Using Wearable Sensors to Measure Motor Abilities following Stroke

Todd Hester[1], Richard Hughes[1], Delsey M. Sherrill[1], Bethany Knorr[2],
Metin Akay[3], Joel Stein[1], and Paolo Bonato[1,4]
[1]*Dept of PM&R, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA*
[2]*Thayer School of Engineering, Dartmouth College, Hanover NH*
[3]*Dept of Bioengineering, Arizona State University, Tempe AZ*
[4]*The Harvard-MIT Division of Health Science and Technology, Cambridge MA*
tahester@partners.org, rhughes1@partners.org, delsey.sherrill@gmail.com,
bethany.r.knorr@dartmouth.edu, metin.akay@asu.edu, jstein@partners.org,
pbonato@partners.org

Abstract

Motor abilities of stroke survivors are often severely affected. Post-stroke rehabilitation is guided by the use of clinical assessments of motor abilities. Clinical assessment scores can be predicted by models based on features extracted from the wearable sensor data. Wearable sensors would allow monitoring of subjects in the home and provide accurate assessments to guide the rehabilitation process. We propose the use of a wearable sensor system to assess the motor abilities of stroke victims. Preliminary results from twelve subjects show the ability of this system to predict clinical scores of motor abilities.

Keywords: Wearable Sensors, Stroke, Clinical Assessment

1. Introduction

Approximately 700,000 people are affected by stroke each year in the United States and about 275,000 die from stroke each year [1]. Strokes affect a person's cognitive, language, perceptual, sensory, and motor abilities [2]. More than 1,100,000 Americans have reported difficulties with functional limitations following stroke [3]. Recovery from stroke is a long process that continues beyond the hospital stay and into the home setting. The rehabilitation process is guided by clinical assessments of motor abilities.

Accurate assessment of motor abilities is important in selecting the best therapies for stroke survivors. These assessments are based on observations of subjects' motor behavior using standardized clinical rating scales. The accuracy and consistency of observational assessments may vary greatly across clinicians [4]. Wearable sensors could be used to provide more accurate measures or could be used in addition to observational clinical tools. Wearable systems have the ability to measure motor behavior at home and for longer periods than could be observed in a clinical setting. Accelerometers can capture specific patterns of movement relating to motor disabilities. We propose that wearable systems can be used to predict clinical scores of motor abilities and we present an initial analysis of data demonstrating an association between accelerometer data and clinical scores.

2. Methods

Twelve subjects who had a stroke within the past 2 to 24 months were recruited for the study. Each subject was evaluated by a clinician using standardized clinical motor performance scales, including the Fugl-Meyer Assessment of Sensorimotor Recovery after Stroke, Chedoke-McMaster Stroke Assessment, Wolf Motor Performance Test, and the Reaching Performance Scale. These scales measure dimensions of upper limb motor behavior including movement quality, stage of motor recovery, use of compensatory movement strategies, and the ability to perform functional tasks. All testing was performed at Spaulding Rehabilitation Hospital. Subjects provided informed consent approved by the hospital's research review board. Accelerometers were attached to the affected arm and the trunk (Figure 1).

Sensor data was recorded using the Vitaport 3 (Temec BV, The Netherlands) ambulatory recorder, which was worn on the waist. Subjects performed multiple repetitions of tasks requiring reaching and prehension, selected from the clinical scales. The tasks included reaching to close and distant objects, placing the hand or forearm from lap to a table, pushing and pulling a weight across a table, drinking from a beverage can, lifting a pencil, flipping a card, and turning a key.

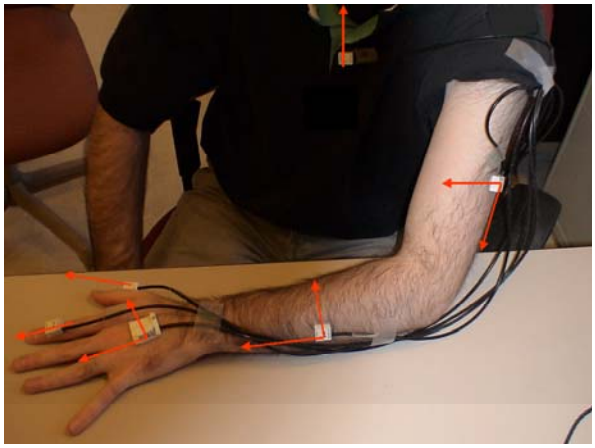

Figure 1 Sensor setup and orientation of the axes of the accelerometers.

The accelerometer data was digitally low-pass filtered in Matlab with a cutoff frequency of 15 Hz to remove high frequency noise. Both this low-pass filtered and a high-pass filtered version of the data were utilized in the analysis. The high-pass filtered version of the data was derived in an attempt to isolate actual acceleration components from gross postural adjustments. A 1 Hz cut-off frequency was used.

The signals were marked manually during testing using a 5 V pulse and marks were checked via visual inspection of the data. The marks were used to segment the data by task through an automated software procedure based on threshold detection of the manual markings. Manipulation tasks such as card flipping were also segmented within the task using a touch sensor. They were segmented into a reaching epoch, a manipulation epoch, and a release/return epoch. Subjects performed between 10 and 20 repetitions of each task, resulting in an average of 109 segments for each subject. The epochs ranged from 0.30 s to 23.3 s, with a mean length of 2.54 s and a standard deviation of 1.90 s. The following features were extracted from each epoch of accelerometer data for later analysis:

- Mean value of the low-pass filtered data
- Root-Mean-Square value (this feature and the following ones were all derived from the high-pass filtered version of the accelerometer data)
- Dominant frequency
- Ratio of energy in 0.2 Hz bin around the dominant frequency to total energy (measure of periodicity)
- Range of autocovariance
- Root-Mean-Square value of the derivative of acceleration (i.e. jerk time series)
- Dominant frequency of the jerk time series
- Ratio of energy in 0.2 Hz bin around the dominant frequency of jerk to total energy (measure of periodicity)
- Peak velocity
- Jerk metric (i.e. the RMS jerk normalized by the peak velocity
- Approximate entropy (nonlinear measure of complexity)
- Correlation at zero lag between selected pairs of accelerometer time series
- Peak correlation within a 1 s window between selected pairs of accelerometer time series
- Lag time of the peak correlation between selected pairs of accelerometer time series

Features from each task were imported into the Waikato Environment for Knowledge Analysis (WEKA) for exploratory analysis [5]. Initially, we used WEKA to look at scatterplots of features and clinical scores to assess the suitability of the current sensors, tasks, and features to predict the subjects' clinical scores. Next, we built linear regression models in WEKA to explore their ability to predict the clinical scores as well as to examine which feature sets were useful in predicting the scores. Features for the linear regression models were selected by the M5 method, which performs a backward stepwise regression using the Akaike information criterion [6]. The models provided feature sets for each clinical score. Then, linear regression models were built in Matlab to predict clinical scores using a leave-one-subject-out method (i.e. clinical scores for each subject were predicted based on a model built with data from all other subjects). All of the features were normalized to have a mean of zero and a standard deviation of one. Table 1 lists the clinical scores predicted by the model. Scores were predicted based on analysis of features from 16 different segments from 8 tasks. We looked at the forearm to table, hand to table, pushing a weight, and retrieving a weight tasks in their entirety and at the

can lifting, pencil lifting, card-flipping, and key-turning tasks in three segments each.

3. Results

Figure 2 shows a scatterplot of the peak correlation within a 1 s window between the index finger and hand accelerometers and the peak velocity of the thumb accelerometer during the manipulation epoch of the can lifting task. The colors represent the scores on the Chedoke-McMaster Hand Stage. The figure shows that higher peak velocities for the thumb accelerometer data correlate to higher clinical scores, while higher correlations between the index finger and hand data correlate with lower scores.

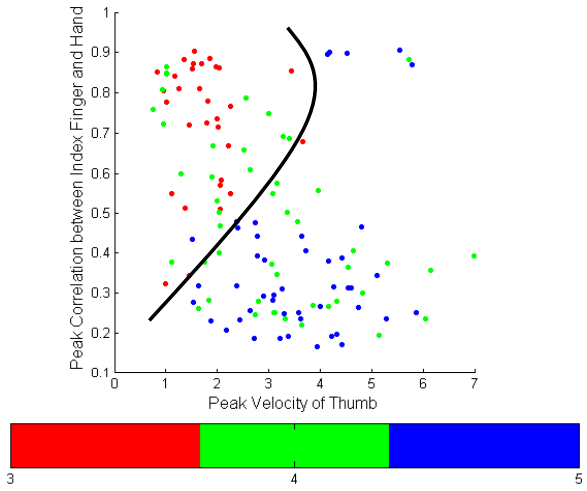

Figure 2 Scatterplot of peak correlation between accelerometer data from index finger and hand and peak velocity derived from thumb accelerometer data in comparison to Chedoke-McMaster Hand Stage scores. A line separates well samples associated with a score of 3 from samples associated with a score of 5. Samples for a score of 4 are in between and overlap with the rest of the data.

Models predicting all seven clinical scores with features from all 16 different task segments were built. Table 1 shows the root mean square error of the best model for each clinical score along with the range for each clinical score. The models were most successful in predicting the Chedoke-McMaster Hand Stage and the shoulder and elbow portion of the Fugl-Meyer scale, with relative errors close to 10%. The models were less successful in predicting other clinical scores.

Table 2 shows the RMS error of the prediction of the shoulder and elbow portion of the Fugl-Meyer scale and the Chedoke-McMaster Hand Stage using each task. The best predictors of the Hand Stage were models built with features extracted from the forearm to table task and the manipulation segments of the can lifting and card flipping tasks. The best predictors of the shoulder and elbow portion of the Fugl-Meyer were the "reaching" segments of the manipulation tasks. The worst predictor of both clinical scores was the "releasing" segment of the pencil lifting task.

Table 1. Errors in predicting clinical scores

Clinical Score	RMS Error	Score Range	Range for Subjects Tested
Chedoke-McMaster Hand Stage	0.42	1-7	3-5
Chedoke-McMaster Arm Stage	1.27	1-7	3-7
Wolf Test – Median Time	1.30	0-120	1.57-9.66
Fugl-Meyer Shoulder-Elbow	2.35	0-30	19-30
Fugl-Meyer Shoulder-Elbow	3.32	0-24	4-22
Fugl-Meyer Total Score	10.01	0-66	29-63

Table 2. RMS Errors in predicting clinical scores for different tasks.

Task	**Chedoke-McMaster Hand Stage**	**Fugl-Meyer Shoulder-Elbow**
Can - Segment 1	0.58	3.66
Can - Segment 2	0.50	6.02
Can - Segment 3	0.67	5.13
Card - Segment 1	0.58	2.35
Card - Segment 2	0.50	3.74
Card - Segment 3	0.67	5.23
Forearm To Table	0.67	4.90
Hand To Table	0.42	6.23
Key - Segment 1	0.86	2.99
Key - Segment 2	0.90	3.75
Key - Segment 3	0.67	4.80
Pencil - Segment 1	0.62	3.10
Pencil - Segment 2	0.67	5.22
Pencil - Segment 3	0.93	7.76
Push Weight	0.71	6.19
Retrieve Weight	0.74	5.16

Table 3 shows the linear regression model used to predict the shoulder and elbow portion of the Fugl-Meyer score using features from the reaching segment of the card flipping task. Table 4 shows the actual scores, predicted scores, and the standard deviation of the predicted scores for the shoulder and elbow portion of the Fugl-Meyer scale using the model shown in Table 3. Five of the predicted scores were within 1 point of the actual score, and the closest was within 0.02. The worst prediction was 35.81 for a subject with a clinical score of 30.

Table 3. Linear regression model for prediction of the shoulder and elbow portion of the Fugl-Meyer score based on the reaching segment of the card flipping task

Coefficient		Feature
1.18	*	Mean of Forearm X acc
1.81	*	Mean of Forearm Y acc
0.65	*	Mean of Upper Arm X acc
1.50	*	Mean of Upper Arm Y acc
-0.86	*	Mean of sternum acc
0.99	*	RMS of forearm X acc
1.66	*	RMS of forearm Y acc
0.31	*	RMS of Upper Arm X acc
-1.31	*	RMS of Upper Arm Y acc
-0.16	*	RMS of sternum acc
-0.28	*	Peak Corr. of Thumb and Hand

Table 4. Scores for the shoulder and elbow portion of the Fugl-Meyer score

Subject	Actual Score	Predicted Score	STD of Prediction
A	23	22.15	0.86
B	26	26.50	1.40
C	19	20.66	0.40
D	26	22.00	0.63
E	30	35.81	2.47
F	30	29.98	2.83
G	24	24.97	0.18
H	24	22.63	0.99
I	30	26.04	1.57
J	30	28.79	0.75
K	20	21.16	1.11
L	27	27.40	1.38

4. Discussion and Conclusion

The results of the linear regression models have been promising so far. Our models predicted two clinical scores within 10% of the average score. Our sensor system showed the ability to detect specific movement patterns related to clinical scores of movement ability. For example, a negative coefficient was associated with the root mean square value of the sternum accelerometer channel in most of the linear regression models, indicating that the method was able to detect compensatory trunk movements related to lower clinical scores. The jerk metric and the root mean square of the jerk time series usually had a high coefficient in the models, showing that the system determined that smooth movement was significantly related to the clinical scores. Scores such as the median time on the Wolf Motor Function Test may be difficult to predict because of non-linear relationships between post-stroke motor ability and performance scores on this test. To predict these scores more accurately, it may be necessary to include non-linear parameters or create a non-linear model to predict the scores. We are currently collecting data from more subjects, which will allow us to improve the linear regression models and explore the use of nonlinear models. The small size of the current dataset limits the number of features we can use in the models. Including more features is expected to improve the model. Collecting data from subjects with a wider range of motor abilities and clinical scores will allow us to develop a more accurate linear regression model.

Acknowledgments

This study was supported by the grant entitled "Field Measures of Functional Tasks for CIT Intervention", #R21HD045873-01, NIH-NICHD.

References

[1] American Heart Association, Heart Disease and Stroke Statistics – 2005 Update, American Heart Association, Dallas, TX, 2004.

[2] National Institute of Neurological Disorders and Stroke (NINDS), Stroke: Hope Through Research, NINDS, Washington, DC, 2004.

[3] Center for Disease Control, Morbidity and Mortality Weekly Report, CDC, Atlanta, GA, 2001.

[4] V. Pomeroy, A. Pramanik, L. Sykes, J. Richards, and E. Hill, "Agreement between physiotherapists on quality of movement rated via videotape", Clinical Rehabilitation, May 2003, pp. 264-272.

[5] IH Witten and E. Frank, Data Mining: Practical machine learning tools and techniques. Morgan Kaufman, San Francisco, 2005.

[6] H Akaike, "A New Look at Statistical Model Identification." IEEE Transactions on Automatic Control, IEEE, 1974, pp. 716-722.