Computational Modeling to Predict and Affect Recovery After Stroke

by

Yazan Abdel Majeed
B.Sc., Princess Sumaya University for Technology, 2010

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Bioengineering
in the Graduate College of the
University of Illinois at Chicago, 2020

Chicago, Illinois

Defense Committee:
James Patton, PhD, Chair and Advisor
Saria Awadalla, PhD, School of Public Health
William 'Zev' Rymer, MD, PhD, Shirley Ryan AbilityLab
Robert Kenyon, PhD, Computer Science
Yang Dai, PhD

To the best people ever... family does not have to be blood.

ACKNOWLEDGMENT

First and foremost, I would like to thank my thesis advisor, Prof. James Patton, of the Richard and Loan Hill Department of Bioengineering at the University of Illinois at Chicago. Prof. Patton was always available whenever I needed direction or help with my research. His feedback was instrumental in vastly improving my skills in communicating my work. He always asked the hard questions, and thanks to him I learned to always be ready with the answers.

This work would not have been possible without Dr. Saria Awadalla's invaluable expertise. Dr. Awadalla introduced me to the statistical tools and algorithms that became the cornerstone of this thesis. His consistent feedback and guidance were instrumental in getting this research to stand on strong mathematical foundations.

I would like to thank my committee members for challenging me to become a better scientist. Their thought-provoking questions and advice were vital in improving this research.

I would also like to thank the stroke survivors at the Shirley Ryan AbilityLab (formerly the Rehabilitation Institute of Chicago), for participating in this research. Their commitment to improving after their stroke and to the common good inspired me to work harder. The most gratifying part of this work was watching their progress as they persevered to improve.

Next, I would like to thank the amazing people, past and present, at the Robotics lab. We had so many conversations, collaborations, and happy hours. Some fruitful, others informative, and a few plain silly. You all added immensely to my experience these past years. Thank you

ACKNOWLEDGMENT (Continued)

Sandro, Felix, Zack, Keon, Fabio, Dalia, Pritesh, Partha, Linda, Claudio, Sasha, Courtney, Eyad, Amit, Justin, Emily, Carmen, Jemina, Emily, Kelly, Molly, and Farnaz.

Last, but certainly not least, I would like to thank my friends and support system: Justin, Theresa, Manu, Jeanette, Jennifer, Dave, Melissa, Fabio, Ivan, Dalia, Linda (Laura), Peter, Claudio, Alexis, Sarah, Tasha, Jenna, Genevieve, Warren, Maynard, Luke, Carmen, and Yuri. You have enriched my life and taught me so much! I appreciate you, and always will.

YAM

CONTRIBUTIONS OF AUTHORS

Chapter 1 is an overview of the current state of the research and motivations for my main thesis questions. Chapter 2 is a published manuscript "Regression techniques employing feature selection to predict clinical outcomes in stroke" published in PloS one, 13(10), e0205639 by Majeed, Y. A., Awadalla, S. S., & Patton, J. L. (2018). I was the primary author and conducting the data cleanup and statistical analysis. Dr. Saria Awadalla gave crucial input on the models and algorithms I used and helped me understand how to properly implement them. My advisor, Prof. James Patton contributed to writing the manuscript and gave me feedback and advice on the general direction of the work. Chapter 3 is a manuscript under review "Effects of Robot Viscous Forces on Arm Movements in Chronic Stroke Survivors: A Randomized Crossover Study" by Majeed, Yazan A., Awadalla, Saria S., & Patton, James L. (Journal of Neuroengineering and Rehabilitation, 2019). I was the primary author, I managed the IRB and collected the data from stroke patients. I ran the statistical analysis with Dr. Awadalla's feedback and wrote the document with Dr. Patton's help. Chapter 4 is a submitted case study "Training With Negative Viscosity Improves Clinical Outcomes After Stroke: A Case Study" by Yazan Abdel Majeed, Courtney Celian, and James Patton. The case study implements the treatment we arrived at after analyzing the data from Chapters 2 and 3 (Clinical Rehabilitation, 2020). I collected the data from a stroke survivor and did the statistical analysis. Courtney Celian performed screenings and clinical assessments for the study, and Dr. Patton contributed

CONTRIBUTIONS OF AUTHORS (Continued)

to writing the manuscript. In Chapter 5, I connect the results from the work presented in this thesis and describe limitations, possible future directions, and lay out my general conclusions.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

APR	Arrest Period Ratio
ARAT	Action Research Arm Test
EXCITE	Extremity Constraint Induced Therapy Evaluation
LARS	Least Angle Regression
LASSO	Least Absolute Shrinkage and Selection Operator
MCID	Minimal Clinically Important Difference
PCA	Principal Componenent Analysis
PMTD	Percentage of Movement in the Target Direction
RMSE	Root Mean Square Error
UEFM	Upper Extremity Fugl-Meyer
VRROOM	Virtual Reality Robotic and Optical Operations Machine
WMFT	Wolf Motor Function Test
WREX	Wilmington Robotic Exoskeleton

SUMMARY

There is currently no direct association between patient movement kinematics and rehabilitation outcomes for stroke survivors. Research has demonstrated minor successes in predicting some clinical measures. Recent advancements in machine learning algorithms and robotic technology enable us to parse through a large amount of patient movement data to find appropriate targets for treatment. This thesis examines the predictability of chronic stroke survivor clinical outcomes following neurological injury using patient movement and demographic data, and whether predictive movement metrics can be influenced to improve recovery.

Building on a recent two-week study with stroke survivors, we construct a machine learning model to predict clinical changes, and use the model to determine an appropriate intervention target. We then establish how this target should be altered in a crossover trial with chronic stroke survivors. Lastly, we test a two-week treatment in a chronic stroke survivor case study.

Our predictive models determined that movement speed was a proper focus for an intervention. In our crossover trial, we showed that negative viscosity had the strongest effect on movement speed. Treating a patient with negative viscosity for two weeks showed significant improvements in multiple movement metrics, and small gains in clinical outcomes.

These results show that designing a rehabilitation intervention based on negative viscosity can lead to improvements in chronic stroke survivors. Our success also establishes the framework we used for this patient population as a potentially generalizable pipeline to identify and test

interventions for other populations. Further research could focus on testing treatments for some of the other model predictors, as they could present more potential intervention targets.

CHAPTER 1

INTRODUCTION

1.1 The Knowledge Gap

Stroke is one of the primary causes of disability in the 21st century, due to improvements in medical practices resulting in increased chances of survival from a stroke incident (1). Stroke survivors commonly suffer from movement deficits due to the high likelihood that one of the many regions of the brain involved in motion will be within the area affected (2). Recovery of these motor skills is quantified and evaluated using a host of clinical assessments (such as Fugl-Meyer (3), Wolf Motor Function (4), and Action Research Arm Test (ARAT) (5; 6)). Yet there is no clear connection between patient movement performance and their functioning on these clinical tests. Even though clinical assessments have been shown to roughly correlate to patient recovery, we argue that more effective therapy can be designed and customized for patients if we know the relationship between movement metrics and the standard clinical tests. Recent advances in computation allow for the design and implementation of novel and powerful statistical and machine learning tools, while recent advances in robotics technology allow the collection of vast amounts of data on patient movements during robotic therapy. There is currently poor understanding of *how* these clinical assessments correlate with patient movement metrics or their demographic characteristics.

Robotic therapy implementing force fields designed to affect specific aspects of movement has proven effective in changing behaviors of neurologically-injured patients (7; 8). On the other hand, popular statistical and machine learning algorithms like LASSO and Random Forests have proven effective at making predictions in various applications (9; 10). Current robotic therapy and sensing technologies allow for the measurement of a host of kinematic and dynamic metrics, and there is a need for parsing metrics to identify the best foci for more effective therapy. Recovery is normally measured using a single clinical scale, and explaining changes can be difficult given the complexity of human movement and the multiplicity of abilities that may change during the course of a single intervention. Not every metric that changes is relevant when explaining clinical results and identifying those important for prediction is a key advancement in our understanding of neurological injury and the interventions that can improve recovery.

Our long-term goal is to develop a tool to *personalize therapy* and *maximize therapeutic outcomes* for better performance both within and outside the clinical environment. This would include both understanding what behaviors correlate with better recovery and identify any specific patient sub-populations that may be more or less affected by treatments. This thesis presents an initial foray towards this goal.

The first step is to develop a framework that uses patient movement and demographic information to predict changes in clinical outcomes, and rank the predictive features to identify potential targets for therapy. This framework could have a profound impact on stroke rehabilitation paradigms, since it can highlight movement metrics that more strongly predict clinical recovery. While researchers typically compare movement of neurologically-injured patients to

those of healthy subjects to propose interventions, our framework uses robust statistical tools and cross-validation to systematically identify intervention targets that are correlated with clinical change. Within the scope of stroke rehabilitation, our central hypothesis is that a framework identifying a statistically robust, minimal feature space underlying clinical change will help design future therapies, and customize therapies based on relevant patient movement metrics.

Using this framework, we will determine a handful of key factors to test as new interventions. We will run a crossover study with stroke survivors to determined the most effective method to affect one of these key factors, and we will test the use of that method as a treatment in a stroke case study.

1.2 Significance of This Work

Identification of proper therapy targets is the foundation of effective rehabilitation and positive training outcomes. Neural injury or disease causes changes to both the ability to plan and to execute motor tasks. On the execution side, this can manifest as affecting muscle strength (11; 12), muscle synergies (13; 14), and limb movements (15; 16). Rehabilitation interventions face many challenges due to the host of possible impairments at varying severities and the lack of direct relationships between these impairments and standard clinical assessments.

Current diagnostic tools for rehabilitation assess patient progress from various perspectives. From motor ability (3) and function (4; 17) to MRI, CT, and elastography imaging to assess tissue damage (18; 19). While these tools provide an indispensable and somewhat thorough

evaluation of a stroke survivor's condition, the picture they provide cannot be easily linked back to patient movement performance.

The area of predicting patient recovery is that of active research and undergoes constant development. Many researchers employ imaging techniques to correlate changes in brain structure and perfusion patterns to clinical outcomes (20; 21; 22). However, these tools are predictive of recovery only when the brain is imaged immediately after stroke. Recent computational work has demonstrated promise in the ability to predict some clinical measures (23). However, the complex algorithms used were able to explain ≈60% of the clinical outcome variance at best. The simple statistical tools outlined in this proposal promise to be more robust, especially under cross-validation, and easier to interpret using the original predictor set.

The study of features that predict stroke recovery, especially movement metrics we can directly affect, would lead to controlled therapy prescriptions individualized to patient needs and perhaps even tailored to specific daily tasks. We know that interacting with haptic disturbances and vision/proprioception manipulations in virtual reality environments can modify the way healthy subjects and stroke survivors move (7; 24). There is also evidence that simple statistical and machine learning methods like LASSO and Random Forests are incredibly robust and are capable of making excellent predictions in many computational problems (25; 26; 9; 27; 10). These models are especially effective when the predictor space has many more dimensions than the number of observations available in each dimension (a class of problems mathematically referred to as the high $\mathtt{p}$ low $\mathtt{N}$ problems (28)). The majority of prediction algorithms lack interpretability as they use abstractions constructed from the input features that are then used

to make the predictions. The resulting models are therefore presented in terms of these abstractions and not the input space. The framework proposed here will outperform current modeling efforts and set a new standard for effective, interpretable prediction of clinical outcomes to serve as a guide for future rehabilitation therapies.

This thesis will lay down the foundation for a robust system to pinpoint factors that are associated with recovery, determine the methods that need to be used in order to influence these factors, and determine how these methods may be leveraged to maximize clinical outcomes. Our validation approaches also tell us about the certainty of our findings. This work undertakes a crucial line of inquiry of whether novel statistical tools can reveal insights about optimal rehabilitation interventions.

1.3 Methodology

1.3.1 Statistical Framework

We will lay out an intelligible framework to predict clinical outcomes, particularly after neurological injury. At this stage, our framework will focus on stroke survivors and two popular clinical assessments used to assess recovery as a test case. Rather than use dimensionality reduction of the input feature space, our approach uses the high-dimensional input space to directly make predictions. This leverages penalized regression (elastic nets, LASSO, ridge, Least Angle Regression (LARS)), and decision trees (of which we chose the Random Forests implementation) to construct models that are easily interpretable while maintaining prediction power.

Our framework has the ability to systematically shrink the list of predictors to identify the features most important for the prediction. Rather than use healthy controls as a reference to recommend interventions, our data-driven approach relies on robust statistical tools to identify predictors with stronger influence on clinical change. These tools will suggest interventions that may not be easily discernible using current modeling techniques. While existing interventions derived from comparison to healthy subject movements can show some effectiveness, there is no guarantee that more similarity to healthy movements correlates with better clinical outcomes. Our unique analytic framework focuses on strong statistical tools to detect these indirect correlations with clinical evaluations. This powerful approach allows for principled reduction in the dimensionality of the predictor space while maintaining interpretability. Hence, it is particularly suited for these problems where the number of possible predictors far exceeds the number of stroke survivors enrolled in a particular study. Ultimately, our goal is to create a structured approach to forecast clinical changes and to determine the subset of predictors and interventions most responsible for these changes. When these important predictors are identified, we can test them in a clinical setting to determine the most effective interventions. This work will introduce simple models that emphasize aspects of the patient's condition that can be influenced to improve recovery, leading to more effective treatments and a better quality of life following neurological injury.

Our method uses cross-validated parsing of high-dimensional feature spaces and uniquely applies cutting-edge statistical methods to understand movement changes underlying the most common clinical measures. We intend to show that the high-dimensionality of the input space

can be reduced and still predict clinical changes. Interestingly, recent findings suggest that at least one less-popular assessment can be predicted using patient demographics (23).

Our framework adapts established modeling techniques (29; 30; 31) for the purpose of predicting stroke rehabilitation outcomes and identifying targets for future interventions. This allows us to navigate the complex array of movement metrics to focus on the small subset underlying clinical recovery. Movement metrics were extracted from patients' early interactions with a virtual reality training paradigm, then used to predict clinical changes. Models successful at making these predictions were then "unpacked" to determine metrics most critical to the predictions. As opposed to common dimensionality reduction methods like Principal Componenent Analysis (PCA) (32), which lack interpretability of the most predictive components, we used a comprehensible framework that systematically contracted the list of predictors to identify the most important ones, while maintaining interpretability. Only a small subset of the feature space was needed to make accurate predictions. Additionally, this framework was useful in determining whether the intervention itself was predictive of clinical change, by including the treatment as one of the possible predictors.

1.3.2 Intervention Testing

Robotics have been used to augment, assist, and manipulate human movement since the late 1940s and used specifically for rehabilitation since the 1960s (33). It is possible and somewhat common to manipulate movement metrics in a virtual reality environment as both haptic and visual feedback have been used to affect people's movement in these settings (34; 35). Motor learning and rehabilitation literature supports three main approaches to alter movement

kinematics. The first is assistive, where the intervention aims to simplify the task dynamics for the patient, making it easier to complete the task (36; 37; 38). There is evidence that this method can be effective in improving clinical outcomes (39; 40).

The second method is resistive, where the intervention aims to make it more difficult for the patient to complete the task by implementing a disturbance either visually or haptically. Such implementations can include curl fields and error augmentation (41; 42; 24; 43). There is support that this approach causes overcompensation in the opposite direction of the intervention, thereby leveraging the training's aftereffect to improve patient performance (44). An offshoot of the resistive approach is combining resistive forces with a reward mechanism (45; 46; 47; 48), where patients are rewarded for effectively countering the resistive forces by removing the robot resistance and thereby making the task easier. This would ideally leverage human reward-seeking behaviors to bias patients towards the desired movements.

The last of these methods is self-directed training, where the patient is given feedback about their performance and encouraged to alter certain movement kinematics without outside intervention. This method requires more active involvement on the subject's part, and may therefore be more difficult to implement for patients with higher impairment levels. There is evidence, though, that active patient involvement might aid in achieving reliable clinical changes (49; 50).

Once we use our statistical framework to identify a suitable target for rehabilitation, we will directly compare the three modalities – assistive, resistive, and self-directed – in a single day

crossover study. The most effective approach will be used as the basis for our intervention in the fourth chapter of this thesis.

To test the most successful approach of altering the movement metric identified by our data-informed framework, we will train a chronic stroke survivor using this approach for two weeks. We will test their performance before, during, and after the treatment to determine its effectiveness. This will lay the foundations for a future randomized controlled clinical trial.

1.4 Outline

1.4.1 Part 1: Determine the predictability of standard clinical outcomes and identify strong predictors.

Current approaches to stroke rehabilitation evaluate deficient movement metrics based on comparisons to healthy movement without an understanding of how these metrics correspond to outcomes. Identifying strong predictors of clinical change could lead to more effective interventions following neurological injury. Clinical assessments are the current therapeutic standard for evaluating a patient's condition, performance, and recovery (51; 52). These assessments rely on subjective observations of patient ability by the therapist, and are therefore susceptible to intra- and interrater variability (53; 54).

While these clinical assessments have demonstrated success in tracking patient recovery, our framework provides patients and clinicians with a clear therapy target. Our goal is to develop frameworks connecting these clinical assessments to quantifiable metrics of the patient's movement. Neurological injury, such as stroke, affects multiple aspects of a patient's movement, from perception to motor ability (55), and recovery involves changes on both the structural and

functional levels (56; 57). The ability to independently evaluate the effects of all these changes on recovery and how they are associated with widely-accepted clinical outcomes will advance rehabilitation interventions. On the computational side, advances in prediction algorithms that reduce the dimensionality of the input space while maintaining interpretability are providing insights into mechanisms underlying observable metrics in many fields. Researchers have employed both penalized regression (58; 10) and robust decision trees (9; 27) to make predictions and sort through large numbers of predictors to identify the most important features.

Identification of this much smaller list of the best predictive features has some important advantages. We are able to discard a variety of predictors that may have been a logical choice but are not essential to the predictions, simplifying the data collection process. The effect of each strong predictor on the quality of the prediction can be determined analytically. Though this framework may be generalized to tackle any prediction task, the choice of features going into the algorithms is critical. Different therapy tasks lend themselves to varying sets of metrics to define movement and its quality. We demonstrate the ability to predict clinical changes using a rationally chosen set of predictive features, guided by the literature. With this, is it possible to construct prediction models where we are confident about the relationships between features and to describe clinical changes using a minimal set of predictors that can then be pursued for subsequent therapy interventions.

Applying these techniques to model stroke recovery can help clarify the interdependencies between the different levels of impairment and pinpoint the most worthwhile targets for treatment to improve therapy outcomes. Moreover, framework can reliably determine whether the

treatment in any clinical intervention was effective by including the treatment itself as a predictor in the model. Effectual treatments will be chosen by the model as it searches through the feature space to detect the best predictors. Consequently, these tools can be employed to determine the statistical relevance of an intervention. In Chapter 2, we show the framework we constructed using LASSO penalized regression and Random Forest decision trees. We applied the framework to predict two clinical outcomes from a three-week study and determined the important predictors. Our framework revealed a small set of valuable targets for therapy that explain most of the variability in clinical outcomes.

1.4.2 Part 2: Identify the most effective method to shape movement metrics.

Predicting clinical changes based on movement metrics is an important result of our modeling efforts. However, our goal is to use these metrics to improve patient performance. Since there are multiple ways to intervene, we first need to determine the most effective approach of modifying the movement metrics deemed predictive in Chapter 2. Thus, we compare three different methods for influencing movement – assistive, resistive, and self-directed – in a single day crossover study in Chapter 3. We show that one of these approaches will be more effective at shaping the way patients move, and that approach will be used as the basis for our intervention in the third part of this thesis.

1.4.3 Part 3: Determine how targeting strongly-predictive movement metrics affects clinical outcomes.

While typical stroke rehabilitation designs interventions to encourage stroke survivor movements that more closely resemble healthy movement, we target specific movement aspects that

are *predictive* of clinical recovery. Using the strongest predictors of recovery identified in Chapter 2, we designed an intervention to change these movement metrics using the approach deemed most effective by our crossover study in Chapter 3. In Chapter 4, we include the results of a pilot case study with a chronic stroke survivor, demonstrating the effect of altering the patient's movement in the fashion suggested by our models in Chapter 2. We compare clinical performance before and after the treatment on the same assessments used in Chapter 2. We also compare several movement metrics directly across treatment days to highlight progress due to our chosen training.

CHAPTER 2

REGRESSION TECHNIQUES EMPLOYING FEATURE SELECTION TO PREDICT CLINICAL OUTCOMES IN STROKE

This work is published, under the same title, as: Majeed Y. A., Awadalla S. S., and Patton J. L. in Plos one, 13(10):e0205639, 2018.

2.1 Abstract

It is not fully clear which measurable factors can reliably predict chronic stroke patients' recovery of motor ability. In this analysis, we investigate the impact of patient demographic characteristics, movement features, and a three-week upper-extremity intervention on the post-treatment change in two widely used clinical outcomes – the Upper Extremity portion of the Fugl-Meyer and the Wolf Motor Function Test. Models based on LASSO, which in validation tests account for 65% and 86% of the variability in Fugl-Meyer and Wolf, respectively, were used to identify the set of salient demographic and movement features. We found that age, affected limb, and several measures describing the patient's ability to efficiently direct motions with a single burst of speed were the most consequential in predicting clinical recovery. On the other hand, the upper-extremity intervention was not a significant predictor of recovery. Beyond a simple prognostic tool, these results suggest that focusing therapy on the more important features is likely to improve recovery. Such validation-intensive methods are a novel approach

to determining the relative importance of patient-specific metrics and may help guide the design of customized therapy.

2.2 Introduction

Recovering from stroke is a highly variable process(59) that is difficult to influence or predict. There are many clinical assessments to evaluate the state of a patient and gauge his or her long-term prognosis. Some assessments are sufficiently reliable(60), though there is no widely accepted gold standard(61; 62). In practice, a battery of clinical evaluations are conducted, each used to assess a different aspect of a patient's condition. Common assessment areas include: (1) motor ability, such as Fugl-Meyer(3); (2) functional performance, such as Wolf Motor Function Test(4); and (3) self-reported motor activity, as in the case of the Motor Activity Log(63) and the Functional Independence Measure(64). There is no consolidated outcome measure that encompasses these disparate evaluations, and there is general consensus that a combination of assessments provides the best profile of a patient (65).

The relationship between these clinical assessments and a patient's movement features while performing a task is not fully understood, nor whether or how they are impacted by non-movement variables such as socio-demographic characteristics. Prediction of patient recovery has been an area of active research where much of the recent developments have focused on using imaging techniques to correlate changes in brain structure and perfusion patterns to clinical outcomes(22; 21; 66; 67; 68; 69), or to use other neurophysiological and neuroimaging *biomarkers* to predict recovery(70; 71). However, these approaches offer reasonable predictions of recovery only when the brain is imaged immediately following a stroke.

Recent computational work has shown promise predicting some clinical measures. However, the complex algorithms used were able to explain approximately 60% of the clinical outcome variability at best. To date, few and relatively recent studies used robots to explore the relationship between patient progress and clinical outcomes(72; 23). There has also been recent success in using psychological priming to influence patient recovery from stroke without directly controlling for aspects of movement(73).

One difficulty associated with exploring the relationship between patient progress and clinical outcomes is the reliability of clinical outcomes (74; 75), especially in attempting to identify small changes in a patient's condition or small differences between patients. These changes are often within the test-retest and inter-rater variability ranges for the clinical measures, making them difficult to use under these conditions. Some researchers have looked into using robots to obtain a more comprehensive set of clinical assessments (76; 77). However, these works did not attempt to predict the most widely-accepted clinical outcomes of UEFM and WMFT.

Another computational challenge is the low number of patients in many of these studies. Combined with the high number of measurable assessments (features) available, few methods are available that produce reliable predictions while also pinpointing the most important features. The field of machine learning has recently offered robust tools to address these challenges. Here, we compared the top three candidate algorithms best known for their abilities to predict in this type of scenario of few observations and many features (mathematically referred to as the high p low n problems). Importantly, we then used cross-validation to assess the *certainty* of such predictive power, allowing us to gauge confidence in our results.

In this study, we investigated the relationships between the changes in two typically-used clinical outcomes (UEFM and WMFT) following a three-week bimanual self-telerehabilitation (Figure 1A) randomized placebo-controlled intervention. We trained $\mathsf{N} = 26$ chronic stroke survivors for a two-week period (six 1-hour training sessions, Figure 1B) and extracted variables pertaining to three domains: patient movement, clinical state and demographics (Table I). For more details regarding the features listed in the table, please refer to the "Construction of the Feature Set" heading in our Materials and Methods section.

TABLE I: BASELINE DEMOGRAPHIC AND MOVEMENT FEATURES FOR N = 26 RANDOMIZED STUDY SUBJECTS, COLLECTED AT THE REHABILITATION INSTITUTE OF CHICAGO (NOW SHIRLEY RYAN ABILITYLAB) IN THE YEARS 2012 – 2014.

	Control Arm (n = 13)			Treatment Arm (n = 13)			
Movement Features[a]	$\overline{mean} \pm$ sd	$\overline{max} \pm$ sd	$\overline{s^2} \pm$ sd	$\overline{mean} \pm$ sd	$\overline{max} \pm$ sd	$\overline{s^2} \pm$ sd[b]	Abbr.
Reaction Time (s)	0.121 ± 0.063	0.508 ± 0.215	0.028 ± 0.023	0.196 ± 0.145	0.704 ± 0.504	0.075 ± 0.129	
Trial Time (s)	8.748 ± 2.131	9.690 ± 1.107	1.090 ± 1.266	8.630 ± 1.735	9.984 ± 0.078	2.186 ± 2.515	
Initial Direction Error (rad)	0.806 ± 0.293	2.495 ± 0.407	0.913 ± 0.336	0.915 ± 0.152	2.524 ± 0.268	1.007 ± 0.257	IDE
Pre-Movement Speed (m/s)	0.032 ± 0.021	0.120 ± 0.079	0.001 ± 0.002	0.024 ± 0.013	0.107 ± 0.074	0.001 ± 0.001	PMS
Maximum Speed (m/s)	0.264 ± 0.054	0.386 ± 0.101	0.005 ± 0.004	0.273 ± 0.069	0.407 ± 0.071	0.004 ± 0.002	
Initial Movement Ratio	0.264 ± 0.157	0.791 ± 0.153	0.080 ± 0.040	0.280 ± 0.139	0.803 ± 0.146	0.081 ± 0.031	IMR
Speed Ratio	0.571 ± 0.204	1.00 ± 0.00	0.094 ± 0.046	0.615 ± 0.167	1.00 ± 0.00	0.112 ± 0.027	
Path Length Ratio	3.448 ± 1.178	5.613 ± 2.595	1.249 ± 1.701	3.307 ± 0.999	5.913 ± 2.533	1.557 ± 1.639	PLR
Number of Speed Peaks (count)	12.81 ± 5.073	19.85 ± 7.105	16.63 ± 14.01	11.45 ± 3.220	18.38 ± 3.595	15.81 ± 9.676	NSP
Maximum Perpendicular Distance (m)	0.099 ± 0.030	0.158 ± 0.053	0.001 ± 0.001	0.089 ± 0.035	0.151 ± 0.060	0.001 ± 0.001	MPD
Percentage of Movement in the Target Direction (%)	44.4 ± 15.4	57.7 ± 14.1	0.60 ± 0.30	46.3 ± 10.5	63.5 ± 13.6	0.80 ± 0.50	PMTD
Arrest Period Ratio	0.375 ± 0.101	0.654 ± 0.138	0.022 ± 0.009	0.403 ± 0.120	0.674 ± 0.122	0.022 ± 0.011	APR
Patient Characteristics[c]							
Age (yrs)		55.54 ± 12.63			55.23 ± 9.11		
Height (in)		67.62 ± 3.36			69.85 ± 4.01		
Mass (lbs)		190.08 ± 27.56			214.31 ± 47.41		
Months Since Stroke (months)		65.15 ± 70.32			64.00 ± 40.96		
Females (count)		5(38.5%)			5(38.5%)		
Left Dominant Arm (count)		3(23.1%)			2(15.4%)		
Left Affected Arm (count)		9(69.2%)			5(38.5%)		
Affected Arm = Dominant Arm (count true)		5(38.5%)			6(46.2%)		
Hemorrhagic Stroke (count)		5(38.5%)			4(30.8%)		
Cortical Lesion (count)		5(38.5%)			8(61.5%)		
Subcortical Lesion (count)		9(69.2%)			6(46.2%)		
Brainstem Lesion (count)		1(7.7%)			3(23.1%)		
Initial Fugl-Meyer Score (Fugl-Meyer Units)		38.31 ± 6.77			36.08 ± 6.86		Initial UEFM
Initial Wolf Motor Function Time Score (sec)[d]		12.35 ± 16.40			8.59 ± 6.99		Initial WMFT
Initial Box-and-Blocks Score (number of blocks)		27.54 ± 15.15			27.00 ± 9.06		Initial BB

[a] Features are based on 20 trials per subject

[b] s^2 denotes the variance of a feature, $\overline{s^2}$ represents the mean of this variance for a feature

[c] $\pm$ notation refers to mean $\pm$ sd

[d] WMFT is timed and therefore inversely related to ability

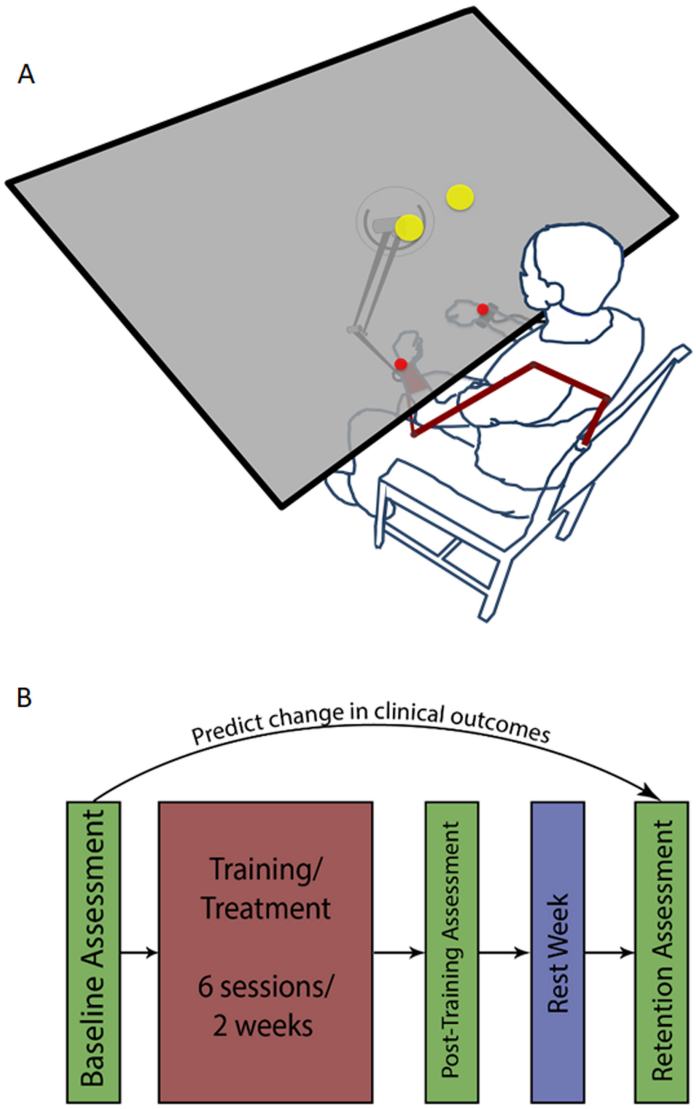


Figure 1: **Experiment design.** (A) Patients reach bimanually to two targets, pseudorandomly placed at one of four possible locations in the workspace. Patients return to a central "Home" position after every center-out reach. Patient's wrists are represented by red spheres. Their task is to get the red spheres inside the yellow targets at the same time. (B) Patients underwent six treatment sessions over two weeks. They were evaluated prior to and immediately after training, as well as one week post-training. Our goal is to use the initial assessment, clinical, and demographic information (Table I) to predict the change in outcome measures between the baseline assessment and the final (retention) assessment.

2.3 Results/Discussion

Only some models were able to effectively predict clinical outcomes (WMFT and UEFM) using quadratic polynomials of the features given in Table I, and pairwise interactions (see Methods). Next, we used only those successful models to identify and rank salient predictors of these clinical outcomes, and these rankings were consistent in 4-fold cross-validation with 100 repeats. These two steps are described in more detail in the sections below.

2.3.1 Regression Performance

Of the regression models we tested, we found that LASSO (29) models performed best. We also employed *elastic nets*(78), which generalize the LASSO method, Random Forests(30), and LARS (31) to simultaneously conduct a sensitivity analysis of our choice of LASSO penalty and establish benchmarks for predictive ability. We relied on both the RMSE; Table I) and the coefficient of determination (R^2; Figure 7) to quantify model performance.

LASSO's success may be unsurprising, because it has the advantage of being able to narrow down the high-dimensional feature space to identify important features in cross-validation, and demonstrate the impact of those features on clinical outcomes. Prediction using LASSO was comparable to both the performances of LARS and a range of elastic net models with varying parameterizations (Figure 7). However, unlike elastic nets and LARS, LASSO *shrinks* the coefficient of features deemed not consequential and, thus, leads to a more parsimonious model compared to the other methods. We decided to further examine LASSO models more closely to determine the smallest subset of features that can produce the same high performance as the other methods (which tend to use many more features).

Our models predicted the change in patients' WMFT with better coefficient of determination ($\mathtt{mean} \pm \mathtt{sd}$: $86.07\% \pm 5.26\%$) than the change in Upper Extremity Fugl Meyer ($65.34\% \pm 17.45\%$). Interestingly, first-order LASSO models performed better for predicting WMFT change, while second-order models (using the base 51 features, their interactions, and quadratic terms) performed better in predicting UEFM change.

The poorer prediction performance with UEFM may be partly due to its coarse, discrete nature. That is, the continuous nature of the model prediction is more precise than the discretely reported UEFM. This may inflate the resulting RMSE values. While categorization of UEFM and the subsequent use of logistic and multinomial models may offer a remedy, there are no clear guidelines for establishing thresholds for discretizing the measures.

We found that for WMFT, first-order models mostly performed better than the more complex second-order ones (Figure 6). This is likely due to the increased likelihood of over-fitting in the second-order case, leading to poorer performance under cross-validation. The pairwise interactions and second-order terms are also likely to magnify the multicollinearity problem. These issues are well known to degrade performance of the LASSO algorithm (79; 29). The lack of an advantage to using second-order models leads us to conclude that first-order LASSO models are sufficient for making predictions in the WMFT case. We are less confident in recommending this for the UEFM predictions because of lower predictability. An added advantage of first-order models is that they are easier to interpret and understand, and relationships between predictive features and the outcomes can be more readily translated into actionable clinical interventions.

In contrast, Random Forests exhibited poor performance with very low R^2 (for both first- and second-order models). R^2 was $< 2.24\%$ and $< 4.68\%$ for UEFM and WMFT, respectively. These R^2 values were consistent with high RMSE (≈ 5 UEFM and ≈ 6.2 WMFT). This is likely due to either collinearity or sparsity (or both). The large 51-feature input space makes this problem implicitly sparse and the sparsity of the outcomes can be observed in Figure 2, where several extreme values are represented by only a single subject. Cross validation removed points that often led to complete exclusion of some regions of the input/output spaces, worsening predictions. Collinearity by itself has little effect on a Random Forests(80; 81) because repeated random re-sampling can discriminate between collinear features; however, collinearity is likely to compound the effects of sparsity.

Another issue is that Random Forests tend to under-perform when the proportion of features that should have been selected (also called *consequential covariates*) is small ((82), section 15.3.4). Poor performance is not unreasonable if only some features are pertinent to clinical outcome, and the likelihood of randomly selecting any consequential feature at each split in a decision tree is lower. However, our lack of *a priori* knowledge of salient features motivated this analysis. Nevertheless, because of poor prediction from Random Forests, our subsequent analytic approach below focuses only on the LASSO approach.

It is important to note that UEFM and WMFT measure different aspects of movement difficulties. Not only do UEFM and WMFT have an inverse relationship, UEFM measures motor ability while WMFT measures function. Concretely, where UEFM evaluates how well patients perform fundamental actions such as bending their elbow, WMFT measures the time it takes

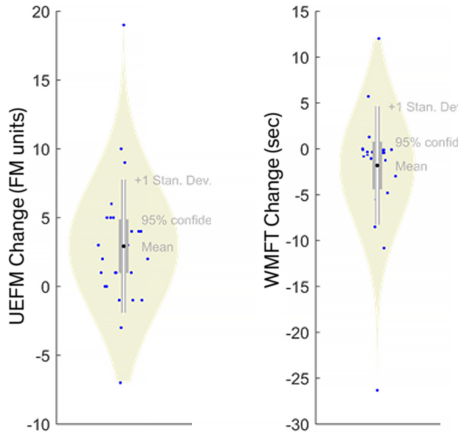


Figure 2: **Model output distributions showing sparsity.** Both UEFM and WMFT include few (sometimes one) subject(s) representing extreme values of clinical change. Clinical changes for these subjects will be difficult to predict under cross validation.

patients to effectively perform functional tasks such as grasping and transporting an object. In the sample of patients involved in this study, while participants with higher impairment levels (low UEFM) needed more time to complete functional tasks (high WMFT) (Figure 10).

2.3.2 Feature Importance

To identify a reduced feature space we were interested in the relative importance of predictor features. We used 4-fold cross-validation enumerate how often a feature was selected (feature sufficiency), and feature omission allowed us to measure the impact of removing a feature on model prediction ability (feature necessity). The features selected most frequently by LASSO to predict UEFM change (red diamonds on Figure 3) were age (younger patients improved more), height (taller patients improved more), and affected arm (non-dominant arm improved more). Next on the list were several *movement* features related to either speed or stability:

variance of speed ratio (higher variance improved more), number of speed peaks (fewer speed peaks improved more), and maximum speed (lower speeds did better). While Figure 3 shows the top 10 features, a full list of all features is shown in Supporting Information (Figure 9).

These results are consistent with previous research. It is known that cognitive performance declines with age (83). Since height correlates with arm length(84), we posit that taller patients had an easier time reaching their virtual targets. Our results also agree with the effect of handedness on stroke recovery discussed in(85). Higher variability leads to more comprehensive and often better learning (86). Stroke survivors had fewer submovements as they improved(87).

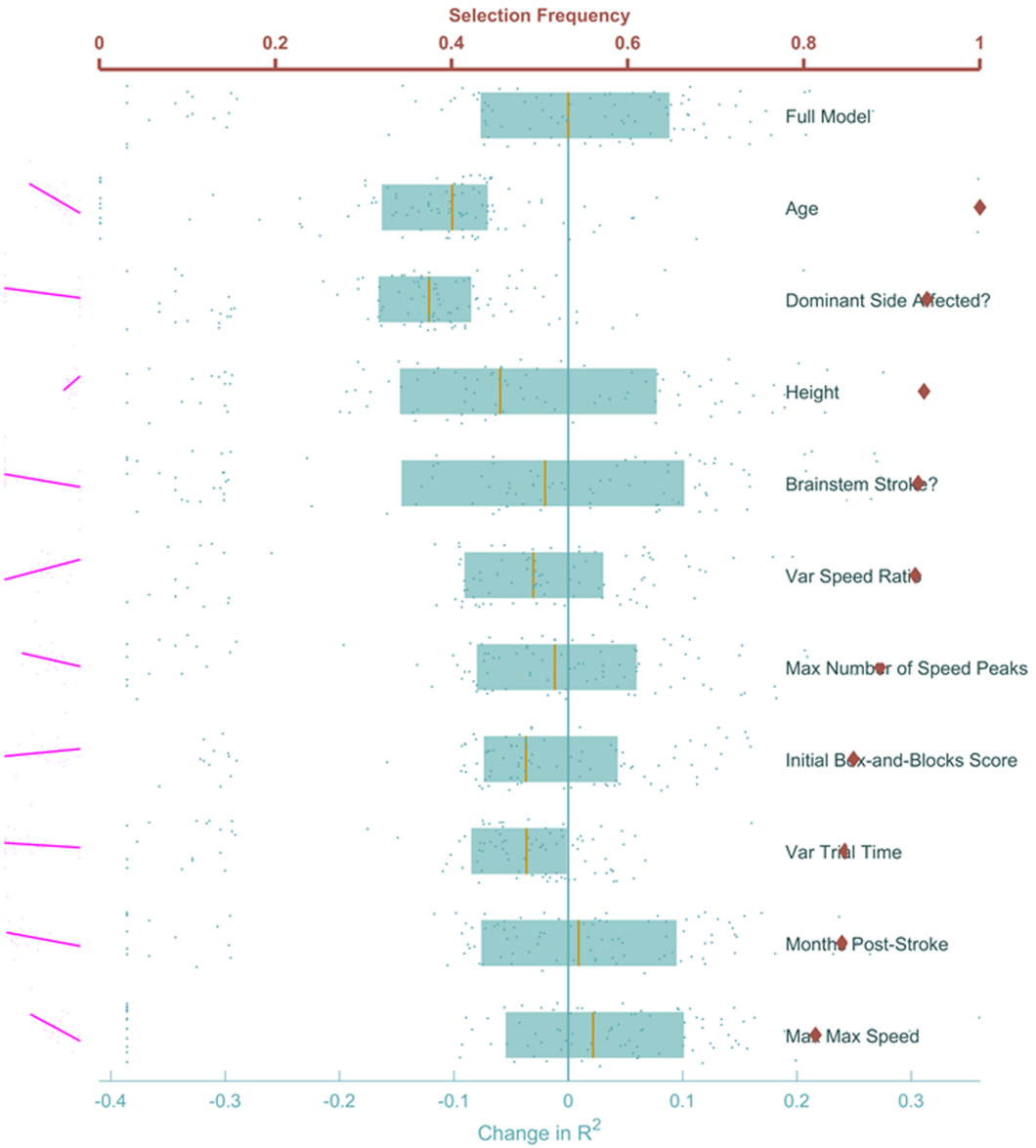


Figure 3: **Feature ranking to predict UEFM change.** Red diamonds mark the proportion of times during cross-validation where each feature was selected, with the red horizontal axis on top showing the range. The effect of removing each feature on the adjusted coefficient of determination R^2 is shown in blue, each dot represents a single cross-validation run. Blue boxes show the lower quartile, median, and upper quartile of the R^2 for each feature. The bottom horizontal axis measures the change in this R^2 with respect to the median R^2 of the full model, which is represented by the vertical blue line. The full model is shown at the top for comparison. None of the features stood out as clearly redundant or clearly essential for the model. Pairwise correlations of each feature with the outcome are shown in magenta to the left of each row

Another way to gauge importance was to see how model fit was influenced by excluding a feature (blue on Figure 3). The model fit median R^2 (over the 100 cross-validation runs) was most negatively impacted when either age, height or dominant side were removed, consistent with rank results above. Interestingly, removing some features resulted in changes in R^2 that spanned wide ranges and even had median *improvement* (such as when maximum trial time and sex were removed), suggesting that it was better to exclude these features from consideration. With this amount of variable data, no concrete statements can be made on the importance of these features on UEFM change.

Feature importances were more distinguishable when we inspected WMFT changes, with all of the top 10 selected by LASSO in nearly 100% of the cross-validation runs (Figure 4, Red). These were initial WMFT score (severely impaired improved more, which had the strongest correlation with WMFT change, $r = 0.78$), affected hemisphere (left side affected improved more), max path length ratio (higher ratios improved more), variance in number of speed peaks (more consistent improved more), mean max speed (this time higher speeds improved more), arrest period ratio (less time moving improved more), and age (younger improved more). It is also important to point out that mean maximum speed was next in correlation strength ($r = 0.24$).

That change in WMFT scores was best predicted by patients' initial performance shows that the Wolf Motor test itself is a robust, consistent measure of functional recovery. The other features deemed important to the prediction indicate possible interventions to improve these WMFT changes.

These results were also supported by feature exclusions. Predictions were most negatively impacted when initial WMFT was excluded, as well as the affected hemisphere (Figure 4, blue). Because our evaluations were on the cross-validation data, the model sometimes improved when excluding a feature. This was particularly true with max and variance of the speed ratio (Figure 4). Unlike UEFM, the effects on the WMFT R^2 had a smaller variance over cross-validation runs. The high variance and improvement of the mode upon their exclusion are key indicators that more data is necessary before conclusive statements regarding their importance may be made. The full list of feature ranks for predicting WMFT change is shown in the Supporting Information section (Figure 8).

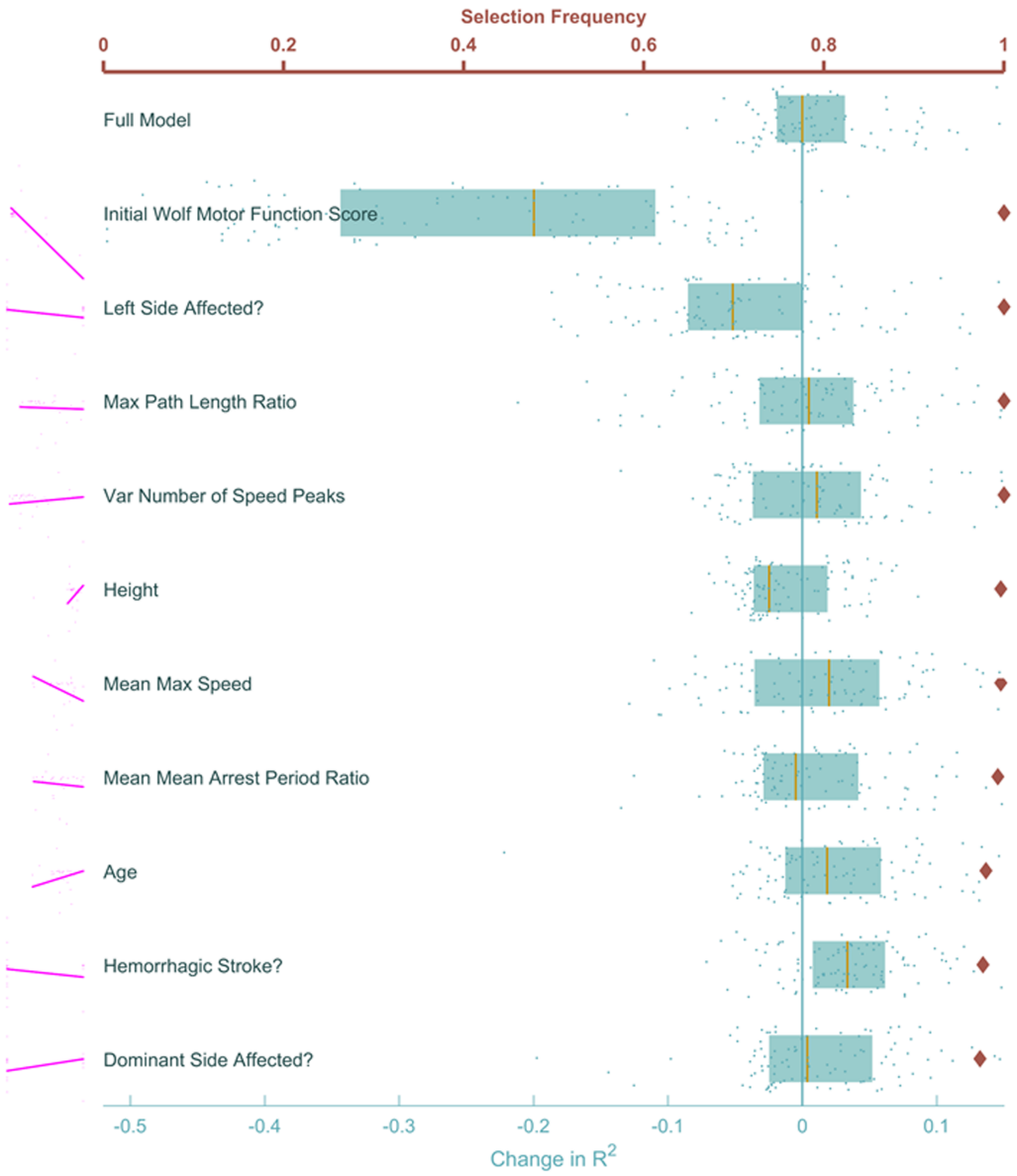


Figure 4: **Feature ranking to predict WMFT change.** Similar to Figure 3, proportion of cross-validations each feature was selected is shown in red. The blue points and boxplots show the effect of excluding each feature and rerunning the LASSO models with cross-validation. A patient's initial WMFT score and whether their left side was affected by the stroke are the two features whose removal most negatively impacts the prediction. Conversely, removing information about the patient's mass, stroke type and location was most helpful to the model, improving the adjusted R^2. Notable among the top ten features is mean max speed, which showed a strong correlation with the outcome, indicating patients who were faster on the first day improved more on the WMFT scale.

In any case, features nearly always selected by LASSO were deemed essential to prediction, while features never selected were deemed ineffectual. In the central region (diamonds roughly between 0.3 and 0.7 in Figure 3 and Figure 4) were features the LASSO model could not reliably determine were important to the model. Generally, selection of these features was contingent on which data were available to the model at the time they are considered by LASSO.

As mentioned at the beginning of this section, each of these ranking methods, though powerful, has weaknesses. When shrinking the list of features, the LASSO algorithm elects to keep the first strongly predictive features it comes across, and shrinks all features highly correlated with those to zero, as their impact is negligible given the first features are already in the model. Therefore, features chosen using this method should be treated as motifs, whereby each is interpreted not as the raw quantity it represents, but as a thematic property of a patient's condition or ability to move that is helpful in predicting the outcome.

On the other hand, measuring the consequence of excluding an individual feature is not always accurate, because we cannot control for how other correlated features compensate for the drop in R^2. One or more other factors may compensate, resulting in a small or negligible effect on R^2. It is reasonable to expect that, for at least some of the features, the change in R^2 reported using our method was small in spite of the importance of that feature.

Random Forests provide an estimate of feature importance that is robust against high feature correlations and not vulnerable to some of the weaknesses of LASSO for feature ranking. We were unable to use that algorithm to rank features, however, due to small number of samples.

As more clinical data appears as rehabilitation science matures, such tree methods may better inform our understanding of outcome predictions and feature importance.

It is important to note that our intervention (whether the patient received Error-Augmentation treatment) was not deemed useful or important in predicting our clinical outcome scores. Surprisingly, this is in spite of showing a significant benefit to the treatment type in this randomized, controlled clinical study. If large set of features used here demonstrates that other (perhaps superfluous and uncontrolled) features are much better predictors of outcome, one questions the meaning of the classic clinical test. Truly effective interventions should appear as consistent predictors. We posit here that our validation-intensive methods can verify if detected effects are confounded by other factors.

One might be concerned about whether we would have used all the movement data across the 20 trials undertaken by each patient rather than construct our features from summary statistics, where information might be lost or hidden. However, we believe there is an advantages to using summary statistics because they can robustly resist the spurious influences of random measurement error, while also allowing for easy interpretation. Our chosen breadth of different types of summary statistics (mean, maximum, and variance) led to effective prediction models that correlate with the clinical interpretation of central tendencies, best/worst performance, and consistency.

Another concern is that only main effects and not interactions were chosen from our data. However, our fundamental goal was to identify single factors that were related to outcomes, therefore we focused our feature ranking analysis on the linear models. Interaction models not

only require more data, they also are more difficult to interpret in a predictive model with no a priori selected primary exposure variable.

What is most important is the implications these results might have for patients. Younger, taller, non-dominant-affected arm individuals were more likely to improve their abilities (UEFM). Younger, more severely impaired, left arm affected individuals were more likely to improve their function (WMFT). What is also important is what these results might suggest for altering treatment strategies. While fewer speed peaks and lower maximum speeds were loosely related to ability as measured by UEFM gains, UEFM models were less successful and therefore not as reliable as the WMFT models. For the WMFT predictions, the set of highly ranked features included movement *speed*. Speed also had a strong correlation with outcome and suggests that interventions focusing on speed might improve prognosis. It remains to be seen whether a therapy that encourages faster movements might lead to better functional recovery.

2.4 Conclusion

Changes in motor ability (UEFM) and motor function (WMFT) can both be predicted by our models. Change in WMFT is easier to predict since it is a continuous measure. Both changes in UEFM and WMFT can be linked to specific movement features as well as patient demographics and clinical characteristics. Our validation approach also allowed us to measure the certainty of our findings. Since we are unable to affect demographics or clinical characteristics, features that we can influence during rehabilitation are the most critical. This work suggests that speed would be a good first target for further study.

2.5 Materials and Methods

2.5.1 Ethics Statement

This work was approved by the University of Illinois at Chicago's Institutional Review Board and Northwestern University's Institutional Review Board. This work conforms to the Declaration of Helsinki for research involving human subjects. All participants provided written consent to participate in the study using consent procedures approved by both Institutional Review Boards.

2.5.2 Patient Selection & Initial Evaluation

We enrolled twenty-six chronic hemiparetic stroke survivors in our study. Participants had mild to moderate impairment, determined by their intake Fugl-Meyer scores (range 25-49) and were selected according to the criteria outlined in (44; 88). Patients' stereoscopic vision was tested using the Stereo Fly Test. Their reaching abilities were then evaluated before starting the study under similar conditions. Patients were instructed to reach with both arms in parallel, without crossing the midline, to two targets in a three-dimensional virtual reality environment (Figure 1A). Each subject underwent baseline and post-intervention evaluations consisting of a battery of clinical assessments performed by a therapist, followed by 20 bimanual reaches (trials) in the virtual reality environment, each to one of four target locations chosen pseudorandomly.

2.5.3 Intervention

Patients were block-randomized controlling for age and impairment as closely as possible. Both patient groups trained for two weeks using a Phantom® 3.0 robot arm. The control group received no intervention and used a passive robot arm, while the treatment group experienced

disturbance to their paretic arm, in the form of visual and haptic Error Augmentation (EA) (89; 44). All patients were evaluated again immediately after the end of training, with final evaluation taking place one week later to assess longer-term recovery effects. Our main outcome measures were changes in the patients' clinical scores, as evaluated by a therapist, between the first evaluation (prior to beginning the study) and the final evaluation three weeks later. Specifically, our clinical outcomes were the patients' UEFM, which measures motor ability, and the WMFT, which measures completion time for functional tasks. This protocol is summarized in Figure 1B.

2.5.4 Construction of the Feature Set

We gathered a total of 51 features from two sources, either demographic/physiological characteristics, and descriptive statistics of movement (Table I). Demographic and physiological features were denoted Z_{il}, $i = 1, \ldots, N$, $l = 1, \ldots, q$, where q is the number of variables, were noted at baseline. A battery of p measured movements were observed in $T = 20$ trials for each i-th subject. These measurements, X_{ijk}, $j = 1, \ldots, p$, $k = 1, \ldots, T$, in $T = 20$, were used to compute baseline summary features across trials for each study participant: The mean ($\bar{X}_{ij.}$), maximum ($M_{ij.}$), and variance($V_{ij.}$) movement features. A descriptive summary of demographic and moment variables is provided in Table I.

Features were primarily based on common metrics or were reported in previous research (90; 72; 91); a few of the features were newly explored in this work. These included: *(a)* performance-related measures evaluating error, speed, and reaction time, *(b)* descriptive features such as hand path length and trial time, and *(c)* patients' demographic and clinical characteristics such as

height, weight, stroke location, affected side, and initial clinical scores. Ultimately, our first-order feature set contained 51 features, and used summaries of the movement features across the 20 evaluation reaches (mean, maximum, and variance of each feature).

Most of the features we used (as detailed in Table I) are fairly straightforward. A few are, however, somewhat ambiguous. For trial k, we defined speed ratio as the speed of the first launch divided by the maximum speed, while path length ratio is the distance traveled by the subject's arm divided by the straight line distance from the home position to the target. We defined reaction time as the time between the appearance of the target and the subject crossing our pre-defined threshold of $0.06 ms^{-1}$. Mean Arrest Period Ratio (APR) is the time the subject spent below 10% of their maximum speed for that trial divided by the total trial time. Finally, Percentage of Movement in the Target Direction (PMTD) is defined as the proportion of the distance traveled during a trial in the effective direction to the target, defined formally as the sum of the projections of the distance traveled between two time samples onto the straight line path to the target, divided by the total hand path length.

2.5.5 Predictive Models

Movement features and patient characteristics were used to predict change in clinical outcomes: UEFM and WMFT. Since the number of possible predictive features is larger than the number of observations (patients), the most likely models to succeed used algorithms that shrink the number of features to avoid overfitting. These models included Elastic Net(78). Elastic Net employs penalized linear regression with a parameter ($0 \leq \alpha \leq 1$) that balances l_1 and l_2 norm penalties (more on that below). On one end ($\alpha = 0$) there is ridge regression(92),

purely penalizing the l_2 norm of the coefficients in the model, and on the other ($\alpha = 1$) is LASSO (29), purely penalizing the l_1 norm of the coefficients. We employed LARS (31)), a less greedy forward feature selection algorithm. Finally, we tried an algorithm that utilizes decision trees to make predictions (namely Random Forests (30)).

Our first-order prediction models used 51 features (many more for second-order case) to predict changes in clinical scores for 26 stroke survivors. Since this is an overdetermined problem that would be guaranteed to cause overfitting, the most likely algorithms to succeed would have to reduce the number of features used by the model. One such algorithm that we implemented was Elastic Nets, represented by the following formula:

$$\min_{\beta_0,\beta}\left\{\frac{1}{N}\sum_{i=1}^{N}(y_i-\beta_0-\mathbf{x_i}^T\beta)^2\right\},\ \text{such that}\ \sum_{j=1}^{p*}|\beta_j|\le t_1, \sum_{j=1}^{p*}|\beta_j|^2\le t_2 \tag{2.1}$$

where $\mathbf{x_i} = \left[\bar{X}_{i1\cdot},\ldots,\bar{X}_{ip\cdot},M_{i1\cdot},\ldots,M_{ip\cdot},V_{i1\cdot},\ldots,M_{ip\cdot}Z_{i1},\ldots,Z_{iq}\right]$, $p* = 3p+q$, t_1 and t_2 are regularization terms related to the penalty, placing an upper limit on the sum of the first and second norms of predictor coefficients, and β_j is the coefficient of the j-th feature. In Lagrangian form:

$$\min_{\beta\in\mathbb{R}^{p*}}\left\{\frac{1}{N}\parallel \mathbf{y}-\mathbf{X}\beta\parallel_2^2+\alpha\parallel\beta\parallel_1+(1-\alpha)\parallel\beta\parallel_2\right\}\text{s.t. }\alpha=\frac{\lambda_1}{\lambda_1+\lambda_2}, \tag{2.2}$$

where $\lambda_1 and \lambda_2$ are the penalty for the sum of l_1 and l_2 norms of the coefficients, respectively. We tested a range of α values from 0 to 1 with increments of 0.1 – where $\alpha = 0$ corresponds to a penalty purely based on the l_2 norm of the coefficients (ridge regression) and $\alpha = 1$ corresponds to a penalty based purely on the l_1 norm (LASSO). LASSO drives a lot of the predictor coefficients to zero and simplifies the resulting model(29) but sometimes over-regularizes. On the other hand, ridge regression keeps all coefficients in the model and drives the less useful ones close to zero, but is more difficult to interpret since it does not remove any features from the model(92). Intermediate α values attempt to balance removing predictors from the model and driving their coefficients close to zero.

LARS adds coefficients to the model in a stepwise manner starting with the most correlated with the outcome being modeled, then adding coefficients in order of correlation with the residual from the previous step. LARS may behave in a more stable fashion than regularized regression in some cases but is also highly affected by noise[(31), discussion by Weisberg]

The chosen model in the case of regularized regression and LARS was the maximum λ where mean cross-validation error was within 1 standard error of its minimum value, as in (29).

Random Forests constructs an ensemble of decision trees, each based on a randomly chosen subset of the observations and features (therefore providing validation via random sampling, and ensuring the trees do not overfit). Each decision tree is constructed to minimize the mean square error at each split, and contributes a "vote" to the ensemble. The value predicted by the random forest for a new observation is either a majority vote (classification problems) or a mean predicted value (regression problems).

For Random Forests, we built large ensembles (50,000 regression trees) with 100 repeats of 4-fold cross validation. The large number of trees was used to ensure the algorithm has adequate usage of each feature to assess its importance, especially in the case of second-order models (> 1300 predictors).

Estimation was performed under cross-validated to avoid over-fitting and reduce the influence of outliers. Over-fitting was a concern because of the limited sample size and the eventual inclusion of second-order terms comprised of quadratic and pairwise interaction variables. We evaluated both first- and second-order models by looking at their variance explained (measured using adjusted R^2) and prediction error distribution (RMSE). The process was identical for predicting both UEFM and WMFT changes. We repeated 4-fold cross-validation 100 times to obtain a range for each prediction quality metric, and ensuring we took into account different data splits.

2.5.6 Ranking the Features

Since LASSO regression has no built-in method for ranking model predictors and reducing the dimensionality of the input feature space needed to predict UEFM and WMFT changes, we devised two methods to unpack our LASSO models and determine the relative importance of features instrumental to clinical outcome prediction.

First, we examined the shrunk feature set resulting from each of our 100 4-fold cross-validation repeats. We used the proportion of these repeats each feature was selected as our main measure of that feature's importance for prediction. This gave each feature a rank from 0 (always shrunk, unimportant for prediction) to 1 (never shrunk, essential for prediction).

Second, we excluded individual features and calculated the difference in prediction R^2 when the feature was present and the $R^2_{(-j)}$ when the feature was removed. We again used 4-fold cross-validation with 100 repeats. To avoid making a distributional assumption for the change in the coefficient of variation, $\Delta_j = R^2 - R^2_{(-j)}$, we used its median as a measure of relative importance. Features whose removal resulted in larger Δ_j were deemed more important to the model.

Random forests are a powerful method(30) for ranking features in high-dimensional data by their relative importance in predicting an outcome variable. However, we saw only limited success in their ability to identify important features in this data set. This is primarily due to their failure to effectively predict changes in clinical outcomes, likely due to the small yet highly sparse data.

2.6 Supporting Information

2.6.1 Feature Balance

Cross-validation folds and repeats were balanced when considering individual features. There were no obvious outliers when examining each feature's mean for each fold during cross-validation. Cross-validation means averaged to zero across all folds/repeats. This is more complex when considering second-order models, but the basic sampling in our cross-validation was balanced.

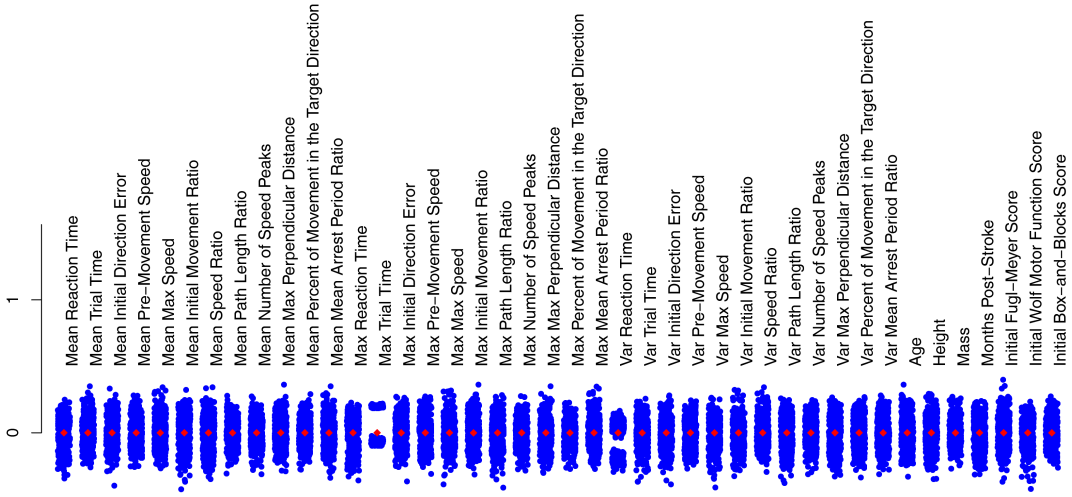


Figure 5: Distribution of model features under cross validation

2.6.2 Results for prediction RMSE

Prediction model root mean square error (RMSE). Models were successful at predicting changes in clinical outcomes, models predicting WMFT had lower errors than those predicting UEFM, Elastic Net (including LASSO and second-order Ridge) models were successful, as was LARS, while Random Forests failed. Second-order models generally did not provide an advantage over first-order models. (A) RMSE results for predicting WMFT change. (B) RMSE results for predicting UEFM change.

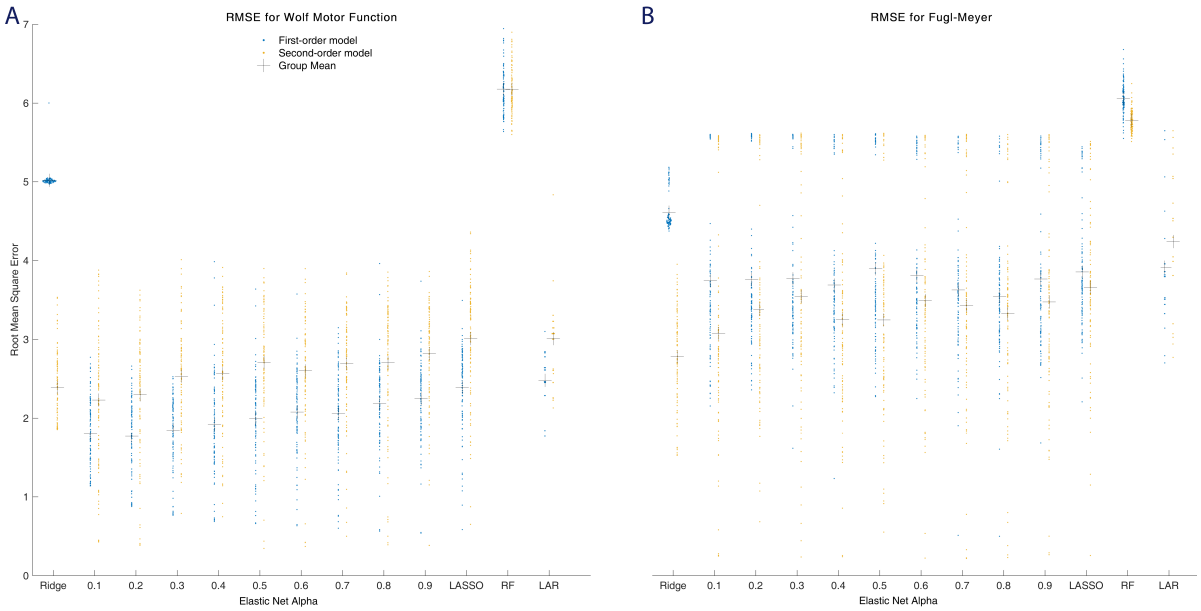


Figure 6: RMSE results of cross-validated models

2.6.3 Results for prediction coefficient of determination

Prediction model coefficient of determination. Models were successful at predicting changes in clinical outcomes, WMFT models performed better than UEFM, Elastic Net (including LASSO and second-order Ridge) models were successful, as was LARS, while Random Forests failed. Second-order models generally did not provide an advantage over first-order models. (A) Adjusted coefficient of determination R^2 results for predicting WMFT change. (B) Adjusted coefficient of determination results for predicting UEFM change. These results are consistent with RMSE finding (subsection 2.6.2). We saw higher mean R^2 with first-order than

second-order Elastic Net and LARS models, and second-order models tended to have higher variance, especially when predicting change in UEFM.

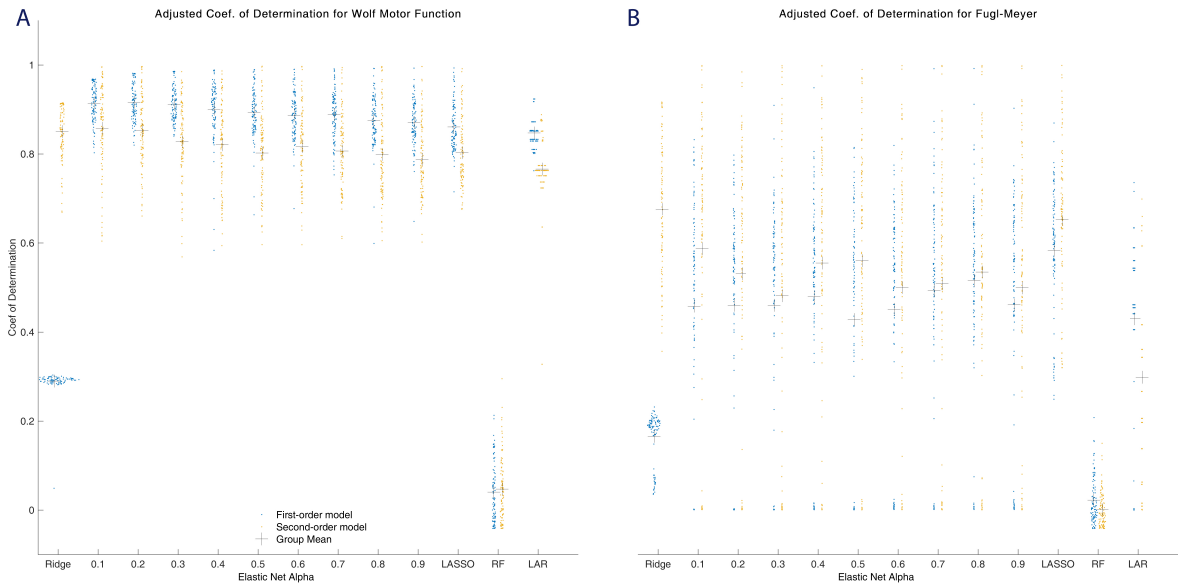


Figure 7: Coefficient of determination of cross-validated models

2.6.4 Complete feature ranks for predicting WMFT change

Full list of feature ranks predicting WMFT change. Proportion of cross-validations each feature was selected is shown in red. The blue points and boxplots show the effect of excluding each feature and rerunning the LASSO models with cross-validation. A patient's initial WMFT score and whether their left side was affected by the stroke are the two features

whose removal most negatively impacts the prediction. Conversely, removing information about the patient's mass, stroke type and location was most helpful to the model, improving the adjusted R^2. Notable among the top ten features is mean max speed, which showed a strong correlation with the outcome, indicating patients who were faster on the first day improved more on the WMFT scale.

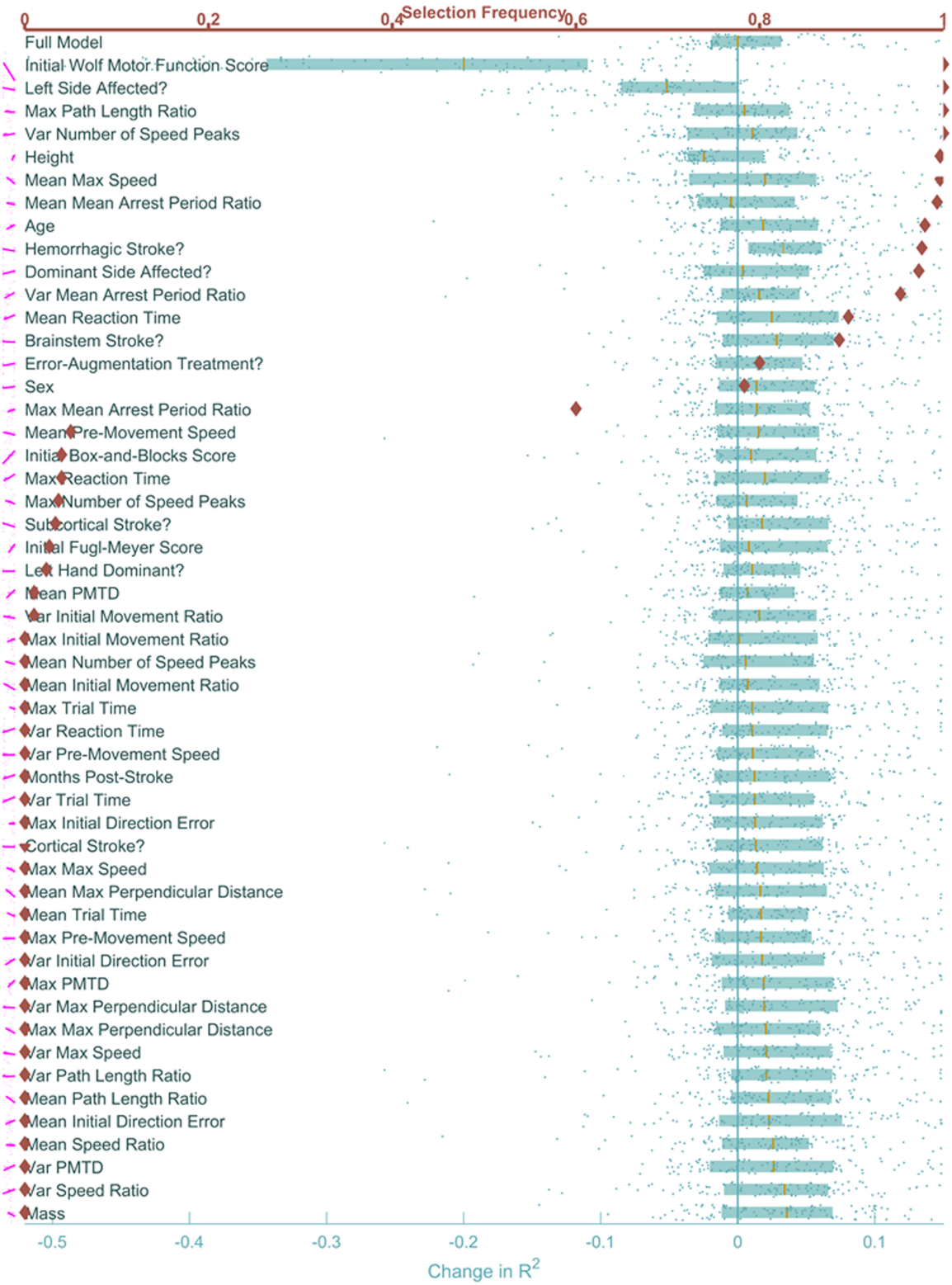


Figure 8: Feature ranking for WMFT LASSO models

2.6.5 Complete feature ranks for predicting UEFM change

Full list of feature ranks predicting UEFM change. Red diamonds mark the proportion of times during cross-validation where each feature was selected, with the red horizontal axis on top showing the range. The effect of removing each feature on the adjusted coefficient of determination R^2 is shown in blue, each dot represents a single cross-validation run. Blue boxes show the lower quartile, median, and upper quartile of the R^2 for each feature. The bottom horizontal axis measures the change in this R^2 with respect to the median R^2 of the full model, which is represented by the vertical blue line. The full model is shown at the top for comparison. None of the features stood out as clearly redundant or clearly essential for the model. Pairwise correlations of each feature with the outcome are shown in magenta to the left of each row.

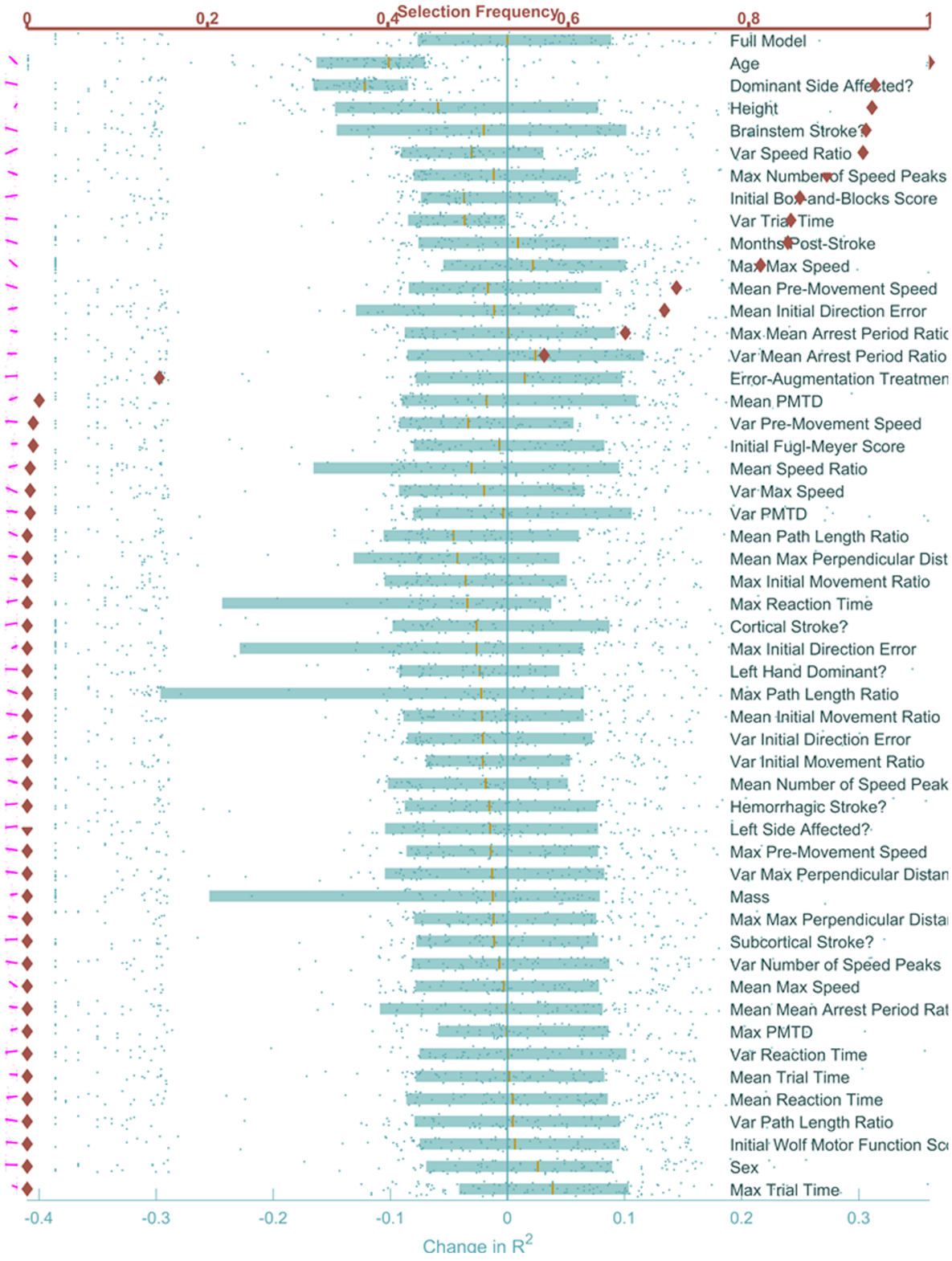


Figure 9: Feature ranking for UEFM LASSO models

2.6.6 Relationship between UEFM and WMFT

UEFM and WMFT have an inverse relationship. UEFM was more sensitive to patients with relatively higher functional ability, while WMFT was more sensitive to those with lower functional ability. WMFT scores plateaued for patients showing larger UEFM changes. This relationship between UEFM and WMFT may explain our observation that slower speed predicted better recovery for UEFM while higher speeds were predictive of faster WMFT times. Changes in clinical scores were not statistically significant after the intervention.

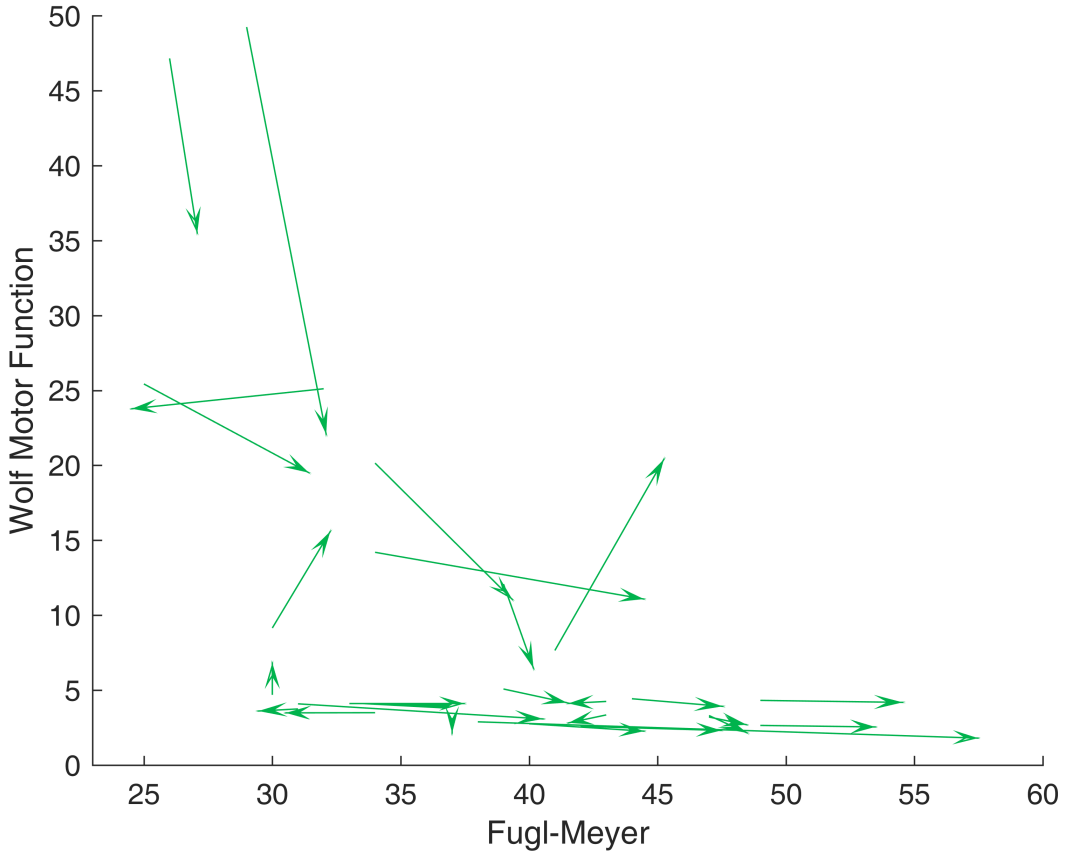


Figure 10: An inverse relationship between UEFM and WMFT

2.7 Acknowledgments

This work was made possible with funding from NIDILRR MARS3 grant 90RE5010-01-01

CHAPTER 3

INCREASING MOVEMENT SPEED IN STROKE SURVIVORS, A RANDOMIZED CROSSOVER STUDY

This work is submitted for publication as: Yazan Abdel Majeed, Saria Awadalla, and James L. Patton, "Effects of Robot Viscous Forces on Arm Movements in Chronic Stroke Survivors: A Randomized Crossover Study", Journal of NeuroEngineering and Rehabilitation, 2019.

3.1 Abstract

As speed may be linked to the ability to recover in chronic stroke survivors, we examined the effects of three candidate speed-modifying fields in a crossover design: negative viscosity, positive viscosity, and a "breakthrough" force that vanishes after speed exceeds an individualized threshold. Negative viscosity resulted in a significant speed increase when it was on. No lasting after effects on movement speed were observed from any of these treatments, however, training with negative viscosity led to significant improvements in movement accuracy and smoothness. Ultimately, our results suggest that negative viscosity may be used as a treatment to augment the training process while still allowing patients to make their own volitional motions in practice.

3.2 Introduction

Stroke neurorehabilitation often uses the unique aspects of technology to improve motor recovery. While some researchers endeavored to simply assist movement to more closely resemble

healthy patterns (93; 94; 95), others have attempted to exploit unique capabilities of robotics or graphic feedback to encourage neuroplasticity by augmenting error (96; 97; 43; 98; 99). Even some traditional physical therapy exercises use mirrors to get the paretic side of the body to imitate the non-paretic side (100). These are beneficial but far from a complete cure, and it remains to be seen what strategies emerge as optimal and what might still be left undiscovered.

An alternative strategy is to first uncover the attributes associated with better clinical movement outcomes, and then target training around these (23; 101). Our previous work (102) employed a data-driven approach to model patient improvement using metrics derived from the movements themselves. We found that patient movement speed during the initial evaluation was most predictive of clinical changes. This speed was also the most strongly correlated with changes in the WMFT, making heightened speed a possible intervention for stroke. However, before such an intervention might be tested in clinical trials, we need to establish effective methods for speeding up patients.

There are multiple possible training conditions that may achieve this increase, and here we compare three candidate classes of conditions. One approach to affect movement speed is to directly increase it with a negative viscous field; previous work (103; 104; 105; 7) showed that training with negative viscosity can improve patient movement and movement generalization abilities. Another possibility is to leverage the motor control mechanisms of error augmentation and after effects. Under this paradigm, patients would train with positive viscosity, under the expectation that their speed would increase as an aftereffect of that training when these resistive forces are removed (43; 89). Finally, some research has shown that combining a resistive

paradigm with a reward mechanism (48) may help patients learn better. In this case, patients will move in a positive viscosity field that attempts to slow them down, but moving above a certain speed is rewarded by a "breakthrough" where resistance vanishes. Subject may bias movements towards higher speeds to avoid the resistance.

In this preliminary clinical study, we simply compared the effects of these three paradigms on subject speed. Chronic stroke survivors participated in a single-visit crossover trial, where they trained for a short time under these three conditions. While we were mainly interested in the direct- and after-effects of these force paradigms on patient movement speed, we examined their effects on other movement metrics as well, such as error, efficiency, and smoothness.

3.3 Methods

3.3.1 Subject Population

We recruited 14 chronic hemiparetic stroke survivors (one patient was not able to participate after consenting due to a second stroke), 15-50 on the Upper Extremity Fugl-Meyer scale. Participants were excluded if they had multiple lesions or multiple stroke events, if they had bilateral paresis, or if they had Botox® treatments to the upper limbs within the last six months. Stroke survivors were recruited through the Clinical Neuroscience Research Registry of the Shirley Ryan AbilityLab and provided informed consent. This study was approved by the Institutional Review Boards at Northwestern University (STU00206579) and the University of Illinois at Chicago (2018-1251), and follows the guidelines of the Declaration of Helsinki. Figure 11 shows a CONSORT-style diagram (**CON**solidated **S**tandards Of **R**eporting **T**rials)

for our crossover study, and Table II shows patient information at the time of participation in the study.

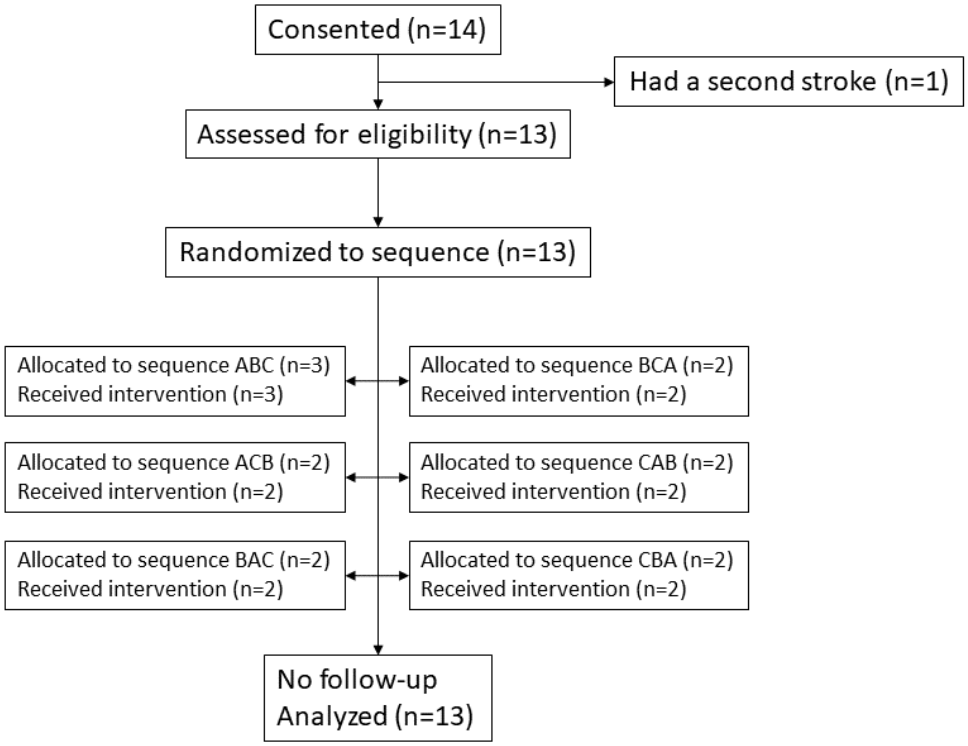


Figure 11: **CONSORT-style diagram for speed crossover.** One patient dropped from the study after having a second stroke, which was disqualifying. Data was collected and analyzed for 13 total patients.

TABLE II: PATIENT INFORMATION FOR THIS STUDY (N = 14)

Subject ID	Age (yrs)	Time Since Stroke (months)	Sex	Dominant Side	Affected Side	UEFM Score
SP01	59	29	M	right	right	26
SP02	52	70	M	right	left	44
SP03[a]	64		M			39
SP04	59	45	M	left	right	37
SP05	37	52	M	right	right	32
SP06	64	95	M	right	left	35
SP07	83	32	M	right	right	46
SP08	45	75	F	right	left	23
SP09	67	11	M	right	left	37
SP10	50	157	F	right	left	33
SP11	58	50	F	right	right	40
SP12	57	29	F	right	left	24
SP13	67	49	M	right	right	23
SP14	46	43	F	right	left	34

[a] SP03 consented but did not perform the study due to a second stroke

3.3.2 Experiment Design

Participants completed a single-visit crossover study. Subjects performed a targeted reaching task with their paretic arm attached to the Proficio® 3-DoF robot from Barrett Technologies (Figure 12). The Proficio allows three dimensional movement in a large workspace, approximating the normal range of human motion.

First, we evaluated baseline performance. Subjects reached for five minutes to targets distributed in a quarter-sphere under no robotic forces. Subjects then alternated between five minutes of reaching under each of the three experimental conditions, and five minutes of reaching with no forces to evaluate the aftereffects. We presented the three conditions to each participant in a pseudo-random order, to control for any possible ordering effect. A simplified 1-D representation of the three force types for our experimental conditions is shown in Figure 13 (A-C) and the timeline of the experiment is shown in Figure 13 (D).

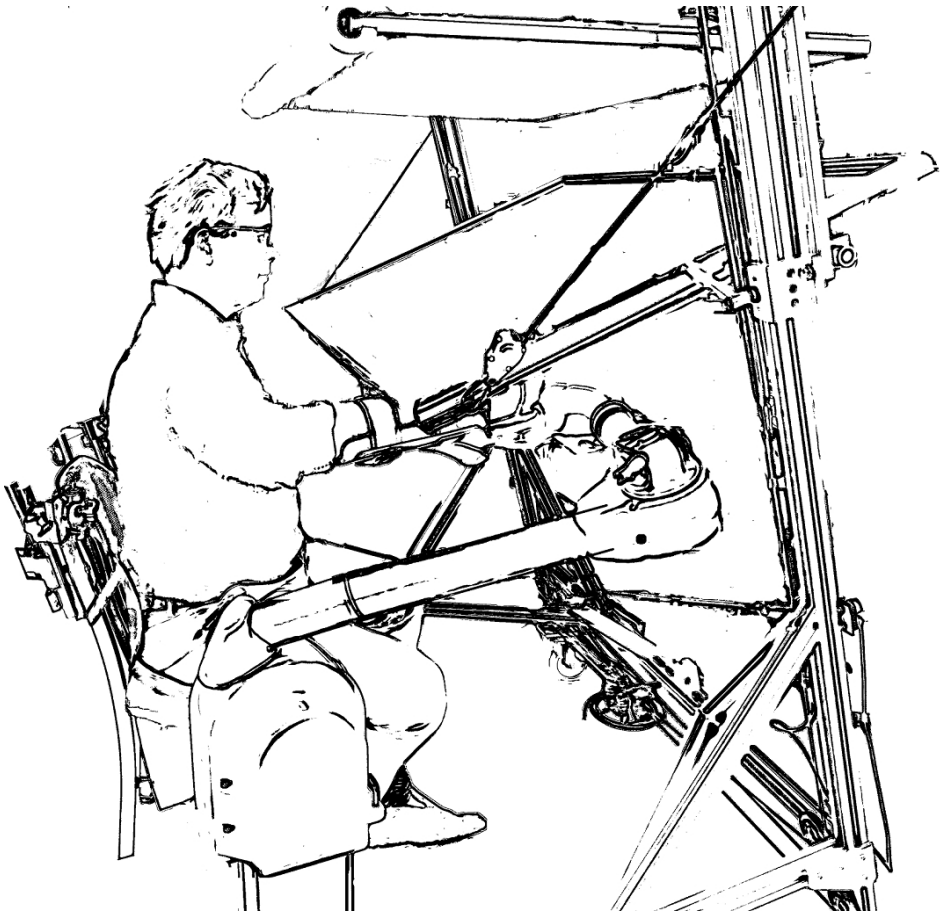


Figure 12: **Experiment Setup.** Patients reached unimanually, alternating between a central home position and randomly to one of eight different target locations. The study was conducted using the Looking Glass virtual reality system and the Barrett Proficio robot arm.

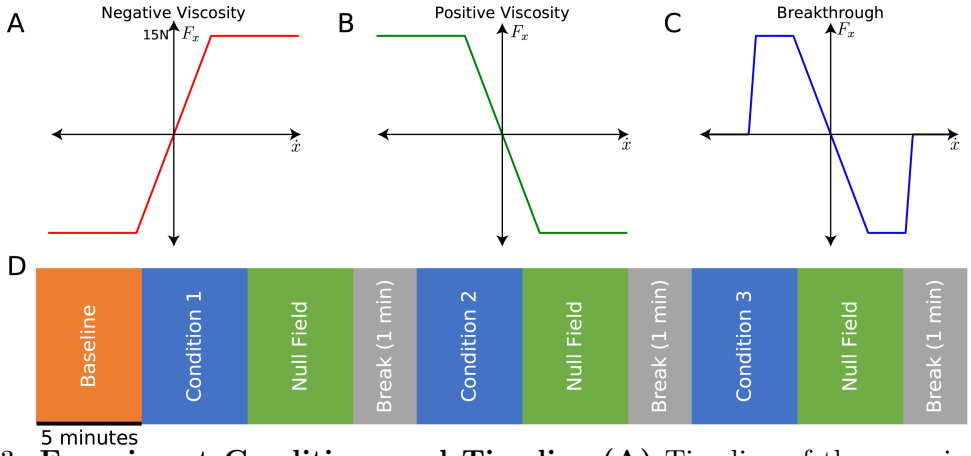


Figure 13: **Experiment Conditions and Timeline (A)** Timeline of the experiment, each "block" of trials lasted for five minutes, and there was a short break (30-60 seconds) in after each block. **(B)** Negative viscosity. Forces were proportional to velocity in all three directions, we drove the forces to zero at higher speeds for safety. **(C)** Positive viscosity. Similar to negative viscosity but the forces acted opposite to the direction of motion, slowing the participants. **(D)** Breakthrough. Forces were proportional to velocity until the participants reached 75% of their baseline speed, forces were then removed as a reward for reaching faster speeds.

3.3.3 Data Analysis

We resampled patient position data at 100Hz, then filtered it using a 7th order Butterworth filter with a 9Hz cutoff frequency. We extracted movement mean and maximum speed, and the various movement metrics that quantify accuracy (perpendicular error), efficiency (movement path distance), and smoothness (number of submovements)(102). We focused our analysis for each condition on groups of five trials: prior to exposure to the robotic forces (pre-exposure), when the forces were initially experienced (early exposure), at the end of the five minute block experiencing the forces (late exposure), the reaction to the forces being turned off (aftereffects), and the end of the null field block before the next condition is experienced (retention). We modeled study outcomes using linear mixed effects regression of the fixed conditions (vs baseline), time, and random subject effects. We compared pairwise differences between conditions using post-hoc contrasts with Tukey's adjustment for multiple comparisons.

3.4 Results

3.4.1 Movement Speed

Negative viscosity produced significant speed increases during early exposure that persisted through late exposure ($p < 0.0004$, Figure 14). There was a significant difference between the effect of negative and positive viscosity forces ($p < 0.001$), and between negative viscosity and breakthrough forces ($p < 0.03$), but not between positive viscosity and breakthrough forces ($p > 0.48$). The order with which the force conditions were presented to the patients did not have a significant effect on speed ($p > 0.16$). The aftereffect of negative viscosity was a

significant decrease in speed ($p < 0.02$) which did not persist at the end of washout (retention $p > 0.09$).

During early exposure, neither breakthrough nor positive viscosity had any significant effects on movement speed ($p > 0.11$). Both had a significant slowing effect on maximum speed ($p < 0.03$) but not on mean speed ($p > 0.055$) during late exposure. Neither had any significant effect on speed during washout ($p > 0.07$). There was no significant difference in the effects of positive viscosity and breakthrough forces during any stage of the experiment ($p > 0.48$) after correcting for multiple comparisons. The effects of negative and positive viscosity forces were significantly different ($p < 0.02$) for all but the retention stage ($p > 0.61$).

3.4.2 Other Movement Metrics

Movement error increased significantly during early exposure to negative viscosity ($p = 0.004$, Figure 15-A). The effect was no longer significant by late exposure ($p = 0.064$). However, both the aftereffect and retention had a significant reduction in movement error ($p < 0.023$). Training with positive viscosity forces yielded a significant increase in movement error during exposure ($p < 0.01$) but no significant aftereffects. Breakthrough forces had no significant effect on error during exposure, but showed a significant error reduction as an aftereffect ($p < 0.02$). Interestingly, the retention movement error was affected by the order the force types were presented to patients ($p = 0.04$), other stages were not affected by the order ($p > 0.06$).

Movement distance showed behavior similar to movement error (Figure 15-B), where negative viscosity resulted in worse performance during exposure ($p < 0.0001$) but significantly improved performance as an after effect ($p \leq 0.001$). Positive viscosity had no significant ef-

fects during any stage, while breakthrough showed a significant improvement only during the aftereffect ($p \leq 0.001$).

Movement smoothness, as quantified by the number of speed peaks, was significantly worse during exposure ($p \leq 0.0099$) and significantly better during washout ($p \leq 0.033$, Figure 15-C). Interestingly, all force types showed a significant improvement in movement smoothness in early aftereffects ($p \leq 0.031$), though only after training with negative viscosity did patients retain that improvement.

Finally, there was a significant increase in pre-movement speed during exposure to negative viscosity (Figure 15-D). There was no significant effect for any of the three force types on this pre-movement tremor by late washout, though training with negative viscosity showed an average reduction in pre-movement speed.

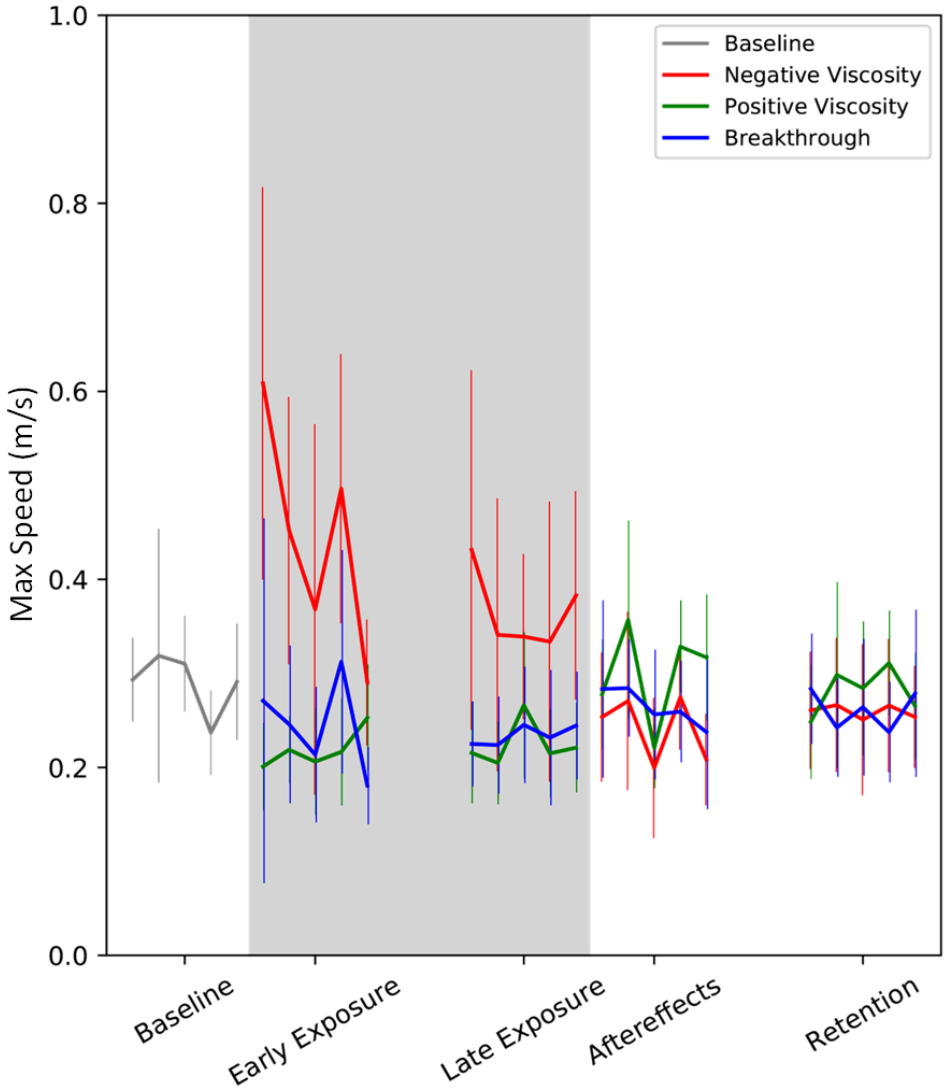


Figure 14: **Effect of experiment conditions on altering patient speed.** Error bars represent 95% confidence interval. Negative viscosity had the strongest effect on participant speed, though the aftereffect was opposite to what we hoped to achieve (participants slowed down, on average). There was no significant difference between the effects of positive viscosity and breakthrough, and no significant change from baseline.

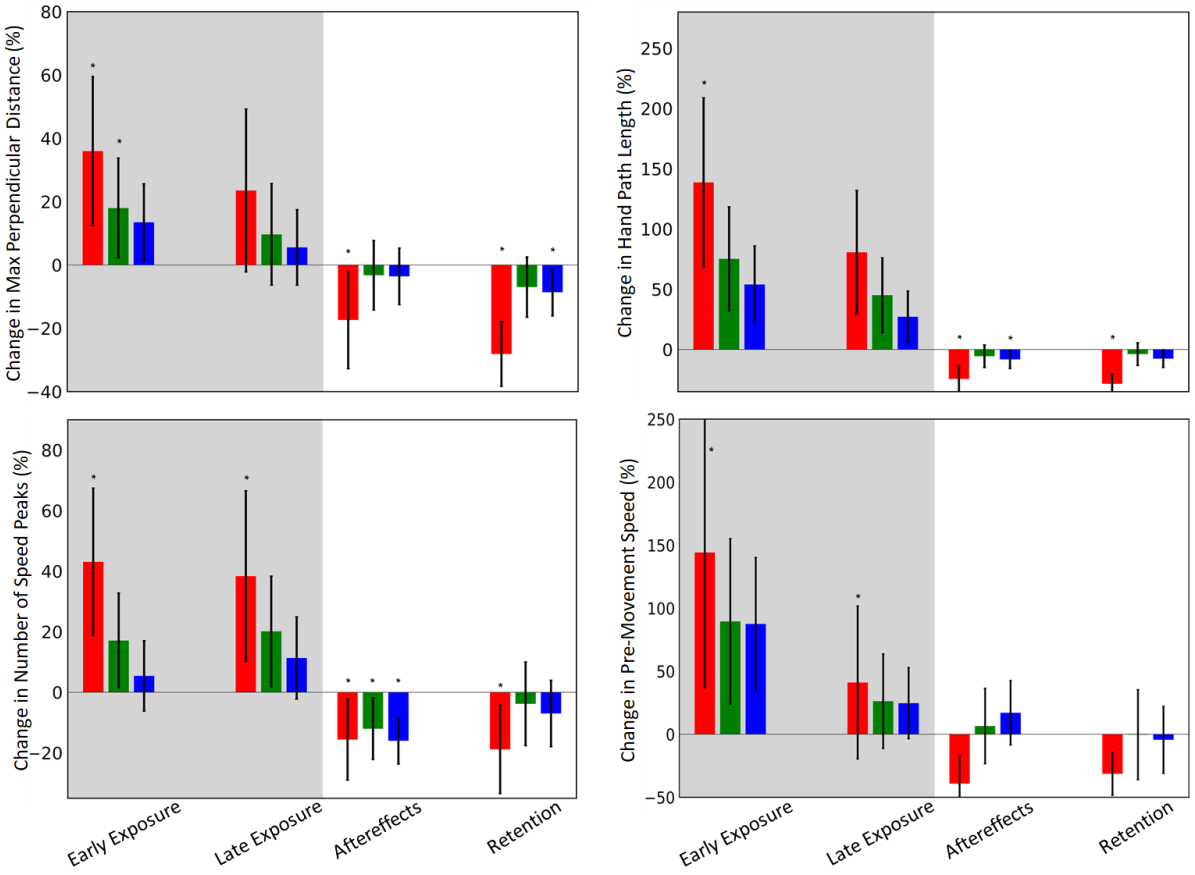


Figure 15: **Effect of experiment conditions on various movement features.** (A) Movement error measured using maximum perpendicular distance. (B) Movement length measured as the magnitude of the distance covered by the subject's arm. (C) Movement smoothness measured as the number of peaks in the speed profile. (D) Pre-movement speed measued as the average speed after target is displayed and before subjects begin moving.

3.5 Discussion

We looked for evidence of speeding patients up in both the direct effects when we turn on robotic forces and after effects after minutes of exposure then turning the forces off. As we expected, negative viscosity produced the largest direct effect (an increase in speed), since it was directly pushing patients to move faster. Speeds decreased by the end of exposure as they adapted, although they remained faster than baseline – a prolonged direct effect. Interestingly, we observed no speed-related after effects from any of our force treatments and any mismatch from baseline behaviour quickly dissipated.

During exposure to positive viscosity, patients exhibited the slower movements we expected, though this slowdown was not statistically significant in most cases. There was a minor increase in movement speed as an aftereffect, which was also not significant. Ultimately, patients behaved as we expected during and after experiencing positive viscosity, but the effects were very small compared to their baseline movements.

We expected that breakthrough forces would bias towards faster movements during late exposure and the early aftereffect due to this condition rewarding faster movements. We instead saw a reduction in movement speed. This can be due to incomplete learning of the forces, where patients did not have enough experience with the breakthrough condition to sufficiently bias their movements. Another explanation can be that breakthrough forces do not have the desired effect on patient movement speed. Either way, in a direct comparison with negative and positive viscosity under similar force magnitudes and length of exposure, breakthrough forces did not succeed in increasing patient speed.

Negative viscosity reduces patients' ability to control their arm movements during the initial ballistic phase. This led to a significant reduction in movement accuracy, effectiveness, and smoothness, as patients tried to counteract the destabilising force. These effects were reversed, as expected, once the forces were turned off, and patients significantly improved their reaches.

There were smaller, mostly non-significant effects of the other force conditions on these movement metrics. Overall, patients did not improve their reaches after training with positive viscosity or breakthrough forces.

We expected incomplete learning and/or incomplete washout due to our short exposure and null field blocks. Hence, we presented the forces to our participants in a randomized order and tested for any ordering effect. Out of 30 linear mixed effects models, the order the forces were experienced was significant in only two cases: late exposure movement distance, and retention movement error. This could be due to a few outliers in our small sample of patients, since in both cases the p-values were just below the significance level ($p \approx 0.04$).

It is important to note that this study was not powered. Our goal was to perform a crossover exploration into the effects of different types of robotic forces, while at the same time accounting for the possibility of ordering effects. Since there were $3! = 6$ possible order sequences for the three force types, we wanted to examine data from at least two patients per order sequence (Figure 11) to reduce the effects of possible outliers, which increased our number of participants to 12. We did not examine the data before all participants had completed the study. Due to the first patient having some initial difficulty moving the robot arm, we enrolled one additional patient in case parts of the movement data from the first patient was unusable. We were

ultimately able to utilize the data we collected from all patients, and had a grand total of 13 participants.

This study showed that negative viscosity had a strong direct effect at increasing patient speed, and that its usefulness as a treatment for chronic stroke survivors was bolstered by an improvement in movement accuracy, efficiency, and smoothness after the training. We are currently piloting training with negative viscosity over multiple sessions to improve clinical outcomes in chronic stroke survivors, as a prelude to a full clinical trial.

3.6 Conclusion

In a direct comparison between the three force conditions that may increase movement speed, patients significantly increased their speed only as a direct effect of negative viscosity. Positive viscosity and breakthrough forces had no effect on patient speed. After the forces were removed, only negative viscosity showed significant improvements in other movement metrics measuring accuracy, efficiency, and smoothness. We conclude that training to increase movement speed should be conducted using negative viscosity. Even though we did not achieve the desired lasting effect on movement speed, the improvements in other movement parameters shows promise for negative viscosity as a potential treatment.

3.7 Acknowledgement

We would like to thank Jennifer Kahn, PT/DPT/NCS at the Shirley Ryan AbilityLab for her help in accessing and using the Clinical Neuroscience Research Registry (STU00013948). We would like to also thank Courtney Celian, OTR/L for her help in recruiting patients for this study.

This study was made possible with funding from the American Heart Association Predoctoral Fellowship (18PRE34080333).

CHAPTER 4

TRAINING WITH NEGATIVE VISCOSITY IMPROVES CLINICAL OUTCOMES AFTER STROKE: A CASE STUDY

This work is submitted for publication, under the same title, to the Neurorehabilitation and Neural Repair, 2019, by Yazan Abdel Majeed, Courtney Celian, and James L. Patton.

4.1 Abstract

Movement speed is a strong predictor of clinical improvement in chronic stroke survivors. Negative viscosity shows the strongest effects in increasing patient speed during exposure and leads to significant arm motion improvements after training. This offered an opportunity to test the hypothesis that faster practice leads to better recovery. In this work, we present a case study using a negative viscosity treatment for a chronic stroke survivor. The patient trained with negative viscosity for two weeks, and clinical assessments were performed before and after the training. The patient showed small improvements as measued by clinical assessments. Movement speed significantly decreased across training sessions, while accuracy, efficiency, and smoothness significantly increased. This case study presents promising potential for the use of negative viscosity as a treatment for chronic stroke survivors, and provides solid confirmation that basing clinical interventions on the results of machine learning models can lead to greater clinical improvements where traditional rehabilitation ceases to have positive effects.

4.2 Introduction

Rehabilitation robotics have made great advances since the turn of the century (106; 107; 108). Robots provide the ability to customize a patient's environment in ways not possible otherwise. And though some research doubted the effectiveness of using robots for rehabilitation after neurological injury (109; 110) and showed that simple robotic assistance to patients is not sufficient for any significant clinical improvements, there is ample evidence that robots aid recovery in both sub-chronic and chronic patients (111; 112; 113; 114). Additionally, combining the use of robots with virtual reality environments leads to a more immersive therapy experience and increases brain activity after stroke (115).

Assistive active impedance has been used successfully in exoskeletons (116; 117) with the goal of increasing movement speed and improving control. The forces produced by a motor, such as in a robot, are controlled using the user's movement kinematics. This encourages the user to be more active during the task, which was shown to produce improvements in walking for both healthy participants and stroke survivors in a pilot study (118) and correlated with EEG patterns significantly different from passive walking (119).

In our previous work, we created a statistical framework for a data-driven approach at pinpointing targets for chronic stroke survivors (102). We used data collected from patients on the first day of a two-week treatment to predict clinical changes over that period. LASSO models resulted in the best predictions, and patient movement speed was strongly predictive and robustly correlated with clinical outcomes.

Our followup exploration compared the effectiveness of three types of robotic forces in changing movement speed in a crossover study with chronic patients. Our study revealed that negative viscosity provided the strongest effects to help patients generate faster movements during training, and significantly improved the accuracy and smoothness of their movements after a short training (120).

Here, we present the results of a pilot clinical intervention using a robot arm to apply negative viscosity, with the goal of improving clinical outcomes in arm movements for a chronic stroke survivor. The patient had six training sessions of approximately one hour over two weeks. We used the robot to assess the patient's progress at the beginning of each session, and a clinician evaluated the patient before and after training.

4.3 Methods

4.3.1 Participant

The patient recruited in this study needed to be over 18 years of age, be classified as chronic (more than 6 months since they had their stroke). They needed to score between 15 and 50 points on the Fugl-Meyer scale. The patient needed to pass the Stereo Fly Test (121) for 3D vision, and be able to perform the basic task (reaching to targets in a virtual reality environment). Patients were excluded if they had multiple lesions or bilateral paresis, if they had significant arm or shoulder pain, or if they had upper extremity Botox® treatments in the past six months. We also excluded participants who had participated in any other upper extremity intervention. This study was approved by the Northwestern University Institutional Review Board and followed the guidelines in the Declaration of Helsinki.

Two patients were screened for this case study. The first patient failed the Stereo Fly Test and was not able to perceive the task properly in three dimensions. The second patient fit our inclusion/exclusion criteria and provided informed consent. The patient is male, 54 years old, right-side dominant with a right hemiparesis. The patient's stroke occurred in February 2010 (i.e. treatment was 114 months post-stroke).

4.3.2 Study Setting

The study took place on the Arms and Hands floor of the Shirley Ryan Ability Lab in November 2019. We used the Looking Glass immersive virtual reality system based on the Virtual Reality Robotic and Optical Operations Machine (VRROOM) technology (122). The patient performed a unimanual point-to-point reaching task in three dimensions using the Proficio® 3-DoF robotic arm from Barrett® Technologies Figure 16A. The patient's paretic arm was supported against gravity using a device based on the Wilmington Robotic Exoskeleton (WREX) (123), and we strapped the robot around the patient's forearm such that the robot's end effector was near the wrist.

4.3.3 Experiment Protocol

The patient completed eight visits at the Arms and Hands lab at Shirley Ryan AbilityLab. The visits included six training sessions over two weeks. Each session lasted approximately 60 minutes. Clinical assessments were performed on visits 1, 7, and 8. Each of the eight sessions started with five minutes of targeted reaching with no robotic forces to assess progress. Our primary clinical outcome was the WMFT, and our secondary outcome was the UEFM. The training in visit 7 was followed by clinical assessments and a five-minute robot assessment to

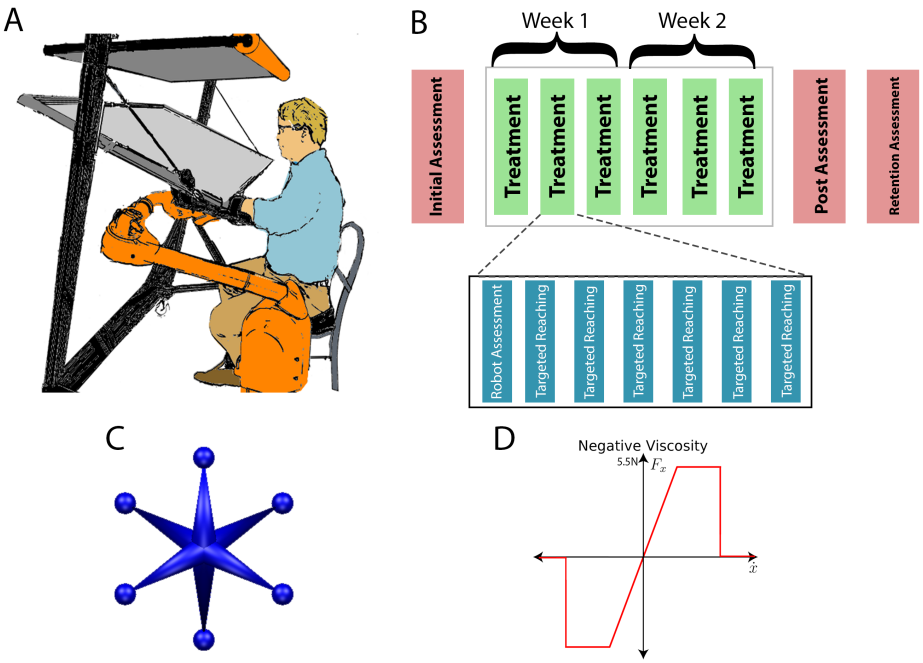


Figure 16: **Experiment Setup and Timeline.** The patient performed targeted reaching in three dimensions using the Looking Glass and the Barrett Proficio medical robot (**A**). The patient completed 8 visits. Clinical assessments were conducted during visits 1, 7, and 8, with the post assessment taking place after a short rest on the last treatment day (visit 7). Robot measurements were done with no forces during each visit to assess progress (**B**). We designed the patient's cursor in the virtual environment to help with 3D depth perception (**C**). During training, the robot applied negative viscosity in the same direction as the movement (**D**).

evaluate the immediate effects of the treatment. Visit 8 took place one week after the last treatment to assess retention Figure 16B. The patient did not participate in any other upper extremity training while taking part in this study.

During training, the patient saw a target in one of eight pseudo-randomized position around a quarter sphere of radius 15cm from the home position, which was updated at the beginning of each visit. The patient was to move a cursor representing their wrist inside the target sphere.

The cursor was shaped like an octahedral molecule Figure 16C to help with 3D perception. After the five-minute reaching assessment, the patient reached for six 5-minute segments under negative viscosity Figure 16D, with short breaks in between. Explicitly, the patient experienced forces according to:

$$F(x) = \begin{cases} 10\dot{x} & |\dot{x}| \leq 0.55 \\ 5.5 & 0.55 \leq |\dot{x}| \leq 2.0 \quad N \\ 0 & 2.0 \leq |\dot{x}| \end{cases}$$

These forces were independently calculated for each direction and combined such that the patient experienced the resultant 3D force at the end effector.

4.3.4 Data Analysis

We examined the effect of the negative viscosity treatment on three aspects of the patient's movements: clinical performance, movement speed, and other movement metrics.

An occupational therapist evaluated the patient's clinical performance during visits 1, 7, and 8. Our main outcome measure was the WMFT, and the secondary measure was the UEFM. Since this was a case study with one patient, no statistical analysis was performed on the clinical scores.

For movement speed and other metrics evaluating accuracy, smoothness, and efficiency, we performed an analysis of variance (ANOVA), with the single factor being the treatment session. We also performed paired t-tests with Bonferroni corrections to evaluate significance of changes between sessions. We used p and α values of 0.05 throughout our analysis.

4.4 Results

4.4.1 Clinical Change

Our primary outcome in this study was the WMFT. The patient completed certain tasks on the timed WMFT much faster during the follow-up assessment than the initial evaluation, the average time over the 17 tasks on the WMFT went down by ≈ 0.36 seconds. Additionally, the patient improved by 3 points on the UEFM scale by the end of training, and some of that improvement was preserved at the follow-up assessment one week later (2 points). There was no WMFT assessment on the last day of treatment (Figure 17A). As the patient trained by moving in the negative viscosity field, they were able to complete a higher number of reaches as time progressed (Figure 17B).

The clinical improvement was not evenly accrued over all subtasks of the assessments. Of note, the patient initially needed 6.00 seconds to extend their elbow with a weight attached, that time went down to 2.79 seconds during the follow-up assessment. The time they took to reach and retrieve an object decreased from 3.25 to 1.28 seconds, and the time to lift a paper clip went from 11.43 to 5.00 seconds. Some other tasks on the WMFT took the patient longer to complete, such as lifting a can (4.88 to 5.72 seconds). Similarly, on the UEFM, some synergies, and wrist movement in general, improved, while other motions became slightly more difficult for the patient.

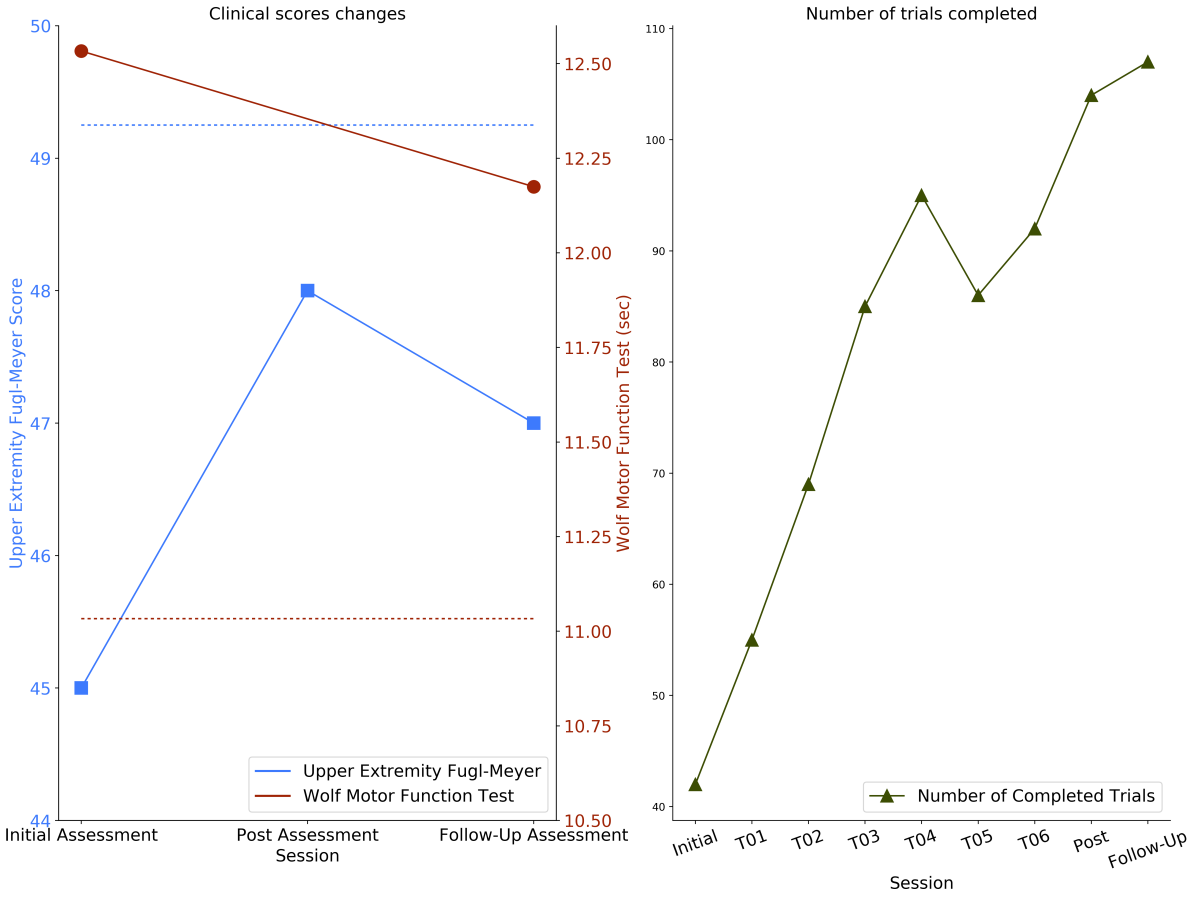


Figure 17: **Patient improved clinically and completed more reaches over sessions.** Study participant improved by two points on the UEFM and reduced their average time on WMFT by 0.36 seconds (**A**). The patient completed more targeted reaches during the robot assessment blocks at the beginning of each session over the two week study (**B**).

4.4.2 Changes in Movement Speed

Consistent with our findings in (120), we saw that both movement maximum and mean speeds significantly decreased with training ($F(8, 726) = 26.24, p < 0.001$ and $F(8, 726) = 28.95, p < 0.001$, respectively). These reductions persisted during the follow-up assessment (Figure 18).

There were significant changes in movement speed between the first two sessions, before experiencing negative viscosity, and all the other sessions ($p < 0.001$ for maximum speed and $p < 0.001$ for mean speed). In particular, looking at the difference between the initial and follow-up assessments, mean speed showed a significant reduction (average dropped by 0.044 ms^{-1}, $p < 0.001$) and maximum speed showed a decrease of 0.142 ms^{-1} on average, $p < 0.001$. In contrast, there was no significant change in speed between the post assessment and the follow-up one week later ($p = 0.008$ for max speed and $p = 0.71$ for mean speed, Bonferroni significance level $= 0.00139$)

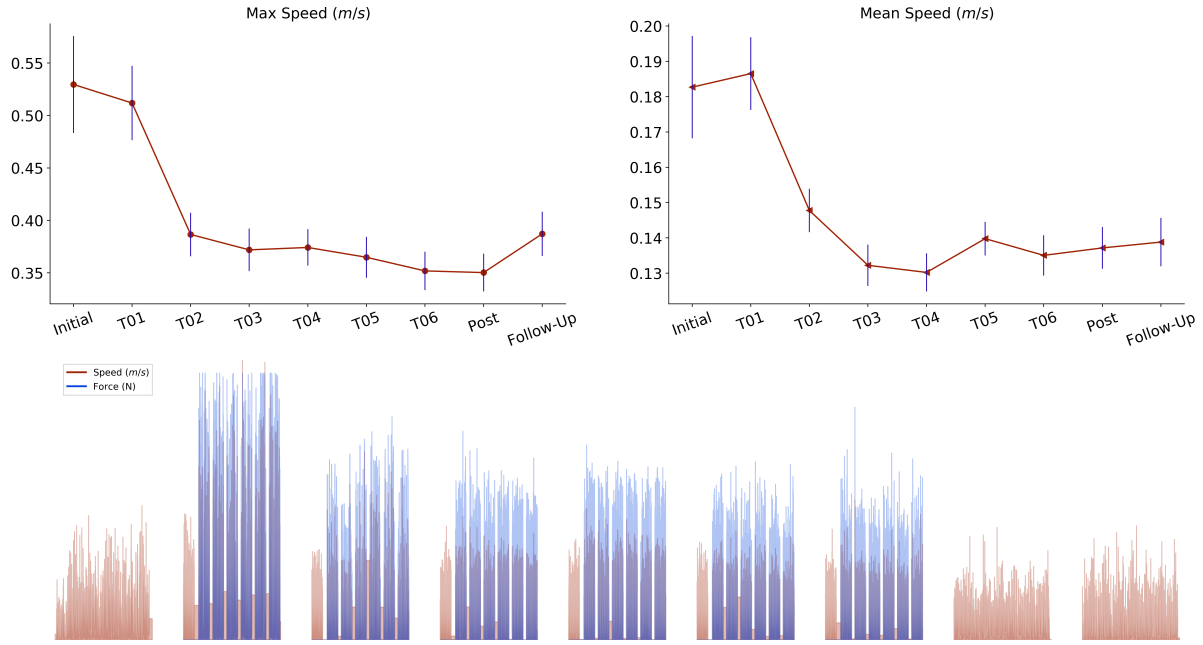


Figure 18: **Both mean and maximum speed decreased significantly with training.** Patient maximum speed (**A**) and mean speed (**B**) decreased over sessions. There was no significant difference in either mean or maximum speed prior to exposure to the treatment (sessions 1 & 2), or between any of the later sessions. Both robot forces and movement speed gradually decreased over time (**C**)

4.4.3 Changes in Other Movement Metrics

Movement error (maximum perpendicular distance from the straight line trajectory to target) significantly decreased over the course of the study ($F = 48.46, p < 0.001$). The patient generally reached more directly towards the target with smaller corrections (Figure 19B).

As expected with greater accuracy, movement distance significantly decreased ($F = 67.64, p < 0.001$, Figure 19C), and the patient needed significantly less time to complete each trial ($F = 35.48, p < 0.001$, Figure 19A). This also resulted in the increase in the number of trials com-

pleted during each robot assessment block at the beginning of each session, shown in Figure 17B. As the patient reached more accurately, fewer corrections were needed, and movement smoothness (as measured by the number of peaks in each trial's speed profile) significantly improved ($F = 27.66, p < 0.001$, Figure 19D).

Finally, in agreement with a non-significant result from our previous crossover trial (120), the patient's speed prior to launching in each trial (what we termed pre-movement speed) significantly decreased ($F(8, 549) = 10.67, p < 0.001$, Figure 19E).

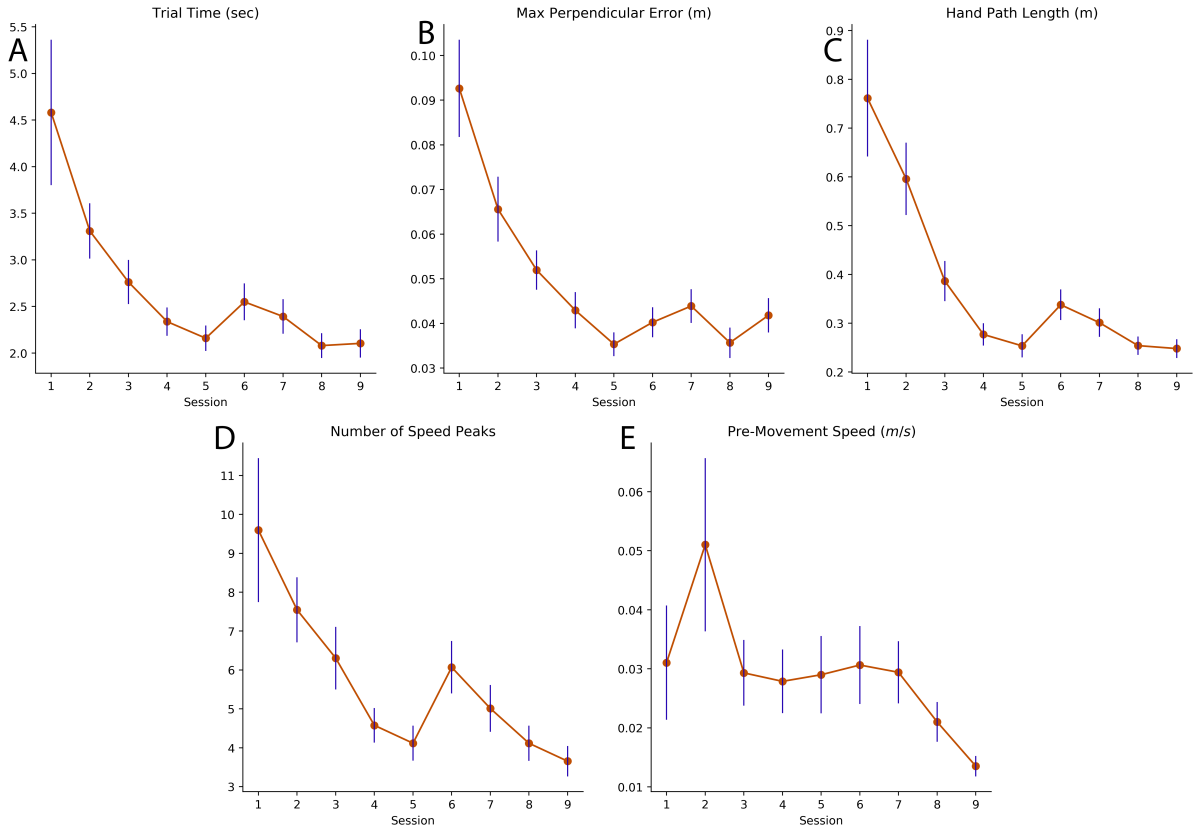


Figure 19: **The patient significantly improved their movements as they trained with negative viscosity.** Many movement metrics improved across sessions: Trial time decreased (**A**). Movement accuracy - measured as maximum perpendicular distance from a straight trajectory - improved (**B**). Movement efficiency - defined as the distance the hand moves through each trial to get to the target - decreased (**C**). Movement smoothness improved, evidenced by the lower number of peaks in the speed profile of the assessment trials, which indicates a lower number of submovements/corrections (**D**). Finally, the patient had less activity between the time the target appeared and the time they launched their reaching, termed pre-movement speed (**E**). All these changes were significant between the initial assessment and the follow-up. The metrics were not significantly different between the initial assessment and the first treatment day (sessions 1 & 2), and most were not significantly different after exposure to negative viscosity (session 3-9).

4.5 Discussion

The patient improved on both clinical assessments between the initial evaluation and the follow-up. Even though the clinical improvement did not reach the Minimal Clinically Important Difference (MCID) (124; 125), there was detectable progress on both scales due to our intervention. The fact that this progress was not evenly distributed over the subtasks of each clinical assessment might indicate a propensity for this intervention to affect certain types of impairment, which could help with the larger rehabilitation goal of customizing robotic therapy to the patient group that would most benefit from this particular intervention.

Treatment dosage could account for the lack of clinical significance in our patient's improvement. The patient completed a total of six hours of training over two weeks. By contrast, the Extremity Constraint Induced Therapy Evaluation (EXCITE) trial (126; 127) tested a very rigorous treatment schedule of up to six hours per day for 14 days in subacute patients. Our intervention achieved $\approx$ 25% of the MCID in WMFT with less than 10% of the treatment dosage, while working with chronic patients. This bolsters our intervention as a potentially significant treatment for chronic stroke survivors.

During our previous crossover trial (120), patients exhibited faster movements when first exposed to negative viscosity. That speed up effect remained significant but tapered off towards the end of a five-minute exposure. Here, we see from Figure 18C that the patient's speed during the first hour of training was higher than the subsequent sessions. Speeds then tapered off to a lower level that was maintained though subsequent days, even after one week of no exposure to negative viscosity. The speed after treatment and during later evaluations was significantly

slower than during the initial assessment. Comparatively, in (120), the slowdown after training was not significant. This could mean that stronger and more lasting results could be achieved with higher treatment dosages.

Multiple movement metrics improved significantly after training with negative viscosity (Figure 19). A lot of these changes were evident after a single treatment session. By the fourth session, almost all of the metrics had stabilized at an improved value, which was maintained in subsequent sessions. This follows the trend we saw in our crossover trial (120). Though, similar to speed after training with negative viscosity, pre-movement speed showed significant differences in this case study. The average value followed the same trend (lower pre-movement speed after exposure), but was now significant with the longer training.

Walking studies in spinal cord injury patients (128) showed that when participants moved in a destabilizing field (such as with negative viscosity), they improved control over their movements to counteract the destabilizing forces, and these improvements were preserved after the disruptive field was removed. One interpretation is that having to stabilize movements against these forces requires more patient involvement, both in terms of muscle activity and brain engagement. Other modes of robotic training, such as assistive fields, could cause patients to depend on the robot to do most of the work, and thereby not be as involved in their own therapy process.

One possible caveat for the data we showed in this case study is that this treatment may not work for every type of chronic stroke impairment. Patients exhibit a wide range of movement difficulties due to the large number of brain structures that could be affected. In this particular

case study, the patient had Ataxia and exhibited reduced control over their arm movements prior to training with negative viscosity. It is possible, following the point about patient involvement, that this type of treatment is most useful for patients exhibiting lower levels of control over their movements.

We can devise two methods to possibly boost the effects of our intervention in the clinical environment to reach the MCID levels. One is to increase the intervention intensity, adding more treatment sessions during the two-week training period. This has the advantage of being easy to implement, and could encourage patients to be more engaged in their recovery process. Another method would be to implement a challenge point paradigm, where data from the robot assessment blocks at the beginning of each session (Figure 16B) can be used to determine when the patient's performance is reaching a plateau, and adjust the strength of the robotic forces in the subsequent training blocks accordingly. This would provide an adaptive approach that would keep the patient challenged, and encourages continual motor learning.

Overall, these results indicate that training with negative viscosity improves clinical outcomes. Movement accuracy, smoothness, and efficiency are significantly better after a single one-hour session, but continue to improve marginally for the next two or so sessions before reaching a plateau. Since this is a case study with a single patient, we cannot generalize our conclusions to the larger chronic stroke population. However, the positive outcomes here bolster the use of negative viscosity training, and provide encouraging preliminary data for a larger clinical trial.

CHAPTER 5

DISCUSSION

The main goal of our research was to link patient arm movement kinematics to their performance on traditionally accepted clinical assessments. In doing that, we provided a complete pipeline for designing an effective therapy for chronic stroke survivors. Throughout this work, we ensured that the structure we developed fulfilled the original goals we laid out in Chapter 1: parsimony, interpretability, and clinical relevance.

Our LASSO models were the most parsimonious and best at predicting changes during the three-week study in Chapter 2. These models were easily interpretable, since they produced an equation much like standard least-squares regression, in terms of the predictor features. We then unpacked the best models to examine the aspects of movement most relevant to predicting clinical change (Figure 4). One of the features contributing most to the prediction that we could also directly affect was movement speed.

While our LASSO models succeeded at making predictions, we tested other algorithms that had potential according to the literature. Our random forests models did not have enough data to converge on a good prediction, and while the LARS models showed similar performance to LASSO, they were not as parsimonious. Likewise, elastic net models with a balance between L1 and L2 regularization showed similar performance to the LASSO model, but the final models contained more predictors. Ultimately, the LASSO models most closely fulfilled the requirements we established for a good predictive algorithm at the outset.

Once we identified movement speed as a potential treatment target, we decided to directly compare three of the paradigms with the strongest evidential backing in a crossover study with chronic stroke survivors. Motor control literature contains evidence for these training paradigms having an effect on movement speed (103; 7; 43; 48), and not all were tested in the chronic stroke survivor population. Originally, we wanted to set up this study to compare performance with or without feedback showing the patient how fast they were moving after each trial. Our pilot subjects showed no measurable effect to having that feedback as opposed to removing it, so we decided to keep that feedback on for the entire study. We saw that negative viscosity caused a significant direct increase in movement speed, and significant improvements in movement quality afterwards (error, smoothness, and efficiency; Figure 14, Figure 15). The other two paradigms we tested had much smaller effects that were predominantly not statistically significant.

We were a little surprised that neither positive viscosity, nor breakthrough forces brought about any significant longer-term changes in any of the metrics we examined during this crossover trial. All effects of these force paradigms washed out very quickly. This could lend credence to the idea from (128) that when patients are not as involved, they learn less and are able to retain less from their training once the treatment conditions are removed.

At this stage of the research, we had a successful predictive model from Chapter 2 indicating that movement speed was a suitable intervention target. Further, the results of our crossover study with chronic stroke survivors (Chapter 3) revealed that negative viscosity had a significant effect on movement speed and enhanced other movement attributes. Next, we tested the use of negative viscosity for treating a chronic stroke survivor.

The patient showed significant changes to multiple movement kinematic metrics after a single one-hour training session with negative viscosity. Movement speed decreased, while error, smoothness, and efficiency improved. After two weeks of training (six one-hour sessions), both the patient's WMFT and UEFM improved. It is hard to determine why the level of improvement did not reach clinical significance (MCID), but the treatment showed effectiveness nonetheless, in a patient population that is supposed to be mostly clinically static.

When we engineered movement features to use in our predictive modeling efforts, we knew that a lot of these predictors would have a high level of mutual information. We needed algorithms that were robust to that, and able to discard predictors that do not add new information to the model. The beginnings of our attempts to tackle this high $\mathtt{p}$, low $\mathtt{N}$ problem relied on forward feature selection and PCA (129), without cross-validation. These attempts ultimately failed, but pointed us in the direction of speed being an important factor that had high correlations with the clinical outcomes. Regularized regression was ultimately successful at predicting changes in clinical outcomes, but even though random forests failed, a non-cross-validated random forests model still assigned the highest variable importance to movement speed.

Given the success of our treatment using negative viscosity to directly affect speed in this patient population, it is interesting that removing movement speed from the LASSO models (See section 2.3) did not degrade the performance of the model. This could be because the information carried by this feature (Mean max speed) is also mostly contained in trial time and path length ($\mathtt{speed} = \frac{\mathtt{distance}}{\mathtt{time}}$), so the model was able to compensate for the information loss caused by the removal of mean max speed.

Ultimately, the work we presented in this thesis demonstrated that conducting a treatment based on data-driven recommendations from a machine learning model can help improve clinical outcomes in stroke survivors, long after they stopped improving due to conventional treatments. This provides a method to directly solve the problem of treatment design for chronic stroke survivors, and offers a pathway to perform a similar task with other patient populations.

5.1 Limitations and Future Work

Though robots and sensors can make extremely accurate measurements of patient kinematics and kinetics, one major difficulty with this research is the constraint of traditional clinical measures. Any new metrics will need to be validated against current clinical scales. A completely new robot or sensor-based method of measuring patient performance will not be accepted by the rehabilitation community without contrasting it against current established guidelines. We first needed to establish the link between patient movements and clinical changes. Consequently, this work is limited by the quality and usefulness of established clinical measures like the UEFM and the WMFT.

5.1.1 Predictive Models

Feature engineering is the most important aspect of constructing successful predictive models. Robust statistical algorithms can parse through a large number of predictors to find ones relevant for prediction. The performance of such algorithms is, however, limited by the information content of the features themselves. In this work, we relied on previous motor control literature and observations from earlier studies at our research lab to compile the list of predictors. When generalizing the framework we presented here to other patient populations, care

must be taken to engineer useful features with a meaningful connection to the outcome being predicted.

A major limitation of motor learning studies with patient populations is small sample sizes. Algorithms for tree-based machine learning are built to leverage large amounts of data, which can lead to convergence issues for very small data sets like ours (130). These methods combine a large number of weak predictor trees to achieve consistent and trustworthy predictions. The lack of large datasets for chronic stroke survivors limits the ability to use these first-rate techniques in our proposed framework.

5.1.2 Speed Crossover Trial

Our crossover trial had a short, five-minute exposure to each type of robot forces. This limits the extent of patient learning that can be achieved. Also, based on (128), patients are less involved in their own treatment when robot forces are doing most of the stabilization work. In our case, this occurs when patients are experiencing positive viscosity and breakthrough forces. Longer training sessions may show significant results for these two conditions. However, we wanted to conduct a fair comparison between the force paradigms, so we did not aim for a certain level of learning, but compared the changes in movement given equal exposure to the forces.

We saw that memory of previous learning transfers between sessions in section 4.4. Future work can expand on our crossover trial by including longer training sessions over a minimum of two days. This allows for achieving stronger effects for each of the robot force paradigms and opens the door for a new research question regarding savings with each force type (131; 132; 133).

We believe that negative viscosity will still show the strongest effects and larger savings, and expanding the crossover trial as we described here would answer that question.

The force types we tested in our crossover trial are not an exhaustive list. We selected three robot force paradigms based on motor learning literature. However, other force types may also have a strong effect on patient speed and movement performance. An example of a field that we *could* test is assistive guidance through a haptic 'channel' or 'trough' (134). It was shown to help motor learning, but the effects on speed are not yet understood.

Finally, our sample size of 13 patients was sufficient to determine which of the three force paradigms had the strongest effect on movement speed. A larger crossover study may be able to further stratify this effect on movement speed based on impairment types and other patient characteristics. As this work is an initial step in answering these questions, we needed to answer the most basic question before subdividing the patient population and personalizing treatments. This is a longer-term goal for which we laid the foundations in this thesis.

5.1.3 Treatment Case Study

Our case study in Chapter 4 provided preliminary results to show that training with negative viscosity can improve clinical performance in chronic stroke survivors. This needs to be confirmed through a larger controlled, blinded clinical trial before any firm statements can be made.

We can improve on the protocol we outlined in section 4.3 by including a second baseline assessment before starting treatment. This will account for any spontaneous recovery patients may have (135). Since we conducted our treatment with a chronic stroke survivors, this was

under the assumption that a chronic patient will not exhibit any spontaneous recovery. Having that confirmed by conducting a second baseline would allay any concerns.

Though we saw significant effects on movement quality with negative viscosity training, the improvements on clinical measures did not reach the MCID. This may be related to the training dosage. In our case study, we endeavored to emulate the original study our machine learning model was trained on as closely as possible, to allow for more direct comparisons of results. In contrast, the EXCITE trial had up to 14 times our training intensity and exhibited clinically significant improvements (126). A more intensive treatment protocol may lead to stronger clinical changes.

A clinical trial to determine the appropriate level of negative viscosity and the suitable training dosage would require a large number of participants for proper statistical power. One way of circumventing that issue is organic data collection. Though it would require relaxing inclusion/exclusion criteria for such a study, organic data collection involves placing a robot with a virtual reality setup in the clinic that patients always have access to would yield larger datasets. Having that setup randomize the strength of negative viscosity forces for each patient, along with that larger data volume, and the different choices patients would make on how long to train for, we would be able to determine the most effective form of negative viscosity training for stroke survivors.

5.2 Conclusion

We set out to create a model that establishes a connection between movement kinematics and clinical outcomes in neurologically injured patients, and to determine whether designing

a treatment based on that model can improve these clinical outcomes. Developing this data-supported treatment will lead to more effective patient interventions, which can give more hope to chronic stroke survivors.

Our LASSO model was able to predict clinical changes for a two-week study with chronic stroke survivors. We were more successful at predicting change in patients' WMFT (R^2mean $\pm$ sd: $86.07\% \pm 5.26\%$) than the change in UEFM ($65.34\% \pm 17.45\%$), most likely due to the discrete nature of the latter. When we examined the predictors contributing to the model, we saw that one of the top predictors we could directly affect was patient movement speed. Other strong predictors, such as age and initial impairment may help target the treatment at patients who could most benefit from it.

Since there are multiple means to alter movement speed, we designed a crossover trial to determine the most effective of these, and our results showed that negative viscosity had the strongest effect, with the added advantage of significant improvements in other common movement metrics after only a short, five-minute exposure. Interestingly, the other methods of altering speed we explored in this trial did not have a significant effect on chronic stroke survivors. These results suggest that negative viscosity should be used as the treatment intervention to alter speed in chronic stroke survivors.

Testing this intervention with a chronic stroke participant showed significant and fast improvements in multiple movement metrics, and small gains in patient outcomes. While the therapist-assessed changes were not clinically significant, patient movements were markedly smoother and more accurate after two weeks of training. Further work will need to confirm

these results through a controlled clinical trial, but the preliminary data provided through our case study is encouraging. Anecdotally, the treatment appears to impart its benefit quite quickly. The subject had plateaued in robot metric improvement by the fourth session. A more intensive treatment regimen may result in larger improvements, and "organic data collection" in the clinic can help determine appropriate training dosages.

We presented a framework to determine appropriate therapy targets for chronic stroke survivors. This framework can, however, be expanded to work with any patient population, especially where studies suffer from a small number of participants and a large amount of possible treatment targets. Provided proper feature engineering is performed beforehand, our framework can serve as a standard to identify and test new therapies likely to be effective for the patient population in question.

Based on the results of our work as presented in this thesis, therapists should consider training chronic stroke survivors with negative viscosity. These patients, by definition, have stopped improving with traditional physical and occupational therapies. Negative viscosity could provide a chance for further improvements in these patients' lives, and the possibility of more independent living after what is sometimes a devastating neurological injury.

CITED LITERATURE

1. Donkor, E. S.: Stroke in the century: A snapshot of the burden, epidemiology, and quality of life. Stroke research and treatment, 2018, 2018.

2. Handley, A., Medcalf, P., Hellier, K., and Dutta, D.: Movement disorders after stroke. Age and ageing, 38(3):260–266, 2009.

3. Fugl-Meyer, A. R., Jääskö, L., Leyman, I., Olsson, S., and Steglind, S.: The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scandinavian journal of rehabilitation medicine, 7(1):13–31, 1974.

4. Wolf, S. L., Catlin, P. A., Ellis, M., Archer, A. L., Morgan, B., and Piacentino, A.: Assessing wolf motor function test as outcome measure for research in patients after stroke. Stroke, 32(7):1635–1639, 2001.

5. Adams, R. J., Meador, K. J., Sethi, K. D., Grotta, J. C., and Thomson, D. S.: Graded neurologic scale for use in acute hemispheric stroke treatment protocols. Stroke, 18(3):665–669, 1987.

6. Yozbatiran, N., Der-Yeghiaian, L., and Cramer, S. C.: A standardized approach to performing the action research arm test. Neurorehabilitation and neural repair, 22(1):78–90, 2008.

7. Patton, J. L., Kovic, M., and Mussa-Ivaldi, F. A.: Custom-designed haptic training for restoring reaching ability to individuals with poststroke hemiparesis. Journal of Rehabilitation Research & Development, 43(5), 2006.

8. Hornby, T. G., Campbell, D. D., Kahn, J. H., Demott, T., Moore, J. L., and Roth, H. R.: Enhanced gait-related improvements after therapist- versus robotic-assisted locomotor training in subjects with chronic stroke: A randomized controlled study. Stroke, 39(6):1786–1792, 2008.

9. Wu, J., Liu, H., Duan, X., Ding, Y., Wu, H., Bai, Y., and Sun, X.: Prediction of DNA-binding residues in proteins from amino acid sequences using a random forest model with a hybrid feature. Bioinformatics, 25(1):30–35, 2009.

10. Fujino, Y., Murata, H., Mayama, C., and Asaoka, R.: Applying "Lasso" regression to predict future visual field progression in glaucoma patients. Investigative Ophthalmology and Visual Science, 56(4):2334–2339, 2015.

11. Colebatch, J. G., Gandevia, S. C., and Spira, P. J.: Voluntary muscle strength in hemiparesis: distribution of weakness at the elbow. Journal of Neurology, Neurosurgery and Psychiatry, 49(9):1019–1024, 1986.

12. Mercier, C., Bertrand, A. M., and Bourbonnais, D.: Differences in the magnitude and direction of forces during a submaximal matching task in hemiparetic subjects. Experimental Brain Research, 157(1):32–42, 2004.

13. Vattanasilp, W., Ada, L., and Crosbie, J.: Contribution of thixotropy, spasticity, and contracture to ankle stiffness after stroke. Journal of Neurology, Neurosurgery and Psychiatry, 69(1):34–39, 2000.

14. Hawe, R. L. and Dewald, J. P. A.: Development of a method to quantify inter-limb coupling in individuals with hemiparetic stroke. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 3476–3479. IEEE, aug 2015.

15. Lum, P. S., Burgar, C. G., Kenney, D. E., and Van der Loos, H. M.: Quantification of force abnormalities during passive and active-assisted upper-limb reaching movements in post-stroke hemiparesis. IEEE Transactions on Biomedical Engineering, 46(6):652–662, 1999.

16. Handley, A., Medcalf, P., Hellier, K., and Dutta, D.: Movement disorders after stroke. Age and Ageing, 38(3):260–266, 2009.

17. Barreca, S., Gowland, C. K., Stratford, P., Huijbregts, M., Griffiths, J., Torresin, W., Dunkley, M., Miller, P., Masters, L., (Kelly) Gowland, C., Stratford, P., Huijbregts, M., Griffiths, J., Torresin, W., Dunkley, M., Miller, P., and Masters, L.: Development of the Chedoke Arm and Hand Activity Inventory: Theoretical Constructs, Item Generation, and Selection. Topics in Stroke Rehabilitation, 11(4):31–42, 2004.

18. Chakouch, M. K., Pouletaut, P., Charleux, F., and Bensamoun, S. F.: Viscoelastic shear properties of in vivo thigh muscles measured by MR elastography. Journal of Magnetic Resonance Imaging, 43(6):1423–1433, 2016.

19. Cheung, D. K., Climans, S. A., Black, S. E., Gao, F., Szilagyi, G. M., and Mochizuki, G.: Lesion Characteristics of Individuals with Upper Limb Spasticity after Stroke. Neurorehabilitation and Neural Repair, 30(1):63–70, 2016.

20. Barber, P., Darby, D., Desmond, P., Yang, Q., Gerraty, R., Jolley, D., Donnan, G., Tress, B., and Davis, S.: Prediction of stroke outcome with echoplanar perfusion- and diffusion-weighted MRI. Neurology, 51(2):418–426, 1998.

21. Stinear, C.: Prediction of recovery of motor function after stroke. The Lancet Neurology, 9(12):1228–1232, 2010.

22. Riley, J. D., Le, V., Der-Yeghiaian, L., See, J., Newton, J. M., Ward, N. S., and Cramer, S. C.: Anatomy of stroke injury predicts gains from therapy. Stroke, 42(2):421–426, 2011.

23. Mostafavi, S. M., Glasgow, J. I., Dukelow, S. P., Scott, S. H., and Mousavi, P.: Prediction of stroke-related diagnostic and prognostic measures using robot-based evaluation. In 2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR), pages 1–6. IEEE, 2013.

24. Patton, J. L., Wei, Y. J., Bajaj, P., and Scheidt, R. A.: Visuomotor learning enhanced by augmenting instantaneous trajectory error feedback during reaching. PloS one, 8(1):e46466, 2013.

25. Prasad, A. M., Iverson, L. R., and Liaw, A.: Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystems, 9(2):181–199, 2006.

26. Chen, X. and Wang, L.: Integrating biological knowledge with gene expression profiles for survival prediction of cancer. Journal of computational biology : a journal of computational molecular cell biology, 16(2):265–278, 2009.

27. Ellis, K., Kerr, J., Godbole, S., Lanckriet, G., Wing, D., and Marshall, S.: A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiological Measurement, 35(11):2191–2203, dec 2014.

28. Shen, D., Shen, H., Zhu, H., and Marron, J.: The statistics and mathematics of high dimension low sample size asymptotics. Statistica Sinica, 26(4):1747, 2016.

29. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), pages 267–288, 1996.

30. Breiman, L.: Random forests. Machine learning, 45(1):5–32, 2001.

31. Efron, B., Hastie, T., Johnstone, I., Tibshirani, R., et al.: Least angle regression. The Annals of statistics, 32(2):407–499, 2004.

32. F.R.S., K. P.: Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11):559–572, 1901.

33. Van der Loos, H. M. and Reinkensmeyer, D. J.: Rehabilitation and Health Care Robotics. In Springer Handbook of Robotics, pages 1223–1251. Berlin, Heidelberg, Springer Berlin Heidelberg, 2008.

34. Popescu, V. G., Burdea, G. C., Bouzit, M., and Hentz, V. R.: A virtual-reality-based telerehabilitation system with force feedback. IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society, 4(1):45–51, 2000.

35. Merians, A. S., Jack, D., Boian, R., Tremaine, M., Burdea, G. C., Adamovich, S. V., Recce, M., and Poizner, H.: Virtual reality–augmented rehabilitation for patients following stroke. Physical therapy, 82(9):898–915, 2002.

36. Ferraro, M., Palazzolo, J. J., Krol, J., Krebs, H. I., Hogan, N., and Volpe, B. T.: Robot-aided sensorimotor arm training improves outcome in patients with chronic stroke. Neurology, 61(11):1604–1607, 2003.

37. Eriksson, J., Mataric, M. J., and Winstein, C. J.: Hands-off assistive robotics for post-stroke arm rehabilitation. Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics, 2005:21–24, 2005.

38. Song, R., Tong, K. Y., Hu, X., and Li, L.: Assistive control system using continuous myoelectric signal in robot-aided arm training for patients after stroke. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 16(4):371–379, 2008.

39. Lum, P., Reinkensmeyer, D., Mahoney, R., Rymer, W. Z., and Burgar, C.: Robotic Devices for Movement Therapy After Stroke: Current Status and Challenges to Clinical Acceptance. Topics in Stroke Rehabilitation, 8(4):40–53, jan 2002.

40. Kahn, L. E., Zygman, M. L., Rymer, W. Z., and Reinkensmeyer, D. J.: Robot-assisted reaching exercise promotes arm movement recovery in chronic hemiparetic stroke: a randomized controlled pilot study. Journal of neuroengineering and rehabilitation, 3:12, 2006.

41. Emken, J. L. and Reinkensmeyer, D. J.: Robot-enhanced motor learning: Accelerating internal model formation during locomotion by transient dynamic amplification. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 13(1):33–39, 2005.

42. Kao, P.-C., Srivastava, S., Agrawal, S. K., and Scholz, J. P.: Effect of robotic performance-based error-augmentation versus error-reduction training on the gait of healthy individuals. Gait & Posture, 37(1):113–120, jan 2013.

43. Abdollahi, F., Case Lazarro, E. D., Listenberger, M., Kenyon, R. V., Kovic, M., Bogey, R. A., Hedeker, D., Jovanovic, B. D., and Patton, J. L.: Error augmentation enhancing arm recovery in individuals with chronic stroke: a randomized crossover design. Neurorehabilitation and neural repair, 28(2):120–128, 2014.

44. Abdollahi, F., Rozario, S. V., Kenyon, R. V., Patton, J. L., Case, E., Kovic, M., and Listenberger, M.: Arm control recovery enhanced by error augmentation. In Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on, pages 1–6. IEEE, 2011.

45. Nikooyan, A. A. and Ahmed, A. A.: Reward feedback accelerates motor learning. Journal of neurophysiology, 113(2):633–646, 2014.

46. Summerside, E. M., Shadmehr, R., and Ahmed, A. A.: Vigor of reaching movements: reward discounts the cost of effort. Journal of neurophysiology, 119(6):2347–2357, 2018.

47. Pekny, S. E., Izawa, J., and Shadmehr, R.: Reward-dependent modulation of movement variability. Journal of Neuroscience, 35(9):4015–4024, 2015.

48. Shah, A. K., Sharp, I., Hajissa, E., and Patton, J. L.: Reshaping movement distributions with limit-push robotic training. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(11):2134–2144, 2018.

49. Lotze, M., Braun, C., Birbaumer, N., Anders, S., and Cohen, L. G.: Motor learning elicited by voluntary drive. Brain, 126(4):866–872, 2003.

50. Marchal-Crespo, L. and Reinkensmeyer, D. J.: Review of control strategies for robotic movement training after neurologic injury. Journal of NeuroEngineering and Rehabilitation, 6(1):20, 2009.

51. Fujimoto, S. T., Longhi, L., Saatman, K. E., and McIntosh, T. K.: Motor and cognitive function evaluation following experimental traumatic brain injury. Neuroscience and Biobehavioral Reviews, 28(4):365–378, 2004.

52. Sadek, J. R., Stricker, N., Adair, J. C., and Haaland, K. Y.: Performance-based everyday functioning after stroke: relationship with IADL questionnaire and neurocognitive performance. Journal of the International Neuropsychological Society : JINS, 17(5):832–40, 2011.

53. Coderre, A. M., Zeid, A. A., Dukelow, S. P., Demmer, M. J., Moore, K. D., Demers, M. J., Bretzke, H., Herter, T. M., Glasgow, J. I., Norman, K. E., Bagg, S. D., Scott, S. H., Amr Abou Zeid, Dukelow, S. P., Demmer, M. J., Moore, K. D., Demers, M. J., Bretzke, H., Herter, T. M., Glasgow, J. I., Norman, K. E., Bagg, S. D., and Scott, S. H.: Assessment of Upper-Limb Sensorimotor Function of Subacute Stroke Patients Using Visually Guided Reaching. Neurorehabilitation and Neural Repair, 24(6):528–541, 2010.

54. Prabhakaran, S., Zarahn, E., Riley, C., Speizer, A., Chong, J. Y., Lazar, R. M., Marshall, R. S., and Krakauer, J. W.: Inter-individual variability in the capacity for motor recovery after ischemic stroke. Neurorehabilitation and neural repair, 22(1):64–71, 2015.

55. Mercier, L., Audet, T., Hébert, R., Rochette, a., and Dubois, M. F.: Impact of motor, cognitive, and perceptual disorders on ability to perform activities of daily living after stroke. Stroke; a journal of cerebral circulation, 32(11):2602–2608, 2001.

56. Chollet, F., DiPiero, V., Wise, R. J., Brooks, D. J., Dolan, R. J., and Frackowiak, R. S.: The functional anatomy of motor recovery after stroke in humans: a study with positron emission tomography. Ann Neurol, 29(1):63–71, 1991.

57. Nudo, R. J.: Functional and structural plasticity in motor cortex: Implications for stroke recovery. Physical Medicine and Rehabilitation Clinics of North America, 14(1 SUPPL.):57–76, 2003.

58. Tibshirani, R.: The lasso method for variable selection in the cox model. Statistics in Medicine, 16(4):385–395, 1997.

59. Duncan, P. W., Lai, S. M., and Keighley, J.: Defining post-stroke recovery: implications for design and interpretation of drug trials. Neuropharmacology, 39(5):835–841, 2000.

60. Platz, T., Pinkowski, C., van Wijck, F., Kim, I.-H., Di Bella, P., and Johnson, G.: Reliability and validity of arm function assessment with standardized guidelines for the fugl-meyer test, action research arm test and box and block test: a multicentre study. Clinical Rehabilitation, 19(4):404–411, 2005.

61. Adams, H. P., Del Zoppo, G., Alberts, M. J., Bhatt, D. L., Brass, L., Furlan, A., Grubb, R. L., Higashida, R. T., Jauch, E. C., Kidwell, C., et al.: Guidelines for the early management of adults with ischemic stroke. Circulation, 115(20):e478–e534, 2007.

62. Duncan, P. W., Goldstein, L. B., Matchar, D., Divine, G. W., and Feussner, J.: Measurement of motor recovery after stroke. outcome assessment and sample size requirements. Stroke, 23(8):1084–1089, 1992.

63. Uswatte, G., Taub, E., Morris, D., Vignolo, M., and McCulloch, K.: Reliability and validity of the upper-extremity motor activity log-14 for measuring real-world arm use. Stroke, 36(11):2493–2496, 2005.

64. Keith, R., Granger, C., Hamilton, B., and Sherwin, F.: The functional independence measure. Adv Clin Rehabil, 1:6–18, 1987.

65. Reuben, D. B., Magasi, S., McCreath, H. E., Bohannon, R. W., Wang, Y.-C., Bubela, D. J., Rymer, W. Z., Beaumont, J., Rine, R. M., Lai, J.-S., et al.: Motor assessment using the nih toolbox. Neurology, 80(11 Supplement 3):S65–S75, 2013.

66. Barber, P., Darby, D., Desmond, P., Yang, Q., Gerraty, R., Jolley, D., Donnan, G., Tress, B., and Davis, S.: Prediction of stroke outcome with echoplanar perfusion- and diffusion-weighted MRI. Neurology, 51(2):418–426, 1998.

67. Schiemanck, S. K., Kwakkel, G., Post, M. W., Kappelle, L. J., and Prevo, A. J.: Predicting long-term independency in activities of daily living after middle cerebral artery stroke: does information from mri have added predictive value compared with clinical information? Stroke, 37(4):1050–1054, 2006.

68. Marshall, R. S., Zarahn, E., Alon, L., Minzer, B., Lazar, R. M., and Krakauer, J. W.: Early imaging correlates of subsequent motor recovery after stroke. Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society, 65(5):596–602, 2009.

69. Zarahn, E., Alon, L., Ryan, S. L., Lazar, R. M., Vry, M.-S., Weiller, C., Marshall, R. S., and Krakauer, J. W.: Prediction of motor recovery using initial impairment and fmri 48 h poststroke. Cerebral Cortex, 21(12):2712–2721, 2011.

70. Stinear, C. M.: Prediction of motor recovery after stroke: advances in biomarkers. The Lancet Neurology, 16(10):826–836, 2017.

71. Van de Port, I., Kwakkel, G., Schepers, V., and Lindeman, E.: Predicting mobility outcome one year after stroke: a prospective cohort study. Journal of Rehabilitation Medicine, 38(4):218–223, 2006.

72. Mostafavi, S. M., Mousavi, P., Dukelow, S. P., and Scott, S. H.: Robot-based assessment of motor and proprioceptive function identifies biomarkers for prediction of functional independence measures. Journal of neuroengineering and rehabilitation, 12(1):1, 2015.

73. Stinear, C. M., Byblow, W. D., Ackerley, S. J., Barber, P. A., and Smith, M.-C.: Predicting recovery potential for individual stroke patients increases rehabilitation efficiency. Stroke, 48(4):1011–1019, 2017.

74. Duncan, P. W., Propst, M., and Nelson, S. G.: Reliability of the fugl-meyer assessment of sensorimotor recovery following cerebrovascular accident. Physical therapy, 63(10):1606–1610, 1983.

75. Morris, D. M., Uswatte, G., Crago, J. E., Cook, E. W., and Taub, E.: The reliability of the wolf motor function test for assessing upper extremity function after stroke. Archives of physical medicine and rehabilitation, 82(6):750–755, 2001.

76. Mostafavi, S., Dukelow, S., Scott, S., and Mousavi, P.: Evaluation of similarities between robotic tasks for reduction of stroke assessment time. In Rehabilitation Robotics (ICORR), 2015 IEEE International Conference on, pages 211–216. IEEE, 2015.

77. Mostafavi, S. M., Dukelow, S. P., Scott, S. H., and Mousavi, P.: Hierarchical task ordering for time reduction on kinarm assessment protocol. In Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE, pages 2517–2520. IEEE, 2014.

78. Zou, H. and Hastie, T.: Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67(2):301–320, 2005.

79. Kyung, M., Gill, J., Ghosh, M., Casella, G., et al.: Penalized regression, standard errors, and bayesian lassos. Bayesian Analysis, 5(2):369–411, 2010.

80. Sandri, M. and Zuccolotto, P.: Variable selection using random forests. In Data analysis, classification and the forward search, pages 263–270. Springer, 2006.

81. Siroky, D. S. et al.: Navigating random forests and related advances in algorithmic modeling. Statistics Surveys, 3:147–163, 2009.

82. Friedman, J., Hastie, T., and Tibshirani, R.: The elements of statistical learning, volume 1. Springer series in statistics New York, 2001.

83. Salthouse, T. A.: When does age-related cognitive decline begin? Neurobiology of aging, 30(4):507–514, 2009.

84. Jarzem, P. and Gledhill, R.: Predicting height from arm measurements. Journal of Pediatric Orthopaedics, 13(6):761–765, 1993.

85. Zemke, A. C., Heagerty, P. J., Lee, C., and Cramer, S. C.: Motor cortex organization after stroke is related to side of stroke and level of recovery. Stroke, 34(5):e23–e26, 2003.

86. Barto, A. G., Sutton, R. S., and Anderson, C. W.: Neuron-like adaptive elements that can solve difficult learning control problems. IEEE transactios on Systems, Man and Cybernetics, SMC-13:834–846, 1983.

87. Rohrer, B., Fasoli, S., Krebs, H. I., Volpe, B., Frontera, W. R., Stein, J., and Hogan, N.: Submovements grow larger, fewer, and more blended during stroke recovery. Motor control, 8(4):472–483, 2004.

88. Majeed, Y. A., Abdollahi, F., Awadalla, S., and Patton, J.: Multivariate outcomes in a three week bimanual self-telerehabilitation with error augmentation post-stroke. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 1425–1431. IEEE, 2015.

89. Patton, J. L., Stoykov, M. E., Kovic, M., and Mussa-Ivaldi, F. A.: Evaluation of robotic training forces that either enhance or reduce error in chronic hemiparetic stroke survivors. Experimental brain research, 168(3):368–383, 2006.

90. Mostafavi, S. M., Dukelow, S. P., Glasgow, J. I., Scott, S. H., and Mousavi, P.: Reduction of stroke assessment time for visually guided reaching task on kinarm exoskeleton robot. In 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 5296–5299, Aug 2014.

91. Rohrer, B., Fasoli, S., Krebs, H. I., Hughes, R., Volpe, B., Frontera, W. R., Stein, J., and Hogan, N.: Movement smoothness changes during stroke recovery. The Journal of Neuroscience, 22(18):8297–8304, 2002.

92. Hoerl, A. E. and Kennard, R. W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1):55–67, 1970.

93. Schabowsky, C. N., Godfrey, S. B., Holley, R. J., and Lum, P. S.: Development and pilot testing of hexorr: hand exoskeleton rehabilitation robot. Journal of neuroengineering and rehabilitation, 7(1):36, 2010.

94. Fisher Bittmann, M. F.: Customized Robotic Training Approaches Using the Statistics of Reaching Errors. Doctoral dissertation, University of Illinois at Chicago, 2016.

95. Konidaris, G., Kuindersma, S., Grupen, R., and Barto, A.: Robot learning from demonstration by constructing skill trees. The International Journal of Robotics Research, 31(3):360–375, 2012.

96. Wright, Z. A., Lazzaro, E., Thielbar, K. O., Patton, J. L., and Huang, F. C.: Robot training with vector fields based on stroke survivors' individual movement statistics. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(2):307–323, 2017.

97. Fisher, M. E., Huang, F. C., Klamroth-Marganska, V., Riener, R., and Patton, J. L.: Haptic error fields for robotic training. In 2015 IEEE World Haptics Conference (WHC), pages 434–439. IEEE, 2015.

98. Patton, J. L. and Huang, F. C.: Sensory-motor interactions and error augmentation. In Neurorehabilitation Technology, pages 79–95. Springer, 2016.

99. Marchal-Crespo, L., Michels, L., Jaeger, L., López-Olóriz, J., and Riener, R.: Effect of error augmentation on brain activation and motor learning of a complex locomotor task. Frontiers in neuroscience, 11:526, 2017.

100. Altschuler, E. L., Wisdom, S. B., Stone, L., Foster, C., Galasko, D., Llewellyn, D. M. E., and Ramachandran, V. S.: Rehabilitation of hemiparesis after stroke with a mirror. The Lancet, 353(9169):2035–2036, 1999.

101. Yu, L., Xiong, D., Guo, L., and Wang, J.: A remote quantitative fugl-meyer assessment framework for stroke patients based on wearable sensor networks. Computer methods and programs in biomedicine, 128:100–110, 2016.

102. Majeed, Y. A., Awadalla, S. S., and Patton, J. L.: Regression techniques employing feature selection to predict clinical outcomes in stroke. PloS one, 13(10):e0205639, 2018.

103. Dancause, N., Ptitob, A., and Levin, M. F.: Error correction strategies for motor behavior after unilateral brain damage: short-term motor learning processes. Neuropsychologia, 40(8):1313–1323, 2002.

104. Huang, F. C. and Patton, J. L.: Augmented dynamics and motor exploration as training for stroke. IEEE Transactions on Biomedical Engineering, 60(3):838–844, 2012.

105. Israely, S. and Carmeli, E.: Error augmentation as a possible technique for improving upper extremity motor performance after a stroke–a systematic review. Topics in stroke rehabilitation, 23(2):116–125, 2016.

106. Krebs, H. I., Palazzolo, J. J., Dipietro, L., Ferraro, M., Krol, J., Rannekleiv, K., Volpe, B. T., and Hogan, N.: Rehabilitation robotics: Performance-based progressive robot-assisted therapy. Autonomous robots, 15(1):7–20, 2003.

107. Krebs, H. and Volpe, B.: Rehabilitation robotics. In Handbook of clinical neurology, volume 110, pages 283–294. Elsevier, 2013.

108. Huang, V. S. and Krakauer, J. W.: Robotic neurorehabilitation: a computational motor learning perspective. Journal of neuroengineering and rehabilitation, 6(1):5, 2009.

109. Lo, A. C., Guarino, P. D., Richards, L. G., Haselkorn, J. K., Wittenberg, G. F., Federman, D. G., Ringer, R. J., Wagner, T. H., Krebs, H. I., Volpe, B. T., et al.: Robot-assisted therapy for long-term upper-limb impairment after stroke. New England Journal of Medicine, 362(19):1772–1783, 2010.

110. Rodgers, H., Bosomworth, H., Krebs, H. I., van Wijck, F., Howel, D., Wilson, N., Aird, L., Alvarado, N., Andole, S., Cohen, D. L., et al.: Robot assisted training for the upper limb after stroke (ratuls): a multicentre randomised controlled trial. The Lancet, 2019.

111. Masiero, S., Celia, A., Rosati, G., and Armani, M.: Robotic-assisted rehabilitation of the upper limb after acute stroke. Archives of physical medicine and rehabilitation, 88(2):142–149, 2007.

112. Hesse, S., Mehrholz, J., and Werner, C.: Robot-assisted upper and lower limb rehabilitation after stroke: walking and arm/hand function. Deutsches Ärzteblatt International, 105(18):330, 2008.

113. Díaz, I., Gil, J. J., and Sánchez, E.: Lower-limb robotic rehabilitation: literature review and challenges. Journal of Robotics, 2011, 2011.

114. Liao, W.-w., Wu, C.-y., Hsieh, Y.-w., Lin, K.-c., and Chang, W.-y.: Effects of robot-assisted upper limb rehabilitation on daily function and real-world arm activity in patients with chronic stroke: a randomized controlled trial. Clinical rehabilitation, 26(2):111–120, 2012.

115. Lee, S.-H., Kim, Y.-M., and Lee, B.-H.: Effects of virtual reality-based bilateral upper-extremity training on brain activity in post-stroke patients. Journal of physical therapy science, 27(7):2285–2287, 2015.

116. Aguirre-Ollinger, G., Colgate, J. E., Peshkin, M. A., and Goswami, A.: Active-impedance control of a lower-limb assistive exoskeleton. In 2007 IEEE 10th international conference on rehabilitation robotics, pages 188–195. IEEE, 2007.

117. Kazerooni, H., Racine, J.-L., Huang, L., and Steger, R.: On the control of the berkeley lower extremity exoskeleton (bleex). In Proceedings of the 2005 IEEE

international conference on robotics and automation, pages 4353–4360. IEEE, 2005.

118. Krishnan, C., Ranganathan, R., Dhaher, Y. Y., and Rymer, W. Z.: A pilot study on the feasibility of robot-aided leg motor training to facilitate active participation. PloS one, 8(10):e77370, 2013.

119. Wagner, J., Solis-Escalante, T., Grieshofer, P., Neuper, C., Müller-Putz, G., and Scherer, R.: Level of participation in robotic-assisted treadmill walking modulates midline sensorimotor eeg rhythms in able-bodied subjects. Neuroimage, 63(3):1203–1211, 2012.

120. Abdel Majeed, Y., Awadalla, S. S., and Patton, J. L.: Effects of robot viscous forces on arm movements in chronic stroke survivors: A randomized crossover study. Journal of Neuroengineering and Rehabilitation, submitted.

121. Stereo Optical: Original Stereo Fly Stereotest. `https://www.stereooptical.com/products/stereotests-color-tests/original-stereo-fly/`. Accessed: 2019-11-14.

122. Patton, J., Dawe, G., Scharver, C., Mussa-Ivaldi, F., and Kenyon, R.: Robotics and virtual reality: a perfect marriage for motor control research and rehabilitation. Assistive Technology, 18(2):181–195, 2006.

123. Rahman, T., Sample, W., Seliktar, R., Alexander, M., and Scavina, M.: A body-powered functional upper limb orthosis. Journal of rehabilitation research and development, 37(6):675–680, 2000.

124. Lin, K.-c. C., Hsieh, Y.-w. W., Wu, C.-y. Y., Chen, C.-l. L., Jang, Y., and Liu, J.-s. S.: Minimal Detectable Change and Clinically Important Difference of the Wolf Motor Function Test in Stroke Patients. Neurorehabilitation and Neural Repair, 23(5):429–434, 2009.

125. Page, S. J., Fulk, G. D., and Boyne, P.: Clinically important differences for the upper-extremity fugl-meyer scale in people with minimal to moderate impairment due to chronic stroke. Physical therapy, 92(6):791–798, 2012.

126. Wolf, S. L., Thompson, P. A., Morris, D. M., Rose, D. K., Winstein, C. J., Taub, E., Giuliani, C., and Pearson, S. L.: The excite trial: attributes of the wolf motor

function test in patients with subacute stroke. Neurorehabilitation and Neural Repair, 19(3):194–205, 2005.

127. Winstein, C. J., Miller, J. P., Blanton, S., Taub, E., Uswatte, G., Morris, D., Nichols, D., and Wolf, S.: Methods for a Multisite Randomized Trial to Investigate the Effect of Constraint-Induced Movement Therapy in Improving Upper Extremity Function among Adults Recovering from a Cerebrovascular Stroke. Neurorehabilitation and Neural Repair, 17(3):137–152, 2003.

128. Brown, G., Gordon, K. E., et al.: Control of locomotor stability in stabilizing and destabilizing environments. Gait & posture, 55:191–198, 2017.

129. Majeed, Y. A., Abdollahi, F., Awadalla, S., and Patton, J.: Multivariate outcomes in a three week bimanual self-telerehabilitation with error augmentation post-stroke. In 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 1425–1431, Aug 2015.

130. Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G.: API design for machine learning software: experiences from the scikit-learn project. In ECML PKDD Workshop: Languages for Data Mining and Machine Learning, pages 108–122, 2013.

131. Herzfeld, D. J., Vaswani, P. A., Marko, M. K., and Shadmehr, R.: A memory of errors in sensorimotor learning. Science, 345(6202):1349–1353, 2014.

132. Haith, A. M. and Krakauer, J. W.: Model-based and model-free mechanisms of human motor learning. In Progress in motor control, pages 1–21. Springer, 2013.

133. Smith, M. A., Ghazizadeh, A., and Shadmehr, R.: Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS biology, 4(6):e179, 2006.

134. Grindlay, G.: Haptic guidance benefits musical motor learning. In 2008 Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pages 397–404. IEEE, 2008.

135. Cramer, S. C.: Repairing the human brain after stroke: I. mechanisms of spontaneous recovery. Annals of neurology, 63(3):272–287, 2008.

VITA

NAME Yazan 'Yaz' Abdel Majeed

EDUCATION B.Sc., Electrical Engineering, Princess Sumaya University for Technology, Amman, Jordan, 2010

TEACHING Teaching and Research Assistant, course designer, Biomedical Engineering Department, German Jordanian University (2010-2011)

Teaching Assistant, Biomedical Instrumentation Lab (BIOE 431, several semesters between 2012-2018)

Teaching Assistant, Materials in Bioengineering (BIOE 460)

Teaching Assistant, FDA & ISO Requirements for the Development & Manufacturing of Medical Devices (BIOE 410)

Teaching Assistant, Modeling Physiological Data & Systems (BIOE 240)

PUBLICATIONS Yazan Abdel Majeed, Farnaz Abdollahi, Saria Awadalla, and James Patton. "Multivariate outcomes in a three week bimanual self-telerehabilitation with error augmentation post-stroke" In Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC, 2015).

Yazan Abdel Majeed, Saria Awadalla, and James Patton. "Regression techniques employing feature selection to predict clinical outcomes in stroke" PloS one, 13(10), e0205639.

Justin Horowitz, Yazan Abdel Majeed, and James Patton. "A fresh perspective on dissecting action into discrete submotions" In Proceedings of the 38th annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC, 2016).

Eyad Hajissa, Courtney Celian, Kelly O Thielbar, Francois Kade, Yazan Abdel Majeed, and James L Patton. "Stroke Rehabilitation with Distorted Vision Perceived as Forces"In Proceedings of the IEEE

16th International Conference on Rehabilitation Robotics (ICORR, 2019).

Zachary Wright, Yazan Abdel Majeed, James Patton, and Felix Huang. "Components of Mechanical Work Predict Outcomes in Robotic Stroke Therapy"Journal of Neuroengineering and Rehabilitation, revision submitted for review August 2019.

Yazan Abdel Majeed, Saria Awadalla, and James Patton. "Effects of Robot Viscous Forces on Arm Movements in Chronic Stroke Survivors: A Randomized Crossover Study" Journal of Neuroengineering and Rehabilitation, 2019. *In Review.*

Yazan Abdel Majeed, Courtney Celian, and James Patton. "Training With Negative Viscosity Improves Clinical Outcomes After Stroke: A Case Study" Neurorehabilitation and Neural Repair, 2019. *Submitted.*

ABSTRACTS

Y. ABDEL MAJEED, J. L. PATTON; A predictive model pointing to speed for post-stroke neurorehabilitation. Society for Neuroscience, 2015

Y. ABDEL MAJEED, S. AWADALLA, J. L. PATTON; Building prediction models for clinical recovery after stroke. Society for Neuroscience, 2016

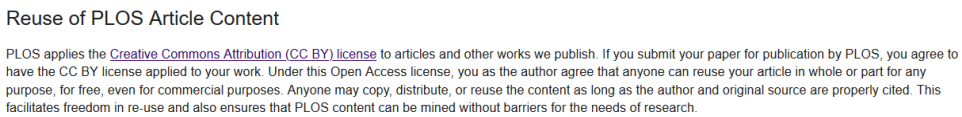
Y. ABDEL MAJEED, S. AWADALLA, J. L. PATTON; Effects of brief exposure to velocity-based perturbations on movement performance in stroke survivors. Society for Neuroscience, 2019

APPENDIX

JOURNAL COPYRIGHT INFORMATION

A.1 PLOS One Copyright Agreement

"PLOS applies the Creative Commons Attribution (CC BY) license to articles and other works we publish. If you submit your paper for publication by PLOS, you agree to have the CC BY license applied to your work. Under this Open Access license, you as the author agree that anyone can reuse your article in whole or part for any purpose, for free, even for commercial purposes. Anyone may copy, distribute, or reuse the content as long as the author and original source are properly cited. This facilitates freedom in re-use and also ensures that PLOS content can be mined without barriers for the needs of research."



Reuse of PLOS Article Content

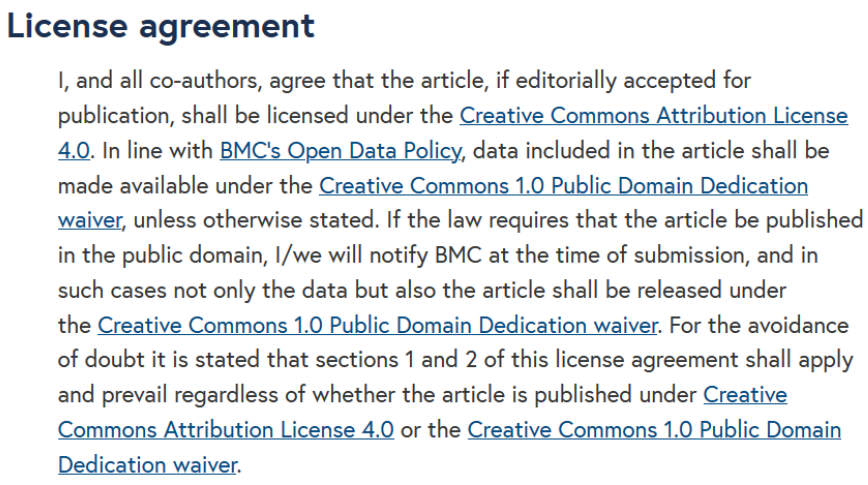
PLOS applies the Creative Commons Attribution (CC BY) license to articles and other works we publish. If you submit your paper for publication by PLOS, you agree to have the CC BY license applied to your work. Under this Open Access license, you as the author agree that anyone can reuse your article in whole or part for any purpose, for free, even for commercial purposes. Anyone may copy, distribute, or reuse the content as long as the author and original source are properly cited. This facilitates freedom in re-use and also ensures that PLOS content can be mined without barriers for the needs of research.

A.2 Journal of Neuroengineering and Rehabilitation License Agreement

"I, and all co-authors, agree that the article, if editorially accepted for publication, shall be licensed under the Creative Commons Attribution License 4.0. In line with BMC's Open Data Policy, data included in the article shall be made available under the Creative Commons 1.0

APPENDIX (Continued)

Public Domain Dedication waiver, unless otherwise stated. If the law requires that the article be published in the public domain, I/we will notify BMC at the time of submission, and in such cases not only the data but also the article shall be released under the Creative Commons 1.0 Public Domain Dedication waiver. For the avoidance of doubt it is stated that sections 1 and 2 of this license agreement shall apply and prevail regardless of whether the article is published under Creative Commons Attribution License 4.0 or the Creative Commons 1.0 Public Domain Dedication waiver."



License agreement

I, and all co-authors, agree that the article, if editorially accepted for publication, shall be licensed under the Creative Commons Attribution License 4.0. In line with BMC's Open Data Policy, data included in the article shall be made available under the Creative Commons 1.0 Public Domain Dedication waiver, unless otherwise stated. If the law requires that the article be published in the public domain, I/we will notify BMC at the time of submission, and in such cases not only the data but also the article shall be released under the Creative Commons 1.0 Public Domain Dedication waiver. For the avoidance of doubt it is stated that sections 1 and 2 of this license agreement shall apply and prevail regardless of whether the article is published under Creative Commons Attribution License 4.0 or the Creative Commons 1.0 Public Domain Dedication waiver.