of NeuroEngineering and Rehabilitation, 13*(1), 84. doi:10.1186/s12984-016-0188-8

Balasubramanian, C. K., Bowden, M. G., Neptune, R. R., & Kautz, S. A. (2007). Relationship between step length asymmetry and walking performance in subjects with chronic hemiparesis. *Arch Phys Med Rehabil, 88*(1), 43-49. doi:10.1016/j.apmr.2006.10.004

Barroso, F. O., Torricelli, D., Molina-Rueda, F., Alguacil-Diego, I. M., Cano-de-la-Cuerda, R., Santos, C., . . . Pons, J. L. (2017). Combining muscle synergies and biomechanical analysis to assess gait in stroke patients. *Journal of Biomechanics, 63*, 98-103. doi:10.1016/j.jbiomech.2017.08.006

Benjamin, E. J., Virani, S. S., Callaway, C. W., Chamberlain, A. M., Chang, A. R., Cheng, S., . . . Muntner, P. (2018). Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. *Circulation, 137*(12), e67-e492. doi:10.1161/cir.0000000000000558

Bernhardt, J., Hayward, K. S., Kwakkel, G., Ward, N. S., Wolf, S. L., Borschmann, K., . . . Cramer, S. C. (2017). Agreed definitions and a shared vision for new standards in stroke recovery research: The Stroke

Recovery and Rehabilitation Roundtable taskforce. *Int J Stroke, 12*(5), 444-450. doi:10.1177/1747493017711816

Boggio, P. S., Nunes, A., Rigonatti, S. P., Nitsche, M. A., Pascual-Leone, A., & Fregni, F. (2007). Repeated sessions of noninvasive brain DC stimulation is associated with motor function improvement in stroke patients. *Restorative neurology and neuroscience, 25*(2), 123-129.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (3rd ed.). New York and London: Routledge.

Bowden, M. G., Balasubramanian, C. K., Neptune, R. R., & Kautz, S. A. (2006). Anterior-posterior ground reaction forces as a measure of paretic leg contribution in hemiparetic walking. *Stroke, 37*(3), 872-876. doi:10.1161/01.STR.0000204063.75779.8d

Bowden, M. G., Behrman, A. L., Woodbury, M., Gregory, C. M., Velozo, C. A., & Kautz, S. A. (2012). Advancing measurement of locomotor rehabilitation outcomes to optimize interventions and differentiate between recovery versus compensation. *J Neurol Phys Ther, 36*(1), 38-44. doi:10.1097/NPT.0b013e3182472cf6

Bowden, M. G., Clark, D. J., & Kautz, S. A. (2010). Evaluation of abnormal synergy patterns poststroke: relationship of the Fugl-Meyer Assessment to hemiparetic locomotion. *Neurorehabil Neural Repair, 24*(4), 328-337. doi:10.1177/1545968309343215

Bowden, M. G., Embry, A. E., & Gregory, C. M. (2011). Physical therapy adjuvants to promote optimization of walking recovery after stroke. *Stroke Res Treat, 2011*, 601416. doi:10.4061/2011/601416

Bowden, M. G., Hannold, E. M., Nair, P. M., Fuller, L. B., & Behrman, A. L. (2008). Beyond gait speed: a case report of a multidimensional approach to locomotor rehabilitation outcomes in incomplete spinal cord injury. *J Neurol Phys Ther, 32*(3), 129-138. doi:10.1097/NPT.0b013e3181838291

Cappellini, G., Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2006). Motor patterns in human walking and running. *J Neurophysiol, 95*(6), 3426-3437. doi:10.1152/jn.00081.2006

Caronni, A., Picardi, M., Gilardone, G., & Corbo, M. (2021). The McNemar Change Index worked better than the Minimal Detectable Change in demonstrating the change at a single subject level. *J Clin Epidemiol, 131*, 79-88. doi:https://doi.org/10.1016/j.jclinepi.2020.11.015

Clark, D. J., Subramanian, S., Neptune, R. R., & SA, K. (2008). *Fewer basic activation patterns account for lower extremity EMG during walking in adults post-stroke compared to healthy controls.* Paper presented at the Neural Control of Movement Annual Meeting, Naples, FL.

Clark, D. J., Ting, L. H., Zajac, F. E., Neptune, R. R., & Kautz, S. A. (2010). Merging of healthy motor modules predicts reduced locomotor performance and muscle coordination complexity post-stroke. *J Neurophysiol, 103*(2), 844-857. doi:10.1152/jn.00825.2009

Coscia, M., Monaco, V., Martelloni, C., Rossi, B., Chisari, C., & Micera, S. (2015). Muscle synergies and spinal maps are sensitive to the asymmetry induced by a unilateral stroke. *J Neuroeng Rehabil, 12*, 39. doi:10.1186/s12984-015-0031-7

Cramer, S. C., Wolf, S. L., Adams, H. P., Jr., Chen, D., Dromerick, A. W., Dunning, K., . . . Broderick, J. P. (2017). Stroke Recovery and Rehabilitation Research: Issues, Opportunities, and the National Institutes of Health StrokeNet. *Stroke, 48*(3), 813-819. doi:10.1161/strokeaha.116.015501

d'Avella, A. (2016) Modularity for motor control and motor learning. In*: Vol. 957. Advances in Experimental Medicine and Biology* (pp. 3-19).

Denny, J. C., & Collins, F. S. (2021). Precision medicine in 2030—seven ways to transform healthcare. *Cell, 184*(6), 1415-1419. doi:https://doi.org/10.1016/j.cell.2021.01.015

Duncan, P. (2012). Locomotor Experience Applied Post-Stroke (LEAPS). *National Institute of Neurologic Disease and Stroke's Archived Clinical Research Datasets*.

Duncan, P. W., Sullivan, K. J., Behrman, A. L., Azen, S. P., Wu, S. S., Nadeau, S. E., . . . Tilson, J. K. (2007). Protocol for the Locomotor Experience Applied Post-stroke (LEAPS) trial: a randomized controlled trial. *BMC Neurol, 7*, 39. doi:10.1186/1471-2377-7-39

Duncan, P. W., Sullivan, K. J., Behrman, A. L., Azen, S. P., Wu, S. S., Nadeau, S. E., . . . Hayden, S. K. (2011). Body-Weight–Supported Treadmill

Rehabilitation after Stroke. *New England Journal of Medicine, 364*(21), 2026-2036. doi:10.1056/NEJMoa1010790

Ferrante, S., Chia Bejarano, N., Ambrosini, E., Nardone, A., Turcato, A. M., Monticone, M., . . . Pedrocchi, A. (2016). A Personalized Multi-Channel FES Controller Based on Muscle Synergies to Support Gait Rehabilitation after Stroke. *Front Neurosci, 10*, 425. doi:10.3389/fnins.2016.00425

Field-Fote, E. C., & Tepavac, D. (2002). Improved intralimb coordination in people with incomplete spinal cord injury following training with body weight support and electrical stimulation. *Phys Ther, 82*(7), 707-715.

Fox, J. (2019). polycor: Polychoric and Polyserial Correlations. R package version 0.7-10.

Fregni, F., Boggio, P. S., Mansur, C. G., Wagner, T., Ferreira, M. J., Lima, M. C., . . . Pascual-Leone, A. (2005). Transcranial direct current stimulation of the unaffected hemisphere in stroke patients. *Neuroreport, 16*(14), 1551-1555. doi:10.1097/01.wnr.0000177010.44602.5e

Fridriksson, J., Basilakos, A., Stark, B. C., Rorden, C., Elm, J., Gottfried, M., . . . Bonilha, L. (2019). Transcranial direct current stimulation to treat aphasia: Longitudinal analysis of a randomized controlled trial. *Brain Stimul, 12*(1), 190-191. doi:10.1016/j.brs.2018.09.016

Fridriksson, J., Rorden, C., Elm, J., Sen, S., George, M. S., & Bonilha, L. (2018). Transcranial Direct Current Stimulation vs Sham Stimulation to Treat Aphasia After Stroke: A Randomized Clinical Trial. *JAMA Neurol, 75*(12), 1470-1476. doi:10.1001/jamaneurol.2018.2287

Gizzi, L., Nielsen, J. F., Felici, F., Ivanenko, Y. P., & Farina, D. (2011). Impulses of activation but not motor modules are preserved in the locomotion of subacute stroke patients. *J Neurophysiol, 106*(1), 202-210. doi:10.1152/jn.00727.2010

Gottlieb, G. L. (1998). Muscle activation patterns during two types of voluntary single-joint movement. *J Neurophysiol, 80*(4), 1860-1867.

Grattan, E. S., Velozo, C. A., Skidmore, E. R., Page, S. J., & Woodbury, M. L. (2019). Interpreting Action Research Arm Test Assessment Scores to Plan Treatment. *OTJR: Occupation, Participation and Health, 39*(1), 64-73. doi:10.1177/1539449218757740

Haley, S. M., & Fragala-Pinkham, M. A. (2006). Interpreting change scores of tests and measures used in physical therapy. *Phys Ther, 86*(5), 735-743.

Hashiguchi, Y., Ohata, K., Kitatani, R., Yamakami, N., Sakuma, K., Osako, S., . . . Yamada, S. (2016). Merging and Fractionation of Muscle Synergy Indicate the Recovery Process in Patients with Hemiplegia: The First Study of Patients after Subacute Stroke. *Neural Plast, 2016*, 5282957. doi:10.1155/2016/5282957

Hermens, H. J., Freriks, B., Disselhorst-Klug, C., & Rau, G. (2000). Development of recommendations for SEMG sensors and sensor placement procedures. *J Electromyogr Kinesiol, 10*(5), 361-374. doi:10.1016/s1050-6411(00)00027-4

HHS.gov. (2017). Revised Common Rule.

Hummel, F., & Cohen, L. G. (2005). Improvement of motor function with noninvasive cortical stimulation in a patient with chronic stroke. *Neurorehabilitation and neural repair, 19*(1), 14-19. doi:19/1/14 [pii]10.1177/1545968304272698

Ivanenko, Y. P., Cappellini, G., Dominici, N., Poppele, R. E., & Lacquaniti, F. (2005). Coordination of locomotion with voluntary movements in humans. *J Neurosci, 25*(31), 7238-7253. doi:10.1523/jneurosci.1327-05.2005

Ivanenko, Y. P., Poppele, R. E., & Lacquaniti, F. (2004). Five basic muscle activation patterns account for muscle activity during human locomotion. *J Physiol, 556*(Pt 1), 267-282. doi:10.1113/jphysiol.2003.057174

Jette, D. U., Halbert, J., Iverson, C., Miceli, E., & Shah, P. (2009). Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Phys Ther, 89*(2), 125-135. doi:10.2522/ptj.20080234

Kahler, E., Rogausch, A., Brunner, E., & Himmel, W. (2008). A parametric analysis of ordinal quality-of-life data can lead to erroneous results. *J Clin Epidemiol, 61*(5), 475-480. doi:10.1016/j.jclinepi.2007.05.019

Kautz, S. A., Bowden, M. G., Clark, D. J., & Neptune, R. R. (2011). Comparison of motor control deficits during treadmill and overground walking poststroke. *Neurorehabil Neural Repair, 25*(8), 756-765. doi:10.1177/1545968311407515

Kielhofner, G., Dobria, L., Forsyth, K., & Basu, S. (2005). The Construction of Keyforms for Obtaining Instantaneous Measures from the Occupational

Performance History Interview Rating Scales. *OTJR: Occupation, Participation and Health, 25*(1), 23-32. doi:10.1177/153944920502500104

Kim, D. Y., Lim, J. Y., Kang, E. K., You, D. S., Oh, M. K., Oh, B. M., & Paik, N. J. (2010). Effect of transcranial direct current stimulation on motor recovery in patients with subacute stroke. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists, 89*(11), 879-886. doi:10.1097/PHM.0b013e3181f70aa700002060-201011000-00004 [pii]

Knutsson, E., & Richards, C. (1979). Different types of disturbed motor control in gait of hemiparetic patients. *Brain, 102*(2), 405-430.

Kozlowski, A. J., Cella, D., Nitsch, K. P., Heinemann, A. W. D. o. P. M., & Rehabilitation, F. S. o. M. N. U. C. I. L. (2016). Evaluating Individual Change With the Quality of Life in Neurological Disorders (Neuro-QoL) Short Forms. *Arch Phys Med Rehabil, 97*(4), 650-654. doi:10.1016/j.apmr.2015.12.010

Krasovsky, T., & Levin, M. F. (2010). Review: toward a better understanding of coordination in healthy and poststroke gait. *Neurorehabil Neural Repair, 24*(3), 213-224. doi:10.1177/1545968309348509

Lamontagne, A., & Fung, J. (2004). Faster is better: implications for speed-intensive gait training after stroke. *Stroke, 35*(11), 2543-2548. doi:10.1161/01.STR.0000144685.88760.d7

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature, 401*(6755), 788-791. doi:10.1038/44565

Lefaucheur, J.-P., Antal, A., Ayache, S. S., Benninger, D. H., Brunelin, J., Cogiamanian, F., . . . Paulus, W. (2017). Evidence-based guidelines on the therapeutic use of transcranial direct current stimulation (tDCS). *Clinical Neurophysiology, 128*(1), 56-92. doi:https://doi.org/10.1016/j.clinph.2016.10.087

Lefebvre, S., & Liew, S. L. (2017). Anatomical Parameters of tDCS to Modulate the Motor System after Stroke: A Review. *Front Neurol, 8*, 29. doi:10.3389/fneur.2017.00029

Li, Y., Fan, J., Yang, J., He, C., & Li, S. (2018). Effects of transcranial direct current stimulation on walking ability after stroke: A systematic review and meta-analysis. *Restor Neurol Neurosci, 36*(1), 59-71. doi:10.3233/rnn-170770

Lim, C. R., Harris, K., Dawson, J., Beard, D. J., Fitzpatrick, R., & Price, A. J. (2015). Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open, 5*(7), e007765. doi:10.1136/bmjopen-2015-007765

Linacre, J. M. (1997). Instantaneous measurement and diagnosis. *Physical Medicine and Rehabilitation., 11*(2), 315-324.

Linacre, J. M. (2003). Rasch power analysis: size vs. significance: standardized chi-square fit statistic. *Rasch Measurement Transactions, 17*, 918.

Linacre, J. M. (2007). Standard Errors and Reliabilities: Rasch and Raw Score. *Rasch Measurement Transactions, 20*(4).

Madhavan, S., & Stinear, J. W. (2010). Focal and bi-directional modulation of lower limb motor cortex using anodal transcranial direct current stimulation. *Brain Stimul, 3*(1), 42. doi:10.1016/j.brs.2009.06.005

Madhavan, S., Weber, K. A., 2nd, & Stinear, J. W. (2011). Non-invasive brain stimulation enhances fine motor control of the hemiparetic ankle: implications for rehabilitation. *Exp Brain Res, 209*(1), 9-17. doi:10.1007/s00221-010-2511-0

McGowan, C. P., Neptune, R. R., Clark, D. J., & Kautz, S. A. (2010). Modular control of human walking: Adaptations to altered mechanical demands. *J Biomech, 43*(3), 412-419. doi:10.1016/j.jbiomech.2009.10.009

Mulroy, S., Gronley, J., Weiss, W., Newsam, C., & Perry, J. (2003). Use of cluster analysis for gait pattern classification of patients in the early and late recovery phases following stroke. *Gait Posture, 18*(1), 114-125. doi:10.1016/s0966-6362(02)00165-0

Neptune, R. R., Clark, D. J., & Kautz, S. A. (2009). Modular control of human walking: a simulation study. *J Biomech, 42*(9), 1282-1287. doi:10.1016/j.jbiomech.2009.03.009

Noetscher, G. M., Yanamadala, J., Makarov, S. N., & Pascual-Leone, A. (2014). Comparison of cephalic and extracephalic montages for transcranial direct current stimulation--a numerical study. *IEEE Trans Biomed Eng, 61*(9), 2488-2498. doi:10.1109/tbme.2014.2322774

Nowak, D. A., Grefkes, C., Ameli, M., & Fink, G. R. (2009). Interhemispheric competition after stroke: brain stimulation to enhance recovery of function

of the affected hand. *Neurorehabil Neural Repair, 23*(7), 641-656. doi:10.1177/1545968309336661

Ojardias, E., Aze, O. D., Luneau, D., Mednieks, J., Condemine, A., Rimaud, D., . . . Giraux, P. (2019). The Effects of Anodal Transcranial Direct Current Stimulation on the Walking Performance of Chronic Hemiplegic Patients. *Neuromodulation*. doi:10.1111/ner.12962

Perry, J., Garrett, M., Gronley, J. K., & Mulroy, S. J. (1995). Classification of walking handicap in the stroke population. *Stroke, 26*(6), 982-989.

Plotnik, M., Giladi, N., & Hausdorff, J. M. (2008). Bilateral coordination of walking and freezing of gait in Parkinson's disease. *Eur J Neurosci, 27*(8), 1999-2006. doi:10.1111/j.1460-9568.2008.06167.x

Reeve, B. B., Hays, R. D., Bjorner, J. B., Cook, K. F., Crane, P. K., Teresi, J. A., . . . Cella, D. (2007). Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care, 45*(5 Suppl 1), S22-31. doi:10.1097/01.mlr.0000250483.85507.04

Reis, J., & Fritsch, B. (2011). Modulation of motor performance and motor learning by transcranial direct current stimulation. *Current opinion in neurology, 24*(6), 590-596. doi:10.1097/WCO.0b013e32834c3db0

Revelle, W. (2020). psych: Procedures for Psychological, Psychometric, and Personality Research. *Northwestern University, Evanston, Illinois, USA*.

Riddle, D. L., & Stratford, P. W. (2013). *Is This Change Real? : Interpreting Patient Outcomes in Physical Therapy*. In F.A. Davis PT Collection.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36.

Routson, R. L., Clark, D. J., Bowden, M. G., Kautz, S. A., & Neptune, R. R. (2013). The influence of locomotor rehabilitation on module quality and post-stroke hemiparetic walking performance. *Gait Posture, 38*(3), 511-517. doi:10.1016/j.gaitpost.2013.01.020

Routson, R. L., Kautz, S. A., & Neptune, R. R. (2014). Modular organization across changing task demands in healthy and poststroke gait. *Physiol Rep, 2*(6). doi:10.14814/phy2.12055

Schlaug, G., Renga, V., & Nair, D. (2008). Transcranial direct current stimulation in stroke recovery. *Archives of neurology, 65*(12), 1571-1576. doi:65/12/1571 [pii]10.1001/archneur.65.12.1571

Seamon, B. A., Neptune, R. R., & Kautz, S. A. (2018). Using a Module-Based Analysis Framework for Investigating Muscle Coordination during Walking in Individuals Poststroke: A Literature Review and Synthesis. *Appl Bionics Biomech, 2018*, 3795754. doi:10.1155/2018/3795754

Seamon, B. A., Neptune, R. R., & Kautz, S. A. (2018). Using a Module-Based Analysis Framework for Investigating Muscle Coordination during Walking in Individuals Poststroke: A Literature Review and Synthesis. *Applied Bionics and Biomechanics, 2018*, 16. doi:10.1155/2018/3795754

Stevens, J. G., & Beurskens, A. J. (2010). Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. *Phys Ther, 90*(6), 953-961. doi:10.2522/ptj.20090105

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*(2684), 677-680.

Stratford, P. W., Binkley, J., Solomon, P., Finch, E., Gill, C., & Moreland, J. (1996). Defining the Minimum Level of Detectable Change for the Roland-Morris Questionnaire. *Phys Ther, 76*(4), 359-365. doi:10.1093/ptj/76.4.359

Stratford, P. W., & Binkley, J. M. (1999). Applying the results of self-report measures to individual patients: an example using the Roland-Morris Questionnaire. *J Orthop Sports Phys Ther, 29*(4), 232-239. doi:10.2519/jospt.1999.29.4.232

Stratford, P. W., Binkley, J. M., Riddle, D. L., & Guyatt, G. H. (1998). Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. *Phys Ther, 78*(11), 1186-1196. doi:10.1093/ptj/78.11.1186

Sullivan, K. J., Knowlton, B. J., & Dobkin, B. H. (2002). Step training with body weight support: effect of treadmill speed and practice paradigms on poststroke locomotor recovery. *Arch Phys Med Rehabil, 83*(5), 683-691. doi:10.1053/apmr.2002.32488

Tahtis, V., Kaski, D., & Seemungal, B. M. (2014). The effect of single session bi-cephalic transcranial direct current stimulation on gait performance in sub-acute stroke: A pilot study. *Restor Neurol Neurosci, 32*(4), 527-532. doi:10.3233/rnn-140393

Tanaka, S., Takeda, K., Otaka, Y., Kita, K., Osu, R., Honda, M., . . . Watanabe, K. (2011). Single session of transcranial direct current stimulation transiently increases knee extensor force in patients with hemiparetic

stroke. *Neurorehabil Neural Repair, 25*(6), 565-569. doi:10.1177/1545968311402091

Tatemoto, T., Yamaguchi, T., Otaka, Y., Kondo, K., & Tanaka, S. (2013). *Anodal Transcranial Direct Current Stimulation over the Lower Limb Motor Cortex Increases the Cortical Excitability with Extracephalic Reference Electrodes*, Berlin, Heidelberg.

Team, R. C. (2019). *R: A Language and Environment for Statistical Computing*. In. Retrieved from http://www.R-project.org/

Thair, H., Holloway, A. L., Newport, R., & Smith, A. D. (2017). Transcranial Direct Current Stimulation (tDCS): A Beginner's Guide for Design and Implementation. *Front Neurosci, 11*, 641. doi:10.3389/fnins.2017.00641

Thissen, D. E., & Wainer, H. E. (2001). *Test scoring*: Lawrence Erlbaum Associates Publishers.

Ting, L. H., Chiel, H. J., Trumbower, R. D., Allen, J. L., McKay, J. L., Hackney, M. E., & Kesar, T. M. (2015). Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron, 86*(1), 38-54. doi:10.1016/j.neuron.2015.02.042

Ting, L. H., & Macpherson, J. M. (2005). A limited set of muscle synergies for force control during a postural task. *J Neurophysiol, 93*(1), 609-613. doi:10.1152/jn.00681.2004

Tresch, M. C., Cheung, V. C., & d'Avella, A. (2006). Matrix factorization algorithms for the identification of muscle synergies: evaluation on

simulated and experimental data sets. *J Neurophysiol, 95*(4), 2199-2212. doi:10.1152/jn.00222.2005

van Asseldonk, E. H., & Boonstra, T. A. (2016). Transcranial Direct Current Stimulation of the Leg Motor Cortex Enhances Coordinated Motor Output During Walking With a Large Inter-Individual Variability. *Brain Stimul, 9*(2), 182-190. doi:10.1016/j.brs.2015.10.001

Vaz, P. G., Salazar, A., Stein, C., Marchese, R. R., Lukrafka, J. L., Plentz, R. D. M., & Pagnussat, A. S. (2019). Noninvasive brain stimulation combined with other therapies improves gait speed after stroke: a systematic review and meta-analysis. *Top Stroke Rehabil, 26*(3), 201-213. doi:10.1080/10749357.2019.1565696

Velozo, C. A., Kielhofner, G., & Lai, J. S. (1999). The use of Rasch analysis to produce scale-free measurement of functional ability. *Am J Occup Ther, 53*(1), 83-90. doi:10.5014/ajot.53.1.83

Velozo, C. A., & Woodbury, M. L. (2011). Translating measurement findings into rehabilitation practice: an example using Fugl-Meyer Assessment-Upper Extremity with patients following stroke. *J Rehabil Res Dev, 48*(10), 1211-1222. doi:10.1682/jrrd.2010.10.0203

Wade, D. T., Langton-Hewer, R., Wood, V. A., Skilbeck, C. E., & Ismail, H. M. (1983). The hemiplegic arm after stroke: measurement and recovery. *J Neurol Neurosurg Psychiatry, 46*(6), 521-524. doi:10.1136/jnnp.46.6.521

Winstein, C. J., Stein, J., Arena, R., Bates, B., Cherney, L. R., Cramer, S. C., . . . Zorowitz, R. D. (2016). Guidelines for Adult Stroke Rehabilitation and

Recovery: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. *Stroke, 47*(6), e98-e169. doi:10.1161/str.0000000000000098

Woodbury, M. L., Anderson, K., Finetto, C., Fortune, A., Dellenbach, B., Grattan, E., & Hutchison, S. (2016). Matching Task Difficulty to Patient Ability During Task Practice Improves Upper Extremity Motor Skill After Stroke: A Proof-of-Concept Study. *Arch Phys Med Rehabil, 97*(11), 1863-1871. doi:10.1016/j.apmr.2016.03.022

Woods, A. J., Antal, A., Bikson, M., Boggio, P. S., Brunoni, A. R., Celnik, P., . . . Nitsche, M. A. (2016). A technical guide to tDCS, and related non-invasive brain stimulation tools. *Clin Neurophysiol, 127*(2), 1031-1048. doi:10.1016/j.clinph.2015.11.012

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions, 8*(3), 370.

Wright, B. D., & Masters, G. N. (2002). Number of Person or Item Strata: (4*Separation + 1)/3. *Rasch Measurement Transactions, 16*(3), 888.

Wyrwich, K. W. (2004). Minimal important difference thresholds and the standard error of measurement: is there a connection? *J Biopharm Stat, 14*(1), 97-110. doi:10.1081/bip-120028508

Zajac, F. E., Neptune, R. R., & Kautz, S. A. (2003). Biomechanics and muscle coordination of human walking: part II: lessons from dynamical simulations and clinical implications. *Gait Posture, 17*(1), 1-17.