Distribution Analysis of Motor Actions for the Design of Customized Robot-Assisted Therapy

BY

ZACHARY A. WRIGHT
B.S., University of Illinois at Chicago, 2009
M.S., University of Illinois at Chicago, 2012

THESIS

Submitted as 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, 2019

Chicago, Illinois

Defense Committee:

James Patton, PhD, Chair and Advisor
Felix Huang, PhD, Co-Advisor, Tufts University
Yang Dai, PhD, Bioengineering
Max Berniker, PhD, Mechanical and Industrial Engineering
Marc Slutzky, MD, PhD, Northwestern University

For my wife, Kendell, and our girls, Geri and Laney

Love you

ACKNOWLEDGEMENTS

Foremost, I would like to thank my family and friends for their continued support and motivation throughout my many academic years. My most heart-felt gratitude is extended to my parents, Gail and Howard, and my wife, Kendell, for their unwavering support, unconditional love, encouragement, patience and sacrifice.

I would like to express my sincerest appreciation for the invaluable guidance and mentorship for which my advisors, Dr. Jim Patton and Dr. Felix Huang, have provided since I started in the Robotics Lab as an intern in 2007. Thank you for always challenging me to become a better scientific investigator and engineer.

I would also like to acknowledge all members, past and present, of the Robotics Lab at the Shirley Ryan AbilityLab, and formerly, the Rehabilitation Institute of Chicago who have contributed and provided inspiration towards my work. In addition, I would like to thank all support staff at the University of Illinois at Chicago and at the AbilityLab/RIC.

I want to give a special thanks to my committee members – Dr. Yang Dai, Dr. Max Berniker and Dr. Marc Slutzky for their time and willingness to serve on my committee and for providing feedback and direction on my research.

~ZAW

CONTRIBUTION OF AUTHORS

Chapter 1 is an introduction that highlights the significance of the research aims presented in the dissertation for which I am the sole author.

Chapter 2 is a published manuscript in *IEEE Transactions on Neural Systems and Rehabilitation Engineering* (2017). I am the primary author and the major driver of the research; including experimental design, data collection, analysis and writing. Co-authors include my advisors, Dr. James Patton and Dr. Felix Huang, who assisted in experimental design, data analysis and writing. Also included, Emily Lazzaro, DPT and Kelly Thielbar, DPT, who conducted patient assessments and assisted in writing.

Chapter 3 is an unpublished manuscript re-submitted for peer-review in the *Journal of Neural Engineering and Rehabilitation* (August 2019). It is an extension of preliminary findings published in the *Proceedings of IEEE Engineering in Medicine and Biology Conference* (2018). I am the primary author and the major contributor to the research; including, data analysis and writing. Co-authors include Yazan Majeed, Dr. James Patton, Dr. Felix Huang who contributed to data analysis, interpretation of results and writing.

Chapter 4 is an unpublished manuscript submitted for review to the *Journal of Biomedical and Health Informatics* (November 2019) for which I am the primary author and the major contributor to the research. Co-authors include Dr. Yang Dai, Dr. Felix Huang and Dr. James Patton who contributed to study design, interpretation of results and writing.

Chapter 5 is a general discussion of the implications of the work presented in the dissertation for which I am the sole author.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF FIGURES

PAGE

LIST OF TABLES

PAGE

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
ARAT	Action Research Arm Test
CI	Confidence Interval
CMSA-A	Chedoke McMaster Stroke Assessment- Arm
FMA-UE	Fugl-Meyer Assessment – Upper Extremity
LASSO	Least Absolute Shrinkage and Selection Operator
LOOCV	Leave-One-Out Cross-Validation
MAS	Modified Ashworth Scale
MSE	Mean Squared Error
PCA	Principal Component Analysis
PLS	Partial Least Squares
PLS	Partial Least Squares Regression
PRESS	Predicted Residual Error Sum of Squares
ROM	Range of Motion
UEFM	Upper Extremity Fugl-Meyer
VAF	Variance Accounted For
2-D	Two-Dimensional

SUMMARY

The wide variation in upper limb motor impairments among stroke survivors presents a significant challenge to therapy. One approach is to customize treatment based on each individual's particular movement capabilities. Past work in our lab successfully allowed patients to move any way they wanted, freely exploring while being facilitated by robot-applied forces that amplify their movement velocities. This thesis builds upon this framework by introducing a statistical approach, *distribution analysis*, for characterizing each patient's patterns of movement during a special paradigm, *free exploration,* such that forces can be applied in a customized manner. Distribution analysis first builds a model of each individual's unique *motor deficits*, which then informs the design of training forces that push each patient's hand away from their typical movement velocities (i.e. higher probability bins) and towards their less visited velocity deficits (i.e. lower probability bins). We tracked the recovery of patients across weeks of such training using both clinical assessments and engineering metrics (Chapter II). As the success of any robotic intervention is often determined by whether patients are actively moving their affected limb, we relate their energetic contributions (quantified in terms of mechanical work) during training to their recovery outcomes and combine advanced multiple regression techniques to identify the most important biomechanical components of work (Chapter III). Lastly, we apply distribution analysis across a wider domain of variables (endpoint and joint kinematics, kinetics) and relate them to clinical measures, use them to classify stroke survivors and healthy individuals and describe the individual differences between stroke and healthy (Chapter IV). These findings represent a powerful set of new statistical modeling approaches for stroke therapy.

I. INTRODUCTION

Humans are highly variable in the way we move yet we can perform a wide range of movements in a similar manner. As such, movement ability should be described not only by our natural tendencies, but also our ability to express our full repertoire of available movements. The desired patterns of movement considerably change in people who have survived a brain injury, such as a stroke. They not only appear drastically different compared to neurally intact individuals, but also appear different from each other. This would mean that encouraging motions beyond the patient limits while still allowing them to freely move any way they would like could reveal their individual movement patterns. This thesis presents a data-driven approach for identifying each stroke survivor's unique *motor deficits* based on the statistical analysis of their free movement, called *distribution analysis.*

Clinical researchers are often searching for new strategies to improve upper limb therapy for stroke survivors. Despite a range of therapy approaches, many fail to account for the wide variation in their movement abilities. While traditional clinical assessments can guide the decisions of therapists, there is often no direct link to possible treatment options. Finding better strategies to describe individual differences and pinpoint the specific needs of each patient can potentially be more informative for therapy design. This thesis also offers an innovative method of using distribution analysis to inform the design of a therapy customized for each individual patient.

In recent years, rehabilitation engineers have developed several robotic devices which can easily measure upper limb motion. Clinical researchers have proposed several robot-based assessments to characterize movement abilities and track recovery. Some directly relate to clinically relevant measures of movement capabilities (e.g. range of motion, joint coupling),

while others rely on simple engineering metrics to describe performance on a particular motor task (e.g. reaching error). While functional movement skills are important in everyday life, current robot measures fail to capture the natural variability of patients' movements. As advancements in computing processing have made data cheaper to store and easier to access, one option is to record more data over a broader range of movements. An alternative strategy is to allow patients to move any way they want by freely exploring. In contrast to performing a set of prescribed motor tasks, self-directed motor exploration allows patients to practice a variety of movements throughout the entire workspace. As such, motor exploration provides vast amounts of data that offer more opportunities than small discrete tasks used in prior experiments for characterizing movement capabilities.

Scientists are constantly developing new analysis techniques to better visualize and identify useful patterns in such "big data". Recent work in our lab has investigated the value of analyzing motor exploration data to characterize stroke survivors' movement patterns. Our approach is to construct statistical distributions which provide a probabilistic view of each patient's movement tendencies. Probability distributions not only reflect an individual's typical movement patterns, but also over and under representation (F. C. Huang and Patton 2016). For example, less frequent or absent movement behaviors can point to a patient's specific *motor deficits*. In fact, effective robot assessments should be able to identify individual differences in motor capabilities. We previously showed that probability distributions of endpoint kinematic (i.e. hand motion) data recorded during motor exploration are unique to each stroke survivor (F. C. Huang and Patton 2016). Thus, each patient's specific motor deficits could be used to determine the motions which should be encouraged during therapy.

Beyond characterization of individual motor deficits, effective therapy also depends on key decisions about the type of intervention. Therapy that encourages active participation from the patient has shown to result in better recovery of motor function (Mark, Taub, and Morris 2006; Arya et al. 2011). In many cases the therapist will assist the patient's affected limb while the patient also attempts to actively contribute to the movement. However, the shared control during assisted movement therapy presents many challenges. As robots can provide forces as a function of anything, they offer new ways to precisely administer therapy. However, there are many possible strategies in which robot-applied forces can be programmed.

Many current strategies for robotic therapy are not flexible enough to facilitate movement during continuous practice such as motor exploration. Past work in our lab demonstrated how exploration training with negative damping (i.e. forces applied in the direction of movement velocities) can expand overall movement capabilities (F. C. Huang, Patton, and Mussa-Ivaldi 2010). One possible way to improve this method is to customize based on each patient's typical exploration movement patterns in the velocity domain. As such, this thesis builds upon and unifies our statistical approach to characterizing stroke survivors' individual *motor deficits* with our past successes in movement exploration force training. In Chapter II, we offer an innovative approach where the mathematical structure of patients' individual movement distributions forms the design of customized training forces to encourage them to explore their more deficient motions.

Clinical studies involving robotic therapy can provide new, large datasets where researchers can ask retrospective questions. While some methods of robotic forces have failed, the successful ones share a common feature - active participation by the patient while the robot facilitates movement only when needed (Blank et al. 2014). Hence, the level of active

participation may be an important factor in determining the success of any intervention. There are many ways to measure how active a patient is during therapy including heart rate monitors, oxygen uptake devices, or wearable fitness trackers to measure energy expenditure or metabolic cost or more direct measurements of muscle activity using electromyography (EMG). However, robots provide a convenient way to evaluate how much patients are actively contributing to movement during therapy since they can measure both motions and forces. In Chapter III, we present a computational approach to quantify patient *involvement* in terms of their energetic contributions (mechanical work) to limb movement during training.

The combination of motions and forces also provides a suite of additional variables (endpoint kinematics, joint kinematics and kinetics) which can be used to describe exploratory movement behaviors. While each variable could provide a unique description of patients' motor capabilities, it is unclear which variable(s) best capture individual differences which is critical to the design of customized therapy. If any variable is valuable, it should (1) be related to clinical assessments, (2) be able to easily differentiate a stroke survivor from a healthy individual, and (3) be able to pinpoint unique differences between stroke survivors. In Chapter IV, we rely on advanced statistical techniques and machine learning methods to investigate the best descriptors of motor deficits among several distribution variables.

Accordingly, this thesis determines how probability distributions of motor exploration data can be used to identify the motor deficits of each stroke survivor and how their deficits can be influenced by training with customized robotic forces. We first present the results of a clinical study where we hypothesized that patients who trained with customized forces would show better recovery (in terms of clinical outcomes and engineering metrics) compared to a control group who trained without forces (Chapter II). We then present the findings of a secondary

analysis where we tested the hypothesis that the amount of patient active involvement during robotic training, which we quantified in terms of mechanical work, dictates the extent of patient recovery (Chapter III). Finally, we revisit our distribution analysis to fully characterize stroke survivor's motor deficits across a range of variables and determine which set of distribution variables best predict clinical assessments, classify stroke survivors and healthy individuals and describe the variation between stroke survivors in how they are different from healthy individuals (Chapter IV). We hypothesized that distributions of kinetic variables would offer a better description of individual differences. Taken together, the findings of these three studies establish a complete package for rehabilitation that makes use of individual patient information to link diagnosis to personalized treatment.

II. ROBOT TRAINING WITH VECTOR FIELDS BASED ON STORKE SURVIVORS' INDIVIDUAL MOVEMENT STATISTICS[1][2]

Zachary A. Wright, Emily Lazzaro, Kelly O. Thielbar, James L. Patton and Felix C. Huang

Abstract: The wide variation in upper extremity motor impairments among stroke survivors necessitates more intelligent methods of customized therapy. However, current strategies for characterizing individual motor impairments are limited by the use of traditional clinical assessments (e.g. Fugl-Meyer) and simple engineering metrics (e.g. goal-directed performance). Our overall approach is to statistically identify the range of volitional movement capabilities, and then apply a robot-applied force vector field intervention that encourages under-expressed movements. We investigated whether explorative training with such customized force fields would improve stroke survivors' (n = 11) movement patterns in comparison to a control group that trained without forces (n = 11). Force and Control groups increased Fugl-Meyer UE scores (average of 1.0 and 1.1, respectively), which is not considered clinically meaningful. Interestingly, participants from both groups demonstrated dramatic increases in their range of velocity during exploration following only six days of training (average increase of 166.4% and 153.7% for the Force and Control group, respectively). While both groups showed evidence of improvement, we also found evidence that customized forces affected learning in a systematic way. When customized forces were active, we observed broader distributions of velocity that were not present in the controls. Secondly, we found that these changes led to specific changes in unassisted motion. In addition, while the shape of movement distributions changed significantly

[2] partial results published in *Proceedings of International Conference on Rehabilitation Robotics* (Z.A. Wright et al. 2015)

for both groups, detailed analysis of the velocity distributions revealed that customized forces promoted a greater proportion of favorable changes. Taken together, these results provide encouraging evidence that patient-specific force fields based on individuals' movement statistics can be used to create new movement patterns and shape them in a customized manner. To our knowledge, this study is the first to directly link engineering assessments of stroke survivors' exploration movement behaviors to the design of customized robot therapy.

A. INTRODUCTION

The thoughtful application of robotics to stroke rehabilitation could offer powerful tools to complement traditional approaches in therapy. The success of such strategies, however, demands that the design of therapy account for the wide variation in motor impairments that exist across stroke survivors. Clinical assessments (e.g. Fugl-Meyer Assessment (Fugl-Meyer et al. 1975), Wolf Motor Function Test (S. L. Wolf et al. 2001), Action Arm Reaching Test (Lyle 1981)) currently in use already recognize the differences in the level and types of impairment in stroke patients; including, abnormal muscle synergies (Dewald et al. 1995; Dewald and Beer 2001), muscle weakness (Lum, Burgar, and Shor 2003; Bohannon and Smith 1987), spasticity (Dietz and Sinkjaer 2007). Some assessments offer guidance to individualize treatment plans, while others are more reliable tools for tracking patients' recovery in response to therapy (Gladstone, Danells, and Black 2002; Sanford et al. 1993; Rabadi and Rabadi 2006). Of course, such clinical assessments are time-consuming and rely on subjective scoring that suffers from inter-rater variation and repeatability issues. One key emerging trend in therapy is the characterization of individual motor capabilities through the use of robotic devices, which afford high fidelity spatial and temporal recording of objective measures in a wide variety of tasks

(Krebs et al. 2014; Bosecker et al. 2010; Balasubramanian et al. 2012; L Zollo et al. 2011; Germanotta et al. 2015). However, it has not been obvious how such robot-based assessments can reflect a patient's clinically relevant capabilities, or how such information can inform the design of therapy.

While experience and judgment guide the assertions of clinicians in diagnosing patients, robotic tools offer the unrelenting and dispassionate facility to organize vast amounts of data. Current strategies for high-resolution automated assessment of upper-limb motor function have been limited to simple engineering metrics related to patients' performance on specific tasks (e.g. reaching accuracy) (Zollo et al. 2011; Germanotta et al. 2015; Bosecker et al. 2010; Krebs et al. 1999). Recently, Huang et al. demonstrated that statistical distributions of movement constructed from kinematic data during a self-directed motor exploration task provide a probabilistic view of stroke survivors' individual movement tendencies (F. C. Huang and Patton 2016). Interestingly, Huang et al. showed that each stroke survivor's patterns were distinct, reflecting individualized impairments, while neurally-intact individuals' patterns were similar to each other. We assert that free exploration encourages the full expression of movement and also provides a comprehensive description of an individual's motor deficits. Moreover, this unique feature of movement distributions provides a basis for the design of therapy that is customized to each individual.

Beyond characterization of motor deficits, effective robot therapy relies on key decisions about the form of intervention. One critical aspect relates to the amount of assistance the robot should provide to a patient. Therapy that encourages patient-mediated motions have shown to have better functional outcomes than guidance-based strategies in which the robot moves the limb (Lotze 2003; Marchal-Crespo and Reinkensmeyer 2009). Treatment that promotes affected

limb *action*, such as constraint-induced therapy that allows only motion of the affected limb, has been shown to strengthen muscle activity and promote neural growth (J. H. van der Lee et al. 1999; Edward Taub, Uswatte, and Pidikiti 1999; S L Wolf et al. 1989; Steven L Wolf et al. 2015; Mark, Taub, and Morris 2006). Other strategies demonstrate assistance as needed or gradual assistance based on real-time measures of performance (Emken, Benitez, and Reinkensmeyer 2007; Kahn, Rymer, and Reinkensmeyer 2004; Colombo et al. 2012). These strategies supply forces sufficient to overcome existing barriers or deficits while still requiring effort and control from the patient. However, one challenge with these approaches is that assistance is generally applied to specific movement types and goals (e.g. reaching to a target) (Kahn, Rymer, and Reinkensmeyer 2004; Patton et al. 2006; Lum et al. 2002) or motor impairments (e.g. muscle weakness). These approaches fail to account for the wide variety of movements in daily life. Motor exploration, on the other hand, enables practice over a broader range of movements. Broader exploratory movement patterns could serve as the foundation to improve functional skills. The ability of robots to provide forces as a function of movement makes them a potential candidate to reshape patients' movement patterns.

Our approach employed a robotic device to both characterize and retrain the unique movement patterns of stroke survivors. We first characterized individuals' typical distributions of movement from a motor exploratory task. A model of these distributions was incorporated into the design of patient-specific force fields that push their hands toward their less frequent motions. In this study, stroke survivors trained with customized forces, while a control group trained without forces. To assess changes in motor capability, we evaluated clinical measures as well as engineering metrics describing the range of motion during motor exploration. We previously presented a portion of this study which revealed evidence of increased velocity range

(Zachary A. Wright et al. 2015). Here, with a fully powered cohort, we further investigated how resulting movement patterns are related to the design of customized forces. Our findings demonstrate preliminary evidence of patient-specific force therapy reshaping the exploratory movement patterns of stroke survivors.

B. METHODS

1. Experimental Participants

Twenty-five stroke survivors participated in this study (Table I). All participants were screened prior to participation by a physical therapist ("rater"). Two participants did not pass the screening criteria and one participant dropped out of the study due to shoulder discomfort during force training. We excluded the data from these participants in our analysis. The main inclusion criteria were 1) chronic stroke (8+ months post-stroke) 2) hemiparesis with moderate to severe arm impairment measured by the upper extremity portion of the Fugl-Meyer Assessment (FMA-UE score of 15-50, (Fugl-Meyer et al. 1975)) 3) primary cortex involvement. The exclusion criteria included 1) severe sensory deficits in the limb 2) severe spasticity (Modified Ashworth of 4 preventing movement (Ashworth 1964)) 3) aphasia, cognitive impairment or visual deficits that would influence their ability to perform the experiment tasks and 4) Botox injection in the past 4 months. Individuals gave informed consent in accordance with the Northwestern University Institutional Review Boards (IRB) and were paid for their participation. Recruitment of stroke participants for this study was primarily through an institutional database that stroke survivors consent to join and be contacted regarding studies held at the Rehabilitation Institute of Chicago. This study is registered with the U.S. National Institutes of Health's clinical study database, ClinicalTrials.gov (Identifier: NCT02570256).

We planned our experiment as a randomized, double-blinded, parallel group, two-arm design. We randomly assigned participants to either the *Force* or *Control* groups, each having eleven participants. We chose to power this study based on group differences in our primary outcome measure, changes in FMA-UE scores. Power analysis for a two-sample t-test included an alpha of 0.05, a power of 0.80, and an effect size of 3.5. Note, our recruitment procedures did not prevent the participants with unstable baseline clinical measurements completing training. However, we removed the data of these participants from the main analysis (see section F.). We employed a Block randomization with two participants per block using the initial assessment of FMA-UE scores. A third party researcher other than the clinical rater and the experimenter assigned participants to the groups by flipping a coin. The experimenter did not explicitly state to the participants which group they were assigned to prior to the first session. However, details regarding training and the potential risks involved were stated upon receiving consent. The clinical rater was not present during training sessions and was blind to group assignments until the conclusion of the study. Two different raters performed the clinical assessments during the course of the study. The same rater evaluated a given participant for each of their four clinical evaluation sessions (see Table I).

Table I: Individual Participant Clinical Data

Participant ID	Age	Gender	Side Affected	Handedness	Stroke Type	Months Post Stroke	FMA-UE Score‡	ARAT Times (seconds)‡	ARAT Score‡	CMSA-A‡	Elbow ROM Flexion (degrees)‡	Elbow ROM Extension (degrees)‡	MAS Biceps‡	MAS Triceps‡
Force group														
F1	42	F	L	R	I	37	18.0 (+2.0)	1083.8 (-0.3)	8.0 (+4.0)	3 (+0)	145.0 (+0)	-31.0 (+5.0)	2.0 (+0)	1.75 (+0.25)
F2	75	F	R	R	I	310	19.5 (+0.5)	1082.4 (-0.7)	14.5 (-0.5)	3 (+0)	141.0 (+3.0)	-61.5 (+0.5)	3.0 (+0)	0 (+1.5)
F3	62	M	R	R	I	44	22.5 (+1.5)	748.9 (-80.7)	16.5 (+0.5)	3 (+0)	116.5 (-4.5)	-18.5 (-0.5)	1.75 (-0.25)	1.5 (+0.5)
F4	55	M	R	R	I	81	26.5 (+0.5)	92.1 (-34.3)	39.5 (+1.5)	3 (+0)	139.5 (-2.5)	-43.5 (+10.5)	1.5 (+0)	1.0 (+0)
F5	68	M	L	R	H	149	27.0 (+0)	564.5 (-69.4)	19.0 (+5.0)	3 (+0)	138.0 (-13)	-12.5 (-7.5)	1.0 (+0.5)	2.0 (+0)
F6	57	M	R	R	B	138	27.5 (+0.5)	550.3 (+15.4)	29.0 (-4.0)	3 (+0)	146.5 (-1.5)	-6.0 (-2.0)	0 (+0)	0 (+0)
F7*†	64	M	R	R	I	107	26.5 (+3.5)	370.1 (-34.5)	34.5 (+0.5)	3 (+0)	142.0 (-2.0)	-36.5 (+6.5)	1.5 (+0)	0 (+0)
F8*	60	F	L	R	I	23	37.5 (+4.5)	158.3 (-65.9)	36.5 (+1.5)	3 (+0)	135.0 (-4.0)	-7.5 (+7.5)	1.75 (+0.25)	0 (+1.0)
F9*	49	M	L	R	I	29	39.0 (-1.0)	225.4 (-68.4)	34.0 (+2.0)	3 (+1)	126.5 (+10.5)	-6.5 (+2.5)	1.75 (-0.25)	1.25 (-1.25)
F10*	61	M	R	L	I	149	45.5 (+0.5)	69.4 (+2.6)	38.0 (+0)	4.5 (-0.5)	130.5 (+4.5)	-3.0 (+0)	1.25 (-0.25)	0 (+0)
F11	56	F	R	R	I	76	48.0 (+2.0)	48.3 (-5.3)	55.5 (+0.5)	7 (+0)	138.5 (-3.5)	-4.0 (+0)	0 (+0)	0 (+0)
Control group														
C1	66	F	R	R	H	136	16.5 (+0.5)	1085.1 (-0.5)	9.5 (-0.5)	3 (+0)	122.0 (+0)	-31.5 (+10.5)	1.75 (+0.25)	1.25 (+0.25)
C2	54	F	R	R	I	30	23.5 (+2.5)	888.8 (-43.0)	18.5 (+1.5)	3 (+0)	143.0 (-3.0)	-17.0 (+4.0)	1.0 (+0.5)	2.0 (-1.0)
C3*	59	M	L	R	UI	69	24.5 (-2.5)	930.9 (-101.6)	22.5 (+1.5)	3 (+0)	134.0 (-2.0)	-42.5 (+9.5)	2.0 (+0)	1.5 (-0.5)
C4	64	F	L	R	UI	16	23.5 (-0.5)	1071.7 (-206.5)	17.0 (+1.0)	3 (+0)	150.5 (-4.5)	-18.5 (-26.5)	1.5 (+0)	0.75 (-0.25)
C5	45	F	L	R	I	105	25.0 (+3.0)	815.2 (+66.5)	13.0 (-1.0)	3 (+0)	130.0 (-2.0)	-8.0 (+1.0)	1.0 (-1.0)	2.0 (-1.0)
C6	56	M	R	L	H	80	27.0 (+2.0)	545.5 (+111.4)	23.0 (+0)	3 (+0)	136.0 (+4.0)	-11.0 (+3.0)	1.5 (+0)	2.0 (-1.0)
C7*†	71	F	L	R	I	15	34.5 (-2.5)	171.6 (+83.8)	36.5 (-3.5)	3.5 (-0.5)	132.0 (-8.0)	-9.0 (+5.0)	2.0 (-0.5)	0 (+0)
C8*†	41	M	R	R	UI	24	33.0 (+1.0)	180.1 (-30.6)	36.0 (+1.0)	3 (+0)	128.0 (+5.0)	-2.0 (-7.0)	1.75 (+0.25)	0.5 (-0.5)
C9	49	F	R	R	I	43	38.0 (+0)	307.2 (-104.6)	37.0 (-3.0)	3 (+0)	144.5 (-0.5)	0 (+0)	2.0 (+0)	2.0 (-1.0)
C10	55	M	L	L	UI	232	49.0 (+2.0)	73.5 (-4.2)	49.0 (+4.0)	5 (+0)	146.0 (+6.0)	0 (+0)	0 (+0)	0 (+0)
C11	48	F	L	R	H	127	50.0 (+2.0)	67.9 (-25.1)	54.0 (+2.0)	7 (+0)	143.5 (-2.5)	-4.0 (-4.0)	1.75 (-0.75)	0.5 (-0.5)

M – Male F – Female R – Right L – Left I – Ischemic H – Hemorrhagic B – Both Uk – Unknown SD – Standard Deviation

* A different clinical rater performed assessments for each of the evaluation sessions

† Data removed from the statistical analysis

‡ Reported as the average between baseline 1 and baseline 2 measurements (± change in Post relative to average baseline)

2. Experimental Apparatus

Participants performed planar motor tasks using a planar robotic device (manipulandum) and a custom video display system (Fig. 1), presented previously (Z. A. Wright et al. 2015; Gandolfo, Mussa-Ivaldi, and Bizzi 1996). The robotic device has two degrees-of-freedom allowing participants to move within the transverse plane. It is equipped two encoders at each of its joints to record the position and velocity of its end-effector. A force sensor attached to the end-effector measures the human-robot interaction. At the base of the manipulandum resides two torque motors capable of generating programmable forces. Participants viewed down into a nontransparent mirror which overlaid their arm. Their hand was not visible and arm was partially covered. We provided real-time feedback of the robot's end-effector position (green cursor) which overlaid directly on top of the participants' hand. Visual feedback also included instructions and measures of performance specific to each motor task participants performed.

Participants operated the robot's end-effector through a wrist brace attached to a revolute joint which allowed them to focus training on forearm and upper arm coordination. Participants also rested their forearm on an arm support which provided gravity assistance. Participants were situated with respect to the robot such that their shoulder lined up with the center of the experimental workspace (0.6 m x 0.4 m). The workspace boundaries (white outline) were visible to the participants. Participants were able to comfortably reach the bottom edge of the workspace. Due to constraints of the experimental set-up, two participants required modifications to the workspace area. For these particular circumstances, the bottom boundary of the workspace was shifted away from the participants so that their body did not overlap with the workspace (see Fig. 1).

The robot control and instrumentation was mediated with a Simulink-based XPC Target

computer, with a basic rate of 1kHz. Data was collected at 200 Hz and filtered using a 5th order Butterworth low pass filter with a 12 Hz cutoff. The robot produced endpoint forces through the two torque motors, and the controller compensated for inertial effects of the robot arm during all experiment phases.

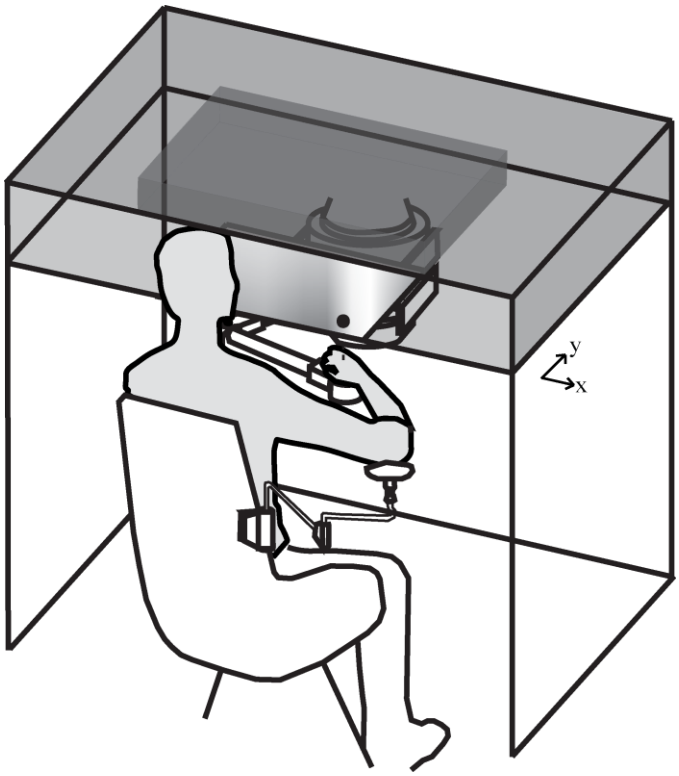


Figure 1: Experimental Apparatus. Participants performed a motor exploration task by controlling the arm of a planar robotic device.

3. Experimental Protocol

Each participant completed nine sessions across five weeks (Fig. 2). The first session (Baseline 1) and the second session (Baseline 2) served as initial evaluations and were separated by two weeks to establish baseline. Each evaluation included a clinical assessment followed by a performance assessment (described in detail below). We also evaluated participants two to three days (session 8; Post Evaluation) and six to eight days (session 9;

Follow-up Evaluation) following the final training session. Each participant trained three days per week for two weeks (sessions 2-7). At the beginning of each training session, participants completed the performance assessment. Training began on the same day as the Baseline 2 evaluation (session 2).

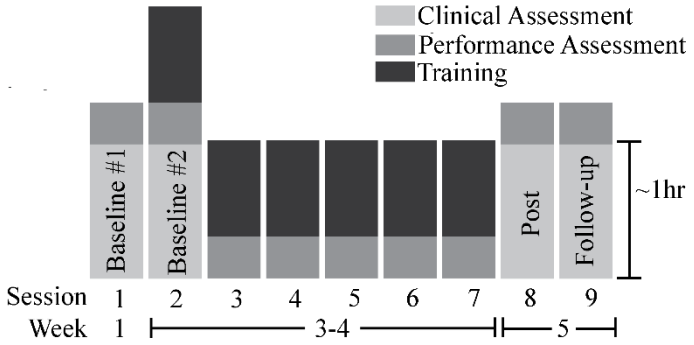


Figure 2: Training Schedule. Participants completed two weeks of motor exploration training in the presence of a customized force field.

a. Clinical Assessment

At the beginning of each evaluation session, a physical therapist administered a clinical assessment of participants' sensorimotor impairments of their affected arm. The clinical assessment included the FMA-UE (Fugl-Meyer et al. 1975), Action Research Arm Test (ARAT) (Lang et al. 2006), Modified Ashworth Scale (MAS) (Ashworth 1964), Chedoke McMaster Stroke Assessment- Arm (CMSA-A) (Gowland et al. 1993), and elbow range of motion (ROM).

b. Performance Assessment

For each session, participants completed three separate tasks using the robotic device:

Goal-Directed Reaching Task: Each session's performance assessment started with a goal-directed reaching task (Fig. 3A). The reaching task tested participants' ability to make straight-

line movements to circular visual targets (blue, 0.1 cm radius). Starting from a center target position, participants attempted to move a visual cursor (green, 0.05 cm radius) in a straight-line to a target in one of five outward target directions (0°, 45°, 90°, 135° or 180° relative to the positive x direction) located 15cm from the center target. Each outward movement was accompanied with a corresponding inward movement back to the center target (180°, -135°, -90°, -45° or 0° relative to the positive x direction). Participants were instructed to hold the cursor within each target for 0.5s. Participants attempted three movements to each target direction (i.e. 30 trials). The target locations were presented in block order and the order of targets remained the same across sessions. The center target was located anterior to the participants' shoulder; however, the distance from the shoulder to the hand varied depending on participants' arm length and range of motion. Some participants were unable to reach all the targets. If a movement attempt lasted longer than 8s, the experimental software advanced to the next target location. Following each reaching movement, participants received visual feedback on movement time. We defined movement time as the time from movement onset (speed > 0.04 m/s) to the time the cursor reached the target. Determination of movement onset was derived from previous experimental measurements of signal noise of the robot encoders when the robot handle is at rest. Applying a speed threshold of 0.04 m/s reasonably separates user-intended movement from resting noise. Movement times within a predetermined range of 400 and 750ms resulted in the appearance of the text "Good" in green color (Kawato 1999). The appearance of "Too Slow" in blue text and "Too fast" in red text indicated movement times slower than 750 ms and faster than 400 ms, respectively. Participants had difficulty achieving the task constraints on movement time. Thus, we provided additional encouragement by instructing participants to move as fast and accurate as possible to each target.

Goal-Directed Circular Movement Task: Each performance assessment also included a circular movement task (Fig. 3A). The circular movement task tested participants' ability to coordinate movement in a cyclical fashion. Participants attempted to make repetitive circles around a visual circular track (blue dotted-line, 10 cm radius). Starting from a target located at the top of the circular track, we instructed participants to move as fast and accurate as possible around the track until the track disappeared. The disappearance of the track marked the end of the trial. The track disappeared after the robot's endpoint traveled a total distance equivalent to four times the circular track's circumference. Participants attempted three movements in the clockwise and counterclockwise directions (six trials). Similar to the reaching task, we provided visual feedback regarding movement time. Movement time was defined as the time from movement onset to the time the cursor traveled the total distance specified. Participants attempted to achieve movement times within a predetermined range of 3-6s.

Movement Exploration Task: The primary portion of the performance assessment included a self-directed motor exploration task. For this task, participants were instructed to move the robot handle to all reachable positions within the robot workspace, at various speeds and movement directions. We also encouraged participants to avoid repeating the same movements continuously. Participants were undisturbed by the robot while performing the task. We informed participants they could rest at any time throughout the experiment. Each motor exploration trial (six trials for 12 minutes total) ended after two cumulative minutes of movement within the workspace. Movement speed below a threshold of 0.04 m/s was considered resting or no movement and the time points did not count towards the total movement time. We previously

determined that 12 minutes of motor exploration is a sufficient amount of data to accurately characterize a stroke survivor's motor behavior during the same task (Z. a Wright et al. 2014). Upon completing a trial, we provided participants with *Post-Trial feedback* related to their motor exploration performance (see section D.).

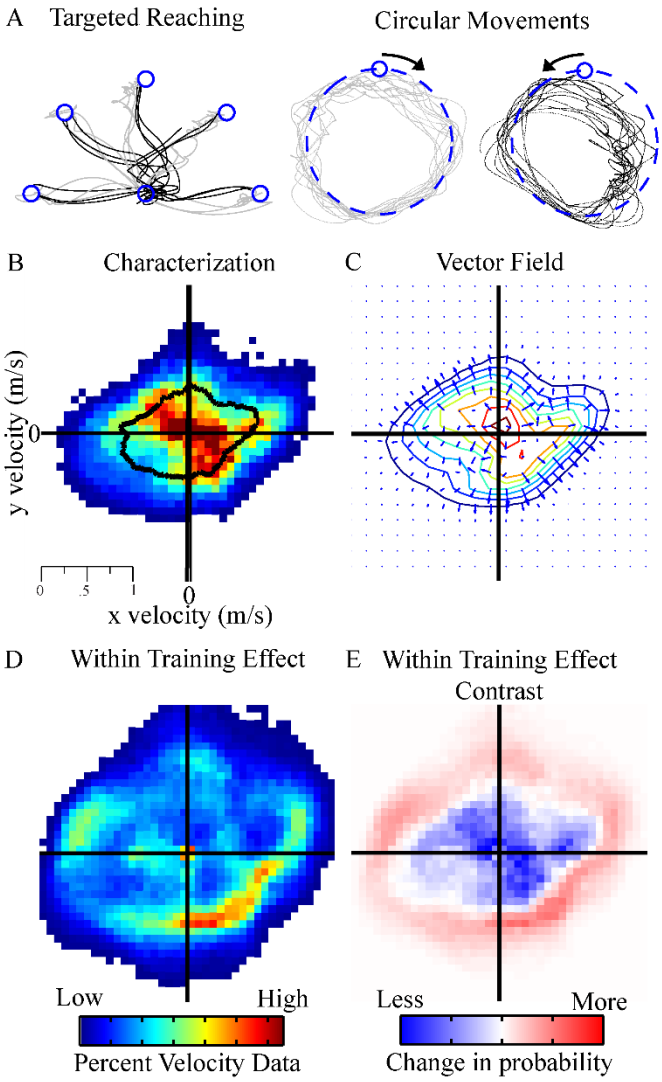


Figure 3: Movement Tasks. (A) Participants completed a goal-directed reaching task and a circular movement task at the beginning of each session. Typical participant's baseline movement trajectories are shown. (B) A typical Force participant's two dimensional probability distribution of velocity data tabulated across six trials of motor exploration during characterization, corresponding to 12 minutes of data. The black outline represents the 50th percentile contour of velocity data. The area of the contour corresponds to velocity coverage. (C) Customized force field designed by fitting a 2-D Gaussian model (colored contours) to the velocity data in (A) then calculating the gradient. The resulting vector field (blue arrows) represents the direction and relative magnitude of force applied during motor exploration training. (D) Training within a customized vector field pushed participants' movement patterns in (A) from high probability areas to low probability areas. (E) Contrast plot shows the relative change in probability between within training effect and characterization distributions.

c. Training

For each training session (sessions 2-7), participants completed an additional two blocks of eight, two-minute motor exploration trials (32 minutes in total) separated by a rest period (1-3 minutes). Participants performed the 32 minutes of motor exploration training either without forces (Control group) or within a customized force field (Force group) for all training sessions (see section E.)We informed participants that they could rest at any time during training. We also provided participants with *Post-Trial feedback* after each trial (see section D.)

d. Post-Trial Feedback on Motor Exploration

At the end of each two-minute motor exploration trial within the performance assessments and following each training session, participants received a score measuring the randomness of their movements, presented previously (Zachary A. Wright et al. 2015). We used a heuristic measure of randomness to encourage participants to express more variety in their movement patterns. The score was calculated by first dividing the experimental workspace into an 8 x 6 grid of two-dimensional (2-D) velocity-based histograms. Each 2-D histogram contained 25 bins in a 5 bin x 5 bin arrangement. Bin counts of individual histograms were based on the velocity of each data point located within the respective position of the workspace. Each histogram ranged from -1.25 to 1.25 m/s along both the x and y axes (lateral and fore-aft axes relative to the body) with each bin having a height and width of 0.5 m/s. For two minutes of motor exploration data (24000 data points), a completely uniform space (i.e. each bin having the same number of counts) equals 20 counts per bin. This was the maximum number of counts each bin could accumulate. The randomness score was determined by dividing the total number of counts across each bin by the total number of data points, displayed as a percentage. Following each completed trial for each participant, we displayed on the screen both their “Current” score

(score from the most recent trial) and "Best" score (highest score across all trials within a given session). We explained to the participants that the scores reflected how well they varied their movements during the task and we encouraged them to attempt to achieve the highest score possible (i.e. 100 percent).

e. Design of Vector Field

The exploration portion of the performance assessment (Characterization) during each training session served as a basis for the design of the customized force field used within each training session. More specifically, we first extracted the 2-D velocity data accumulated across the six trials of motor exploration during characterization (12 minutes of data in total). A 2-D histogram of velocity data offers a detailed view of how participants' movement patterns varied during motor exploration (Fig. 3B, Characterization). We express histograms as probability distributions by dividing each bin by the sum of the number of data tabulated in each histogram. A typical movement distribution from a stroke survivor exhibits areas (i.e. bins) of higher probability (red) and lower probability (blue). Fig. 3B shows a representative movement distribution constructed from 12 minutes of velocity data during characterization within a single session. Note, to visualize the probability distributions, we presented velocity histograms with 40 x 40 equally sized bins, scaled according to each participants' maximum (defined as the 99^{th} percentile) absolute velocity during motor exploration across all sessions.

We then fit the 2-D velocity data with a weighted sum of multivariate Gaussian-normal components according to maximum likelihood estimates (using the 'gmdistribution.fit' function in Matlab 2013):

$$f(x_1 \dots x_k) = \sum_{1}^{J} \frac{1}{(2\pi)^{k/2} S^{1/2}} e^{\left(-\frac{1}{2}(x_j-\mu)^T S^{-1}(x_j-\mu)\right)} \quad (1)$$

For the case of 2-D velocity, k = 2 and x1 and x2 represent velocity in the x and y directions, respectively. Each j-th component is associated with a covariance matrix, S, and a center, μ. It has previously been shown that smoothing of velocity data using a multivariate Gaussian kernel with five components accurately describes the complexity of stroke survivor's velocity distributions (F. C. Huang and Patton 2013b). When participants performed the motor exploration task, a high frequency of data accumulated near zero velocity during user-intended periods of rest and changes in movement direction. Thus, prior to fitting the data, we removed data with a speed below a threshold of 0.04 m/s. Each two-minute motor exploration trial contained 24000 data points. An example of a Gaussian distribution obtained from the model fit of 2-D velocity data during characterization (shown in Fig. 3B) is shown in Fig. 3C (colored contour lines).

Computing the gradient of (1) results in a velocity-dependent continuous function whose output are vectors that represent the slope along the 2-D Gaussian distribution. In principle, the direction of the vectors point from higher probabilities towards lower probabilities of the distribution. An example of a vector field derived from calculating the gradient of the multivariate Gaussian distribution is shown in Fig. 3C (blue arrows). The vector field represents the direction and relative magnitude of force applied during motor exploration training. The applied force was updated continuously based on the current velocity of the robot's endpoint while participants performed motor exploration. An example probability distribution of velocity data during motor exploration training within a vector field (shown in Fig. 3C) is shown in Fig. 3D (within training effect). The probability distribution in Fig. 3D was constructed from 32

minutes of velocity data during motor exploration training. The magnitude of the applied force was determined by 1) normalizing the current vector magnitude by the 80th percentile of the vector magnitudes calculated across the velocity data accumulated during characterization and 2) applying a gain equal to 2% of the participant's body weight (i.e. the approximate weight of the arm). We developed this heuristic normalization technique during pilot testing of the vector field. It accounts for differences in the vector magnitudes and differences in participants' arm impedances. For safety, the applied forces smoothly decreased to zero magnitude outside the workspace boundaries.

3. Analysis

a. Clinical Outcomes

Our primary clinical outcome measure to determine the therapeutic benefit of motor exploration training on overall arm function was changes in clinical FMA-UE scores. We compared FMA-UE scores assessed for each of the evaluation sessions (see Fig. 2). We summarize these results in terms of the change in FMA-UE scores relative to the average score between Baseline 1 (session 1) and Baseline 2 (session 2) evaluations. Statistical differences in FMA-UE were analyzed using a 2 (session: Average Baseline, Post) x 2 (training group: force, control) repeated measures Analysis of Variance (ANOVA). We considered statistical differences significant at α of 0.05.

Prior to analysis, we removed data from participants who demonstrated unstable FMA-UE scores between Baseline 1 and Baseline 2 sessions. To determine stability, we calculated the minimal detectable change needed to exceed measurement error (i.e. the 95% confidence interval of the standard deviation for Baseline 1 and Baseline 2 scores) (Wagner and Rhodes 2008).

Among the participants, two Control group participants and one Force group participant showed a change in baseline measurements greater than the calculated minimal detectable change threshold (2.5 points). We removed the data of these three participants for all subsequent metrics.

b. Motor Exploration Performance

We first evaluated whether participants' movement patterns improved following training, in terms of the range of velocities spanned during motor exploration (characterization) for each session. The metric we used, 50^{th} percentile coverage, represents the estimated area of participants' median movement tendencies in the velocity domain. We first calculated the 50^{th} percentile contour of 2-D velocity data, and then calculated the area (m^2/s^2) within the boundary formed by this contour. The boundary was formed by connecting points represented by the median (i.e. 50^{th} percentile) speed within 64 equally spaced bins radially aligned within the range of 0-2π (see Fig. 3B for a representative 50^{th} percentile contour of velocity data, black outline). We summarize these results in terms of the percent change in 50^{th} percentile coverage relative to the Baseline 2 evaluation (session 2). Statistical differences in coverage were analyzed using a 2 (session: Baseline 2, Post) x 2 (training group: force, control) repeated measures Analysis of Variance (ANOVA).

While velocity coverage revealed overall changes in terms of the range of movements, we also wished to quantify changes in the patterns of movement. We first constructed 2-D probability distributions of velocity data from motor exploration during characterization and training within each session. This analysis featured probability distributions with 100x100 equally sized bins, which were scaled to a common maximum range across all participants (±2 m/s). The bin-by-bin difference between two given probability distributions represented the

change in probability (For visualization purposes, the histogram is presented with 40x40 bins. See Fig. 3E). Positive changes (red) and negative changes (blue) in probability within each bin indicate an increase and decrease in data, respectively. To quantify the difference between two probability distributions, we defined the *contrast score* as the total sum of the absolute difference in probability between corresponding bins.

We present some simple metrics to characterize learning as well as the impact of training forces. We first determined the how training affects subsequent unassisted conditions (cumulative transfer effect). We calculated the contrast score between the probability distributions of velocity data from Baseline 2 characterization and that of each successive session (sessions 3-9). To determine group differences in the cumulative transfer effect, we compared scores calculated between Baseline 2 and Post evaluations. Aside from measures of learning, we also wished to characterize differences in the direct experience of training between groups. To do so, we computed a contrast score between the probability distributions of velocity from training data and characterization data within each session (within training effect). We compared the mean within training effect contrast scores between groups. Statistical differences were analyzed using a Student's t-test ($\alpha = 0.05$).

We devised a novel analysis to test whether the robot mediated training promoted changes in learning that corresponded to the design of customized forces. Our approach was to examine whether the changes in each participant's movement behaviors *across* training (cumulative transfer effect) were similar to the changes *within* training (within training effect). We computed the Pearson's correlation between the Baseline 2 and Post contrast (cumulative transfer effect contrast) and the average characterization and training contrasts within each training sessions (average within training effect contrast). As with the contrast score, we

employed a common maximum range and bin density across all participants (±2 m/s, 100x100 bins). Statistical differences between the training groups were analyzed using a Student's t-test ($\alpha = 0.05$).

c. Goal-Directed Performance

To determine the effect of training on goal-directed task performance, we compared changes in movement error, peak speed and duration between Baseline 2 (session 2) and Post (session 8). For the reaching task, our primary error metric was the maximum perpendicular distance along the movement trajectory (from movement onset (speed > 0.04 m/s) to when the cursor reached the target) with respect to the *ideal* straight-line path to the target. Besides this primary metric, we performed supplementary analysis of the *path length ratio* defined as the total distance traveled for each movement normalized with respect to the distance between targets (15cm). We also defined peak speed as the maximum speed along the trajectory and duration as the total time from movement onset to when the cursor reached the target. Statistical differences were analyzed using a 2 (session: Baseline 2, Post) x 2 (training group: force, control) x 10 (movement directions) repeated measures Analysis of Variance (ANOVA).

For the circular movement task, we compared changes in movement error, average speed and duration. Our primary measure of error was the mean radial deviation relative to a reference track defined by the mean radius of the movement trajectory. We first measured the mean center of each movement trajectory (from movement onset to when the cursor traveled a total distance equivalent to four times the circumference of the circular track). Then, we computed the distance between each point along the movement trajectory and the mean center. The average of these distances served as the radius of the circular reference track. We then calculated the mean

distance between each point along the movement trajectory and the reference track. Besides our main metric for circular movement, we computed the average speed in terms of mean speed along the movement trajectory and movement duration at the total time from movement onset to when the cursor traveled the total distance equivalent to four times the circular track's circumference. Statistical differences were analyzed using a 2 (session: Baseline 2, Post) x 2 (training group: force, control) x 2 (movement direction: clockwise and counterclockwise) repeated measured Analysis of Variance (ANOVA). We considered statistical differences significant at α of 0.05.

C. RESULTS

1. *Clinical Outcomes*

Our pre-declared primary outcome measure, change in FMA-UE scores, showed that both training groups improved with training (Post evaluation) compared to the average between baseline evaluations (Fig. 4); however, we failed to detect a significant difference between training groups. After removing the data from three participants with unstable baseline FMA-UE scores, the mean change and 95% confidence interval (CI) in FMA-UE scores were 1.1 (CI: 0.0, 2.2) and 1.0 (CI: -0.3, 2.3) for the Force and Control group, respectively (session: $F(1, 17) = 7.9$, $p = 0.01$; training group: $F(1, 17) = 0.0005$, $p = 0.9$). We also found that increases in FMA-UE scores persisted for one week (six-eight days) following training for five of the Force participants and five of the Control participants (Fig. 4). The mean change in FMA-UE scores from the Average Baseline evaluation to the Follow-up evaluation were 1.0 (CI: -0.4, 2.4) and 1.8 (CI: 0.7, 2.9) for the Force and Control group, respectively (session: $F(1, 17) = 8.8$, $p = 0.009$; training group: $F(1, 17) = 0.0$, $p = 0.99$). We also evaluated changes in our secondary

clinical outcomes; including, ARAT, CMSA-A, MAS and elbow ROM (see Table I for individual participant data). We summarized these results and provided additional analysis that considers all participant data (see Table II).

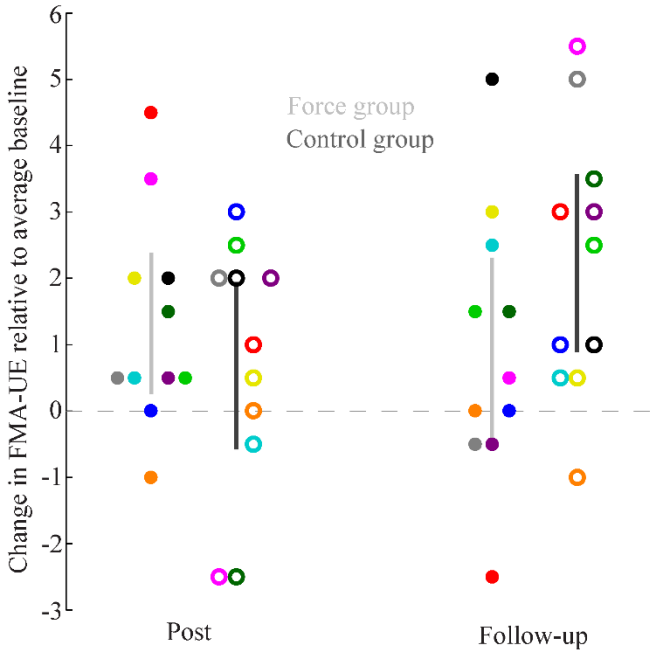


Figure 4: Clinical Outcome. Both training groups improved clinical FMA-UE scores following two weeks of training. Each color represents a stroke participant (●, Force; ○, Control) corresponding to participants' designated color in Table I; data points are staggered horizontally to avoid overlap. Vertical bars represent the mean and 95% confidence interval (gray, Force; black, Control).

2. Motor Exploration Performance

Our analysis of motor exploration first examined the extent to which motor capabilities improved, and then how movement probabilities were redistributed, and finally whether such changes could be attributed to training conditions. We found that the training groups demonstrated similar increases in velocity coverage (See Fig. 5, session: $F(1, 17) = 24.5$, $p = 0.0001$). However, we failed to detect a significant difference between groups; $F(1, 17) = 0.14$, $p = 0.7$. The mean 50^{th} percentile coverage during Baseline 2 was 0.44 (CI: 0.17, 0.72)

$(m/s)^2$ for the Force group and 0.60 (CI: 0.09, 1.1) $(m/s)^2$ for the Control group and during Post evaluation was 0.87 (CI: 0.43, 1.30) $(m/s)^2$ for the Force group and 0.91 (CI: 0.43, 1.40) $(m/s)^2$ for the Control group. Increases in velocity coverage corresponded to increases in distance traveled (group mean of distance traveled averaged across six motor exploration trials during Baseline 2 (Force: 46.3 (CI: 31.5, 61.0) m; Control: 51.4 (CI: 30.8, 71.9) m) and Post characterization (Force: 64.4 (CI: 48.3, 80.5) m; Control: 68.6 (CI: 53.8, 83.3) m). We present each participant's probability distribution from motor exploration during Baseline 2 and Post evaluations (See Fig. 9).

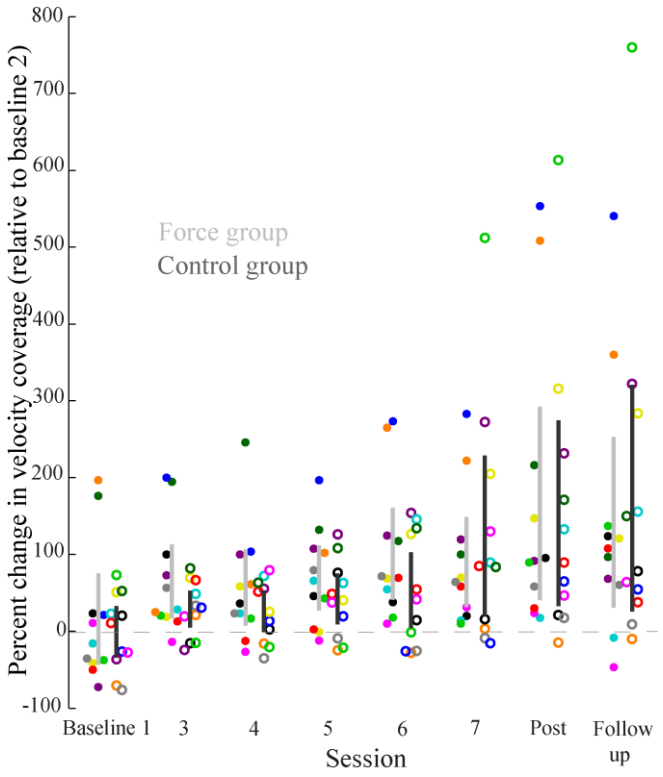


Figure 5: Change in Velocity Coverage. Both training groups improved exploratory movement behaviors in terms of velocity coverage. Each data point (●, Force;○, Control) represents a stroke participants' cumulative transfer effect contrast score across each session. Each stroke participant is represented by a color according to Table I; data points are staggered horizontally to avoid overlap. Vertical bars represent group (gray, Force; black, Control) mean and 95% confidence interval within each session.

Besides changes in velocity coverage, we observed a gradual increase in the degree of change in movement patterns, as indicated by the transfer effect contrast scores across sessions for both training groups (See Fig. 6). Each participant's cumulative transfer effect contrast plot (See Fig. 9) corresponds to the contrast score between Baseline 2 and Post characterizations. We failed to detect a significant difference between groups; $t(17) = 0.5$, $p = 0.63$. Note, to visualize the probability distributions, we presented velocity histograms with 40 x 40 equally sized bins, scaled according to each participants' maximum (defined as the 99th percentile) absolute velocity during motor exploration across all sessions (See Fig. 9).

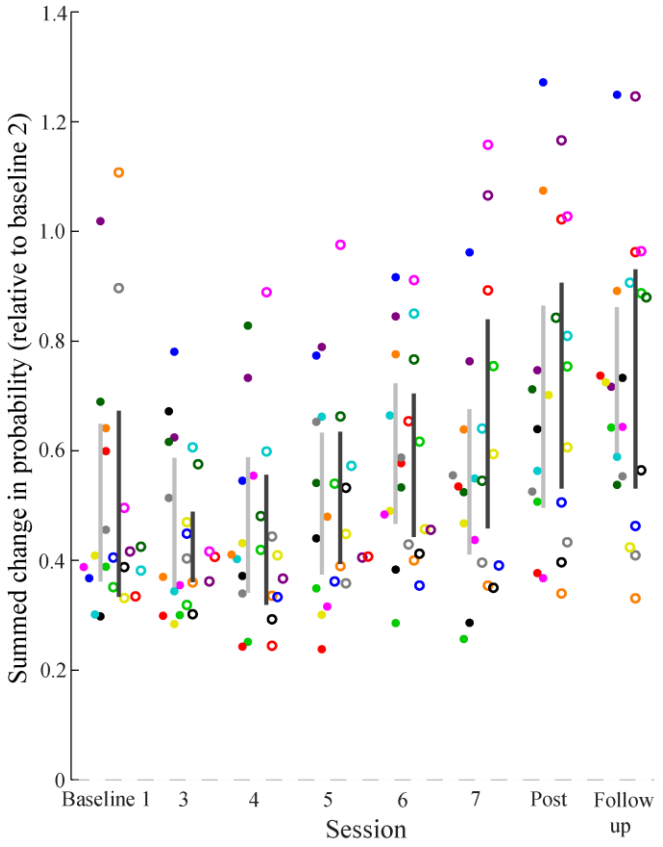


Figure 6: Cumulative Transfer Effect Velocity Contrast Score. Movement behaviors deviated from Baseline 2 characterization across sessions. Each data point (●, Force;○, Control) represents a stroke participants' cumulative transfer effect contrast score across each session. Each stroke participant is represented by a color according to Table I; data points are staggered horizontally to avoid overlap. Vertical bars represent group (gray, Force; black, Control) mean and 95% confidence interval within each session.

The apparent similarities between groups in these changes, however, contrast starkly to the large differences in training conditions. We examined how movement behaviors differed from the initial characterization (null field) to later training within the same session. As expected, Force participants' movement behaviors were drastically altered when training in the presence of forces. Fig. 7A (top) shows a typical Force participant's contrast plots between training and characterization probability distributions within each training session. In contrast, Control participants displayed movement behaviors during training that were similar to that of the beginning of the session (Fig. 7A, bottom). Group means across training sessions revealed significant differences in the within training effect contrast scores; $t(17) = 8.34$, $p < 0.05$ (Fig. 7B). Fig. 9 shows each participant's within training effect contrast plots averaged across training session. Group differences in contrast scores may be explained, in part, by differences in velocity coverage during training. The Force group demonstrated greater velocity coverage (mean across training sessions 2-7, 1.16 (CI: 0.58, 1.75) $(m/s)^2$) compared to the Control group (0.72 (CI: 0.28, 1.17) $(m/s)^2$) which corresponded to a longer average distance traveled during motor exploration training (mean across training sessions 2-7; Force group, 71.1 (CI: 53.9, 88.4) m; Control group, 59.0 (CI: 42.4, 75.8) m).

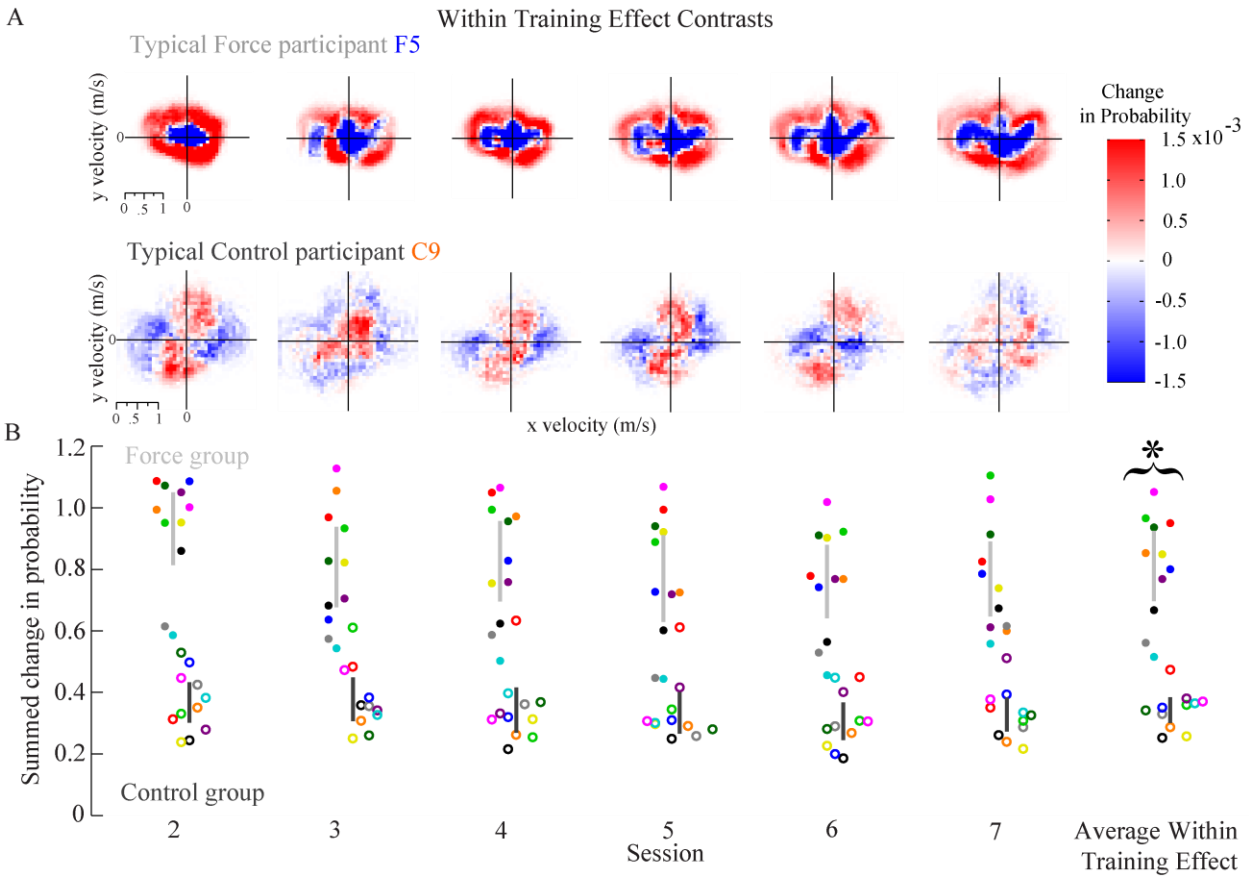


Figure 7: Within Training Effect Velocity Contrast Scores. Velocity distributions were significantly altered during vector field training. (A) Representative contrast plots showing the change between characterization and training velocity distributions within each training session (top row, Force; bottom row, Control). Red and blue shading indicates the relative amount of increase and decrease in velocity data within each bin, respectively. (B) The Force group demonstrated significantly greater within training effect contrast scores compared to the Control group. Each data point (●, Force; ○, Control) represents a stroke participant. Each stroke participant is represented by a color according to Table I; data points are staggered horizontally to avoid overlap. Vertical bars represent group (gray, Force; black, Control) mean and 95% confidence interval. The asterisk represents significance between training groups ($\alpha < 0.05$).

Beyond the general changes described above, we performed new analyses that were supplementary to our planned metrics to better reveal specific differences in learning due to forces. While the contrast score provides critical information about the degree of movement redistribution, it does not indicate whether such changes necessarily improved motor exploration. We created a metric, the favorability score, which summarized the way that our intervention may

have reversed a person's initial deficits, either by decreasing over-expressed or increasing under-expressed velocities. We examined to what extent the observed velocity states for each participant exhibited favorable increases or decreases in probability. For each session, we tabulated two-dimensional velocity histograms with a common maximum range and high bin density (+/-2 m/s, 100x100 bins). We defined a distribution midpoint as half the peak probability for the reference distribution (See Fig 8A). Using the initial baseline distribution on session 2 as the reference point for each participant, we computed favorable changes as the sum of all increases of each velocity state for which the initial distribution was below the midpoint, as well as all decreases of each velocity state for which the initial distribution was above the midpoint. All other changes in the distribution were then evaluated as unfavorable changes. As a final metric, we evaluated the favorable changes as a proportion of the total change for each participant. We observed similar changes across multiple sessions, and similar session dependence between groups (repeated measures ANOVA; session: $F(5, 85) = 1.38$, $p = 0.24$; session x group: $F(5, 85) = 1.19$, $p = 0.32$). Considering the sessions as a whole, this metric of favorable change was actually greater for the Force group compared to the Control group (Fig. 8B, average of sessions 3-8, $\Delta = 0.10$ (CI: 0.035, 0.16), $t(17) = 3.2$, $p = 0.005$).

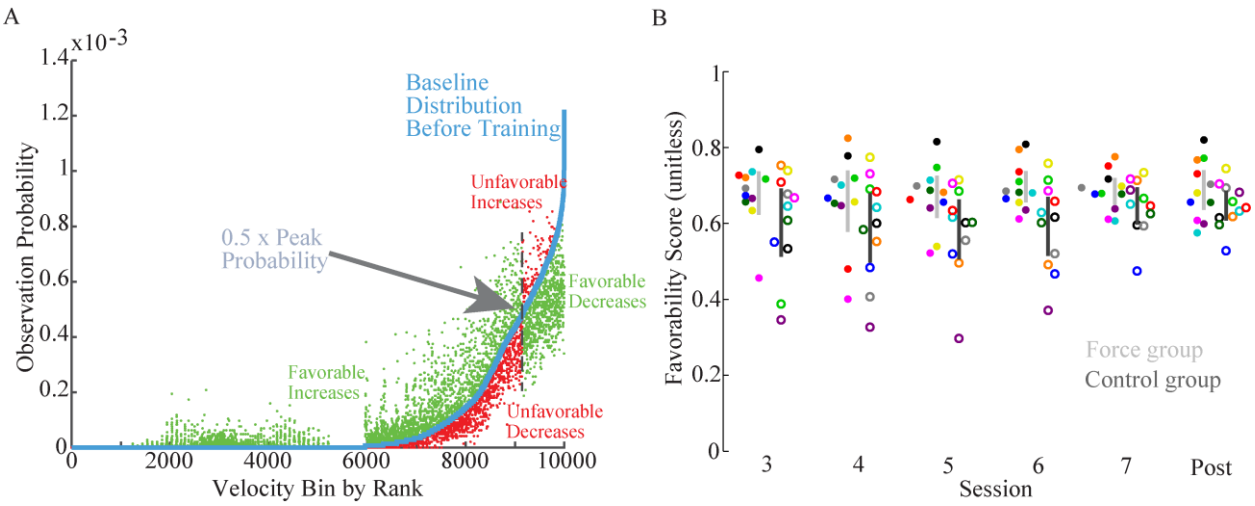


Figure 8: Velocity Favorability Score. (A) A typical baseline velocity distribution for one participant before training (Day 2, blue), and the corresponding probabilities after training (Day 3+, green and red), are shown here each with bins sorted according to the baseline magnitudes (day-2). After training, a new distribution reveals velocities that have exacerbated ("unfavorable changes", red dots) the original trends of under-expressed or over-expressed probabilities (defined operationally as the values separated by the midpoint of 0.5 peak probability). In other cases, the new distribution indicates velocities in which the original trends were reversed ("favorable changes", green dots). (B) We computed a metric as the sum of all favorable changes at each velocity bin as a proportion of all changes. Our results showed that the Force group exhibited significantly higher favorability scores compared to the Control Group (average of sessions 3-8, $\Delta = 0.085$, CI: -0.16, 0.0072, $p = 0.034$).

While the favorability score indicated possible advantages from training with forces, we also devised a novel analysis to test whether changes in movement behavior were consistent with the design of the customized environments. We observed that changes in probability distributions within training were similar to the changes between Baseline 2 and Post (Fig. 9, Cumulative transfer effect contrast compared to Average within training contrast). Increases and decreases in probability from Baseline 2 and Post correlated with their respective probabilities for training, as depicted by red and blue shaded areas, respectively. In other words, velocities that increased their representation each training session also increased by the end of training (Post), and velocities that decreased their representation each training session also decreased by the end of training. This supports the idea that Force participants' preserved the changes in

movement behaviors that were trained using unique training forces. The mean of the Pearson's Correlation for the Force and Control groups were 0.60 (CI: 0.41, 0.79) and 0.36 (CI: 0.09, 0.62), respectively. Interestingly, we failed to detect significant differences between the groups ($t(17) = 1.76$, $p = 0.1$), which suggests that short term changes in movement distribution are predictive of longer term learning. Fig. 9 shows each participant's cumulative transfer effect contrast (Baseline 2-to-Post) and average within training contrast with the corresponding correlation. Supplementary analysis of motor exploration performance, considering all participant data, yielded similar results (see Table III).

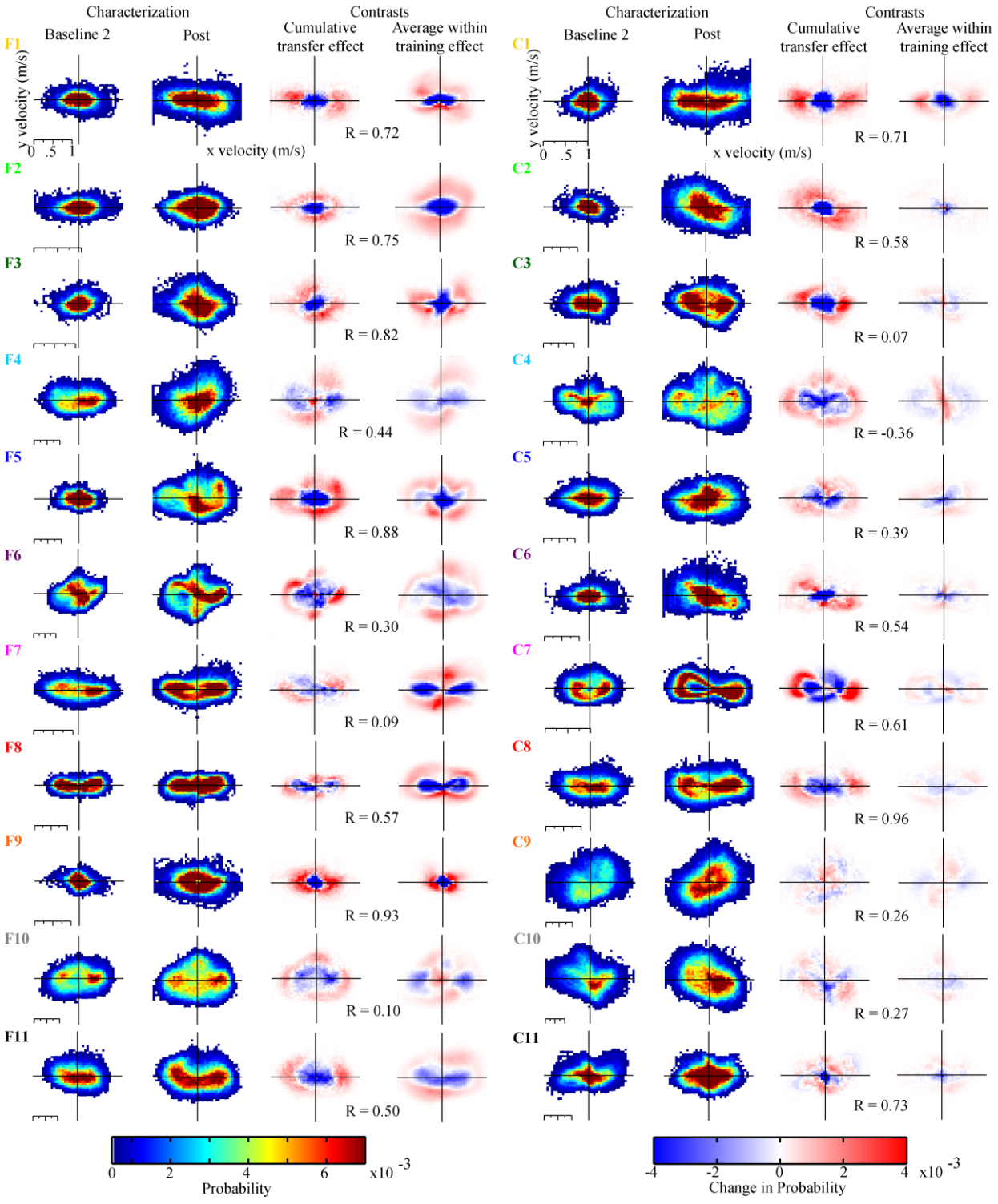


Figure 9: Individual Participant Velocity Distributions. Individual participants' (left, Force; right, Control) velocity distributions of motor exploration characterization prior to (Baseline 2) and following training (Post). Changes in movement behaviors across training (cumulative transfer effect) were correlated with changes during training (average within training effect contrast).

3. Goal-Directed Performance

We also evaluated the effect of motor exploration training on participants' performance during the goal-directed movement tasks that were not trained within each session. Both training groups reduced movement error in the reaching task from Baseline 2 evaluation to Post evaluation; however, we failed to detect significant differences between groups. Surprisingly, for the circular movements, both training groups increased movement error following training. For the reaching task, the mean difference in maximum perpendicular distance for the Force and Control group was -0.28 (CI: -0.59, 0.4) cm and -0.05 (CI: -0.48, 0.38) cm, respectively (session: $F(1, 323) = 2.9$, $p = 0.09$; training group: $F(1, 17) = 0.3$, $p = 0.6$). The mean difference in path length ratio for the Force and Control group was -0.11 (CI: -0.26, 0.03) and -0.02 (CI: = -0.19, 0.14), respectively (session: $F(1, 323) = 6.7$, $p = 0.01$; training group: $F(1, 17) = 0.02$, $p = 0.9$). For the circular movements, the mean difference in average radial deviation for the Force and Control group was 0.13 (CI: -0.09, 0.34) cm and 0.22 (CI: -0.24, 0.69) cm, respectively (session: $F(1, 51) = 5.2$, $p = 0.03$; training groups: $F(1, 17) = 1.9$, $p = 0.2$).

Beside the planned analyses above, we performed a post-hoc analysis of movement duration and speed during the goal-directed tasks and found changes consistent with the increases in velocity coverage observed during motor exploration. For each task, both training groups significantly decreased movement duration and increased speed; however, we failed to detect significant differences between groups. For the reaching task, the mean difference in movement duration for the Force and Control group was -0.31 (CI: -0.65, 0.02) s and -0.07 (CI: -0.24, 0.10) s, respectively (sessions: $F(1, 323) = 8.9$, $p = 0.003$; training group: $F(1, 17) = 0.09$, $p = 0.8$). The mean difference in peak speed for the Force and Control group was 0.006 (CI: -0.03, 0.05) m/s and 0.04 (CI: -0.03, 0.10) m/s, respectively (session: $F(1, 323) = 7.8$, $p < 0.05$; training

group: $F(1, 17) = 0.20$, $p = 0.70$). For the circular movement task, the mean change in movement duration for the Force and Control group was -5.38 (CI: -8.75, -2.01) s and -7.68 (CI: -13.43, -1.92) s, respectively (session: $F(1, 51) = 36.6$, $p < 0.05$; training group: $F(1, 17) = 0.30$, $p = 0.6$). The mean difference in mean speed for the Force and Control group was 0.11 (CI: 0.04, 0.17) m/s and 0.20 (CI: 0.11, 0.29) m/s, respectively (session: $F(1, 51) = 76.2$, $p < 0.05$; training group: $F(1, 17) = 1.5$, $p = 0.2$). Additional analysis of goal-directed performance, which considered all participant data, yielded similar results (see Table IV).

Table II: Statistical Analysis Summary – Clinical Outcomes

	Participant data with unstable baseline FMA-UE removed from analysis (Force group, n = 10; Control group, n = 9)		All participant data included in analysis (both groups, n = 11)	
	Force group mean (CI)	**Control group mean (CI)**	**Force group mean (CI)**	**Control group mean (CI)**
Clinical Outcomes				
FMA-UE score				
Change from Baseline* to Post	1.1 (0.0, 2.2)	1.0 (-0.3, 2.3)	1.3 (0.3, 2.4)	0.7 (-0.6, 1.9)
	session: F(1, 17) = 7.9, p = 0.01 training group: F(1, 17) = 0.0005, p = 0.9		session: F(1, 20) = 7.91, p = 0.015 training group: F(1, 20) = 0.005, p = 0.94	
Change from Baseline* to Follow-up	1.0 (-0.4, 2.4)	1.8 (0.7, 2.9)	1.0 (0.4, 2.4)	2.2 (0.9, 3.5)
	session: F(1, 17) = 8.76, p = 0.009 training group: F(1, 17) = 0, p = 0.99		session: F(1, 20) = 13.51, p = 0.002 training group: F(1, 20) = 0.08, p = 0.80	
ARAT times(seconds)				
Change from Baseline* to Post	-30.7 (-56.8, -4.6)	-34.2 (-106.1, 37.8)	-31.1 (-54.4, -7.7)	-23.1 (-84.4, 38.2)
	session: F(1, 17) = 4.0, p = 0.063 training group: F(1, 17) = 0.93, p = 0.35		session: F(1, 20) = 3.3, p = 0.08 training group: F(1, 20) = 0.4, p = 0.5	
ARAT score				
Change from Baseline* to Post	1.1 (-0.7, 2.8)	0.6 (-0.9, 2.1)	1.0 (-0.6, 2.6)	0.3 (-1.2, 1.7)
	session: F(1, 17) = 2.6, p = 0.12 training group: F(1, 17) = 0.1, p = 0.76		session: F(1, 20) = 1.7, p = 0.2 training group: F(1, 20) = 0.04, p = 0.84	
CMSA-A				
Change from Baseline* to Post	0.05 (-0.21, 0.31)	0.0 (0.0, 0.0)	0.05 (-0.19, 0.28)	-0.05 (-0.15, 0.05)
	session: F(1, 17) = 0.18, p = 0.68 training group: F(1, 17) = 0.02, p = 0.88		session: F(1, 20) = 0.0, p = 1.0 training group: F(1, 20) = 0.007, p = 0.9	
Elbow ROM - Flexion (degrees)				
Change from Baseline* to Post	-1.1 (-5.5, 3.3)	0.06 (-2.55, 2.66)	-1.2 (-5.1, 2.8)	-0.2 (-3.1, 2.6)
	session: F(1, 17) = 0.22, p = 0.65 training group: F(1, 17) = 0.78, p = 0.39		session: F(1, 20) = 0.4, p = 0.5 training group: F(1, 20) = 0.1, p = 0.7	
Elbow ROM - Extension (degrees)				
Change from Baseline* to Post	1.6 (-2.0, 5.2)	-0.3 (-8.5, 7.9)	2.0 (-1.3, 5.4)	-0.4 (-7.1, 6.3)
	session: F(1, 17) = 0.14, p = 0.71 training group: F(1, 17) = 0.23, p = 0.64		session: F(1, 20) = 0.2, p = 0.6 training group: F(1, 20) = 0.95, p = 0.3	
MAS- Biceps				
Change from Baseline* to Post	0.0 (-0.2, 0.2)	-0.1 (-0.5, 0.2)	0.0 (-0.1, 0.1)	-0.1 (-0.4, 0.2)
	session: F(1, 17) = 0.40, p = 0.54 training group: F(1, 17) = 0.03, p = 0.86		session: F(1, 20) = 0.6, p = 0.46 training group: F(1, 20) = 0.001, p = 0.97	
MAS- Triceps				
Change from Baseline* to Post	0.2 (-0.3, 0.7)	-0.5 (-0.9, -0.1)	0.2 (-0.3, 0.6)	-0.5 (-0.8, -0.1)
	session: F(1, 17) = 0.79, p = 0.34 training group: F(1, 17) = 0.53, p = 0.48 training group x session: F(1, 20) = 5.56, p = 0.03		session: F(1, 20) = 1.1, p = 0.3 training group: F(1, 20) = 0.20, p = 0.66 training group x session: F(1, 20) = 6.05, p = 0.02	

*Average of Baseline 1 and Baseline 2 sessions

Repeated Measures ANOVA (significant interaction effects included only)

Table III: Statistical Analysis Summary – Motor Exploration Performance

	Participant data with unstable baseline FMA-UE removed from analysis (Force group, n = 10; Control group, n = 9)		All participant data included in analysis (both groups, n = 11)		
	Force group mean (CI)	**Control group mean (CI)**	**Force group mean (CI)**	**Control group mean (CI)**	**mean**
Motor Exploration Performance					
50th percentile velocity coverage (meters^2/second^2)					
Baseline 2	0.44 (0.17, 0.72)	0.60 (0.09, 1.1)	0.43 (0.18, 0.68)	0.55 (0.14, 0.96)	
Post	0.87 (0.43, 1.3)	0.91 (0.43, 1.4)	0.83 (0.42, 1.23)	0.85 (0.45, 1.24)	
	session: F(1, 17) = 24.5, p = 0.0001 training group: F(1, 17) = 0.14, p = 0.71		session: F(1, 20) = 26.7, p < 0.0001 training group: F(1, 20) = 0.09, p = 0.77		
Average within training	1.16 (0.58, 1.75)	0.72 (0.28, 1.17)	1.12 (0.58, 1.65)	0.69 (0.33, 1.04)	
	training group: t(17) = 1.3, p = 0.2		training group: t(20) = 1.46, p = 0.16		
Distance Traveled (meters)					
Baseline 2	46.3 (31.5, 61.0)	51.4 (30.8, 71.9)	46.8 (33.6, 60.0)	50.0 (33.6, 66.4)	
Post	64.4 (48.3, 80.5)	68.6 (53.8, 83.3)	63.7 (49.2, 78.2)	67.5 (55.7, 79.3)	
	session: F(1, 17) = 34.5, p < 0.05 training group: F(1, 17) = 0.2, p = 0.7		session: F(1, 20) = 42.5, p < 0.05 training group: F(1, 20) = 0.2, p = 0.7		
Average across training	71.1 (53.9, 88.4)	59.0 (42.4, 75.8)	69.6 (53.8, 85.4)	58.8 (45.7, 71.9)	
	training group: t(17) = 1.13, p = 0.28		training group: t(20) = 1.15, p = 0.27		
Contrast score(summed change in probability)					
Cumulative transfer effect	0.71 (0.52, 0.90)	0.65 (0.45, 0.85)	0.68 (0.5, 0.87)	0.72 (0.53, 0.91)	
	training group: t(17) = 0.5, p = 0.63		training group: t(20) = 0.32, p = 0.75		
Average within training effect	0.79 (0.67, 0.90)	0.32 (0.29, 0.36)	0.81 (0.70, 0.92)	0.34 (0.30, 0.38)	
	training group: t(17) = 8.34, p < 0.05		training group: t(20) = 8.5, p < 0.05		
Favorability Metric					
Change from session 3 to Post	-0.006 (-0.052, 0.040)	0.058 (-0.055, 0.171)	0.008 (-0.043, 0.060)	0.045 (-0.048, 0.138)	
	session: F(5, 85) = 1.38, p = 0.24 training group x session: F(5, 85) = 1.19, p = 0.32		session: F(5, 100) = 1.55, p = 0.18 training group x session: F(5, 100) = 0.76, p = 0.58		
Average across sessions 3-8	0.69 (0.66, 0.73)	0.59 (0.53, 0.66)	0.68 (0.62, 0.74)	0.60 (0.51, 0.69)	
	training group: t(17) = 3.2, p = 0.005		training group: t(20) = 2.2, p = 0.04		
Cumulative Training Contrast and Within Training Contrast Correlation					
Correlation	0.60 (0.41, 0.79)	0.36 (0.09, 0.62)	0.56 (0.36, 0.75)	0.43 (0.19, 0.67)	
	training group: t(17) = 1.76, p = 0.1		training group: t(20) = 0.87, p = 0.39		

Repeated Measures ANOVA (significant interaction effects included only)
Two-Sample t-test

Table IV: Statistical Analysis Summary – Goal-Directed Performance

	Participant data with unstable baseline FMA-UE removed from analysis (Force group, n = 10; Control group, n = 9)		All participant data included in analysis (both groups, n = 11)	
	Force group mean (CI)	**Control group mean (CI)**	**Force group mean (CI)**	**Control group mean (CI)**
Goal-Directed Performance				
Reaching: Maximum Perpendicular distance (centimeters)				
Change from Baseline 2 to Post	-0.28 (-0.59, 0.4)	-0.05 (-0.48, 0.38)	-0.27 (-0.55, 0.01)	-0.01 (-0.37, 0.35)
	session: F(1, 323) = 2.9, p = 0.09 training group: F(1, 17) = 0.3, p = 0.6 movement direction: F(9, 323) = 4.7, p < 0.05		session: F(1, 380) = 2.42, p = 0.12 training group: F(1, 20) = 0.084, p = 0.78 movement direction: F(9, 380) = 5.37, p < 0.05	
Reaching: Path length ratio				
Change from Baseline 2 to Post	-0.11 (-0.26, 0.03)	-0.02 (-0.19, 0.14)	-0.10 (-0.24, 0.04)	-0.005 (-0.15, 0.15),
	session: F(1, 323) = 6.7, p = 0.01 training group: F(1, 17) = 0.02, p = 0.9 movement direction: F(9, 323) = 1.3, p = 0.3		session: F(1, 380) = 4.4, p = 0.04 training group: F(1, 20) = 0.09, p = 0.77) movement direction: F(1, 9) = 1.1, p = 0.3	
Circular Movement: Average radial deviation (centimeters)				
Change from Baseline 2 to Post	0.13 (-0.09, 0.34)	0.22 (-0.24, 0.69)	0.15 (-0.05, 0.35)	0.26 (-0.14, 0.66)
	session: F(1, 51) = 5.2, p = 0.03 training group: F(1, 17) = 0.3, p = 0.6 movement direction: F(1, 17) = 1.9, p = 0.2 group x direction interaction: F(1, 51) = 5.5, p= 0.02		session: F(1, 60) = 9.08, p = 0.004 training group: F(1, 20) = 0.5, p = 0.48). movement direction: F(1, 60) = 2.1, p = 0.15 group x direction interaction: F(1, 60) = 4.5, p = 0.04	
Reaching: Movement duration (seconds)				
Change from Baseline 2 to Post	-0.31 (-0.65, 0.02)	-0.07 (-0.24, 0.10)	-0.29 (-0.59, 0.01)	-0.07 (-0.21, 0.07)
	session: F(1, 323) = 8.9, p = 0.003 training group: F(1, 17) = 0.09, p = 0.8 movement direction: F(9, 323) = 3.6, p = 0.0002		session: F(1, 380) = 9.5, p = 0.002 training group: F(1, 20) = 0.14, p = 0.7 movement direction: F(1, 380) = 4.0, p < 0.05	
Reaching: Peak Speed (meters/second)				
Change from Baseline 2 to Post	0.006 (-0.03, 0.05)	0.04 (-0.03, 0.10)	0.01 (-0.03, 0.05)	0.04 (-0.01, 0.09)
	session: F(1, 323) = 7.8, p = 0.005 training group: F(1, 17) = 0.2, p = 0.7 movement direction: F(9, 323) = 2.9, p = 0.003 group x session interaction: F(1, 323) = 4.1, p = 0.04		session: F(1, 380) = 13.7, p < 0.05 training group: F(1, 20) = 0.10, p = 0.75 movement direction: F(9, 380) = 3.6, p = 0.0003 group x session interaction: F(1, 380) = 5.4, p = 0.02	
Circular Movement: Movement Duration (seconds)				
Change from Baseline 2 to Post	-5.38 (-8.75, -2.01)	-7.68 (-13.43, -1.92)	-6.02 (CI: -9.6, -2.4)	-6.87 (CI: -12.7, -1.0)
	session: F(1, 51) = 36.6, p < 0.05 training group: F(1, 17) = 0.3, p = 0.6 movement direction: F(1, 51) = 2.8, p = 0.1		session: F(1, 60) = 42.0, p < 0.05 training group: F(1, 20) = 0.77, p = 0.39 movement direction: F(1, 60) = 4.0, p = 0.049	
Circular Movement: Mean Speed (meters/second)				
Change from Baseline 2 to Post	0.11 (0.04, 0.17)	0.20 (0.11. 0.29)	0.12 (0.06, 0.18)	0.18 (0.1, 0.26)
	session: F(1, 51) = 76.2, p < 0.05 training group: F(1, 17) = 1.5, p = 0.2 movement direction: F(1, 51) = 3.3, p = 0.07 group x session interaction: F(1, 51) = 7.0, p = 0.01		session: F(1, 60) = 82.31, p < 0.05 training group: F(1, 20) = 1.91, p = 0.18 movement direction: F(1, 60) =401, p = 0.03 group x session interaction: F(1, 60) = 3.3, p = 0.07	

Repeated Measures ANOVA (significant interaction effects included only)

D. DISCUSSION

This study investigated an innovative approach to robot-therapy in which the characterization of participants' motor exploration directly informed the mathematical structure of customized robotic training environments. To the best of our knowledge, this is the first clinical study to apply robot guided characterization of stroke survivors' exploratory motor behaviors towards the design of individually customized therapy. Disappointingly, our results from clinical assessments did not indicate differences between the novel treatment and controls. While changes in FMA-UE scores were modest for both groups, stroke survivors exhibited marked increases in velocity coverage following only two-weeks of training. Beyond measures of overall improvement, we were very interested in whether the influence of interactive forces could be detected in learned movement patterns. Consequently, we devised a novel analysis that revealed significant correlations between induced training behaviors and new patterns of unassisted movement. These results provide preliminary evidence that new movement behaviors can be learned from training with forces that target movement deficits.

Our pre-declared primary clinical result, change in Fugl-Meyer scores, showed that both groups benefitted from training; however, such levels of improvement would not be viewed as clinically relevant (Gladstone, Danells, and Black 2002; Johanna H. Van Der Lee et al. 2001; Page, Fulk, and Boyne 2012). Our clinical results fall short of the Fugl-Meyer gains reported in other chronic stroke robot therapies (Prange et al. 2006; Dewald et al. 1995; Dewald and Beer 2001). However, considering that our intervention lasted only two weeks, compared to 6+ weeks in other interventions, it may not be surprising that our effects were only modest. One benefit from our approach may be that free exploration training with velocity feedback is at higher intensities. Interestingly, some participants displayed even greater improvement upon a follow-

up evaluation after training, which could indicate that new motor capabilities required some time to incorporate into activities of daily living. It is also possible that the inactivity between the final day of training and later evaluation allowed patients some needed rest. It is also possible that the repeated exposures to clinical evaluations had a training effect of "teaching to the test." Anecdotally, many participants stated that our motor exploration paradigm appeared to relax the muscles of their affected arm. Such action could have stretched muscles (Gao et al. 2011) or reduced reflex gains (Liu et al. 2011) due to the reduced mechanical impedance. Overall, participants expressed that the motor exploration task was somewhat tiring and not particularly engaging. However, the participants also reported that the feedback score provided incentive to be creative with expressing movement variety. Future iterations of customized force design might target more degrees of freedom, which would have a greater impact on functional skills (Beekhuis et al. 2013). However, it is also possible that the robot-assisted training promoted learning that is not evident from clinical assessments (G Kwakkel, Kollen, and Krebs 2007).

Our analysis of the changes in motor exploration revealed evidence that participants increased movement capability. As a simple metric of the range of motion, we observed that participants from both groups increased their velocity limits. It is worth emphasizing that due to the nature of characteristic movement behaviors during exploration, the vector fields resulting from velocity data generally tend to push participants' movements towards higher velocities, in a manner similar to destabilizing forces from our previous work (F. C. Huang and Patton 2011). It is possible that improved coverage indicates that participants retained some of the movement patterns acquired through vector field training into their exploration practice evaluated without forces.

Beyond the overall range of motion, we were also interested in measuring the degree to

which the probabilities of observed movements were *redistributed*. Interestingly, our findings showed that training groups exhibited similar amounts of change in movement distribution. This similarity demonstrates that motor exploration practice can induce change even in the absence of external forces. Despite considerable differences between training groups' average within training effect, both groups displayed a gradual change from their original movement behaviors across sessions. This result is consistent with our previous study that compared distributions across multiple days without any intervention (F. C. Huang and Patton 2016). Hence, beyond the use of customized forces, there may be other forms of training intervention, with visual feedback or even verbal instructions, which may prove useful in inducing desirable changes to movement distribution.

Our analysis of the proportion of favorable change in movement distribution (See Fig. 8B) provides evidence, however, that training forces can positively impact on how stroke survivors express movement. While both groups exhibited general improvement in the range of motor exploration, our supplementary analysis indicated group differences in how the initial trends of overly low or high probability velocities changed due to training (See Fig. 8A). Learned non-use in stroke survivors represents an extreme case of how a lack of motor expression can be reinforce (F. C. Huang and Patton 2011; E Taub et al. 2006). In addition, abnormal coordination or involvement of additional degrees of freedom can occur. For example, compensatory trunk motion is typical in reaching (Pain et al. 2015), while circumduction at the hip occurs due to stiff knee gate (Patton 2012). In a rat model of stroke recovery, researchers suggest that "inappropriate gestures may represent motor habits that substitute for, and compete with, successful movements" (Alaverdashvili et al. 2008). It is worth emphasizing that training for the Control group was self-mediated except for the knowledge of results presented at the end

of each trial block. Consequently, without more specific guidance, reinforcement of abnormal movement distributions was possible. The fact that participants of the Force group also exhibited both favorable and unfavorable changes indicates that further refinement is needed in the design of customized forces. The crucial lesson here, however, is that forces evidently provided an additional pressure on motor adaptation that evidently helped to reverse the deficits in movement distributions found prior to training.

We devised an analysis to answer a fundamental question about robotic intervention: can mathematical structure of force field customization be detected in learned motor behaviors? Because the customized robot training produced such dramatic changes in movement distribution during the presentation of forces, it was in many ways surprising that our experiment groups exhibited such similar degrees of improvement. Our supplementary analysis of motor exploration, however, revealed analogous changes between training and new behaviors (See correlation analysis in Motor Exploration Performance section. These correlation analyses suggest customized forces caused specific and persistent changes to movement behaviors. Note that the Force participants experienced drastically altered movement distributions due to force interactions. Yet despite such effects, some of these individuals still demonstrated high correlations, indicating some retention of the movement behaviors learned during training. Interestingly, we also observed similar correlations in the Control group. The key difference, however, is that the movement distributions during the training phase of the Control group was self-mediated and not dictated by customized robot forces. It is perhaps unsurprising changes within day would in some way mirror changes in longer term learning. It is however remarkable that new patterns of movement persisted even when induced from externally applied forces. The learning of new exploratory behavior indicated here differs from typical adaptation to novel

force and visual distortions (Donchin et al. 2012; Rabe et al. 2009; Shadmehr and Mussa-Ivaldi 1994) since participants were not given prescribed movement goals and hence did not rely on explicit error feedback. Instead, it is likely that the repeated exposure to motor exploration with interactive forces induced adaptation in terms of use-dependent learning (Hammerbeck et al. 2014; Diedrichsen et al. 2010; Reinkensmeyer, Guigon, and Maier 2012). Further development is needed for predictive models of how practice behaviors during intervention lead to changes in motor exploration capabilities.

We observed some movement behaviors during training that could indicate unintended consequences of our implementation of vector field training. Specifically, some participants exhibited rapid, repetitive motions in a curved path. While repetitive behavior can appear in stroke patients' distributions during un-assisted motor exploration, it was clearly evident in the distributions during interaction with forces. One possibility is that participants intentionally avoided forces by moving at relatively constant velocities outside their characteristic behavior since this is where force magnitudes were low. Alternatively, the destabilizing nature of vector fields may have constrained participants to repetitive behavior because the forces were continuously active. Such a scenario would have similarities to passively moving the limb (Hornby et al. 2008), resulting in less active involvement—an essential component to recovery (Enzinger et al. 2009). One potential limitation of our current protocol is that the task feedback did not penalize cyclic behavior. Instead of gradual adaptation, some participants exhibited substantial and sudden increases in coverage. This effect suggests changes in task comprehension, or in the strategy for how to work with interactive forces (Chib et al. 2005). Future iterations of force fields could be improved by obtaining characterization data that more faithfully reflect participants' full range of capabilities, and by improved task instructions on the

goals of motor exploration.

Our analysis of the changes in reaching and circular motion performance suggests that learned exploration behaviors might not immediately transfer to skill in goal-directed actions. Both treatment groups only showed a modest reduction of movement error on the goal-directed reaching task and an increase in movement error on the circular movement task following training. Our motor exploration task did not provide feedback of movement errors related to specific movement goals. It does, however, encourage participants' to practice upper-arm coordination over a wider range of movements, which has been shown to facilitate generalization to untrained movements (Conditt, Gandolfo, and Mussa-Ivaldi 1997). Since increases in velocity coverage were a main component of the overall changes in movement distributions, it is possible that participants generalized the ability to move at higher speeds as opposed to the ability to minimize reaching errors. Thus, we further inspected whether analogous changes were present in their goal-directed movements. For both tasks, we observed an increase in peak speed and a decrease in the time to complete each movement. This result might suggest that participants retained increases in movement speed at the expense of decreased accuracy ("Keele1968.pdf" 2016). On the other hand, it is likely that any new motor exploration capabilities require time and experience to incorporate into activities of daily living.

Beyond the potential benefits of customized force fields for upper extremity rehabilitation, our approach could serve as a basis for a wide range of therapeutic applications. Statistical profiling of large data sets is an emerging trend, and analysis of distributions could be derived from a variety of domains relating to human behavior; including, electromyography, joint-space variables and electrocorticography. The framework we have provided here could be applied more generally to determine the optimal strategies to customize treatment.

III. KEY COMPONENTS OF MECHANICAL WORK PREDICT OUTCOMES IN ROBOTIC STROKE THERPY[1][2]

Zachary A. Wright, Yazan A. Majeed, James L. Patton and Felix C. Huang

Background: Clinical practice typically emphasizes active involvement during therapy. However, traditional approaches can offer only general guidance on the form of involvement that would be most helpful to recovery. Beyond assisting movement, robots allow comprehensive methods for measuring practice behaviors, including the energetic input of the learner. Using data from our previous study of robot-assisted therapy, we examined how separate components of mechanical work contribute to predicting training outcomes.

Methods: Stroke survivors (n = 11) completed six sessions in two-weeks of upper extremity motor exploration (self-directed movement practice) training with customized forces, while a control group (n = 11) trained without assistance. We employed multiple regression analysis to predict patient outcomes with computed mechanical work as independent variables, including separate features for elbow versus shoulder joints, positive (concentric) and negative (eccentric), flexion and extension.

Results: Our analysis showed that increases in total mechanical work during therapy were positively correlated with our final outcome metric, velocity range. Further analysis revealed that greater amounts of negative work at the shoulder and positive work at the elbow as the most important predictors of recovery (using cross-validated regression, R^2= 52%). However, the work features were likely mutually correlated, suggesting a prediction model that first removed

[1] revised manuscript submitted for review to *Journal of NueroEngineering and Rehabilitation* (August 2019)
[2] partial results published in *Conference Proceedings of IEEE Engineering Medicine and Biology* (Zachary A. Wright, Patton, and Huang 2018)

shared variance (using PCA, R^2=65-85%).

Conclusions: These results support robotic training for stroke survivors that increases energetic activity in eccentric shoulder and concentric elbow actions.

A. BACKGROUND

Assistance is often provided to aid limb movement during the rehabilitation process of stroke survivors. Many clinical researchers agree that active participation enhances recovery, and the goal of therapy should be to maximize "involvement" (Blank et al. 2014; Hogan et al. 2006). Too much assistance can actually discourage patient effort (Reinkensmeyer, Guigon, and Maier 2012). However, measurement of the degree to which patients are *actually* active is often difficult. Advances in rehabilitation devices allow for the measurement of forces *and* motion to better monitor patient activity. Here we investigate how upper limb mechanics during training relate to recovery.

Current tools for measuring physical activity during therapy offer limited information for describing interaction with the external environment or agent. While studies have shown that the intensity of therapy influences patient improvement, researchers have relied on simple metrics related to experimental conditions (e.g. movement repetitions, time-on-task, and therapy dosage) (Gert Kwakkel 2006; Lohse, Lang, and Boyd 2014). More sophisticated tools have been used to directly measure energetic contributions during therapy, such as oxygen consumption devices to measure metabolic cost (Kafri et al. 2014) or electromyography to measure muscle activity (H. J. Huang, Kram, and Ahmed 2012; Israel et al. 2006). However, such measures do not account for the time-varying force-motion relationships that occur during assisted movement. Robots easily measure both kinematic and kinetic variables facilitating the computation of energetic contributions in terms of mechanical power and work.

While energetic descriptions of movement have been widely studied, it has mainly focused on cyclic (Doke and Kuo 2007b) or sustained movements, such as walking. Researchers have computed work and power to characterize normal and abnormal gait patterns (Olney et al. 1991; Winter 1987), to evaluate robot-assisted locomotion (Neckel et al. 2008), and to reduce energetic costs when using exoskeletons (Ferris, Sawicki, and Daley 2007). Recently our work has focused robotic augmentation of upper limb dynamics to facilitate vigorous movement during practice (F. C. Huang and Patton 2013a; Zachary A Wright et al. 2017). We showed that stroke survivors increase total work output during force training (Zachary A. Wright, Patton, and Huang 2018). Our intervention was fundamentally different than many previous strategies in that patients trained over a broader range of movements. In contrast to reaching studies (Loredana Zollo et al. 2011; Mazzoleni et al. 2014), such self-directed exploration allows for the examination of how energetics might depend on different force and motion states.

To better evaluate the variation in patient energetics, we believe more comprehensive measures are required beyond total expenditure of power or work. Researchers have also examined compartmentalized work and power measures in normal limb behaviors, for example, associating magnitudes of mechanical energy (e.g. positive/concentric and negative/eccentric work) with movement actions (e.g. flexion and extension) at individual joints (Farris et al. 2015). Motor impairments due to stroke are also typically described in the context of motor actions of the limb. For example, stroke survivors exhibit abnormal flexion and extension synergies (Ellis, Lan, et al. 2016) and alterations in concentric and eccentric muscle contractions (Hedlund et al. 2012; Eng, Lomaglio, and Macintyre 2009). As such, impairments can be associated with *subcomponents* of work and power. As patients interact differently in response to forces, subcomponents of work and power could reveal individual differences in involvement.

An emerging trend in rehabilitation is to identify certain factors that predict individual improvement in response to therapy. Researchers have identified patient biomarkers (impairment level, neurophysiological) correlated to patient outcomes providing better recommendations for therapy (Stinear 2017; Mostafavi et al. 2017; Kim and Winstein 2016). Similarly, our goal is to determine if particular types of work are more important to patient recovery. Such evaluation could inform decisions on design strategies and optimize assistance to each individual. In contrast to previous studies which have relied on independent analyses of many individual predictors, our analysis goal necessitates more rigorous statistical methods to deal with potentially related work features. One possible solution is to employ multiple regression analysis which can identify features most important for prediction.

In this paper, we investigate how the energetic contributions of stroke survivors during robot-assisted training relate to upper-limb recovery. We employ well-established methods of inverse dynamics to estimate the torques generated by each patient during self-directed motor exploration training with customized forces. These methods conveniently allow us to quantify the energetic involvement of each individual joint in terms of mechanical work. We then use multiple regression analysis to identify which components of work are most important for predicting recovery. We hypothesize that positive work (concentric) in elbow extension is the best predictor of outcome. This study provides a key preliminary step towards evaluating energetic descriptions of patient involvement which can inform methods for upper limb robotic therapy practice.

B. METHODS

1. Study Participants

This investigation considered data collected from a previous study that featured 22 stroke survivors (Zachary A. Wright et al. 2017). The main inclusion criteria included: 1) chronic stroke (8+ months post-stroke) 2) hemiparesis with moderate to severe arm impairment measured by the upper extremity portion of the Fugl-Meyer Assessment (UEFM score of 15-50) 3) primary cortex involvement. Each individual gave informed consent in accordance with the Northwestern University Institutional Review Board (IRB).

2. Apparatus

Experiment participants were asked to operate a two-degree of freedom robotic device with the affected arm (Fig. 10A). A custom video display system (not shown) provided visual feedback of the location of the wrist as the arm moved in the horizontal plane. During movement, the weight of the arm was supported. Movement data was collected at 200 Hz and filtered using a 5^{th} order Butterworth low pass filter with a 12 Hz cutoff. Using anthropometric measurements recorded from each participant, we computed inverse kinematic relationships to obtain elbow and shoulder joint angles corresponding to endpoint position data. The robot control and instrumentation were mediated by a Simulink-based XPC Target computer, with a basic rate of 1kHz. The robot controller compensated for the dynamics of the robot arm. A force sensor attached to the end-effector measured the human-robot interaction forces.

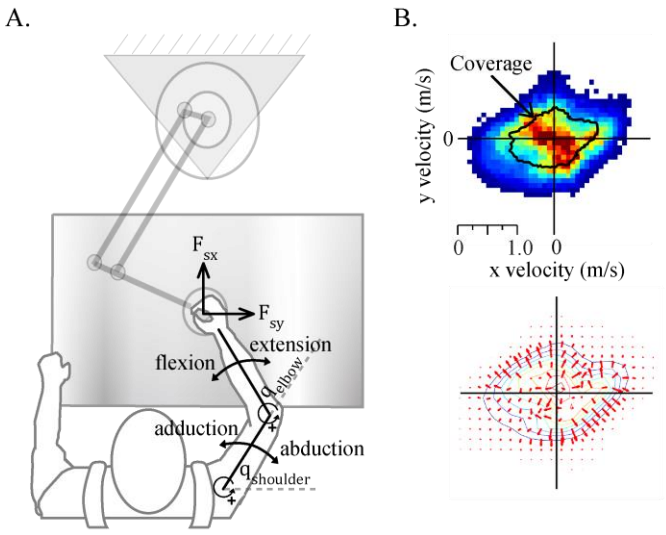


Figure 10: Experimental Design (A) Stroke survivors performed self-directed motor exploration by moving the robot handle in the horizontal plane. Measurements of their limb motion and the interaction forces were used to estimate the positive (concentric) and negative (eccentric) mechanical work exerted in different directions of shoulder and elbow joint motion. (B) The probability distribution of each individual's movement velocities during unassisted motor exploration (top, red indicates higher probability, black contour line represents the 90th percentile velocity coverage) formed the basis for the design of customized training forces (bottom, red arrows, colored contour lines represents Gaussian model fit to velocity data).

3. Experimental Protocol

Each participant completed nine sessions across five weeks, including evaluation (Baseline, sessions 1, 2; Post, sessions 8, 9) and training (sessions 2-7, spanning two weeks) sessions. For each evaluation, participants completed a clinical assessment and a motor performance assessment which included three motor tasks using the robotic device under no forces: point-to-point reaching, circular movements, and six two-minute trials (12 minutes in total) of self-directed motor exploration. For each training session, each participant first completed a performance assessment then completed an additional 16 two-minute trials (32 minutes in total) of motor exploration, either in the presence of a customized force field (Force group, n = 11) or absent forces (Control group, n = 11). This investigation considered only data

from the motor exploration portions of Baseline (session 2) and Post (session 8) evaluation sessions as well as the training sessions.

For the motor exploration task, participants were instructed to move the robot handle to all reachable points within a 0.6 m x 0.4 m workspace, to vary their speed and direction of movement as much as possible and to avoid repetitive movements. Each motor exploration trial ended after two cumulative minutes of movement within the workspace. Movement speed below 0.04 m/s was considered rest so that the time samples did not count towards the total movement time. While we informed participants they could rest at any time, we also provided designated rest periods (1-3 minutes) at the end of the motor performance assessment and prior to the start of training and after the first eight trials during training. After each trial of motor exploration, we provided a *Post-Trial feedback score* summarizing their motor exploration performance, as described previously (Zachary A. Wright et al. 2017). The score was based on a heuristic measure of randomness which was used to encourage more variety in movement patterns. Following each completed trial, we displayed on screen both the "Current" score (score from the most recent trial) and the "Best" score (highest score across all trials) the participant achieved within a given session.

4. Design of Customized Force Field

The motor exploration portion of the performance assessment for each training session served as a basis for the design of a customized force field for training in that session (See Fig. 10). To serve as a model of an individual's typical patterns of movement, this study focused on velocity data accumulated across 12 minutes of motor exploration. We fit this data with a multivariate Gaussian smoothing function. The result of this model fitting procedure can

be visualized by constructing a two-dimensional probability distribution representing the most and least typical movement velocities during exploration (Fig. 10B, colored contours). Next, computing the gradient of the analytical form of this function results in a continuous velocity-dependent function whose output are vectors that represent the slope along the two-dimensional distribution. In principal, the direction of the vectors point from higher probabilities towards lower probabilities of the distribution (Fig. 10B, red arrows). The vector field represents the direction and relative magnitude of robot-applied forces which were updated continuously based on the current velocity of its endpoint. Additional details on the experimental procedures were recently published (Zachary A. Wright et al. 2017).

5. Model of Upper Limb Dynamics

Here we describe the human-robot dynamic interaction for two degree of freedom planar movement (Fig. 10A). We employed established methods of inverse dynamics of upper limb motion to estimate the elbow and shoulder joint torques generated by the human. This analysis considers the human arm as a closed system where an external force at the wrist is available from the force sensor measurements. Thus, the model considers the influence of the torques acting on each joint; including the torque required to move the arm passively (τ_p) which is composed of the torque generated by the human (τ_h) and the torque acting on the arm by the robot (τ_r). The passive load of the arm can be expressed as $\tau_p = M(q)\ddot{q} + C(q, \dot{q})\dot{q}$, where q represents the joint angles of the arm, M is the inertial matrix function and C is the Coriolis-centrifugal matrix function. Anatomical measurements of limb segments, body weight and height for each patient were used to estimate the mass distribution of the arm (Hatze 1980). We computed the torques of the robot acting on the human arm arising from the robot contact forces

according to $\tau_r = J_h^T F_s$, where J_h is the Jacobian matrix of the arm and F_s is the interaction force measured from the force sensor.

6. Model Features

We constructed a set of candidate model features (nine in total) to be used as our model predictors in our regression analysis. These features included a single categorical factor representing training group in addition to eight individual data variables, specifically the components of mechanical work relating to each patient's overall energetic contribution to limb motion during training (Table V). To compute the work features, we first solved for the patient-generated torque at each individual joint and then calculated the mechanical power ($P(t) = \tau_h \dot{q}$) for each two-minute trial of motor exploration within training (96 trials in total across six training sessions). In principal, the integral of power across time represents the total mechanical work ($W = \int P(t)dt$). We divided the time-series calculations of power into four separate components for each joint. Each of these components represented a different combination of the direction of joint torques generated by experiment participants and the relative direction of angular motion at each respective joint. Finally, we computed the numerical integral for each time series of power to obtain work features; including, both the positive (concentric) and negative (eccentric) work performed in elbow flexion and extension and in shoulder horizontal adduction and abduction (Fig. 10A). We represent each individual feature of work as the average across training trials subtracted by the respective average work across unassisted motor exploration trials (six) during Baseline evaluation (session 2).

7. Recovery Outcomes

We evaluated how well the components of mechanical work during training could act as predictors of measures of patient recovery. Our primary clinical outcome included changes in UEFM from Baseline (session 2) to Post (session 8). Beyond standard clinical assessments, we also evaluated changes in motor exploration performance. We employed an engineering metric, previously described in (Zachary A. Wright et al. 2017), which captures the "maximum" range of movement velocities spanned during motor exploration. Velocity coverage is expressed as the estimated area of two dimensional velocity data (in units of m^2/s^2). To determine velocity coverage, we first calculated the 90^{th} percentile speed within 64 equally spaced bins radially aligned from the zero-velocity center within the range of 0-2π. We then calculated the area within the boundary formed by connecting the points representing the 90^{th} percentile speed within each bin. We considered the change in velocity coverage from Baseline (session 2) to Post (session 8) as an additional outcome prediction.

8. Prediction Model

We employed a Least Absolute Shrinkage and Selection Operator (LASSO) predictive model to predict recovery outcomes using our work features (Tibshirani 1996). The LASSO method is a special case of regularized least squares regression which incorporates an additional penalty term on the L_1 norm of the model coefficients. We chose LASSO because it has the advantage over alternative approaches of enhancing the interpretability of the results by reducing the number of features used by the model. We used a first-order LASSO model represented by the following formula in Lagrangian form that determines a set of fitted coefficients such that:

$$\hat{\beta}_{lasso} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{J} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{J} |\beta_j| \right\}$$

where N equals the number of experimental participants (22 in total), J equals the number of features (9 in total), y_i is the outcome measure, $x_i = [X_{i1}, \ldots, X_{ip,}]$ represents the eight components of work features and an additional categorical feature representing training group denoted $X_{ij,}$ ($i = 1,\ldots,N$, $j = 1,\ldots,J$), β_j is the coefficient of the j-th feature and β_0 is the intercept. The model features were standardized to account for relative differences in magnitude between the components of work. λ represents the non-negative penalty term that controls the degree of regularization by effectively driving the coefficients of features that are unhelpful to the predictions to zero. This results in a reduced feature set incorporated by the model. For our model predictions, we chose the largest λ value (λ_{1SE}) where the cross-validated mean-square error MSE_{CV} (i.e. an estimation of how well the model would predict new data) is within one standard error of the minimum MSE_{CV} (Tibshirani 1996).

Our primary goal in using LASSO regression was to determine the most important predictors of recovery outcomes, particularly among the work features (See Ranking the Features). In addition, we used LASSO regression to perform an exhaustive analysis of the sensitivity in the predictive ability of our work features. We relied on the adjusted coefficient of determination (R^2) as our primary metric which measures the proportion of variance in our recovery outcomes that can be explained by the work features selected by LASSO. We determined the sensitivity in R^2 to different data splits using 5-fold cross validation with 231 repeats. Each repeat uniquely split the participant data into the five different folds. For each repeat, we chose the LASSO model where $\lambda = \lambda_{1SE}$, then trained the model using all the data and calculated the R^2 of the resulting predictions. As a secondary metric we calculated the mean

squared error (MSE) of the trained model predictions and compared to the MSE_{CV} in order to assess the degree of overfitting. Our repeated cross-validation method provides a robust estimate of the variability in R^2 and MSE to different choices of λ. As an additional measure of sensitivity, we repeated the cross-validation procedure with six bootstrap datasets. We constructed each of these datasets by randomly resampling, with replacement, the original dataset (all 96 trials). Thus, our measure on the quality of our model predictions included 1,617 R^2 estimates. Our bootstrap method increases confidence in our R^2 and MSE estimates.

9. Ranking the Features

Our primary objective in using LASSO regression was to determine the most important predictors of recovery outcomes, particularly among the work features. LASSO conveniently reduces the number of features used in the model; however, there are no built-in methods for ranking the model predictors. We rely on previous methods in which, as a first step, the ranking of features was based on the frequency each was selected in the model (Majeed, Awadalla, and Patton, n.d.). The selection frequency metric, expressed as a percentage, represents the number of times in total each feature is included in the model (i.e. assigned a non-zero coefficient) across the 231 5-fold cross-validation repeats and each bootstrap dataset. Thus, the maximum possible number of times a given feature could be selected was 1,617. We then evaluated model performance (R^2) with each successive removal of a feature starting with the highest ranked feature (i.e. the most selected) and compared to the model predictions that included the full feature set (Full Model). By excluding features in this order, we determined the extent to which the remaining features were able to compensate for the excluded features and where model performance starts to diminish when they could no longer compensate.

C. RESULTS

We investigated how practice energetics (mechanical work) performed by stroke survivors could explain differences in recovery outcomes. Our use of a multiple regression model (LASSO) revealed more accurate predictions using changes in velocity coverage (coefficient of determination: mean R^2 ± SD; 0.36 ± 0.03), as compared to changes in our primary clinical outcome UEFM (-0.0007 ± 0.005). Our analysis also revealed that 'training group' was the most frequently selected feature in each of these models. This result suggests a significant group effect contributed uncertainty to the model predictions, and more importantly, in determining the work features most important to recovery. Thus, in the following analysis, we performed separate predictions for each group.

1. Energetics Relates to Outcome

Prior to our main analysis, we inspected how training energetics relate to patient outcomes (Fig. 11). We first examine the degree that the total work performed (change from Baseline to Training in Joules, mean ± SE; Force, 617.7 ± 106.2; Control, 87.5 ± 13.6) correlated with changes in our main clinical outcome measure, changes in UEFM scores (change from Baseline to Post, mean ± SE; Force, 1.1 ± 0.4; Control, 0.7 ± 0.6). Because these observed changes in UEFM scores would not be considered "clinically important" (Gladstone, Danells, and Black 2002), it may not be surprising that we failed to detect a trend for both groups (Force, r = -0.29, p = 0.4; Control, r = 0.1, p = 0.77). However, beyond clinical outcomes, we focused our investigation on whether training energetics relates to changes in velocity coverage (change from Baseline to Post; Force, 0.88 ± 0.20 m^2/s^2; Control, 0.66 ± 0.15 m^2/s^2). We expected that

this measure would be more sensitive to any recovery resulting from motor exploration training (Fig. 11A). Interestingly, we found a significant correlation for patients who trained with forces ($r = 0.7$, $p = 0.1$), but not the control group ($r = -0.003$, $p = 0.99$). It is worth noting that the force group performed much greater levels of work than the control group due to the presentation of interactive forces.

Besides the basic link we found between work and recovery, we were interested in pinpointing the specific *components of work* that best predict recovery (Table V). As a preliminary inspection, we correlated changes in work for each component with the recovery outcomes (Fig. 11B). For the force group, we found that negative work during shoulder adduction ($r = 0.73$, $p = 0.01$) and abduction ($r = 0.71$, $p = 0.01$) and positive work during elbow flexion ($r = 0.65$, $p = 0.03$) significantly correlated with changes in velocity coverage, but not positive work during elbow extension ($r = 0.54$, $p = 0.08$), negative work during elbow flexion ($r = 0.45$, $p = 0.17$) and extension ($r = 0.46$, $p = 0.15$), or positive work during shoulder adduction ($r = 0.45$, $p = 0.16$) and abduction ($r = 0.56$, $p = 0.08$). We failed to detect any trends for the control group. Our correlation analysis also did not reveal significant effects for changes in UEFM for both groups.

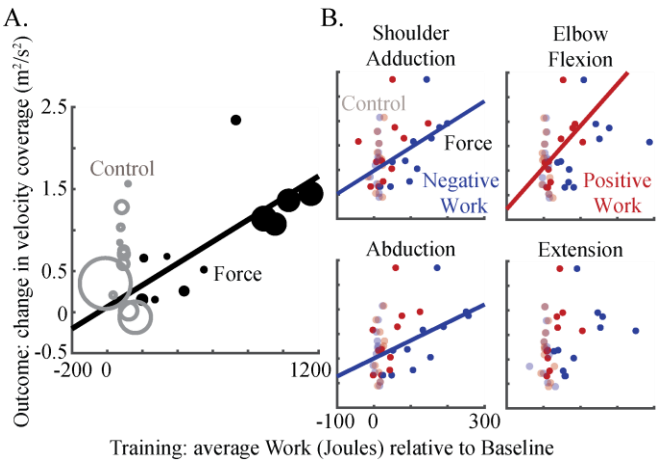


Figure 11: Correlation analysis (A) The total mechanical work performed during motor exploration force training significantly correlated with changes in velocity coverage. Each data point represents an individual participant. The size of the data points is proportional to velocity coverage during Baseline (session 2). (B) The breakdown of work reveals subcomponents that significantly correlated with changes in velocity coverage. A single pair of blue and red data points along the Training axis represents an individual participant. Regression lines only shown for statistically significant correlation ($\alpha < 0.05$).

Table V: Experimental Data

	Outcomes*		Work Features†							
	UEFM	Velocity Coverage (m^2/s^2)	(-) Shoulder Adduction	(+)Shoulder Adduction	(-) Shoulder Abduction	(+)Shoulder Abduction	(-) Elbow Flexion	(+) Elbow Flexion	(-) Elbow Extension	(+) Elbow Extension
Force Group										
1	0.5	1.44	200.9	148.8	251.7	126.6	143.1	86.8	148.4	53.5
2	2	1.36	179.7	58.2	258.0	77.5	179.6	80.4	161.7	37.6
3	2	0.65	51.6	7.9	53.4	13.7	39.2	4.7	30.0	10.1
4	1.5	0.51	119.0	49.9	155.1	46.3	70.9	14.6	83.3	12.9
5	0.5	1.07	108.7	-40.9	134.4	-1.3	288.9	107.7	251.6	106.7
6	0.5	0.15	46.6	13.7	49.7	22.3	59.1	13.2	59.1	10.2
7	0	2.34	144.8	52.0	173.1	60.6	110.5	54.5	92.2	45.6
8	3.5	0.15	18.5	-7.2	26.2	-0.2	83.6	9.4	61.3	10.5
9	-1	0.67	87.2	22.8	92.8	22.7	48.2	16.0	40.7	13.4
10	2.5	0.25	98.2	26.6	107.7	44.7	69.3	20.4	54.7	17.8
11	0.5	1.14	157.0	82.5	189.7	64.8	152.9	48.8	155.5	39.7
Mean	1.1	0.88	110.2	37.7	135.6	43.4	113.2	41.5	103.5	32.5
±SD	1.3	0.67	57.5	50.1	78.7	38.1	74.1	36.3	67.9	29.4
Control Group										
1	0	-0.07	15.5	24.5	17.0	11.8	17.3	26.9	33.5	15.3
2	2	0.01	15.4	6.5	18.9	18.7	15.7	16.4	12.7	20.3
3	2	0.35	-9.0	-15.6	-0.4	5.0	-1.5	18.6	-36.5	31.0
4	3	0.70	14.1	23.0	11.8	23.4	14.9	-1.1	7.2	3.8
5	2.5	1.56	16.6	27.3	23.8	26.8	8.0	2.0	4.6	4.8
6	0.5	0.84	13.7	10.4	11.0	30.3	5.1	-2.0	2.9	2.7
7	2	1.03	15.1	15.1	8.3	18.3	4.8	4.1	7.4	6.5
8	-0.5	1.28	11.5	10.5	13.2	7.4	10.3	8.4	9.3	14.6
9	1	0.58	5.4	24.6	8.4	17.2	14.4	4.5	16.5	3.4
10	-2.5	0.74	12.1	16.7	12.2	27.9	10.8	2.2	10.8	3.7
11	-2.5	0.20	9.1	16.1	4.1	15.4	-0.6	-2.9	1.0	-5.6
Mean	0.7	0.66	10.9	14.5	11.7	18.4	9.0	7.0	6.3	9.1
±SD	1.9	0.51	7.5	12.0	6.7	8.3	6.5	9.6	16.7	10.2

*change from baseline to post †average change from baseline to training in Joules (-) Negative/Eccentric work (+) Positive/Concentric work

2. Model Performance

We performed a rigorous statistical analysis to determine how well components of work during training collectively predict patient outcomes using multiple regression analysis (Fig. 12). Unsurprisingly, we found better predictions for the force group compared to the control group. Our model predicted changes in velocity coverage with a coefficient of determination of 0.52 ± 0.043 (mean R^2 ± SD) for the force group (shown in Fig. 12, Full Model) and 0.34 ± 0.15 for the control group. The mean squared error of the trained model predictions was 0.21 ± 0.02 (MSE ± SD) for the force group and 0.17 ± 0.04 for the control group. While our modeling approach provided a robust estimate of the predictive power of the components of work, we also wanted to evaluate how well the model might perform in predicting new data. We found the estimated predicted mean squared error in cross validation was 0.64 ± 0.09 (MSE_{CV} ± SD) for the force group and 0.46 ± 0.06 for the control group. In contrast to changes in velocity coverage, prediction accuracy for the changes in UEFM scores was substantially lower for both groups (mean R^2 ± SD; force group, 0.16 ± 0.14; control group, 0.25 ± 0.22).

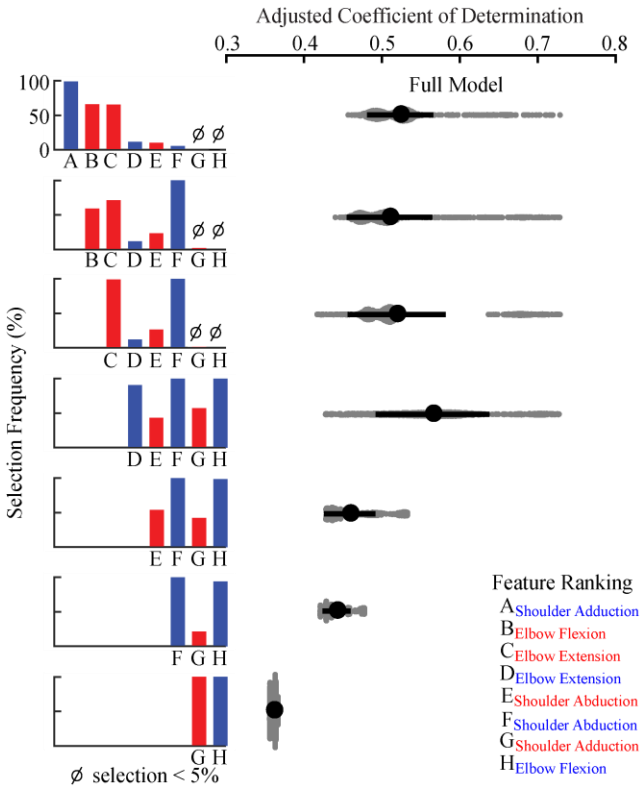


Figure 12: Model Predictions and Feature Selection Predictions of patient recovery, in terms of changes in velocity coverage, using multiple regression analysis. Each gray dot represents a single repeat of a cross-validation staggered for easy visualization by fitting a probability density function. Each black dot and bar represents the mean R^2 ± SD. Positive (concentric) and negative (eccentric) work features are indicated in red and blue, respectively. The negative work in shoulder adduction and positive work in elbow flexion and extension features were selected most often by the LASSO model across the cross-validation repeats. The successive removal of the four most selected features resulted in a diminishing return of model accuracy. The full model equation is represented as $y = [4.85A + 1.46B + 2.75C - 0.41D + 0.16E + 0.09F + 211.0] \times 10^{-3}$, where model coefficients assigned to each feature were averaged across cross-validation repeats.

3. Feature Importance

Beyond proposing an overall predictive model, we used LASSO to determine the relative importance of each work feature to predicting outcomes (Fig. 12). We first examined the selection frequency of each work feature by our LASSO model which will determine the order of feature removal used in our subsequent analysis. Our analysis showed that negative work in shoulder adduction (98.9% selected), positive work in elbow flexion (66.1%) and in extension

(65.5%) were selected most while negative work in elbow flexion was not selected. The remaining features were selected considerably less than the top features (less than 12% each).

As a complement to selection frequency, we analyzed how removing features influenced the accuracy of model predictions. Interestingly, we found that removing the top feature (negative work in shoulder adduction) resulted in little change in model performance (mean R^2 ± SD, 0.51 ± 0.06). However, this apparent lack of sensitivity corresponded to an increase in the selection of negative work in shoulder abduction. Similarly, removing elbow positive work, selection frequency for negative elbow work increased, resulting in modestly improved model performance (0.57 ± 0.07). The apparent replacement of model features that resulted in small changes in model performance indicates that these features are highly correlated and compensated for any loss in variance explained. Model performance started to diminish after removing the fourth ranked feature (negative work in elbow extension, 0.46 ± 0.03) which indicates that the remaining features were unable to compensate for the first four features removed.

4. PCA Model

The results of our feature ranking suggested that several of the subcomponents of work are mutually correlated. To verify the dimensions required for predicting outcome, we performed an additional preprocessing step using principal component analysis (PCA) (See Fig. 13). PCA effectively maps our highly correlated features into an orthogonal space to obtain a set of uncorrelated principal components which we then used as our candidate features to predict recovery (changes in velocity range). Our main observation was that prediction accuracy improved with each additional principal components included in the model. The highest R^2 occurred using all eight possible principal components (mean R^2 ± SD, 0.80 ± 0.05); however,

this trend corresponded with increasing model uncertainty. The most reliable model, however, included the first four principal components (0.66 ± 0.02) since it corresponded with the largest change in R^2 from the preceding three component model and a diminishing return of R^2 on subsequent higher dimensional models. We obtained the coefficients of the four principal component model (See Fig. 13) which revealed that negative work in shoulder adduction, positive work in the elbow, and negative work in elbow flexion contributed the most in terms of relative magnitude, similar to our previous results (compare to our full model in Fig. 12).

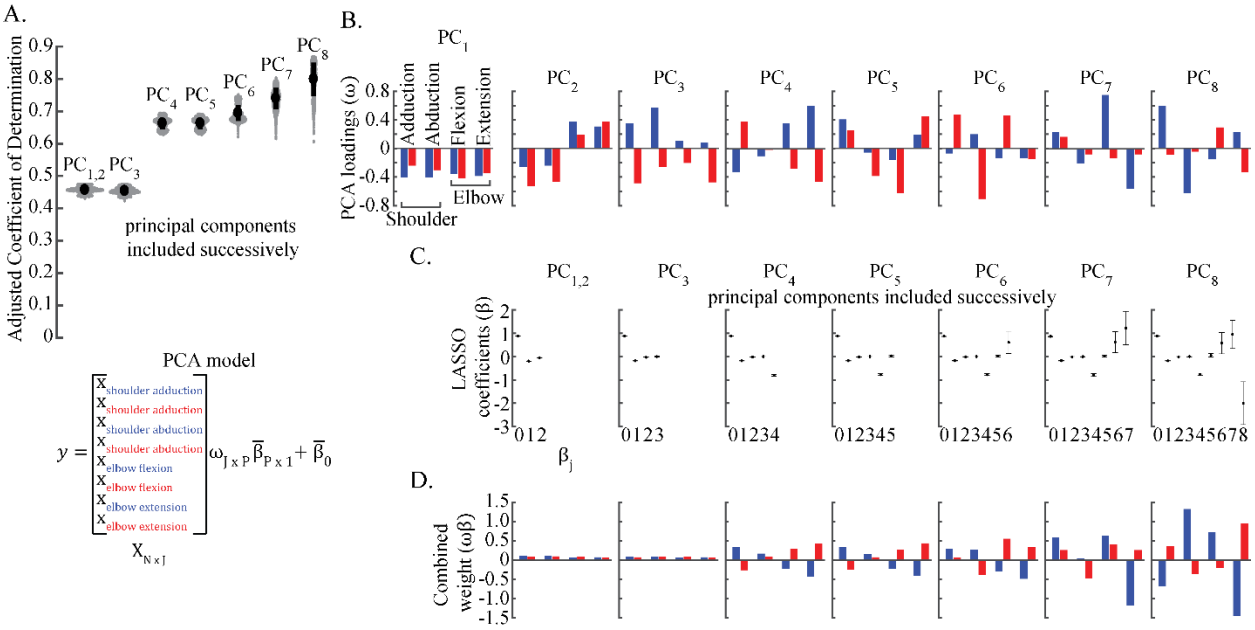


Figure 13: PCA Model Predictions Predicting patient recovery (changes in velocity coverage) due to robotic force training using features obtained from the principal components of the work subcomponents. (A) The successive inclusion of features increased average accuracy of prediction, but with decreased certainty (each gray dot represents a single cross-validation repeat, black dot and bar represents the mean R^2 ± SD). (B) The loadings determined by PCA for each individual principal component applied to the work subcomponents. The first three principal components appeared to correspond to differences between 1) low and high magnitude, 2) shoulder and elbow, 3) positive (concentric) and negative (eccentric) forms of work. (C) The coefficients the LASSO model assigned to each principal component feature (black dot and bar represents the mean ± SD calculated across cross-validation repeats). (D) The combined weights representing the relative contribution of each work subcomponent to model predictions as principal component features were successively included. The PCA model derived using the first four principal components closely resembles the full model in Fig. 12, which suggests a greater predictive importance of negative work in the shoulder adduction and positive work in the elbow.

D. DISCUSSION

The goal of this investigation was to identify the subcomponents of mechanical work that were most associated with recovery. This study provided evidence that the degee of recovery depends on the amount of mechanical work performed while experiencing interactive forces. Our main finding was that negative work (eccentric) during horizontal shoulder adduction and positive work (concentric) during elbow flexion and extension were the most important features in predicting improvement, measured by range of movement velocities. These results imply that more energy exerted in braking (i.e. energy dissipation) shoulder motion and driving (i.e. energy input) elbow motion can heighten recovery.

Beyond identifying these important subcomponents of work, our results showed that the energetics of patient limb motion can account for 52% of the variance in the increases in measures of independent movement capabality among the Force group. Ourfindings highlight the importance of active involvement in recovery, and in particular, how components of energetics may be used to evaluate involvement during robotic training. The differences in predictive power that we found between the subcomponents of work suggest that certain forms of motor expression are important in recovery. Importantly, our results also demonstrate that robotic training forces can provoke increases in both positive (concentric) and negative (eccentric) work compared to the controls that experienced equivalent amounts of repetitive practice without forces (See Fig. 11). Yet, according to our feature-selection procedure, specifically increasing negative work in the shoulder and positive work in the elbow led to corresponding increases in recovery.

Allowing patients to practice using a variety of movement patterns could have introduced several factors that contributed to increases in velocity coverage. Motor exploration differs from

tasks patients typically perform in other robot training studies where the control mechanisms associated with upper limb recovery may have direct clinical interpretations (Zackowski et al. 2004). Our results show that the shoulder contributed the most to energetics overall which is likely due to its importance to workspace area (Lenarcic and Umek 1994). However, the fact that negative work in the shoulder was the strongest predictor suggests that participants actively resisted the destabilizing force fields mostly by using the shoulder to dissipate energy. It is possible that such dissipation of energy in the shoulder must be coordinated with exertion of positive work (concentric) in the elbow (Farris et al. 2015; Ellis et al. 2005). This result would be unsurprising since stroke survivors commonly exhibit abnormal flexion synergy during arm motion (Ellis, Carmona, et al. 2016; Sangani et al. 2007). While appropriate acceleration and deceleration is known to be fundamental in reaching (Flash and Hogan 1985), sustained movement in particular does not engage the motor system to use strategies for dissipating energy (Doke and Kuo 2007a). It is possible that negative work at the shoulder was the preferred mechanism to break out of abnormal synergies; thus, allowing for greater independence of elbow flexion and extension motion. Further analysis is needed to fully inspect how different types of work (negative/eccentric and positive/concentric work) are associated with the specific types of movement patterns patients made during exploration training.

Our model achieved 52% of variance explained even though the features considered were only components of work. However, it is important to note that our model predictions and identification of the important work subcomponents apply only to when robot forces were introduced. Unsurprisingly, our analysis showed that patients who train with the velocity-augmenting forces (Force Group) exhibited substantially more work across training than the control group. The fact that our prediction model identified the training group factor as an

important feature suggests that the physical presence of robotic forces is necessary for predicting recovery from measures of involvement, consistent with previous studies (Loredana Zollo et al. 2011). It is interesting, however, that the control group showed similar gains in velocity coverage (and UEFM scores) despite the absence of training forces. This could point to general therapeutic benefits of motor exploration training with simple feedback that encourages randomness. Alternatively, it is possible that motor exploration without forces reinforces stereotypical movement patterns in stroke, but at greater intensities. As our coverage metric only captures the overall change in patients' ability to sustain greater velocities, additional analysis might elucidate whether the training groups differed in the particular forms of motor behaviors expressed during training that led to similar outcomes.

An important limitation of this study is that our key metric, velocity range, revealed improvements that might not be representative of general motor function. It is possible that the observed changes simply reflect a shift in motion patterns in response to training conditions that represent some combination of long-term adaptation or short-term changes in reward-motivated behavior. This may be one reason why we did not find a relationship between energetics and clinical outcomes. Because power is governed partly by speed, and work is the integral of power, it may be little surprise work is related to velocity coverage. Changes in UEFM scores may have been too small to be considered clinically relevant. Each participant's customized force field was designed to push their hand towards underrepresented velocities, which were nearly always higher. Velocities in the arm observed in this study might not transfer to other forms of motor function. Nevertheless, we previously provided evidence (Zachary A. Wright et al. 2017) that increases in velocity coverage corresponded with faster task execution in "transfer" tasks that were not trained, both important hallmarks of functional gain.

Our PCA analysis highlights a limitation of multiple regression, and feature selection using LASSO. We entertained alternative strategies to deal with the multicollinearity problem; including, re-ranking the features after removing a covariate and removing likely pairs of covariates in random order. The predictive power of our model improved substantially with use of PCA, which suggests that the use of highly correlated metrics introduced a source of noise or unnecessary model complexity. Our results also suggest that a minimum of four dimensions (the first four principal components of the work features) are required in our prediction model of recovery, consistent with the results from our feature removal procedure. It is important to note that, in contrast to our model that excluded covariate features, the PCA-LASSO model relies on the independent contributions of each work feature (See Fig. 13D). This indicates that each individual component of work could play an important role in model predictions when the shared variance is removed with PCA. It is likely that the higher dimensional principal components describe additional noise within the work features, thus resulting in model uncertainty. Machine learning methods have been used recently to not only identify relationships but to also understand levels of uncertainty using bootstrapping and cross validation.

While the focus of this investigation was to understand the outcome relation with mechanical work, additional factors can and should be considered to further enhance prediction of recovery. Individual differences in patient involvement might be described in a variety of ways, for example, movement frequency (Thaut et al. 2002), muscle activations (Dipietro et al. 2005), end-point force production (Mazzoleni et al. 2014), or even considering the energetics of the robot (Guidali et al. 2016). It is also likely that initial patient ability dictates their capacity to do work. In fact, we found trends between the initial characterization of velocity coverage during motor exploration and changes due to training (See Fig. 11). This is consistent with previous

studies that have found that certain features of movement capabilities at the beginning of training can predict changes in clinical outcomes (Mostafavi et al. 2015a), for example, simple measures such as movement speed (Majeed et al. 2015). Further analysis is needed to determine key factors both in terms of patient characteristics and training conditions that impact recovery from robot-assisted therapy. Another shortcoming may be that this model does not consider interaction terms between covariates. We believed that simple linear terms would serve as the best proof of concept, and LASSO provided quality feature selection abilities important for predicting recovery. Adding interaction terms might complicate the model, making it more difficult to interpret. Further work is needed to examine how such interaction effects build upon the main predictive trends found in the current study.

Our implementation of an upper limb inverse dynamic model provides a convenient method for quantifying the energetic contributions of patients to movement. It is worth noting, however, that mechanical work cannot account for physiological preferences established by muscle length-force relationships, joint pain, co-contraction, or other differences that influence force or motion capacities. The interactive forces provided by the robot during training varied as a function of the patients' instantaneous movement velocities without consideration of the differences in effort requirements at different joint angles. For example, it is possible that the forces provided too little assistance at the extremes of motion. Researchers have recommended alternative "weighted" measures of work to account for these important physiological relationships (Andrews 1983) which also require thoughtful consideration for future iterations of force field design.

Last, it is important to clarify that the LASSO methods we present in this paper do not substitute true experimental validation. In fact, our results show that the cross-validated mean

squared error (i.e. the estimated error in predicting future data) was greater than the mean squared error of the trained model predictions which indicates some degree of model overfitting. Our overall approach of applying exhaustive measures (repeated cross-validation and bootstrapping), and subsequent feature removal analysis based on the selection frequency metric, offered a process for identifying the important model features given our limited data set. While LASSO is known to enhance predictive quality and prevent overfitting, it provided a convenient method for examining feature selection in our analysis. Future analysis should consider alternative modelling techniques, including other regularization methods, to determine whether LASSO does in fact improve predictive quality.

E. CONCLUSIONS

This study examined the special case of training with interactive forces and provided evidence that the degree of recovery in stroke survivors depends on the amount of mechanical work performed. In addition, our analysis revealed that the components of work most important to predicting recovery were those associated with eccentric shoulder and concentric elbow motions. We remain cautious about how our predictions might generalize, especially with respect to other types of training (e.g. reaching) that involve different forms of force interactions. However, our results highlight the importance of quantifying patient involvement, as well as revealing how specific forms of involvement should be targeted. Our findings demonstrate the potential for energetic measures in the evaluation and design of robot assisted therapy.

IV. MOVEMENT DISTRIBUTION VARIABLES THAT BEST PREDICT UPPER LIMB IMPAIRMENT IN STROKE SURVIVORS[1]

Zachary A. Wright, Yang Dai, Felix C. Huang and James L. Patton

Abstract: Rehabilitation robots can not only exert forces; they can conveniently measure motor deficits that differ from patient to patient. Here, 22 stroke survivors were asked to freely move at various speeds and directions while trying not to repeat movement patterns, and pairs of two-dimensional probability distributions were used in principal component regression analyses to predict clinical scores and describe differences from healthy (N = 16). We found that partial least squares regression (PLSR) provided the best results, wherein joint and endpoint (Cartesian) kinematics were the best predictors of the clinical Fugl-Meyer score (coefficient of determination, R^2 = 0.45 for endpoint position, 0.44 for joint velocity and 0.29 for endpoint force). These results indicate that easily observed variables (position and velocity) are sufficient predictors of clinical scores. We also examined how these variables can differentiate stroke survivors from healthy, and found that endpoint kinematics provide the best classification accuracy (94.7%) using the leave-one-out cross validation procedure. Principal components reveal that many stroke subjects differ from healthy in a similar way, and several individuals with a unique deficit signature unrelated to clinical score. Taken together, these results provide a concise recipe for reduced datasets to monitor and describe impairment level in this patient population

[1] manuscript submitted for review to *Journal of Biomedical and Health Informatics* (November 2019)

A. INTRODUCTION

The wide variation in functional actions combined with the many differences in upper limb motor impairments among stroke survivors presents many challenges for assessment and therapy. Many clinical researchers agree that more comprehensive and informative measurements of each stroke survivor's movement capabilities are needed to better recommend personalized treatment. In recent years, stroke rehabilitation has focused on advancing intelligent and automated devices such as robots to guide therapy. What makes robots so appealing is their capacity to measure and record vast amounts of data across many different variables; however, it is still unclear which variables are most valuable for characterizing an individual's deficits. While each variable could provide a unique description of impairment, it is perhaps as equally important to identify which variables best describe individual differences in movement capabilities between stroke survivors. In this paper, we investigate how *probability distributions* of several movement variables relate to impairment level.

Recent work has identified the value of looking at a variety of movement actions by asking an individual to fully explore their entire range of available movements (F. C. Huang and Patton 2013b; F. C. Huang and Patton 2013a). Researchers have devised several robot-based measurements to characterize motor performance (L Zollo et al. 2011; Balasubramanian et al. 2012); many of which have known relationships with clinical descriptions of impairment (e.g. range of motion, joint coupling). However, such measurements are typically derived from relatively simple goal-directed (e.g. reaching) or isometric motor tasks which often disregard the natural variability of movement. Our strategy is to allow patients to move any way they want by freely exploring the workspace and their volitional movement capabilities. We then statistically characterize their exploratory movement patterns using *distribution analysis*. Past work in our

lab has shown how probability distributions could provide a detailed description of each patient's unique motor deficits (F. C. Huang and Patton 2016) which directly informed the design of a customized robot intervention (Zachary A Wright et al. 2017). However, our analysis has only focused on distributions of endpoint (i.e. the patient's hand) motion of the limb. As robots can easily track joint motion and measure forces, and data storage has become less expensive, there are additional variables yet to be explored with distribution analysis.

Relating robot-based measures to conventional assessments is one way to test their clinical value. Other researchers have relied on simple correlation and multiple regression analysis to identify measures that best predict impairment (Mostafavi et al. 2015a; Zariffa et al. 2012; Bosecker et al. 2010). Furthermore, recent work has also explored how robotic variables can predict change or recovery (Majeed, Awadalla, and Patton, n.d.; Zachary A. Wright, Patton, and Huang 2018). The focus of this paper is different from our previous work in that it uses the vast probability distribution data to determine the relationship between various robotic metrics and clinical measures. Our goal is to identify the best predictors among several candidate distribution variables.

One key challenge is that the high dimensionality of such distribution analysis offers many possible features which are likely mutually correlated. Multicollinearity among features can make it difficult to interpret the results of multiple regression models. Dimensionality reduction techniques, such as principal component analysis (PCA), have been widely employed to reduce the number of candidate features (Jolliffe, n.d.). PCA also conveniently captures the variation in the data which would seem important for differentiating stroke survivors. Regularized regression techniques, such as Least Absolute Shrinkage and Selection Operator (LASSO) can further enhance predictive power and prevent overfitting by including only the

important features in the model (Tibshirani 1996). Additionally, the method of partial least squares (PLS) uses supervision to enhance the relation between feature variance and regression outcome (Wold 1975). We explore these two regression models to evaluate the predictive ability of the distribution variables.

Because impairments can vary considerably in this heterogeneous population, it is important to initially be inclusive and agnostic to which may lead to success. While endpoint and joint kinematics are often used to describe observable clinical descriptions of impairment (e.g. range of motion and joint coupling), motor impairments often manifest in terms of weak neural drive or changes in muscle force production, which are perhaps more directly attributed to kinetics (i.e. force and energy ; including endpoint force, joint torque and joint power) (Lodha et al. 2010; Ellis et al. 2005). If any of these robotic variables is valuable, they should (1) be related to clinical scores such as Fugl-Meyer (Fugl-Meyer et al. 1975), (2) be able to easily differentiate a patient from a healthy individual, and (3) they should pinpoint unique differences between patients.

Accordingly, we investigated how distributions of self-directed motor exploration data relate to clinical measure of impairment. We combine PCA and PLS with multiple regression analysis to determine the distribution variables that best predict upper-extremity Fugl-Meyer scores (UEFM) and that best differentiate them from healthy individuals and from each other. We hypothesized that kinetic variables best predict individual differences in impairment level. This study provides new information about how patients differ from healthy and from each other, which should guide diagnostic and therapy methods in the future.

B. METHODS

1. Experimental Participants

Our analysis included data collected from stroke survivors (N = 22, Age ± SD = 57 ± 9 years) who were experimental participants in our previous robot therapy study (Zachary A Wright et al. 2017). We only considered data collected during the Baseline session (day 1) which participants completed before the start of intervention. All stroke survivors were screened prior to participation by a physical therapist and completed a clinical assessment (upper extremity portion of the Fugl-Meyer Assessment, UEFM score (Fugl-Meyer et al. 1975); Action Research Arm Test, ARAT (Lang et al. 2006); Modified Ashworth Scale, MAS (ASHWORTH 1964); Chedoke McMaster Stroke Assessment- Arm, CMSA-A (Gowland et al. 1993); and elbow range of motion, ROM). The primary inclusion criteria were 1) chronic stroke, 8+ months post-stroke 2) hemiparesis with moderate to severe arm impairment measured by UEFM, score of 15-50 and 3) primary cortex involvement. This study also considered new experimental data collected from young healthy participants (N = 16, Age Range = 18-24). All participants provided informed consent in accordance with the Northwestern University Institutional Review Boards.

2. Experimental Apparatus

Participants were asked to perform a planar movement task using a two-degree of freedom robotic device (Fig. 14). Stroke survivors performed the task using their affected limb (12 Right, 10 Left). Each healthy participant was right-hand dominant. Each participant was seated in a chair and situated with respect to the robot such that their shoulder lined up with the center of the workspace (0.6 meters x 0.4 meters). The chair was located at distance from the robot where participants could comfortably move their arm within the workspace. Participants

operated the robot's end-effector through a wrist brace attached to a revolute joint which allowed them to focus on forearm and upper arm coordination. The weight of their arm was supported by a gravity assistance device during movement. A custom video display system provided visual feedback (a green cursor) of the location of the wrist as the arm moved. The workspace boundaries were visible to the participants indicated by a white outline. The cursor changed color to red if it moved outside the boundary.

The robot control and instrumentation were mediated by a Simulink-based XPC Target computer, with a basic rate of 1kHz. We programmed the robot to provide compensation for the inertial effects of the weight of the robot arm. The device is equipped with two encoders at each of the joints to record the position of the end-effector and a force sensor attached to the end-effector to measure human-robot interaction. Data was collected at 200 Hz and filtered using a 5th order Butterworth low pass filter with a 12 Hz cutoff.

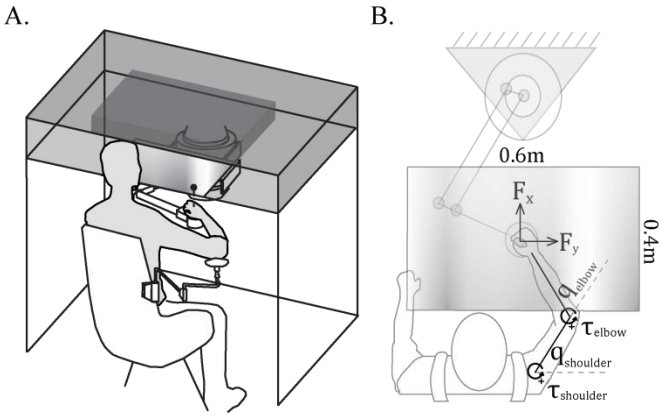


Figure 14: Experiment Apparatus and Movement Task (A) Stroke survivors performed a motor exploration task by controlling the arm of a planar robotic device that allowed them to freely explore different positions and speeds within the workspace. (B) For each participant, the robot recorded the position of their hand and the forces exerted on the handle which allowed for the calculation of a suite of motion variables; including, endpoint kinematics, joint kinematics of the elbow and shoulder and kinetic variables (endpoint force, joint torque and power).

3. Experimental Task

Each participant performed a self-directed movement exploration task for a total of 12 minutes divided into six two-minute trials. The scripted instructions we provided to each participant were as follows: “You are asked to perform self-directed movement in which you will move your arm any way you want while holding onto the robot. We ask that you to try to vary your movement as much as possible by moving your arm any way you want. Try to reach to all different positions within the white boundary and try to move at various speeds and directions. We encourage you to avoid repeating the same movements continuously.” Participants were instructed to begin a trial when they were ready by moving the cursor to the start position (a blue box on the screen). Each trial ended after two cumulative minutes of movement in which movement speed was above a threshold of 0.04 m/s and movement remained within the workspace boundaries. Participants were able to rest for any amount of time during a trial and between trials. At the end of each trial, we provided participants with a Performance Score, described previously (Zachary A Wright et al. 2017). The purpose of the Performance Score was to provide feedback to participants on how well they varied their movement and to encourage participants to continue to vary their movement as much as possible on the next trial.

4. Analysis

a. Motion Variables

Our analysis considered nine different variables that describe upper limb movement of each participant’s exploratory practice behaviors (Fig. 14). We divided them into three domains:

Endpoint kinematic variables consisted of position (meters), velocity (meters/second) and acceleration (meters/second^2) measured in two dimensional Cartesian coordinates. The robot recorded the position of each participant's hand during motor exploration. We computed the first-order and second-order time derivatives of position to obtain velocity and acceleration.

Joint kinematics variables consisted of angular position (radians), velocity (radians/second) and acceleration (radians/second^2) measured at the elbow and shoulder joints. Using anthropometric measurements of each participant's upper limb, we computed inverse kinematic relationships to obtain elbow and shoulder joint angles corresponding to endpoint position data. We then computed the first-order and second-order time derivatives to obtain angular velocity and acceleration for each joint.

Kinetic motion variables consisted of endpoint force (Newtons) measured at the handle of the robot in Cartesian coordinates and elbow and shoulder joint torque (Newtons-radians) and power (Joules/second). We employed established methods of inverse dynamics of upper limb motion to estimate the torque generated by the human at the elbow and shoulder (Bhushan and Shadmehr 1999). The model considers the torque acting on each individual joint; including, the torque from the passive limb dynamics which is equal to sum of the torque acting on the arm by the robot and the torque generated by the human. We computed the torque of the robot acting on the arm using the Jacobian matrix of the arm and the interaction force measured from the force sensor. Using the torque and angular velocity measurements, we computed the mechanical power generated by the patients at each joint.

b. Distribution Analysis

Our analysis considered the overall variation of each participant's movement

behaviors during motor exploration. For each participant, we constructed two dimensional histograms of each motion variable measured during motor exploration. We constructed histograms using 40 x 40 equally sized bins and scaled according to the maximum (defined as the 99.9th percentile) magnitude of the data across participants which were included in the analysis. We express histograms as probability distributions by dividing each bin by the sum of the total number of data tabulated across 12 minutes of motor exploration. It is important to note that since our stroke population included both right and left side affected participants, we flipped the histograms for each left side affected participant along the horizontal axis. Fig. 15 shows the probability distributions of each motion variable constructed from the motor exploration data collected for each stroke survivor and Fig. 16 shows the probability distributions for each healthy participant. Here, we present distributions visually as binned probabilities with bins of higher probability in red, lower probability in blue and zero probability in white. The distributions of some variables result in very large probabilities near the zero crossing which is likely due to the accumulation of data during intermittent short periods of rest or changes in movement direction. For subsequent analysis, we first compute the log of the probability represented by each bin.

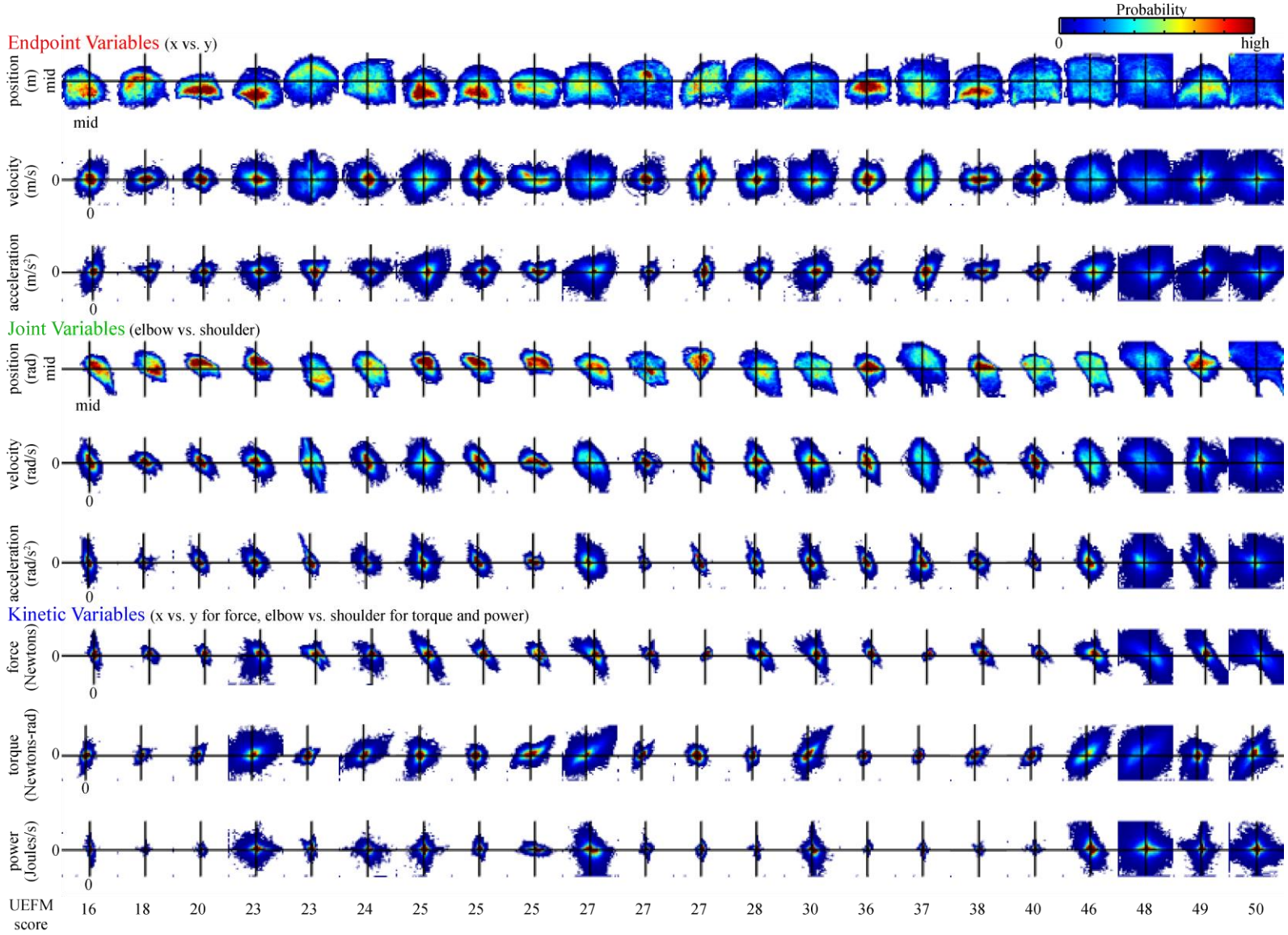


Figure 15: Movement Distributions of Stroke Survivors Probability distributions of several motion variables were constructed from each individual's motor exploration data. Each column represents the distributions for a single participant in order of clinical UEFM scores. Each row represents the distributions for each variable scaled and centered according to the "maximum" magnitude of the data across all stroke survivors.

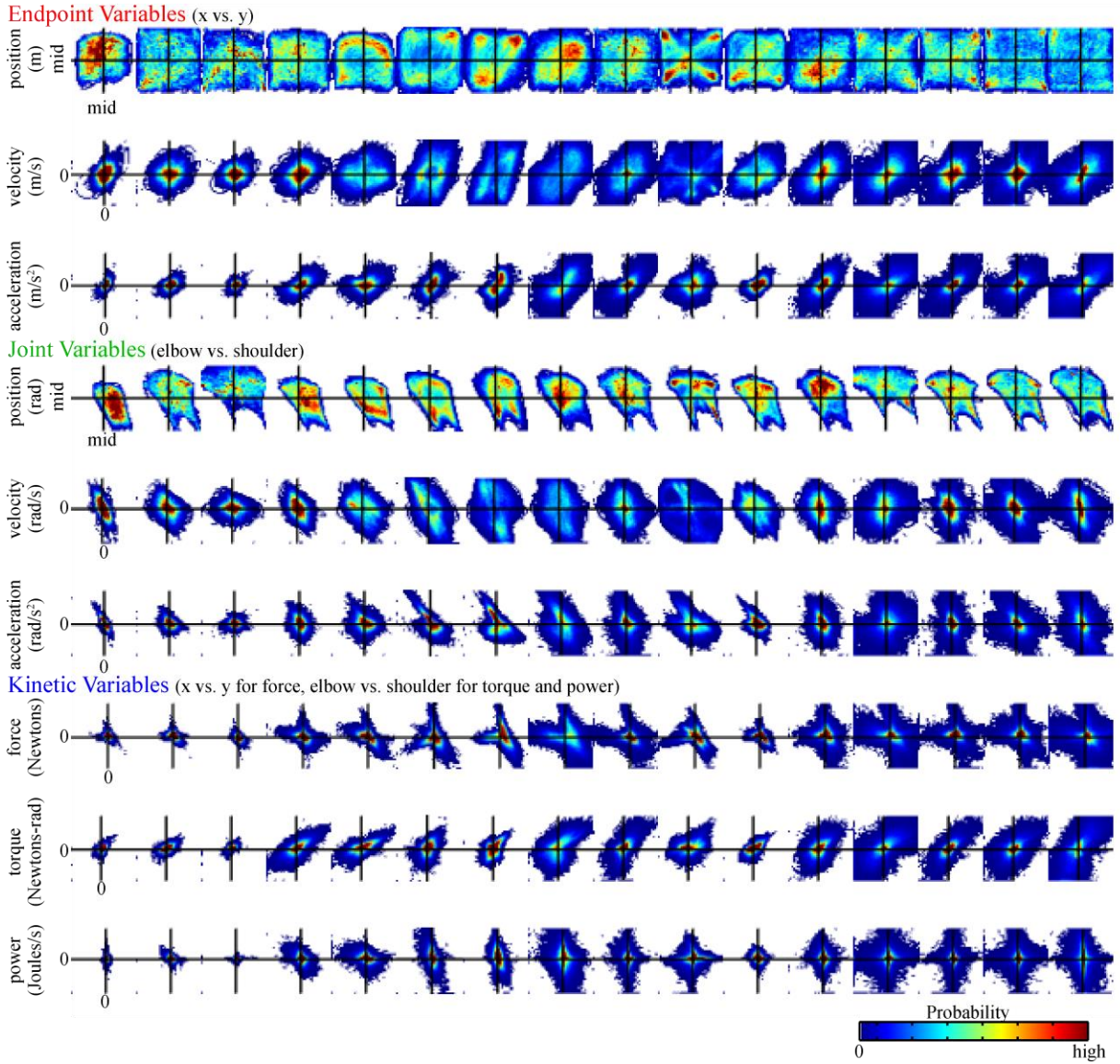


Figure 16: Movement Distributions of Healthy Individuals Probability distributions of several motion variables were constructed from each individual's motor exploration data. Each column represents the distributions for a single participant. Each row represents distributions for a single variable scaled and centered according to the "maximum" magnitude of the data across all stroke and healthy participants.

c. Principal Component Analysis

Our primary goal is to determine which domain of motion variables best predict clinical UEFM scores. Considering that the probability distributions for each individual motion variable consists of a large number of bins (40 bins x 40 bins), many of which are likely mutually correlated, it is impractical to include bins as the features in a regression model (i.e. the number of bins far exceeds the number of participants). Thus, we determined a reduced set of possible model predictors by first applying principal component analysis (PCA) (Jolliffe, n.d.). The advantage of PCA is that it transforms possibly correlated variables into a reduced set of linearly uncorrelated variables (called PC scores) using an orthogonal transformation. This transformation of the binned data into a new coordinate system is represented by the PC loadings which are aligned orthogonally in the directions where the variance in the data is greatest. The first principal component contains the largest possible variance in the probability distributions across participants. It is important to note that the principal components only capture the variance in the bin features and are computed independent of UEFM scores.

We determined a set of PC scores for each variable domain independently by applying PCA to the probability distributions of the three motion variables within each domain. As a preprocessing step to PCA, we first computed the log of the probability represented by each bin and then standardized these values across participants. For each separate PCA, we included all of the bins which were used to construct the probability distributions for each of the three variables (40 bins x 40 bins x 3 variables in total) within each respective domain. Each PCA provides N-1 PC scores which represent the total number of possible features to include in a separate prediction model for each variable domain. In addition to the PCA for each of three variable domains, our analysis considered a separate PCA applied to all of the bins used to construct the

distributions for all nine variables (14,400 bins in total). Fig. 17 shows the variation between stroke survivors described by the first three principal components resulting from PCA on all nine distribution variables. Fig. 17 also shows the relative contribution of each variable domain and of each individual variable to the total variance accounted for (%VAF) in each of the first three principal components.

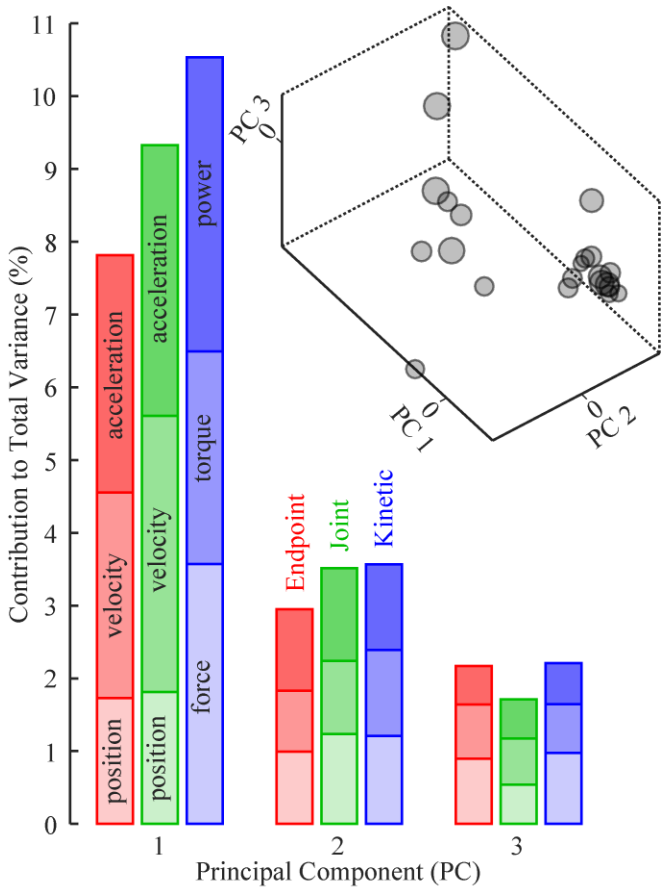


Figure 17: PCA on Movement Distributions of Stroke Survivors PCA applied to movement distributions (including all nine motion variables) across stroke survivors. (top right) PC scores on the first three principal components for each stroke survivor. The size of each data point is scaled according to clinical UEFM score. (bottom left) Kinetic variables and higher kinematic derivatives (i.e. endpoint and joint acceleration variables) contribute more to the variance accounted for in the first principal component. The contribution of each distribution bin to each individual principal component was determined by computing the square of the PC loadings matrix. Thus, the contribution of each variable to a principal component was computed by summing the contributions of all the bins and the contribution of each domain was computed by summing the contributions of each variable. The combined contribution of each domain is equal to the percent variance accounted for by each individual principal component.

d. PCA-LASSO model to predict UEFM scores

To determine which domain of motion variables best predict clinical UEFM scores, we employed a Least Absolute Shrinkage and Selection Operator (LASSO) model using PC scores as our input features (Tibshirani 1996). The LASSO method is a special case of regularized least squares regression which incorporates an additional penalty term on the L_1 norm of the model coefficients. It effectively drives the coefficients assigned to each input feature to zero if they are unhelpful to the predictions. This results in the further reduction of the number of potential features included in the model. The LASSO model is represented by the following formula in Lagrangian form:

$$\underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left(y_i - \beta_0 - \sum_{j=1}^{J} x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^{J} |\beta_j| \right\}$$

where y_i represents the UEFM score of subject i, x_{ij} represents the j-th PC scores of subject i, β_j represents the coefficient for j-th PC score, N equals the number of participants, J equals the number of PC scores included in the model, and λ is the penalty term which determines the degree of regularization and the number of PC scores excluded from the model.

We evaluated how well each domain of distribution variables, transformed into an independent set of PC scores, predicts our chosen clinical measure of impairment level, UEFM scores. We used leave-one-out cross-validation (LOOCV) to train a separate model to the PC scores computed for each domain. For each model, we included only the scores on the first principal components in which the cumulative VAF% in the probability distributions across participants was no greater than 80% (see Supplementary Table VI). Our training procedure included a search of the λ value (λ_{min}) that minimized the cross-validated mean squared error (MSE_{CV}) in the predicted UEFM scores. The range of λ is from 0 to λ_{max} in 1000 steps, where λ

= 0 represents a least squares regression model without regularization and $\lambda = \lambda_{max}$ represents an intercept-only model.

MSE_{CV} is an estimate of the error in predicting unseen data and is equivalent to the predicted residual error sum of squares (PRESS) statistic. The PRESS statistic is a summary measure of model fit commonly used to evaluate predictions resulting from LOOCV where each left-out observation is predicted from a new model fit to the remaining observations (Allen 1971). We compared the MSE_{CV} to the mean squared error of the predictions from the model fit of all the data (MSE_{ALL}) in order to determine the degree of overfitting. To evaluate the overall performance of each prediction model, we computed the PRESS coefficient of determination (R^2_{CV}). We compared the MSE_{CV} and the R^2_{CV} of the respective model fit of the PC scores computed for each variable domain and for a model fit of the PC scores computed from applying PCA to nine distribution variables (Full model).

As an alternative to the PCA-LASSO models, we considered partial least squares regression (PLSR) (Geladi and Kowalski 1986). Similar to PCA, partial least squares (PLS) is a data reduction technique commonly used as a preprocessing step to regression analysis when the independent variables are highly collinear and when the number of independent variables far exceeds the number of observations. PLS also constructs a set of component scores that are linear combinations of the independent variables. Whereas each successive principal component in PCA attempts to explain as much of the variance in the independent variables as possible, each successive component in PLS attempts to maximize the covariance between the independent variables (i.e. distribution bins) and the outcome variable (i.e. UEFM scores). PLSR is considered a supervised modelling approach. The LOOCV procedure was used to evaluate the PLSR models. We successively included components in the model and determined the number of

components which resulted in best model performance according to the PRESS coefficient of determination (R^2_{CV}).

e. PCA-logLASSO model to classify stroke and healthy

As a secondary goal for this study, we investigated the possibility of differentiating between stroke survivors and healthy individuals using the distributions of motor exploration. Specifically, we examined whether certain distribution variables are more useful for classification. For this analysis, we combined PCA with regression analysis to train logistic regression model (PCA-logLASSO) using LOOCV (Tibshirani 1996). For each participant, we constructed probability distributions using 40 x 40 equally sized bins scaled according to the maximum (defined as the 99.9^{th} percentile) magnitude of the data across both healthy and stroke participants. We then computed the log of the probability represented by each bin and standardized across all participants.

Similarly, we included only the scores on the first principal components where the cumulative %VAF was no greater than 80% in training PCA-logLASSO model for each domain (see Supplementary Table VIII). For each domain, the model was selected based on the λ value (λ_{min}) that minimized the cross-validated binomial deviance (D_{CV}). Binomial deviance is the difference in the log-likelihoods between the fitted model and the model that perfectly fits the data (i.e. saturated model). We compared the D_{CV} to the binomial deviance of the model fit of all the data (D_{ALL}) to assess the degree of overfitting. To evaluate the overall classification performance of each PCA-logLASSO model, we computed the Accuracy (ACC_{CV}, expressed as a percentage) in correct classifications (cutoff value = 0.5). The chance of correctly classifying a

stroke survivor was 57.9%. Since our two datasets were unbalanced, we also calculated the Matthews Correlation Coefficient (MCC_{CV}) (Matthews 1975) defined by the formula:

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

where TP, TN, FP and FN are the numbers of true positives, true negatives, false positives and false negatives, respectively.

C. RESULTS

1. Prediction of UEFM scores

The primary goal of this study was to evaluate the utility of distributions of motion variables for predicting impairment level in stroke survivors. Using a combined PCA-LASSO regression model, we compared the prediction performance for three domains of variables: joint kinematic variables, endpoint kinematic and kinetic motion variables (Fig. 18A). Our main finding was that distributions of joint kinematic variables provided the best overall prediction (R^2_{CV} = 0.44) of UEFM scores followed by endpoint kinematic variables (R^2_{CV} = 0.23). Distributions of kinetic variables resulted in considerably worse prediction (R^2_{CV} = -0.05) which corresponded to a larger value of MSE_{CV} compared to those based on the endpoint and joint kinematic variables (Fig. 18B). The MSE_{CV} (± SE) of the model predictions for endpoint kinematic variables, joint kinematic variables and kinetic variables was 86.1 (± 21.2), 62.9 (± 15.5), 117.2 (± 27.1), respectively. Interestingly, the Full model which considered the contribution of each individual distribution variable only predicted UEFM scores with R^2_{CV} = 0.26 and MSE_{CV} ± SE = 93.3 ± 18.2. We have summarized the results of the model predictions for each variable domain and the Full model (see Supplementary Table VI).

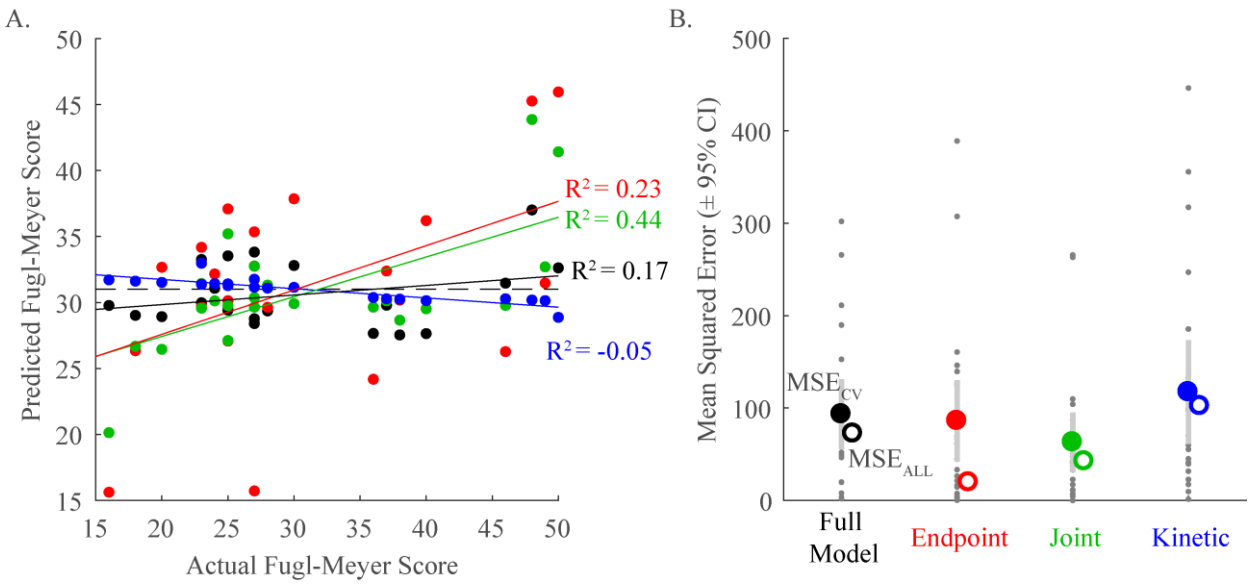


Figure 18: Prediction of UEFM scores using PCA-LASSO Distributions of joint kinematic variables best predict impairment level in stroke survivors. (A) Actual and predicted UEFM scores and corresponding coefficient of determination of model prediction for each variable domain. Each color corresponds to a variable domain. Each colored data point represents a single observation. Each colored line represents the regression line for each respective variable domain. The gray dashed line represents the average UEFM score. (B) The cross-validated predicted mean square error (MSE_{CV}, closed circles) for each variable domain and corresponding mean squared error of the model predictions trained on all the data (MSE_{ALL}, open circles). Each dark gray data point represents the prediction error for a single observation using leave-one-out cross validation. Each colored data point and light gray bar represents the mean and 95% confidence interval of the predicted mean square error for each respective variable domain.

We assessed the degree of overfitting in each of the models by comparing the estimated prediction error of unseen data (MSE_{CV}) to the respective mean prediction errors of the model trained on all the data (Fig. 18B). The model for endpoint kinematic variables resulted in the largest difference between MSE_{CV} and MSE_{ALL} (MSE_{CV} - MSE_{ALL} = +66.5), compared to joint kinematic variables (+20.5) and kinetic variables (+15.0). This suggests that the endpoint kinematics model suffers from overfitting leading to less reliable prediction of new data. It is important to note that the PCA-LASSO model at the selected λ_{min} value for endpoint kinematic variables retained a greater total number of model features (i.e. 10 out of the possible 13 PC scores were retained by the fitted model) compared to the joint kinematics model (3 out of 12

features retained). None of the features were retained in the kinetics model; meaning that none of the corresponding PC scores was unable to improve prediction over the intercept-only model.

We also offer an alternative to the PCA-LASSO models which considered the methods of partial least squares regression (PLSR). The results of the PLSR models showed that endpoint kinematic variables and joint kinematic variables performed better than kinetic distribution variables (Fig. 19). However, in contrast to the PCA-LASSO models, we found similar prediction accuracy between endpoint kinematic variables (R^2_{CV} = 0.39) and joint kinematic variables (R^2_{CV} = 0.39) using PLSR (Fig. 19A). In addition, the prediction accuracy of the PLSR model for the kinetic domain (R^2_{CV} = 0.23) was greater compared to the respective PCA-LASSO model. The MSE_{CV} (± SE) of the model predictions for endpoint kinematic variables, joint kinematic variables and kinetic variables was 68.6 (± 13.5), 68.5 (± 18.3), 87.7 (± 17.0), respectively (Fig. 19B). It is important to highlight that even though MSE_{ALL} for each PLSR model decreased with the successive inclusion of additional components, the predictive power of each model only marginally improved. This suggests that only a few components are required to achieve best model performance. We have provided a summary of the results of the one component PLSR model for each variable domain and the Full model (see Supplementary Table VII).

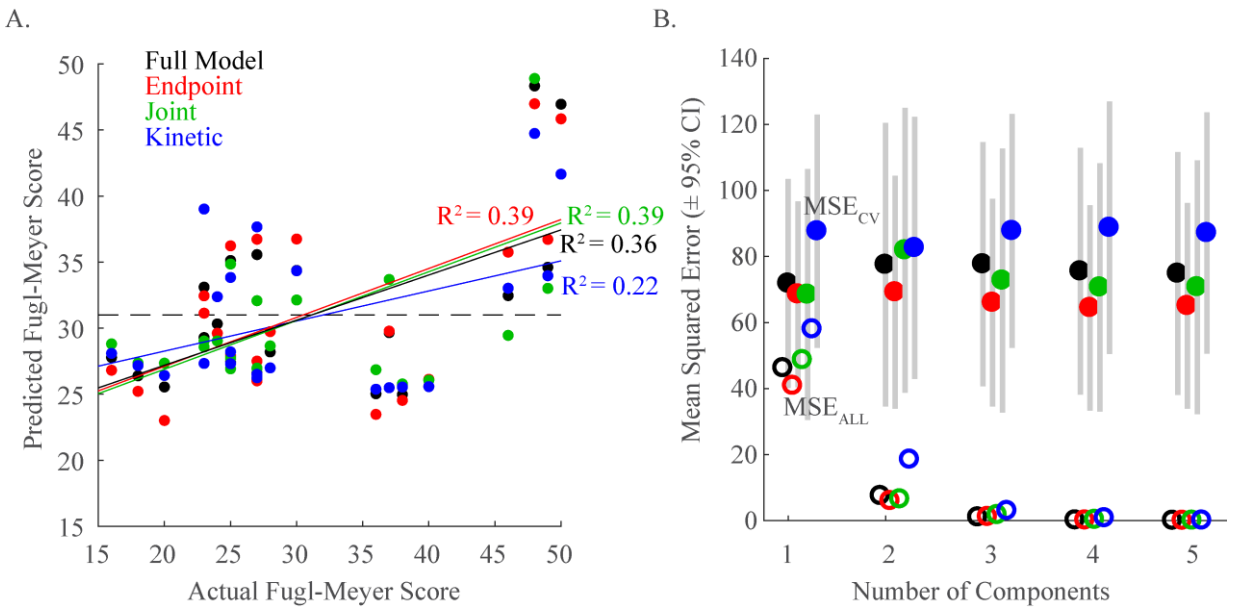


Figure 19: Prediction of UEFM scores using PLSR Distributions of endpoint and joint kinematic variables best predicted impairment level in stroke survivors. (A) Actual and predicted UEFM scores and corresponding coefficient of determination for the one component model for each variable domain. Each color corresponds to a variable domain. Each colored data point represents a single observation. Each colored line represents the regression line for each respective variable domain. The gray dashed line represents the average UEFM score. (B) The cross-validated predicted mean square error (MSE_{CV}, closed circles) for each variable domain and corresponding mean squared error of the model predictions trained on all the data (MSE_{ALL}, open circles) with and increasing number of components included in the model. Each colored data point and light gray bar represents the mean and 95% confidence interval of the predicted mean square error for each respective variable domain.

In addition to identifying the variable domain that best predicts impairment level, we evaluated the predictability of each individual distribution variable. Interestingly, we found that some distribution variables in isolation were more predictive than when combined with other variables. For predictions using PCA-LASSO, we found that distributions of joint angular velocity provide the best overall prediction (R^2_{CV} = 0.48). Also, for both endpoint variables and joint variables, we found that position and velocity distributions were more predictive than acceleration. Among the endpoint variables, distributions of endpoint position (R^2_{CV} = 0.36) performed better than velocity (R^2_{CV} = 0.20) and acceleration (R^2_{CV} = -0.005). Among the joint variables, distributions of angular velocity performed better than angular position (R^2_{CV} = 0.34)

and acceleration (R^2_{CV} = 0.22). Among the kinetic variables, distributions of joint power (R^2_{CV} = 0.30) were more predictive than endpoint force (R^2_{CV} = -0.01) and joint torque (R^2_{CV} = -0.01). We found similar trends among the endpoint variables and joint variables using PLSR, but not kinetic variables. Endpoint force distributions (R^2_{CV} = 0.29) were more predictive than joint torque (R^2_{CV} = 0.01) and power (R^2_{CV} = 0.24). A summary of the results of the model predictions for each individual distribution variable using PCA-LASSO and PLSR can be found in Supplementary Table VI and Supplementary Table VII, respectively.

2. Classification of stroke and healthy

As a secondary goal for this study, we tested the possibility of differentiating between of stroke survivors and healthy individuals using the distributions of motor exploration behaviors. Our analysis demonstrated successful classification between stroke and healthy individuals and revealed that distributions of endpoint variables are the most useful for classification (ACC_{CV} = 94.7%, MCC_{CV} = 0.89) followed closely by joint variables (ACC_{CV} = 92.1%, MCC_{CV} = 0.84). While the model using kinetic distribution variables (ACC_{CV} = 81.6%, MCC_{CV} = 0.62) did not perform as well, it is better than chance which equals57.9%. Similarly, including all distribution variables in the PCA did not improve classification performance (ACC_{CV} = 86.8%, MCC_{CV} = 0.73). Among all individual distribution variables, using endpoint acceleration outperformed the rest the others. A summary of the results of the PCA-logLASSO model can be found in Supplementary Table VIII.

As a supplement to the classification analysis, we inspected which variables contributed the most to the variation in how the movement distributions of stroke survivors differed from healthy individuals (Fig. 20). For each individual distribution variable, we subtracted the respective average distribution across healthy individuals from that of each individual stroke

survivor. We then applied PCA including all bins from each variable. As shown in Fig. 20, we examined the principal component axes that most differentiated stroke survivors in their differences from healthy according to the VAF (variance accounted for). We found that joint kinematic distribution variables collectively contributed the most to the variation captured in the first principal component (9.6 %VAF) and endpoint force distributions (3.8 %VAF) contribute the most individually (Fig. 20).

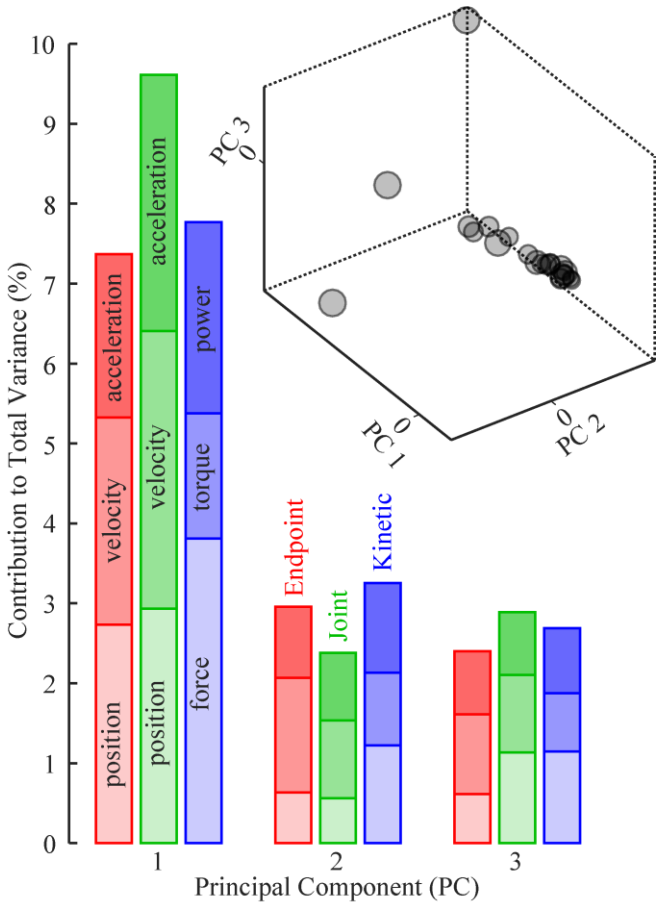


Figure 20: PCA on Stroke minus Average Healthy Movement Distributions(top right) The first three principal components account for 41.3% of the total variation within the contrast between the distributions of stroke survivors and the mean healthy distributions (all nine motion variables were included in a single PCA). Each data point represents the PC scores for each individual stroke survivor. The size of each data point is scaled according to respective clinical UEFM score. (bottom left) Joint variables contribute the most to the variance accounted for by the first principal component compared to endpoint and kinetic variables. Force distributions contribute the most to the first principal component compared to the individual contributions of all other variables.

Supplementary Table VI:: Prediction of UEFM scores using PCA-LASSO

	λ_{min}	# PC Included	# PC Retained	R^2_{CV}	R^2_{ALL}	$MSE_{CV} \pm SE$	MSE_{ALL}
Full Model	3.12	13	1	0.17	0.29	93.3 ± 18.2	72.4
Endpoint Variables	0.76	13	10	0.23	0.81	86.1 ± 21.2	19.6
Position	2.27	11	2	0.36	0.49	71.6 ± 17.3	51.9
Velocity	2.51	12	3	0.20	0.39	90.3 ± 18.7	62.1
Acceleration	3.71	12	1	-0.005	0.14	112.7 ± 23.1	88.2
Joint Variables	1.73	12	3	0.44	0.58	62.9 ± 15.5	42.4
Position	1.53	9	2	0.34	0.44	74.6 ± 23.0	57.0
Velocity	0.96	10	5	0.48	0.60	58.8 ± 15.6	40.6
Acceleration	1.94	12	4	0.22	0.46	87.4 ± 18.9	55.0
Kinetic Variables	4.89	11	0	-0.05	0.00	117.2 ± 27.1	102.2
Force	4.95	8	1	-0.01	0.06	113.4 ± 25.7	95.5
Torque	4.03	12	1	-0.01	0.05	112.7 ± 26.3	96.6
Power	0.25	10	9	0.30	0.75	78.3 ± 22.0	25.4

PC – Principal Components; R^2 – Coefficient of Determination; MSE – Mean Squared Error; SE – Standard Error
CV – Results on Test Set using Leave-One-Out Cross-Validation; ALL – Results on model trained on all data

Supplementary Table VII:: Prediction of UEFM scores using PLSR

	# Components	R^2_{CV}	R^2_{ALL}	$MSE_{CV} \pm SE$	MSE_{ALL}
Full Model	1	0.36	0.55	71.8 ± 15.3	46.2
Endpoint Variables	1	0.39	0.60	68.6 ± 13.5	40.9
Position	1	0.45	0.63	61.8 ± 19.7	37.6
Velocity	1	0.29	0.48	79.6 ± 13.7	52.7
Acceleration	1	0.23	0.44	86.4 ± 17.4	57.3
Joint Variables	1	0.39	0.52	68.5 ± 18.3	48.6
Position	1	0.33	0.45	75.2 ± 25.0	56.2
Velocity	1	0.44	0.55	63.3 ± 15.8	45.8
Acceleration	1	0.30	0.45	79.1 ± 19.6	56.5
Kinetic Variables	1	0.22	0.44	87.7 ± 17.0	58.0
Force	1	0.29	0.46	79.9 ± 16.8	56.0
Torque	1	0.01	0.34	110.8 ± 21.3	68.9
Power	1	0.24	0.43	84.9 ± 16.8	58.8

R^2 – Coefficient of Determination; MSE – Mean Squared Error; SE – Standard Error
CV – Results on Test Set using Leave-One-Out Cross-Validation; ALL – Results on model trained on all data

Supplementary Table VIII:: Classification of Stroke and Healthy using PCA-logLASSO

	λ_{min}	# Features Included	# Features Retained	$D_{CV} \pm SE$	D_{ALL}	ACC_{CV}	MCC_{CV}
Full Model	0.003	21	13	0.48 ± 0.21	0.02 ± 0.001	86.8%	0.73
Endpoint Variables	0.016	20	8	0.45 ± 0.21	0.12 ± 0.004	94.7%	0.89
Position	0.098	20	3	0.93 ± 0.15	0.73 ± 0.03	78.9%	0.57
Velocity	0.131	17	1	0.99 ± 0.11	0.91 ± 0.03	84.2%	0.68
Acceleration	0.027	14	4	0.60 ± 0.19	0.41 ± 0.03	92.1%	0.84
Joint Variables	0.031	18	7	0.54 ± 0.15	0.25 ± 0.01	92.1%	0.84
Position	0.067	21	4	0.80 ± 0.13	0.58 ± 0.02	81.6%	0.62
Velocity	0.000	21	17	0.71 ± 0.34	0 ± 0	86.8%	0.73
Acceleration	0.045	15	5	0.61 ± 0.16	0.35 ± 0.02	86.8%	0.73
Kinetic Variables	0.031	17	10	0.84 ± 0.19	0.39 ± 0.02	81.6%	0.62
Force	0.135	20	2	1.03 ± 0.11	0.91 ± 0.03	73.7%	0.46
Torque	0.036	15	8	0.81 ± 0.21	0.42 ± 0.02	86.8%	0.73
Power	0.116	20	3	0.96 ± 0.12	0.85 ± 0.03	84.2%	0.68

D – Binomial Deviance; ACC – Accuracy; MCC – Matthews Correlation Coefficient; SE – Standard Error
CV – Results on Test Set using Leave-One-Out Cross-Validation; ALL – Results on model trained on all data

D. DISCUSSION

This study investigated the relationship between the probability distributions of motor exploration data and a clinical measure of impairment in stroke survivors. We presented results of principal component regression analyses to predict clinical scores and describe differences from healthy. Kinematic variables were the best predictors of the clinical Fugl-Meyer score, indicating that easily observed variables (position and velocity) are important predictors of clinical scores. We also examined how these variables can differentiate stroke survivors from healthy, and similarly, found that kinematics provide the best classification accuracy. Additionally, most subjects differed from healthy in similar ways, but several individuals' deficits were definitely unique in ways unrelated to their clinical scores. Taken together, these results provide a concise recipe for reduced datasets to monitor and describe impairment in this patient population.

Our results highlight the important relationship between motion and clinical assessments. The fact that kinematics is better at pinpointing impairments than kinetics may be unsurprising since, in practice, clinicians typically track the position of the hand and its relation to the patient body across a variety of tasks while scoring clinical assessments. Deficits in range of motion and abnormal joint coupling of the limb are easier to infer from observing motion than deficits related force and torque generation.

In addition to determining the differences in predictive ability of our variable domains, we were also interested in comparing the predictive power of each individual distribution variable. Our results show that kinematic acceleration variables were not as predictive as position and velocity. Considering the close relationship between acceleration and force, it might have been expected to find similarities in the lack of prediction of acceleration and kinetic variables.

As our motor exploration task allows patients to explore their full range of motion, it does not allow them to vary their full range of force capabilities. It is possible that our free exploration task provided insufficient challenge to reveal the kinetic deficits since patients were only required to overcome their passive limb dynamics. Research has shown that impairments, such as abnormal synergies and weakness, to be task dependent (Beer, Given, and Dewald 1999; Alessandro et al. 2013). Using isometric tasks (Melendez-Calderon et al. 2017)or muscle activity tasks (Aggarwal et al. 2019) would better probe maximal contraction capacities. Alternatively, introducing large resistive robot loads would require patients to increase muscle activity and make maximal deficits more detectable and might improve the predictive power of kinetic variables.

The trends in predictive power among the variable domains generally agreed with principal components (PCA-LASSO), but we found better predictions overall using partial least squares (PLSR), where the components were built relating to outcome. While both regression techniques considered the shared contributions among variables, they are fundamentally different in terms of the variance each attempts to maximize. It is likely that some principal components were only relevant for describing variance within the distributions, but account for too little variance that contains noise not relevant for prediction. However, the PLSR method showed that only a few, and in most cases, model performance required only one component. It is likely that most of the variation accounted for in the first component describes differences in range of motion capabilities among the stroke survivors (see Fig. 15). Despite overall modest predictive performance, it is remarkable that only one component can predict with upwards of 45% accuracy. Still, 45% is not high; the general lack of predictive success could be simply due to a lack of data or differences in range of motion could be the only feature of distributions that is

important for prediction.

Beyond predicting stroke-stroke differences in impairment, our results show that distribution analysis of motor exploration can reliably differentiate stroke from healthy individuals. We did find better classification with kinematic variables compared to kinetic variables. Again, it is likely that the differences in range of motion contributed the most to the differences between stroke and healthy. It is interesting to note, however, that classification performance was more likely to result in a false positive (meaning healthy individual were confused as stroke) than a false negative, but less impaired stroke survivors (i.e. higher UEFM score) were more likely to be confused as a healthy. It is evident that the distributions of healthy individuals (see Fig. 16) can greatly vary, although less so than stroke survivors. Such variation could reflect the open-ended nature of motor exploration. Compared to conventional motor tasks, such as reaching, motor exploration likely invites cognitive processes in addition to sensory and motor processes that could influence the size and shape of the distributions. Whereas movement capabilities described by other robotic metrics, such as movement error, smoothness or peak speed, can be characterized within a single visit, distributions of motor exploration data may require more time (Z. a Wright et al. 2014). In fact, we previously showed that stroke survivors motor exploration patterns considerably change even when training in the absence of interactive forces (Z.A. Wright et al. 2018).

In addition to identifying differences in impairment between stroke survivors or differences between populations, our results highlight the potential for using distribution analysis to track recovery. These results provide a concise procedure of data reduction to monitor and describe impairment level in this patient population. As some stroke subjects are quite unique from others, these methods provide a tool for distinguishing each patient with its own set of

impairment features, allows quantitative assessment for tracking recovery, and most importantly allows technology-based measures to be used to target specific motor deficits in stroke and possibly other populations as well. As rehabilitation continues to search for more cost-effective devices which can be used for at-home therapy, our results suggest that these important questions can be answered best with kinematic features of movement.

V. GENERAL DISCUSSION

The body of work described in this dissertation offers a powerful diagnostic tool for preparing customized therapies. We have presented an innovative approach to robot-therapy in which a statistical characterization of a stroke survivor's unique *motor deficits* directly informed the mathematical structure of customized training environments. Our findings from Chapter II provided evidence that such customizable forces could reshape the exploratory movement patterns of stroke survivors. This thesis also highlighted the importance of measuring and evaluating patient *involvement* during robotic therapy. Chapter III provided two key results: first, we found that the degree of patient recovery depends on how active they were during therapy (total mechanical work), and secondly, we identified the most important biomechanical features of such involvement (i.e. components of mechanical work). Chapter IV offered a prescription for therapy based on the best possible set of motor exploration variables (endpoint kinematics, joint kinematics and kinetics). Our results suggested that distributions of kinematic movement variables best predict clinical impairment level and best classify stroke survivors and healthy individuals. Taken together, this thesis offers a guide for using robotic devices for therapy design *and* evaluation which could impact the future of rehabilitation. We set forth a framework for using data-driven approaches to establish the link between diagnosis and personalized treatment. The techniques we employed can also be generalized to other evaluations and therapies where data-first approaches are becoming more favorable.

To best use distribution statistics for design of therapy an important question is whether to use the initial (first day) or the most recent (daily) distribution. One advantage of robotic technology is that they can reduce assessment time compared to traditional clinical assessments, reportedly up to 91% (Mostafavi et al. 2017). Our methods demonstrate how patient performance

on simple tasks can be characterized within a single visit and how their performance is less susceptible to day-to-day changes without intervention. However, distributions of free movement take time (i.e. amount of data) to fully characterize patient motor deficits and as our control group demonstrated, patient distributions considerably changed across days of training even without forces. The force group also showed gradual increases in velocity *coverage* across training days. Interestingly, we previously showed that more time is needed to characterize distributions of healthy individuals (Z. a Wright et al. 2014). This may not be surprising since healthy individuals are more capable of expressing a wider range of movements and greater coverage as a result. It is likely that *time to characterization* is related to our coverage metric. This would suggest that forces should be customized daily and like coverage, time to characterization could be a useful metric for tracking recovery of patients.

The relationship between motor exploration performance and functional ability is still unknown. Despite the range of therapies available, no strategy for upper limb rehabilitation has had meaningful impact on patient recovery. Task-oriented training has been the most common for stroke rehabilitation (Winstein et al. 2016). More recently, researchers have made the case for non-task-oriented approaches, such as "motor babbling", to increase training intensity and movement variety (Krakauer and Cortés 2018). Here we showed that motor exploration training with simple biofeedback could have therapeutic benefits. However, we were also motivated that practice over a broader range of movements would transfer to better goal-directed task performance which was not the case in our study. We did find that increases in velocity range during motor exploration training lead to faster task execution in "transfer" tasks (reaching and circle drawing). Such movements were not explicitly trained and hence demonstrated the potential for generalization (Conditt, Gandolfo, and Mussa-Ivaldi 1997). It is possible that any

new motor exploration capabilities require time and experience to incorporate into activities of daily living. Alternatively, combining motor exploration with task-oriented training could help reinforce new motor capabilities (Zeiler and Krakauer 2013). Nevertheless, the relationship between exploration behavior and functional ability is now better known as a consequence of the work here.

We speculate that incorporating information about how the movement distributions of stroke survivors differ from that of healthy distributions will enable the design of more effective force field treatment. For reaching movements, the healthy model is known - a straight-line with bell-shaped velocity (Morasso 1981). However, the structure of healthy movement distributions can vary greatly given that individuals may have preferred movement patterns or certain biomechanical constraints (see Fig. 16). Our current force field design is motivated by the idea that a wider range of velocities (both magnitude and directions) would provide clinical value. As such, a healthy model of a velocity distribution might have the appearance of a normal distribution. In Chapter IV we constructed a healthy model by averaging the distributions across participants. Our findings not only identified the movement distributions that best differentiate stroke survivors from healthy, but also provided preliminary evidence that a stroke survivor can be different from the average healthy individual in a variety of ways. Future work should consider alternative force field designs that encourage patient motions to look more like those of healthy individuals.

The clinical value of many robot-based assessments deserves further investigation. Our work is closely related to that of other researchers that employ statistical analysis to predict impairment level (Balasubramanian et al. 2012; Mostafavi et al. 2015b; Krebs et al. 2014; Wood et al. 2018; Bosecker et al. 2010; Zariffa et al. 2012) and recovery outcomes (Majeed, Awadalla,

and Patton, n.d.; Stinear 2017). While some studies have reported high correlations between robot metrics and clinical assessments, many prediction models fail when tested on new data. We were not particularly surprised that movement distributions did not act as good predictors of impairment level. However, our analysis did find that some movement variables outperformed others. Such comparisons could provide clues about underlying mechanisms of impairment or recovery and about which variables could be more informative for customizing therapy. For statistical analyses to have a major impact in rehabilitation the final product should be a model that can be shared with other researchers and tested with their respective experimental platforms.

As is often the case in rehabilitation, the lack of study participants was a key limitation in the research presented. While some studies using robot assessments have reported a remarkable number of stroke participants (Mostafavi et al. 2015b), they were typically in the acute stage of recovery (< 8 months post-stroke). The number of stroke survivors we statistically powered for our clinical study was comparable to other robotic intervention studies (G Kwakkel, Kollen, and Krebs 2007). However, our study design provided a limited dataset for our retrospective analyses that is typically recommended for building predictive models. Hence, we relied on machine learning methods (e.g. bootstrapping and leave-one-out cross validation) to not only identify relationships but to also understand levels of uncertainty in model predictions. These methods were used to gain as much information from the data we collected in our large-scale clinical study. Nevertheless, more patient data is certainly required for a true test of generalization of our predictive models.

Another key limitation is that our robotic intervention lasted only two weeks compared to 6+ weeks in other interventions. Nonetheless, patients in this study still showed signs of improvement in motor function. There is some debate among the clinical research community

regarding what is the optimal amount (i.e. dosage or duration) of therapy for patients or even how to define therapy "intensity" (Gert Kwakkel 2006). Our findings suggest that "intensity" may be better defined by a patient's energetic contributions since increasing total mechanical work resulted in better recovery outcomes. Free exploration training likely requires greater energetic output than repetitive task training and as the differences in total work between the force and control group demonstrated, large interactive forces greatly elevated the intensity of exploration training. Our results support other clinical findings that showed high intensity resistance training can have therapeutic benefits (e.g. increased muscle strength and in our case, increased velocity range), but should be combined with functional skill practice (Eng 2004; Patten et al. 2006).

While there were many specific decisions made in the design of this study, the potential of distribution analysis is quite broadly applicable. First, there was the question of what patients should do in order for us to gather large sets of movement data. In contrast to simple movement tasks, the task instructions for motor exploration are open-ended. Our goal was to have patients move as randomly as possible and we provided a feedback score related to randomness. Considering the cognitive processes that can influence patient choices to freely explore, this work invites a whole other field of research that could add to our understanding of the effects of different task instructions or different types of feedback on movement distributions.

Next, there was the question of which variables we should use to create movement distributions. We limited our distribution analysis to several pairs of two-dimensional movement variables. However, distributions can in principle include an unlimited number of dimensions. For example, a four-dimensional distribution of endpoint position and velocity might reveal whether a patient has particular velocity deficits in certain workspace areas. There is also the

possibility of combining different types of movement variables. For example, distributions of joint torque or power versus joint angle position or velocity could reveal post-stroke deficits in concentric and eccentric contractions (Hedlund et al. 2012). Continuing work in our lab also considers distributions of movement exploration performed in three dimensions. Of course, additional dimensions make distributions more difficult to visualize, but further work is needed to fully investigate all possible multivariate movement distributions and to determine which may be most clinically relevant and most informative for customized therapy.

Another question was how we should in the end make use of distributions in the context of therapy. Our customized force field design is one of many possible strategies. In addition to other applications for robotic therapy, the visual presentation of movement distributions singularly could be informative to a therapist for identifying a patient's particular motor deficits and making decisions about treatment. Finally, there was the question of how we should measure patient activity during therapy. Mechanical work provided a compact measure of effort level and the breakdown of work provided clinical relevance. Besides direct measures of how active patients were during therapy (e.g. muscle activity) which were not recorded in our study, one other possibility would be to apply distribution analysis to visualize power as a function of state-space variables (e.g. endpoint position or joint angles). All of these choices open the door to different or new opportunities to improve diagnosis of patient motor deficits and the design of customized therapy.

Distribution analysis has already shown to have an impact in a wide range of therapeutic applications beyond upper extremity robotic therapy. For example, recent research has investigated distributions of accelerometer data to identify differences between affected and non-affected limb use in stroke survivors (Lang et al. 2017). Another study recorded exploratory

shoulder movements using inertial measurement systems in order to customize the control of a wheelchair for individuals with tetraplegia (Thorp et al. 2016). Beyond kinematic and kinetic motion variables, our statistical approach of characterizing impairment could be applied to other signals relating to human movement behavior. For example, distributions of neurophysiological signals, including electromyography (EMG) and electrocorticography (ECoG), could uncover the underlying causes of impairment. Recent work in our lab investigated distributions of muscle activity to identify differences between stroke survivors and healthy individuals in muscle activation patterns during isometric exploration (Aggarwal et al. 2019). Distribution analysis has also been used to characterize residual muscle activation in lower limb amputees for optimizing control of a prostheses (S. Huang and Huang 2019). The approach in the current study provides a framework for how to better exploit information within distribution analysis to explain phenomena of clinical interest. There are also areas of potential application that represent broad impact. Statistical profiling of large datasets is an emerging trend in healthcare and other industries including sports, banking, finance, insurance, manufacturing, agriculture, education, digital media and entertainment, weather, transportation, etc. Our analysis tool can be more broadly applied to generate detailed distributions of any data.

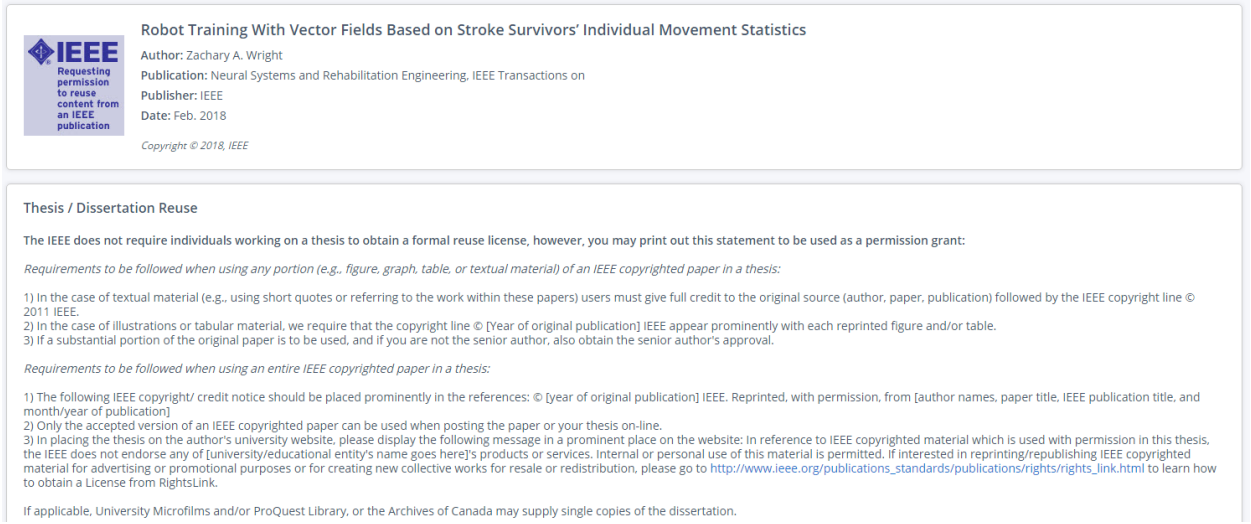
This statistical approach of linking diagnosis to treatment could also serve athletic skill training. The challenge in characterizing complex motions, such as throwing a baseball, is often determining the best signals to measure and where to measure them from. For highly skilled athletes such as pitchers, it is often not a question of whether they are generally capable, but rather the degree of variation expressed in how they are in performing the intended throw. Recent work in our lab showed how statistical profiles of movement errors can describe variability in simple reaching (Fisher et al. 2014). Complex motions, for example those in the

pitching motion, can be decomposed into a set of discrete movements that are combined sequentially (known as the “kinematic sequence”). Distribution analysis could describe variability in the magnitude or timing of each discrete movement (Scarborough et al. 2018). As athletes are constantly exploring new strategies to enhance training and performance, movement distributions could help identify when or which part of the motion is statistically more likely to break down.

The rehabilitation clinic of tomorrow could be guided by the work presented in this thesis. Robots offer endless opportunities to automate the rehabilitation process from diagnosis to treatment to tracking recovery of patients. Such a clinic would make therapy easy to access thus affording patients the possibility of maximizing their time spent towards achieving their fullest recovery. Patients would be able to interact with any available rehabilitation robot. These robots could use a suite of assessments to characterize patient movement, recommend an individualized therapy plan and monitor patient progress. As observational technology becomes more pervasive, these tools may be easy to implement and cost effective to automate. In the future a computer may say, “I have been watching you and I notice that statistically, you are having trouble with your right elbow extension torque whenever you raise your shoulder. Let me recommend a therapy for you.” Essentially, the rehab clinic of the future can be best described as an “automated rehabilitation robot gym”.

APPENDICES

Appendix A: Copyright Policies



IEEE
Requesting permission to reuse content from an IEEE publication

Robot Training With Vector Fields Based on Stroke Survivors' Individual Movement Statistics

Author: Zachary A. Wright
Publication: Neural Systems and Rehabilitation Engineering, IEEE Transactions on
Publisher: IEEE
Date: Feb. 2018

Copyright © 2018, IEEE

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

CITED LITERATURE

Aggarwal, A., Zachary A. Wright, Felix C Huang, and J.L. Patton. 2019. “Post-Stroke Motor Deficits Are Most Evident at Frequencies near 125 Hz in EMG Multivariate Probability Distributions.” *IEEE Engineering in Medicine and Biology Society,* no. July.

Aalverdashvili, M, Afra Foroud, Diana H Lim, and Ian Q Whishaw. 2008. “‘Learned Baduse’ Limits Recovery of Skilled Reaching for Food after Forelimb Motor Cortex Stroke in Rats: A New Analysis of the Effect of Gestures on Success.” *Behavioural Brain Research* 188 (2): 281–90. doi:10.1016/j.bbr.2007.11.007.

Alessandro, Cristiano, Ioannis Delis, Francesco Nori, Stefano Panzeri, and Bastien Berret. 2013. “Muscle Synergies in Neuroscience and Robotics: From Input-Space to Task-Space Perspectives.” *Frontiers in Computational Neuroscience* 7 (April): 1–16. doi:10.3389/fncom.2013.00043.

Allen, David M. 1971. “Mean Square Error of Prediction as a Criterion for Selecting Variables.” *Technometrics* 13 (3): 469. doi:10.2307/1267161.

Andrews, J G. 1983. “Biomechanical Measures of Muscular Effort.” *Medicine and Science in Sports and Exercise* 15 (3): 199–207. http://www.ncbi.nlm.nih.gov/pubmed/6621306.

Arya, Kamal Narayan, Shanta Pandian, Rajesh Verma, and R. K. Garg. 2011. “Movement Therapy Induced Neural Reorganization and Motor Recovery in Stroke: A Review.” *Journal of Bodywork and Movement Therapies* 15 (4). Elsevier Ltd: 528–37. doi:10.1016/j.jbmt.2011.01.023.

Ashworth, B. 1964. “Preliminary Trial of Carisoprodol in Multiple Sclerosis.” *The Practitioner* 192 (April): 540–42. http://www.ncbi.nlm.nih.gov/pubmed/14143329.

Balasubramanian, Sivakumar, Roberto Colombo, Irma Sterpi, Vittorio Sanguineti, and Etienne Burdet. 2012. “Robotic Assessment of Upper Limb Motor Function after Stroke.” *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists* 91 (11 Suppl 3): S255–69. doi:10.1097/PHM.0b013e31826bcdc1.

Beekhuis, J Houdijn, Ard J Westerveld, Herman van der Kooij, and Arno H A Stienen. 2013. “Design of a Self-Aligning 3-DOF Actuated Exoskeleton for Diagnosis and Training of Wrist and Forearm after Stroke.” *IEEE ... International Conference on Rehabilitation Robotics : [Proceedings]* 2013 (June): 6650357. doi:10.1109/ICORR.2013.6650357.

Beer, R F, J D Given, and J P Dewald. 1999. “Task-Dependent Weakness at the Elbow in Patients with Hemiparesis.” *Archives of Physical Medicine and Rehabilitation* 80 (7): 766–72. http://www.ncbi.nlm.nih.gov/pubmed/10414760.

Bhushan, N, and R Shadmehr. 1999. “Computational Nature of Human Adaptive Control during Learning of Reaching Movements in Force Fields.” *Biological Cybernetics* 81 (1): 39–60. doi:10.1007/s004220050543.

Blank, Amy A., James A. French, Ali Utku Pehlivan, and Marcia K. O’Malley. 2014. “Current Trends in Robot-Assisted Upper-Limb Stroke Rehabilitation: Promoting Patient Engagement in Therapy.” *Current Physical Medicine and Rehabilitation Reports* 2 (3): 184–95. doi:10.1007/s40141-014-0056-z.

Bohannon, R W, and M B Smith. 1987. “Assessment of Strength Deficits in Eight Paretic Upper Extremity Muscle Groups of Stroke Patients with Hemiplegia.” *Physical Therapy* 67 (4):

522–25.

Bosecker, Caitlyn, Laura Dipietro, Bruce Volpe, and Hermano Igo Krebs. 2010. “Kinematic Robot-Based Evaluation Scales and Clinical Counterparts to Measure Upper Limb Motor Performance in Patients with Chronic Stroke.” *Neurorehabilitation and Neural Repair* 24 (1): 62–69. doi:10.1177/1545968309343214.

Chib, V S, J L Patton, K M Lynch, and F A Mussa-Ivaldi. 2005. “The Effect of Stiffness and Curvature on the Haptic Identification of Surfaces.” *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, IEEE-WHC 2005*, 126–31. doi:10.1109/WHC.2005.131.

Colombo, Roberto, Irma Sterpi, Alessandra Mazzone, Carmen Delconte, and Fabrizio Pisano. 2012. “Taking a Lesson from Patients’ Recovery Strategies to Optimize Training during Robot-Aided Rehabilitation.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20 (3): 276–85. doi:10.1109/TNSRE.2012.2195679.

Conditt, M A, F Gandolfo, and F A Mussa-Ivaldi. 1997. “The Motor System Does Not Learn the Dynamics of the Arm by Rote Memorization of Past Experience.” *Journal of Neurophysiology* 78 (1): 554–60. http://www.ncbi.nlm.nih.gov/pubmed/9242306.

Dewald, J P, and R F Beer. 2001. “Abnormal Joint Torque Patterns in the Paretic Upper Limb of Subjects with Hemiparesis.” *Muscle & Nerve* 24 (2): 273–83. http://www.ncbi.nlm.nih.gov/pubmed/11180211.

Dewald, J P, P S Pope, J D Given, T S Buchanan, and W Z Rymer. 1995. “Abnormal Muscle Coactivation Patterns during Isometric Torque Generation at the Elbow and Shoulder in Hemiparetic Subjects.” *Brain : A Journal of Neurology* 118 (Pt 2 (April): 495–510. http://www.ncbi.nlm.nih.gov/pubmed/7735890.

Diedrichsen, Jörn, Olivier White, Darren Newman, and Níall Lally. 2010. “Use-Dependent and Error-Based Learning of Motor Behaviors.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 30 (15): 5159–66. doi:10.1523/JNEUROSCI.5406-09.2010.

Dietz, Volker, and Thomas Sinkjaer. 2007. “Spastic Movement Disorder: Impaired Reflex Function and Altered Muscle Mechanics.” *The Lancet. Neurology* 6 (8): 725–33. doi:10.1016/S1474-4422(07)70193-X.

Dipietro, Laura, Mark Ferraro, Jerome Joseph Palazzolo, Hermano Igo Krebs, Bruce T. Volpe, and Neville Hogan. 2005. “Customized Interactive Robotic Treatment for Stroke: EMG-Triggered Therapy.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 13 (3): 325–34. doi:10.1109/TNSRE.2005.850423.

Doke, J., and A. D. Kuo. 2007a. “Energetic Cost of Producing Cyclic Muscle Force, rather than Work, to Swing the Human Leg.” *Journal of Experimental Biology* 210 (13): 2390–98. doi:10.1242/jeb.02782.

Doke, J., and A. D. Kuo. 2007b. “Energetic Cost of Producing Cyclic Muscle Force, rather than Work, to Swing the Human Leg.” *Journal of Experimental Biology* 210 (13): 2390–98. doi:10.1242/jeb.02782.

Donchin, O., K. Rabe, J. Diedrichsen, N. Lally, B. Schoch, E. R. Gizewski, and D. Timmann. 2012. “Cerebellar Regions Involved in Adaptation to Force Field and Visuomotor Perturbation.” *Journal of Neurophysiology* 107 (1): 134–47. doi:10.1152/jn.00007.2011.

Ellis, Michael D., Carolina Carmona, Justin Drogos, Stuart Traxel, and Julius P.A. Dewald.

2016. “Progressive Abduction Loading Therapy Targeting Flexion Synergy to Regain Reaching Function in Chronic Stroke: Preliminary Results from an RCT.” In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016:5837–40. IEEE. doi:10.1109/EMBC.2016.7592055.

Ellis, Michael D., Bradley G. Holubar, Ana Maria Acosta, Randall F. Beer, and Julius P. A. Dewald. 2005. “Modifiability of Abnormal Isometric Elbow and Shoulder Joint Torque Coupling after Stroke.” *Muscle & Nerve* 32 (2): 170–78. doi:10.1002/mus.20343.

Ellis, Michael D., Yiyun Lan, Jun Yao, and Julius P. A. Dewald. 2016. “Robotic Quantification of Upper Extremity Loss of Independent Joint Control or Flexion Synergy in Individuals with Hemiparetic Stroke: A Review of Paradigms Addressing the Effects of Shoulder Abduction Loading.” *Journal of NeuroEngineering and Rehabilitation* 13 (1). Journal of NeuroEngineering and Rehabilitation: 95. doi:10.1186/s12984-016-0203-0.

Emken, Jeremy L, Raul Benitez, and David J Reinkensmeyer. 2007. “Human-Robot Cooperative Movement Training: Learning a Novel Sensory Motor Transformation during Walking with Robotic Assistance-as-Needed.” *Journal of NeuroEngineering and Rehabilitation* 4 (1): 8. doi:10.1186/1743-0003-4-8.

Eng, Janice J. 2004. “Strength Training in Individuals with Stroke.” *Physiotherapy Canada. Physiotherapie Canada* 56 (4): 189–201. http://www.ncbi.nlm.nih.gov/pubmed/23255839.

Eng, Janice J., Melanie J. Lomaglio, and Donna L. Macintyre. 2009. “Muscle Torque Preservation and Physical Activity in Individuals with Stroke.” *Medicine and Science in Sports and Exercise* 41 (7): 1353–60. doi:10.1249/MSS.0b013e31819aaad1.

Enzinger, Christian, Helen Dawes, Heidi Johansen-Berg, Derick Wade, Marko Bogdanovic, Jonathan Collett, Claire Guy, et al. 2009. “Brain Activity Changes Associated with Treadmill Training after Stroke.” *Stroke; a Journal of Cerebral Circulation* 40 (7): 2460–67. doi:10.1161/STROKEAHA.109.550053.

Farris, Dominic, Austin Hampton, Michael D Lewek, and Gregory S Sawicki. 2015. “Revisiting the Mechanics and Energetics of Walking in Individuals with Chronic Hemiparesis Following Stroke: From Individual Limbs to Lower Limb Joints.” *Journal of NeuroEngineering and Rehabilitation* 12 (1): 24. doi:10.1186/s12984-015-0012-x.

Ferris, Daniel P, Gregory S Sawicki, and Monica A Daley. 2007. “A PHYSIOLOGIST’S PERSPECTIVE ON ROBOTIC EXOSKELETONS FOR HUMAN LOCOMOTION.” *International Journal of HR : Humanoid Robotics* 4 (3). NIH Public Access: 507–28. doi:10.1142/S0219843607001138.

Fisher, M.E., F.C. Huang, Z.A. Wright, and J.L. Patton. 2014. “Distributions in the Error Space: Goal-Directed Movements Described in Time and State-Space Representations.” In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2014*. doi:10.1109/EMBC.2014.6945227.

Flash, T, and N Hogan. 1985. “The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 5 (7): 1688–1703. doi:4020415.

Fugl-Meyer, A R, L Jääskö, I Leyman, S Olsson, and S Steglind. 1975. “The Post-Stroke Hemiplegic Patient. 1. a Method for Evaluation of Physical Performance.” *Scandinavian Journal of Rehabilitation Medicine* 7 (1): 13–31. http://www.ncbi.nlm.nih.gov/pubmed/1135616.

Gandolfo, F, F a Mussa-Ivaldi, and E Bizzi. 1996. “Motor Learning by Field Approximation.” *Proceedings of the National Academy of Sciences of the United States of America* 93 (9): 3843–46. doi:10.1073/pnas.93.9.3843.

Gao, Fan, Yupeng Ren, Elliot J Roth, Richard Harvey, and Li-Qun Zhang. 2011. “Effects of Repeated Ankle Stretching on Calf Muscle-Tendon and Ankle Biomechanical Properties in Stroke Survivors.” *Clinical Biomechanics (Bristol, Avon)* 26 (5): 516–22. doi:10.1016/j.clinbiomech.2010.12.003.

Geladi, Paul, and Bruce R Kowalski. 1986. “PARTIAL LEAST-SQUARES REGRESSION: A TUTORIAL.” *Analytica Chimica Acta*. Vol. 186. Elsevier Science Publishers B.V.

Germanotta, Marco, Gessica Vasco, Maurizio Petrarca, Stefano Rossi, Sacha Carniel, Enrico Bertini, Paolo Cappa, and Enrico Castelli. 2015. “Robotic and Clinical Evaluation of Upper Limb Motor Performance in Patients with Friedreich’s Ataxia: An Observational Study.” *Journal of Neuroengineering and Rehabilitation* 12 (January): 41. doi:10.1186/s12984-015-0032-6.

Gladstone, David J, Cynthia J Danells, and Sandra E Black. 2002. “The Fugl-Meyer Assessment of Motor Recovery after Stroke: A Critical Review of Its Measurement Properties.” *Neurorehabilitation and Neural Repair* 16 (3): 232–40. doi:10.1177/154596802401105171.

Gowland, C., P. Stratford, M. Ward, J. Moreland, W. Torresin, S. Van Hullenaar, J. Sanford, S. Barreca, B. Vanspall, and N. Plews. 1993. “Measuring Physical Impairment and Disability with the Chedoke-McMaster Stroke Assessment.” *Stroke* 24 (1): 58–63. doi:10.1161/01.STR.24.1.58.

Guidali, Marco, Urs Keller, Verena Klamroth-marganska, Tobias Nef, and Robert Riener. 2016. “Journal of Rehabilitation Research & Development (JRRD) Estimating the Patient ’ S Contribution during Robot- Assisted Therapy,” 1–14.

Hammerbeck, Ulrike, Nada Yousif, Richard Greenwood, John C Rothwell, and Jörn Diedrichsen. 2014. “Movement Speed Is Biased by Prior Experience.” *Journal of Neurophysiology* 111 (1): 128–34. doi:10.1152/jn.00522.2013.

Hatze, H. 1980. “A Mathematical Model for the Computational Determination of Parameter Values of Anthropomorphic Segments.” *Journal of Biomechanics* 13 (10): 833–43. doi:10.1016/0021-9290(80)90171-2.

Hedlund, Mattias, Peter Sojka, Ronnie Lundström, and Britta Lindström. 2012. “Torque-Angle Relationship Are Better Preserved during Eccentric Compared to Concentric Contractions in Patients with Stroke.” *Isokinetics and Exercise Science* 20 (2). IOS Press: 129–40. doi:10.3233/IES-2012-0455.

Hogan, Neville, Hermano I Krebs, Brandon Rohrer, Jerome J Palazzolo, Laura Dipietro, Susan E Fasoli, Joel Stein, et al. 2006. “Assistance of Motor Recovery” 43 (5): 605–18. doi:10.1682/JRRD.2005.06.0103.

Hornby, T George, Donielle D Campbell, Jennifer H Kahn, Tobey Demott, Jennifer L Moore, and Heidi R Roth. 2008. “Enhanced Gait-Related Improvements after Therapist- versus Robotic-Assisted Locomotor Training in Subjects with Chronic Stroke: A Randomized Controlled Study.” *Stroke; a Journal of Cerebral Circulation* 39 (6): 1786–92. doi:10.1161/STROKEAHA.107.504779.

Huang, Felix C, and James L Patton. 2011. “Evaluation of Negative Viscosity as Upper Extremity Training for Stroke Survivors.” *IEEE ... International Conference on*

Rehabilitation Robotics : [Proceedings] 2011 (January): 5975514. doi:10.1109/ICORR.2011.5975514.

Huang, Felix C, and James L Patton. 2013a. "Augmented Dynamics and Motor Exploration as Training for Stroke." *IEEE Transactions on Bio-Medical Engineering* 60 (3): 838–44. doi:10.1109/TBME.2012.2192116.

Huang, Felix C, and James L Patton. 2013b. "Individual Patterns of Motor Deficits Evident in Movement Distribution Analysis." *IEEE ... International Conference on Rehabilitation Robotics : [Proceedings]* 2013: 1–6. doi:10.1109/ICORR.2013.6650430.

Huang, Felix C, and James L Patton. 2016. "Movement Distributions of Stroke Survivors Exhibit Distinct Patterns That Evolve with Training." *Journal of NeuroEngineering and Rehabilitation*. Journal of NeuroEngineering and Rehabilitation, 1–13. doi:10.1186/s12984-016-0132-y.

Huang, Felix C, James L Patton, and Ferdinando a Mussa-Ivaldi. 2010. "Manual Skill Generalization Enhanced by Negative Viscosity." *Journal of Neurophysiology* 104 (4): 2008–19. doi:10.1152/jn.00433.2009.

Huang, Helen J, Rodger Kram, and Alaa A Ahmed. 2012. "Reduction of Metabolic Cost during Motor Learning of Arm Reaching Dynamics" 32 (6): 2182–90. doi:10.1523/JNEUROSCI.4003-11.2012.

Huang, Stephanie, and He Huang. 2019. "Voluntary Control of Residual Antagonistic Muscles in Transtibial Amputees : Reciprocal Activation , Coactivation , and Implications Lower Limb Prostheses." *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27 (1). IEEE: 85–95. doi:10.1109/TNSRE.2018.2885641.

Israel, J. F, D. D Campbell, J. H Kahn, and T G. Hornby. 2006. "Metabolic Costs and Muscle Activity Patterns During Robotic- and Therapist-Assisted Treadmill Walking in Individuals With Incomplete Spinal Cord Injury." *Physical Therapy* 86 (11): 1466–78. doi:10.2522/ptj.20050266.

Jolliffe, I T. n.d. "Principal Component Analysis, Second Edition."

Kafri, Michal, Mary Jane Myslinski, Venkata K. Gade, and Judith E. Deutsch. 2014. "High Metabolic Cost and Low Energy Expenditure for Typical Motor Activities among Individuals in the Chronic Phase after Stroke." *Journal of Neurologic Physical Therapy* 38 (4): 226–32. doi:10.1097/NPT.0000000000000053.

Kahn, L E, W Z Rymer, and D J Reinkensmeyer. 2004. "Adaptive Assistance for Guided Force Training in Chronic Stroke." *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 4: 2722–25. doi:10.1109/IEMBS.2004.1403780.

Kawato, Mitsuo. 1999. "Internal Models for Motor Control and Trajectory Planning." *Current Opinion in Neurobiology*, no. 6: 718–27. doi:10.1016/S0959-4388(99)00028-8.

"Keele1968.pdf." 2016. Accessed May 10. http://grants.hhp.coe.uh.edu/clayne/HistoryofMC/Keele1968.pdf.

Kim, Bokkyu, and Carolee Winstein. 2016. "Can Neurological Biomarkers of Brain Impairment Be Used to Predict Poststroke Motor Recovery ? A Systematic Review." doi:10.1177/1545968316662708.

Krakauer, John W., and Juan Camilo Cortés. 2018. "A Non-Task-Oriented Approach Based on

High-Dose Playful Movement Exploration for Rehabilitation of the Upper Limb Early after Stroke: A Proposal." *NeuroRehabilitation* 43 (1). IOS Press: 31–40. doi:10.3233/NRE-172411.

Krebs, H I, M L Aisen, B T Volpe, and N Hogan. 1999. "Quantization of Continuous Arm Movements in Humans with Brain Injury." *Proceedings of the National Academy of Sciences of the United States of America* 96 (8): 4645–49. doi:10.1073/pnas.96.8.4645.

Krebs, H. I., M. Krams, D. K. Agrafiotis, A. DiBernardo, J. C. Chavez, G. S. Littman, E. Yang, et al. 2014. "Robotic Measurement of Arm Movements After Stroke Establishes Biomarkers of Motor Recovery." *Stroke* 45 (1): 200–204. doi:10.1161/STROKEAHA.113.002296.

Kwakkel, G, B J Kollen, and H I Krebs. 2007. "Effects of Robot-Assisted Therapy on Upper Limb Recovery after Stroke: A Systematic Review.," 1–11. doi:10.1177/1545968307305457.

Kwakkel, Gert. 2006. "Impact of Intensity of Practice after Stroke: Issues for Consideration." *Disability and Rehabilitation* 28 (13-14): 823–30. doi:10.1080/09638280500534861.

Lang, Catherine E, Joanne M Wagner, Alexander W Dromerick, and Dorothy F Edwards. 2006. "Measurement of Upper-Extremity Function Early after Stroke: Properties of the Action Research Arm Test." *Archives of Physical Medicine and Rehabilitation* 87 (12): 1605–10. doi:10.1016/j.apmr.2006.09.003.

Lang, Catherine E., Kimberly J. Waddell, Joseph W. Klaesner, and Marghuretta D. Bland. 2017. "A Method for Quantifying Upper Limb Performance in Daily Life Using Accelerometers." *Journal of Visualized Experiments* 2017 (122). Journal of Visualized Experiments. doi:10.3791/55673.

Lenarcic, J, and A Umek. 1994. "Simple Model of Human Arm Reachable Workspace." *Systems, Man and Cybernetics, IEEE Transactions on* 24 (8): 1239–46. doi:10.1109/21.299704.

Liu, Jie, Dali Xu, Yupeng Ren, and Li-Qun Zhang. 2011. "Evaluations of Neuromuscular Dynamics of Hyperactive Reflexes Poststroke." *Journal of Rehabilitation Research and Development* 48 (5): 577–86. http://www.ncbi.nlm.nih.gov/pubmed/21674407.

Lodha, Neha, Sagar K. Naik, Stephen A. Coombes, and James H. Cauraugh. 2010. "Force Control and Degree of Motor Impairments in Chronic Stroke." *Clinical Neurophysiology* 121 (11). International Federation of Clinical Neurophysiology: 1952–61. doi:10.1016/j.clinph.2010.04.005.

Lohse, Keith R., Catherine E. Lang, and Lara A. Boyd. 2014. "Is More Better? Using Metadata to Explore Dose–Response Relationships in Stroke Rehabilitation." *Stroke* 45 (7): 2053–58. doi:10.1161/STROKEAHA.114.004695.

Lotze, M. 2003. "Motor Learning Elicited by Voluntary Drive." *Brain* 126 (4): 866–72. doi:10.1093/brain/awg079.

Lum, Peter S, Charles G Burgar, and Peggy C Shor. 2003. "Evidence for Strength Imbalances as a Significant Contributor to Abnormal Synergies in Hemiparetic Subjects." *Muscle & Nerve* 27 (2): 211–21. doi:10.1002/mus.10305.

Lum, Peter S., Charles G. Burgar, Peggy C. Shor, Matra Majmundar, and Machiel Van der Loos. 2002. "Robot-Assisted Movement Training Compared with Conventional Therapy Techniques for the Rehabilitation of Upper-Limb Motor Function after Stroke." *Archives of*

Physical Medicine and Rehabilitation 83 (7): 952–59. doi:10.1053/apmr.2001.33101.

Lyle, R C. 1981. “A Performance Test for Assessment of Upper Limb Function in Physical Rehabilitation Treatment and Research.” *International Journal of Rehabilitation Research. Internationale Zeitschrift Für Rehabilitationsforschung. Revue Internationale de Recherches de Réadaptation* 4 (4): 483–92. http://www.ncbi.nlm.nih.gov/pubmed/7333761.

Majeed, Yazan Abdel, Farnaz Abdollahi, Saria Awadalla, and James Patton. 2015. “Multivariate Outcomes in a Three Week Bimanual Self-Telerehabilitation with Error Augmentation Post-Stroke.” In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2015-Novem:1425–31. doi:10.1109/EMBC.2015.7318637.

Majeed, Yazan Abdel, Saria S Awadalla, and James L Patton. n.d. “Regression Techniques Employing Feature Selection to Predict Clinical Outcomes in Stroke,” 1–15.

Marchal-Crespo, Laura, and David J Reinkensmeyer. 2009. “Review of Control Strategies for Robotic Movement Training after Neurologic Injury.” *Journal of NeuroEngineering and Rehabilitation* 6 (1): 20. doi:10.1186/1743-0003-6-20.

Mark, V W, E Taub, and D M Morris. 2006. “Neuroplasticity and Constraint-Induced Movement Therapy.” *Europa Medicophysica* 42 (3): 269–84.

Matthews, B. W. 1975. “Comparison of the Predicted and Observed Secondary Structure of T4 Phage Lysozyme.” *BBA - Protein Structure* 405 (2): 442–51. doi:10.1016/0005-2795(75)90109-9.

Mazzoleni, S., L. Puzzolante, L. Zollo, P. Dario, and F. Posteraro. 2014. “Mechanisms of Motor Recovery in Chronic and Subacute Stroke Patients Following a Robot-Aided Training.” *IEEE Transactions on Haptics* 7 (2): 175–80. doi:10.1109/TOH.2013.73.

McCrea, P H, Eng, J J, and Hodgson, A J. 2002. “Biomechanics of Reaching: Clinical Implications for Individuals with Acquired Brain Injury.” *Disability and Rehabilitation* 24 (10): 534–41. doi:10.1080/09638280110115393.

Melendez-Calderon, Alejandro, Michael Tan, Moria Fisher Bittmann, Etienne Burdet, and James L Patton. 2017. “Transfer of Dynamic Motor Skills Acquired during Isometric Training to Free Motion.” *Journal of Neurophysiology* 118 (1): 219–33. doi:10.1152/jn.00614.2016.

Morasso, P. 1981. “Spatial Control of Arm Movements.” *Experimental Brain Research* 42 (2). Springer-Verlag: 223–27. doi:10.1007/BF00236911.

Mostafavi, Sayyed Mostafa, Parvin Mousavi, Sean P. Dukelow, and Stephen H. Scott. 2015a. “Robot-Based Assessment of Motor and Proprioceptive Function Identifies Biomarkers for Prediction of Functional Independence Measures.” *Journal of NeuroEngineering and Rehabilitation* 12 (1). Journal of NeuroEngineering and Rehabilitation: 1–12. doi:10.1186/s12984-015-0104-7.

Mostafavi, Sayyed Mostafa, Parvin Mousavi, Sean P. Dukelow, and Stephen H. Scott. 2015b. “Robot-Based Assessment of Motor and Proprioceptive Function Identifies Biomarkers for Prediction of Functional Independence Measures.” *Journal of NeuroEngineering and Rehabilitation* 12 (1). Journal of NeuroEngineering and Rehabilitation. doi:10.1186/s12984-015-0104-7.

Mostafavi, Sayyed Mostafa, Stephen Scott, Sean Dukelow, and Parvin Mousavi. 2017. “Reduction of Assessment Time for Stroke-Related Impairments Using Robotic Evaluation.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25 (7).

Institute of Electrical and Electronics Engineers Inc.: 945–55. doi:10.1109/TNSRE.2017.2669986.

Neckel, Nathan D., Natalie Blonien, Diane Nichols, and Joseph Hidler. 2008. "Abnormal Joint Torque Patterns Exhibited by Chronic Stroke Subjects While Walking with a Prescribed Physiological Gait Pattern." *Journal of NeuroEngineering and Rehabilitation* 5: 1–13. doi:10.1186/1743-0003-5-19.

Olney, S J, M P Griffin, T N Monga, and I D McBride. 1991. "Work and Power in Gait of Stroke Patients." *Archives of Physical Medicine and Rehabilitation* 72 (5): 309–14. http://www.ncbi.nlm.nih.gov/pubmed/2009047.

Page, Stephen J, George D Fulk, and Pierce Boyne. 2012. "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): 791–98. doi:10.2522/ptj.20110009.

Pain, Liza M., Ross Baker, Denyse Richardson, and Anne M. R. Agur. 2015. "Effect of Trunk-Restraint Training on Function and Compensatory Trunk, Shoulder and Elbow Patterns during Post-Stroke Reach: A Systematic Review." *Disability and Rehabilitation* 37 (7): 553–62. doi:10.3109/09638288.2014.932450.

Patten, Carolynn, Jody Dozono, Stephen G. Schmidt, Mary E. Jue, and Peter S. Lum. 2006. "Combined Functional Task Practice and Dynamic High Intensity Resistance Training Promotes Recovery of Upper-Extremity Motor Function in Post-Stroke Hemiparesis: A Case Study." *Journal of Neurologic Physical Therapy* 30 (3): 99–115. doi:10.1097/01.NPT.0000281945.55816.e1.

Patton, James L. 2012. "NIH Public Access" 41 (8): 1709–14. doi:10.1161/STROKEAHA.110.586917.Preswing.

Patton, James L., Mary Ellen Stoykov, Mark Kovic, and Ferdinando a. Mussa-Ivaldi. 2006. "Evaluation of Robotic Training Forces That Either Enhance or Reduce Error in Chronic Hemiparetic Stroke Survivors." *Experimental Brain Research* 168 (3): 368–83. doi:10.1007/s00221-005-0097-8.

Prange, Gerdienke B, Michiel J a Jannink, Catharina G M Groothuis-Oudshoorn, Hermie J Hermens, and Maarten J Ijzerman. 2006. "Systematic Review of the Effect of Robot-Aided Therapy on Recovery of the Hemiparetic Arm after Stroke." *Journal of Rehabilitation Research and Development* 43 (2): 171–84. doi:10.1682/JRRD.2005.04.0076.

Rabadi, Meheroz H, and Freny M Rabadi. 2006. "Comparison of the Action Research Arm Test and the Fugl-Meyer Assessment as Measures of Upper-Extremity Motor Weakness after Stroke." *Archives of Physical Medicine and Rehabilitation* 87 (7): 962–66. doi:10.1016/j.apmr.2006.02.036.

Rabe, K, O Livne, E R Gizewski, V Aurich, A Beck, D Timmann, and O Donchin. 2009. "Adaptation to Visuomotor Rotation and Force Field Perturbation Is Correlated to Different Brain Areas in Patients with Cerebellar Degeneration." *Journal of Neurophysiology* 101 (4): 1961–71. doi:10.1152/jn.91069.2008.

Reinkensmeyer, David J, Emmanuel Guigon, and Marc A Maier. 2012. "A Computational Model of Use-Dependent Motor Recovery Following a Stroke: Optimizing Corticospinal Activations via Reinforcement Learning Can Explain Residual Capacity and Other Strength Recovery Dynamics." *Neural Networks : The Official Journal of the International Neural Network Society* 29-30 (May): 60–69. doi:10.1016/j.neunet.2012.02.002.

Sanford, J, J Moreland, L R Swanson, P W Stratford, and C Gowland. 1993. “Reliability of the Fugl-Meyer Assessment for Testing Motor Performance in Patients Following Stroke.” *Physical Therapy* 73 (7): 447–54. doi:10.1177/1545968304269210.

Sangani, Samir G., Andrew J. Starsky, John R. Mcguire, and Brian D. Schmit. 2007. “Multijoint Reflexes of the Stroke Arm: Neural Coupling of the Elbow and Shoulder.” *Muscle & Nerve* 36 (5): 694–703. doi:10.1002/mus.20852.

Scarborough, Donna Moxley, Ashley J Bassett, Lucas W Mayer, and Eric M Berkson. 2018. “Kinematic Sequence Patterns in the Overhead Baseball Pitch.” *Sports Biomechanics*, September, 1–18. doi:10.1080/14763141.2018.1503321.

Shadmehr, R, and F A Mussa-Ivaldi. 1994. “Adaptive Representation of Dynamics during Learning of a Motor Task.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 14 (5 Pt 2): 3208–24. http://www.ncbi.nlm.nih.gov/pubmed/8182467.

Stinear, Cathy M. 2017. “Prediction of Motor Recovery after Stroke: Advances in Biomarkers.” *The Lancet Neurology* 16 (10). Elsevier Ltd: 826–36. doi:10.1016/S1474-4422(17)30283-1.

Taub, E, G Uswatte, V W Mark, and D M M Morris. 2006. “The Learned Nonuse Phenomenon: Implications for Rehabilitation.” *Europa Medicophysica* 42 (3): 241–56. http://www.ncbi.nlm.nih.gov/pubmed/17039223.

Taub, Edward, Gitendra Uswatte, and Rama Pidikiti. 1999. “Constraint-Induced Movement Therapy: A New Family of Techniques with Broad Application to Physical Rehabilitation - A Clinical Review.” *Journal of Rehabilitation Research and Development* 36: 237–51. http://www.ncbi.nlm.nih.gov/pubmed/10659807.

Thaut, M. H., G. P. Kenyon, C. P. Hurt, G. C. McIntosh, and V. Hoemberg. 2002. “Kinematic Optimization of Spatiotemporal Patterns in Paretic Arm Training with Stroke Patients.” *Neuropsychologia* 40 (7): 1073–81. doi:10.1016/S0028-3932(01)00141-5.

Thorp, Elias B, Farnaz Abdollahi, David Chen, Ali Farshchiansadegh, Mei-Hua Lee, Jessica P Pedersen, Camilla Pierella, Elliot J Roth, Ismael Seanez Gonzalez, and Ferdinando A Mussa-Ivaldi. 2016. “Upper Body-Based Power Wheelchair Control Interface for Individuals With Tetraplegia.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society* 24 (2): 249–60. doi:10.1109/TNSRE.2015.2439240.

Tibshirani, Robert. 1996. “Regression Shrinkage and Selection via the Lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)*. WileyRoyal Statistical Society. doi:10.2307/2346178.

Van Der Lee, J. H., R. C. Wagenaar, G. J. Lankhorst, T. W. Vogelaar, W. L. Deville, and L. M. Bouter. 1999. “Forced Use of the Upper Extremity in Chronic Stroke Patients : Results From a Single-Blind Randomized Clinical Trial.” *Stroke* 30 (11): 2369–75. doi:10.1161/01.STR.30.11.2369.

Van Der Lee, Johanna H., Heleen Beckerman, Gustaaf J. Lankhorst, and Lex M. Bouter. 2001. “The Responsiveness of the Action Research Arm Test and the Fugl-Meyer Assessment Scale in Chronic Stroke Patients.” *Journal of Rehabilitation Medicine* 33 (3): 110–13. doi:10.1080/165019701750165916.

Wagner, Joanne M, and Jennifer A Rhodes. 2008. “Research Report Reproducibility and Minimal Detectable Change of Three- Dimensional Kinematic Analysis of Hemiparesis

After Stroke.” doi:10.2522/ptj.20070255.

Winstein, Carolee J., Steven L. Wolf, Alexander W. Dromerick, Christianne J. Lane, Monica A. Nelsen, Rebecca Lewthwaite, Steven Yong Cen, and Stanley P. Azen. 2016. “Effect of a Task-Oriented Rehabilitation Program on Upper Extremity Recovery Following Motor Stroke the ICARE Randomized Clinical Trial.” *JAMA - Journal of the American Medical Association* 315 (6). American Medical Association: 571–81. doi:10.1001/jama.2016.0276.

Winter, David A. 1987. *The Biomechanics and Motor Control of Human Gait. Waterloo, Ontario, Canada, University of Waterloo Press*. doi:: http://dx.doi.org/10.1016/0021-9290(92)90236-T.

Wold, Herman O. A. 1975. “Soft Modeling by Latent Variables: The Nonlinear Iterative Partial Least Squares Approach.”

Wolf, S L, D E Lecraw, L a Barton, and B B Jann. 1989. “Forced Use of Hemiplegic Upper Extremities to Reverse the Effect of Learned Nonuse among Chronic Stroke and Head-Injured Patients.” *Experimental Neurology* 104 (2): 125–32. doi:10.1016/S0014-4886(89)80005-6.

Wolf, S. L., P. A. Catlin, M. Ellis, A. L. Archer, B. Morgan, and A. Piacentino. 2001. “Assessing Wolf Motor Function Test as Outcome Measure for Research in Patients After Stroke.” *Stroke* 32 (7): 1635–39. doi:10.1161/01.STR.32.7.1635.

Wolf, Steven L, Carolee J Winstein, J Philip Miller, and David Morris. 2015. “Effect of Constraint-Induced Movement” 296 (17): 2095–2104.

Wood, Michael D., Leif E.R. Simmatis, J. Gordon Boyd, Stephen H. Scott, and Jill A. Jacobson. 2018. “Using Principal Component Analysis to Reduce Complex Datasets Produced by Robotic Technology in Healthy Participants.” *Journal of NeuroEngineering and Rehabilitation* 15 (1). BioMed Central Ltd. doi:10.1186/s12984-018-0416-5.

Wright, Z.A., J.L. Patton, F.C. Huang, and E. Lazzaro. 2015. “Evaluation of Force Field Training Customized according to Individual Movement Deficit Patterns.” In *IEEE International Conference on Rehabilitation Robotics*. Vol. 2015-Septe. doi:10.1109/ICORR.2015.7281198.

Wright, Zachary a, Moria E Fisher, Felix C Huang, and James L Patton. 2014. “Data Sample Size Needed for Prediction of Movement Distributions.” *36th Annual International Conference of the IEEE Engineering in Medicince and Biology Society*, 5780–83. doi:10.1109/EMBC.2014.6944941.

Wright, Zachary A., Emily Lazzaro, Kelly O. Thielbar, James L. Patton, and Felix C. Huang. 2017. “Robot Training with Vector Fields Based on Stroke Survivors’ Individual Movement Statistics.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (2): 1–1. doi:10.1109/TNSRE.2017.2763458.

Wright, Zachary A., James L. Patton, Feliix C. Huang, and Emily Lazzaro. 2015. “Customized Force Field Training Based on Stroke Survivors’ Individual Movement Distributions.” *Society for Neuroscience*. doi:10.1017/CBO9781107415324.004.

Wright, Zachary A., James L. Patton, and Felix C. Huang. 2018. “Energetics during Robot-Assisted Training Predicts Recovery in Stroke.” In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2507–10. IEEE. doi:10.1109/EMBC.2018.8512737.

Zackowski, K M, A W Dromerick, S A Sahrmann, W T Thach, A J Bastian, Kathleen

Zackowski, and Kennedy Krieger. 2004. “How Do Strength , Sensation , Spasticity and Joint Individuation Relate to the Reaching de ® Cits of People with Chronic Hemiparesis ?” 127 (5). doi:10.1093/brain/awh116.

Zariffa, José, Naaz Kapadia, John L.K. Kramer, Philippa Taylor, Milad Alizadeh-Meghrazi, Vera Zivanovic, Urs Albisser, et al. 2012. “Relationship between Clinical Assessments of Function and Measurements from an Upper-Limb Robotic Rehabilitation Device in Cervical Spinal Cord Injury.” *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 20 (3): 341–50. doi:10.1109/TNSRE.2011.2181537.

Zeiler, Steven R., and John W. Krakauer. 2013. “The Interaction between Training and Plasticity in the Poststroke Brain.” *Current Opinion in Neurology*. doi:10.1097/WCO.0000000000000025.

Zollo, L, E Gallotta, E Guglielmelli, and S Sterzi. 2011. “Robotic Technologies and Rehabilitation: New Tools for Upper-Limb Therapy and Assessment in Chronic Stroke.” *European Journal of Physical and Rehabilitation Medicine* 47 (2): 223–36. http://www.ncbi.nlm.nih.gov/pubmed/21445028.

Zollo, Loredana, Luca Rossini, Marco Bravi, Giovanni Magrone, Silvia Sterzi, and Eugenio Guglielmelli. 2011. “Quantitative Evaluation of Upper-Limb Motor Control in Robot-Aided Rehabilitation.” *Medical and Biological Engineering and Computing* 49 (10): 1131–44. doi:10.1007/s11517-011-0808-1.

VITA

NAME: Zachary A. Wright

EDUCATION:

PhD, Bioengineering, **University of Illinois at Chicago**, September 2019
Dissertation Title: *Distributions of Motor Actions for the Design of Customized Robot-Assisted Therapy*

MS, Bioengineering, **University of Illinois at Chicago**, March 2012
Thesis Title: *Startle Stimuli Reduces Internal Model Control in Discrete Movements*

BS, Bioengineering, **University of Illinois at Chicago**, March 2009

RESEARCH EXPERIENCE:

Doctoral Research, Arms + Hands Lab, **Shirley Ryan AbilityLab**, Chicago, IL, August 2013 – September 2019

Research Technician, Neurology Department, **Northwestern University**, Evanston, IL, April 2011 – August 2013

Master's Research, SMPP Robotics Lab, **Rehabilitation Institute of Chicago,** June 2008 – March 2011

TEACHING:

Teaching Assistant, Senior Design, Bioengineering, UIC, Spring 2019
Teaching Assistant, Modeling Data and Systems, Bioengineering, UIC, Spring 2018
Guest Lecturer, Intro. to Bioengineering, Bioengineering, UIC, Fall 2016, Fall 2017
Docent, Sensory Motor Performance Program, Rehabilitation Institute at Chicago, Summer 2015
Teaching Assistant, Cell Biology, Biology, UIC, Fall 2009, Spring 2010

PATENTS:

Slutzky, MW, Lindberg, EW, **Wright, ZA** Device and method for treating abnormal muscle function. US Patent 10,175 755. Jan. 8, 2019

PUBLICATIONS:

Wright, ZA, Majeed, YA, Patton, JL, Huang, FC (*under review*) Key components of mechanical work predict outcomes in robotic stroke therapy. Journal of NeuroEngineering and Rehabilitation

Wright, ZA, Lazzaro, E, Thielbar, KO, Patton, JL, Huang, FC (2017) Robot training with vector fields based on stroke survivors' individual movement statistics. IEEE Transactions on Neural Systems and Rehabilitation Engineering. DOI 10.1109/TNSRE.2017.2763458

Flint, RD, Scheid MR, **Wright ZA**, Solla AS, Slutzky MW (2016) Long-term stability of motor cortical activity: implications for brain machine interfaces and optimal feedback control. Journal of Neuroscience. 36(12): 3626-3632 Mar 20016.

Wright, ZA, Carlsen, AN, Mackinnon, CD, Patton, JL (2015) Degraded expression of learned feedforward control in movements released by startle. Experimental Brain Research doi: 10.1007/s00221-015-4298-5

Mugler, E, Patton, J, Flint, R, **Wright, Z**, Schuele, S, Rosenow, J, Shih, J, Krusienski, D, Slutzky, M (2014) Direct classification of all American English phonemes using signals from the functional speech motor cortex. Journal of Neural Engineering. doi:10.1088/1741-2560/11/3/035015

Flint, RD, **Wright, ZA**, Scheid MR, Slutzky, MW (2013) Long term, stable brain machine interface performance using local field potentials and multiunit spikes. Journal of Neural Engineering 10(5):056005

Wright, ZA, Rymer, WZ, Slutzky, MW (2013) Reducing muscle co-activation after stroke using a myoelectric-computer interface: a pilot study. Neurorehabilitation and Neural Repair 10(1177):1545968313517751

CONFERENCE PROCEEDINGS:

Aggarwal, A, **Wright ZA**, Huang FC, and Patton, JL (2019) Post-Stroke Motor Deficits Are Most Evident at Frequencies near 125 Hz in EMG Multivariate Probability Distributions. IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany

Wright, ZA, Patton, JL, Huang FC (2018) Energetics during robot-assisted training predicts recovery in stroke. IEEE Engineering Medicine and Biology Conference (EMBC), Honolulu, HI, USA

Wright, ZA, Lazzaro, E, Huang, F, Patton, JL (2015) Evaluation of force field training customized according to individual deficit patterns. IEEE International Conference on Rehabilitation Robotics (ICORR), Singapore

Fisher, M, **Wright, ZA**, Huang F, Patton, JL (2014) Distributions in the error space: goal-directed movements described in time and state-space representations. IEEE Engineering Medicine and Biology (EMBC), Chicago, IL USA

Wright, ZA, Fisher, M, Huang, F, Patton, JL (2014) Data sample size needed for prediction of movement distributions. IEEE Engineering Medicine and Biology (EMBC), Chicago, IL USA

Mugler, EM, Flint, RD, **Wright, ZA**, Schuele, SU, Roseneow, J, Patton, JL, Slutzky, MW (2013) Decoding Articulatory Properties of Overt Speech from Electrocorticography Proceedings of the fifth international brain computer interface meeting 10.3217/978-3-85125-260-6-38

Scheid, MR, Flint, RD, **Wright, ZA**, Slutzky, MW (2013) Long-term, stable behavior of local field potentials during brain machine interface use. IEEE Engineering in Medicine and Biology (EMBC), Osaka, Japan

Flint, RD, **Wright, ZA**, Slutzky, MW (2012) Control of a biomimetic brain machine interface with local field potentials: performance and stability of a static decoder over 200 days. IEEE Engineering in Medicine and Biology (EMBC), San Diego, CA, USA

Wright, ZA, Rymer, WZ, Slutzky, MW (2012) Myoelectric computer interfaces to reduce co-contraction after stroke. IEEE Engineering in Medicine and Biology in Medicine and Biology Conference (EMBC), San Diego, CA, USA

Wright, ZA, Patton, JL, Ravichandran, V (2011) Startle reduced recall of recently learned internal model. IEEE International Conference on Rehabilitation Robotic (ICORR), Zurich, CH

Wright, ZA, Rogers, MW, MacKinnon, CD, Patton, JL (2009) Startle stimuli reduce the internal model control in discrete movements. IEEE Engineering in Medicine and Biology Conference (EMBC) Minneapolis, MN USA

ABSTRACT/POSTERS:

Wright, ZA, Patton, JL, Huang FC (2018) Distribution analysis variables that best capture individual differences in motor deficits due to stroke. Society for Neuroscience (SfN), San Diego, CA, USA

Wright, ZA, Patton, JL, Huang FC (2017) Detecting state-dependency of patient involvement during force field training. Society for Neuroscience (SfN), Washington, DC, USA

Wright, ZA, Patton, JL, Huang, FC (2016) Customized force field training based on stroke survivors' individual movement distributions. Society for Neuroscience (SfN), Chicago, IL, USA

Wright, ZA, , Lazzaro, E, Thielbar, KO Patton, JL, Huang, FC, Lazzaro, E (2016) Robot training with vector fields of force based on a patient's movement statistics. Arizona State University Rehabilitation Robotics Workshop, Tempe, AZ, USA

Slutzky, MW, **Wright, ZA**, Flint, RD, Sheid, MR (2014) A novel neurofeedback paradigm enables the decoupling of LFP high gamma activity from spike rate. Society for Neuroscience (SfN), Washington, DC, USA

Flint, RD, **Wright, ZA**, Mugler, EM, Dewald, JPA, Brkic, NN, Ripley, DL, Slutzky, MW (2014) Hemicraniectomy as a platform for low risk investigations of brain machine interfaces in patients with brain injuries. Society for Neuroscience (SfN), Washington, DC, USA

Fisher, M, **Wright, ZA**, Huang F, Patton, JL (2014) Parametrization of error in time versus space for goal-directed movements. Society for Neuroscience (SfN), Washington, DC, USA

Wright, ZA, Fisher, M, Huang, F, Patton, JL (2014) Time required to predict movement distributions. Society for Neuroscience (SfN), Washington, DC, USA

Wright, ZA, Rymer, WZ, Slutzky, MW (2012) Using a myoelectric computer interface for rehabilitation of chronic hemiparesis after stroke. Society for Neuroscience (SfN), New Orleans, LA, USA