Reduction of Stroke Assessment Time for Visually Guided Reaching Task on KINARM Exoskeleton Robot

SM Mostafavi, SP Dukelow, JI Glasgow, SH Scott and P Mousavi, *Senior Member IEEE*

Abstract— Robotic technologies provide objective, highly reliable tools for assessment of brain function following stroke. KINARM is an exoskeleton device that quantifies sensorimotor brain function using a visually guided reaching task among many other behavioral tasks. As further tasks are developed to more broadly assess different aspects of behavior using the robot, techniques and approaches are required to reduce the time it takes to complete each task. The present study investigates how the value of robot-measured parameters changes under alternative schemes that significantly reduce assessment time compared to the current assessment protocol for the visually guided reaching task. Results of the study are validated by addressing an important diagnostic question using an SVM classifier, showing that the alternative schemes provide nearly identical performance in terms of classification sensitivity, specificity and accuracy.

Keywords— KINARM, Stroke Assessment, Visually Guided Reaching, Sensorimotor Evaluation

I. INTRODUCTION

Stroke is a leading cause of permanent disability in North America [1]. It can be a result of thrombosis, embolism or hemorrhage and, depending on the size and location of lesions, can lead to death or sensorimotor impairments in upper or lower limbs [2].

Previous research shows that a substantial amount of functional recovery occurs in a short time frame post-stroke, often within the early weeks and months [3]. Thus, clinicians have only a limited time window to assess the damage to the brain, render a prognosis and decide on therapeutic interventions. Prolonged assessment can lead to delays of treatment delivery despite evidence for the importance of early rehabilitation treatment [4].

Classic clinical scores for stroke assessment tend to rely on observer based ordinal scales, many of which have limited inter-rater and intra-rater reliability. Recently, robotic technologies, capable of recording objective, highly reliable data for assessment of sensorimotor impairments have been developed [5]. KINARM (BKIN Technologies, Kingston, ON) is one such robotic device that quantifies sensorimotor performance through a number of behavioral tasks supported by a virtual reality system [7]. KINARM quantifies motor performance using a visually guided reaching task [6]. This task requires involvement of occipital, parietal, and frontal lobes [7]. The current task consists of reaching movements to 8 spatial locations over 64 trials. Under the current protocol, the task takes approximately 8 minutes per arm to complete for each healthy individual attempting the task and can take slightly longer for subjects with stroke depending on their performance. Several other tasks are presently performed along with the reaching task including limb proprioceptive function, bimanual skill, and a rapid target interception capabilities task [5]. As more tasks are incorporated on the robot, the length of time to assess each subject continues to grow. This leads to the question of whether the length of each task can be reduced while still retaining the maximal amount of information to quantify subject performance across a broad range of neurological functions.

In this study, we present, for the first time, the results of time reduction schemes for the visually guided reaching task on the KINARM exoskeleton robot. These schemes involve reductions in both the number of spatial targets and the number of repetitions (trials). In particular, we investigate how the value of, and the variation in, several robotic parameters change as a result of reductions in the number of trials and targets in the task using statistical criteria such as the Standard Error of the Mean and Coefficient of Variation. We further evaluate the results of our findings by using reduced data to address an important diagnostic classification problem, i.e., the separation of stroke and control subjects. Using an SVM classifier, we show that the classification performance, as measured by sensitivity, specificity and accuracy, undergoes small changes in the presence of trial and target reductions. These results confirm the feasibility of the suggested reduction schemes for the visually guided reaching task.

II. MATERIALS AND METHODS

A. Participants

One hundred and fifty seven subjects with stroke are recruited after admission to St. Mary's of the Lake Hospital (Kingston, ON, Canada) and Foothills Hospital (Calgary, AB, Canada) for robotic evaluation using the KINARM exoskeleton robotic device. A group of 196 age-matched control subjects is also recruited for robotic evaluation. The study is approved by the institutional ethics review boards and all subjects for the study provided their informed consent to participate in the study.

This work is supported by the Natural Sciences and Engineering Research Council of Canada, the Ontario Research Foundation, the Canadian Institutes of Health Research and the Heart and Stroke Foundation of Alberta, Northwest Territories and Nunavut Grant-in-Aid.

SM Mostafavi, P Mousavi, and JI Glasgow are with School of Computing, Queen's University, Kingston, Ontario, Canada

SP Dukelow is with the Hotchkiss Brain Institute, and Dept. of Clinical Sciences, University of Calgary, Calgary, Alberta, Canada.

SH Scott is with Centre for Neuroscience Studies, Queen's University, Kingston, Ontario, Canada

Corresponding author: mostafa@cs.queensu.ca

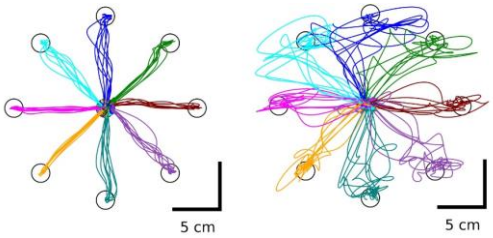

Figure 1: Hand path trajectories for 8 spatial locations (targets) on the visually guided reaching task. Left: Control subject. Right: Stroke subject. The circles on the images show a target.

B. *Visually Guided Reaching Task*

This is a task designed for assessment of sensorimotor performance of the upper limb [6]. With full vision, subjects are asked to reach "quickly and accurately" from a central target to one of eight peripheral targets located 10 *cm* away, distributed around the circumference of a circle. Each trial begins with subjects holding their index finger tip at the central target for 1250-1750 *ms*. Then a peripheral target is illuminated and subjects are given 3000 *ms* to complete the reach. Figure 1 shows hand path trajectories of movement to 8 spatial locations for one stroke subject and one control subject. Each target is presented once and subjects completed eight repetitions for a total of 64 trials. Subjects perform the task with both arms. In this study, we only analyze the data from the affected arm of subjects with stroke. A total of twelve movement parameters are recorded in each trial [6]. The following five parameters were further analyzed in this study: reaction time, posture speed, first movement direction error, total movement time, and number of movement peaks.

C. *Trial Reduction*

We performed an analysis to investigate the effect of trial reduction by computing the values of three statistical measures of the robotics parameters when dropping 0 trials (i.e. keeping the complete set of 64 trials) up to 7 trials, and keeping all 8 reaching directions. These statistical measures include the Standard Error of the Mean (SEM), Relative Standard Error (RSE) and Coefficient of Variation (CV) [8].

In the present study, the sample mean is assumed to be the mean parameter value acquired over a specified (reduced) number of trials/targets and the true mean is assumed to be the mean over 8 trials and 8 targets of the visually guided reaching task.

The relative standard error (RSE) is simply the SEM divided by the mean value of the measured parameter and is often expressed as a percentage (RSE %).

To determine a suitable trade-off between the number of dropped trials and SEM for stroke and control groups, an analysis of the second difference of SEM with respect to the number of trials was performed when dropping 0 to 7 trials. A local minima of the second difference indicates minimal increase in the amount of SEM if a further trial is dropped. We used this minima as a threshold to determine a trade-off between the number of dropped trials and SEM for Stroke and Control groups.

	Dropped trials	SEM	RSE (%)	CV
Stroke	0	0.0171	3.7310	0.3102
	1	0.0186	4.0616	0.3093
	2	0.0202	4.4202	0.3106
	3	0.0225	4.9220	0.3216
	4	0.0253	5.5249	0.3258
	5	0.0290	6.3369	0.3350
	6	0.0392	8.5679	0.3439
	7	0.0564	12.3278	0.3725
Control	0	0.0116	3.0311	0.2363
	1	0.0122	3.1984	0.2344
	2	0.0132	3.4708	0.2357
	3	0.0143	3.7490	0.2328
	4	0.0157	4.1211	0.2304
	5	0.0182	4.7673	0.2300
	6	0.0232	6.0888	0.2345
	7	0.0368	9.6500	0.2337

Table 1: SEM of reduced trials for the Reaction Time parameter.

D. *Classification*

Once a threshold has been established to reduce trials, we evaluated the clinical impact of such reduction by classifying stroke vs. control subjects using the reduced data. We employed an SVM classifier with RBF kernel function and performed a 10-fold cross validation for binary classification of stroke vs. control subjects. This involves using 90% of available subjects for training the classifier, and keeping the remaining 10% for testing. Classification was assessed using sensitivity, specificity and accuracy. The classification procedure was repeated 100 times with different training/testing data splits to produce a distribution for the three aforementioned criteria. Results of our findings are summarized in Table 4.

In order to investigate the effect of trial and reductions on classification accuracy, out of all possible reduction schemes, we considered for this investigation trial reductions of 6, 4, and 2 randomly selected trials (out of a total of 8 trials) and target reductions of *4TargetDiagonal* and *4TargetStraight* as described before and the reductions were compared using ANOVA.

III. RESULTS

Table 1 summarizes the results of our SEM, RSE and CV analysis for the Reaction Time parameter on both Stroke and Control data. It can be observed that RSE ranges from 3.7% to 12.3% for the Stroke group and from 3.0% to 9.6% for the control group. CV ranges from 0.31 to 0.37, in a generally increasing trend with increased number of dropped trials, for the stroke group while it is oscillating at the 0.230-0.236 range for the Control group.

Task parameter	Trial Repeat Threshold	RSE (%)	True RSE (%)	Error increase (%)	Trial Repeat Threshold	RSE (%)	True RSE (%)	Error increase (%)
Stroke					**Control**			
Reaction Time	6	4.42	3.73	0.69	5	3.75	3.03	0.72
Posture Speed	3	5.73	3.52	2.22	4	5.47	3.37	2.09
First Movement Direction Error	4	6.85	5.33	1.52	4	7.46	5.89	1.58
Total Movement Time	4	4.73	3.70	1.03	4	4.37	3.13	1.24
Number of Movement Peaks	4	7.76	5.73	2.02	4	8.37	5.83	2.54

Table 2: SEM error increase for reduced trials.

We performed a second difference analysis to determine a suitable trade-off between the number of dropped trials and SEM. Figure 2 shows the second difference SEM analysis for the results presented in Table 1. It can be seen that the trade-off is determined at 2 dropped trials for the Stroke group and at 3 dropped trials for the Control group. A similar procedure was followed to determine a trial repeat threshold for the remaining four parameters (Posture Speed, First Movement Direction Error, Number of Movement Maximum Speed, and Total Movement Time). Table 2 shows the trial repeat threshold for the five parameters (based on the second difference analysis), the RSE for the determined threshold, and the RSE if the entire set of 64 trials and targets had been used (True RSE). The last column of Table 2 indicates the error increase as a result of trial reduction (at threshold level) for each of the five parameters. The most frequent trial repeat threshold across all parameters for both Stroke and Control groups was 4 repetitions. This happens when the number of repeats is reduced by half, resulting in a reduction of assessment time by half on the visually-guided reaching task. Another situation where this happens is when all 8 trials are kept, but the number of targets is reduced to 4 spatial directions (compared to the original 8 directions).

We performed an analysis to compare SEM and RSE for the three trial and target reduction schemes described above. Target reductions were to 4 straight (0, 90, 180, and 270 degrees) and 4 diagonal (45, 135, 225, and 315 degrees) targets. We refer to these as *4TargetStraight* and *4TargetDiagonal*, respectively. The case where we keep all targets and reduce trials to 4 for each direction is referred to as *8Target4Trials* (similar naming terminology for other number of trials). Results of the analysis are summarized in Table 3.

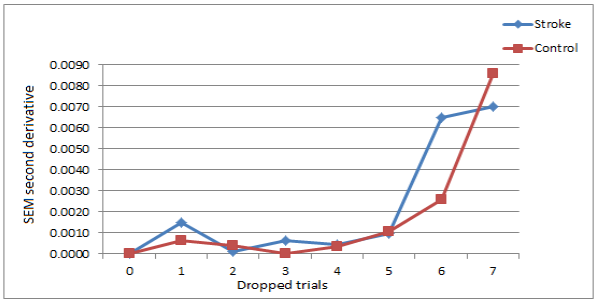

Figure 2: Second difference SEM analysis for the Reaction Time parameter on stroke and control subjects.

For each row in Table 3, we performed an ANOVA test to see whether or not the reductions are significantly different when comparing SEMs across the three experiments and reported the p-values. Out of a total of 10 comparisons (5 stroke and 5 control group parameters), five p-values are significant (smaller than 0.05).

IV. Discussion

Our results lead us to a few important observations regarding the target and trial reductions for the visually guided reaching task. First, SEM analysis for the Reaction Time parameter (Table 1) and the remaining four parameters (not shown here) reveal an increasing pattern as more trials are dropped. This is expected since SEM is inversely proportional to the square root of the total number of trials, and is expected to increase with fewer numbers of trials. A more interesting pattern is observed when analyzing CV values for the two Stroke and Control groups. While CV values are almost constant with fewer trials for the Control group (with small oscillations), the trend is generally increasing for the Stroke population. This can be explained by the fact that the subjects with stroke are generally less consistent with their reaching movements and are expected to exhibit a higher level of variation across different trials.

The second important point is regarding the reductions based on the trial repeat threshold. Reductions for both groups come at the expense of increased SEM for dropped trials. However, error increase was relatively small. The maximum SEM increase was 2.02% for the Stroke group and 2.53% for the Control group (in both cases for the Number of Movement Peaks parameter). The small error increase is a significant finding since it is an indication of the feasibility of reductions based on a trial repeat threshold.

Third, a comparison of the three reduction schemes that reduce the assessment time by half signifies no superior

Task parameter	4Target Straight	4Target Diagonal	8Target 4Trials	p-value	4Target Straight	4Target Diagonal	8Target 4Trials	p-value
Stroke					**Control**			
Reaction Time	0.02 (5.2%)	0.02 (5.2%)	0.03 (5.5%)	**1.21e-5**	0.017 (4.4%)	0.02 (4.4%)	0.02 (4.1%)	0.1
Posture Speed	0.002 (5%)	0.002 (5%)	0.002 (5%)	0.28	0.001 (4.9%)	0.001 (4.7%)	0.002 (4.9%)	0.33
First Movement Direction Error	0.004 (7.8%)	0.004 (7.4%)	0.004 (7.5%)	0.15	0.004 (8.6%)	0.003 (7.8%)	0.004 (7.5%)	**2.34e-7**
Total Movement Time	0.07 (4.7%)	0.08 (5%)	0.07 (4.7%)	**0.001**	0.05 (4.2%)	0.05 (4.4%)	0.05 (4.4%)	0.75
Num Movement Maximum Speed	0.26 (8.5%)	0.22 (7.3%)	0.24 (7.8%)	**2.61e-4**	0.17 (7.9%)	0.16 (7.5%)	0.18 (8.4%)	**0.004**

Table 3: SEM Comparison for three Target and Trial Reduction schemes: 4TargetStraight, 4TargetDiagonal, 8Target4Trials. *p-values* < 0.05 are in bold.

scheme over the other two. Over a total of 10 comparisons (last column of Table 3), 5 comparisons were not statistically significant (p-value > 0.01). Out of the remaining five comparisons, there is not a scheme which is consistently associated with lower SEM values. However, from a clinical standpoint, the *8Target4Trials* scheme is preferred over the other two reductions since it includes 3 different targets in each of the two right and left spatial directions. This is clinically important as it allows covering a wider range of space for subjects with stroke who suffer from spatial neglect.

The last observation is regarding the classification performance in the light of trial and target reduction (Table 4). Separation of Stroke and Control groups is an important question that has been addressed using the KINARM robot as a diagnostic tool [6]. Despite the fact that *p-values* suggest a difference between different reduction schemes, the difference between the largest and smallest set of trials (*All* vs. *8Target-2Trials*) is small. For instance, the difference between the two corresponding Accuracy values is only 0.22% (i.e. two extra misclassified subjects in a pool of 1000 subjects). This finding verifies the feasibility of the reduction of the current 64-trial regime to a much smaller set of trials to address diagnostic questions on the KINARM robot.

V. CONCLUSION

In this work, for the first time, we investigated the effect of trial and target reduction of the KINARM exoskeleton robot for stroke assessment. This is the primary step for a wide scheme aimed at reduction of vital assessment time for stroke survivors. Results of our analysis show that a reduction scheme that reduces the assessment time by half is a safe and suitable replacement for the current scheme when investigating five measured robot parameters on a population of 157 stroke and 196 control subjects. Results also confirm that following reduction to a much smaller set of trials, the measured parameters from the KINARM can still be used for predicting stroke versus control subjects. As a result, the proposed scheme is feasible if the aim is to use the current visually guided reaching task as a simple diagnostic tool.

We wish to extend our analysis to other discrete, trial-based tasks on the KINARM robot in the future. In addition, we plan to investigate the effect of the suggested reduction schemes on the quality of decisions for other, more complex clinical questions such as stroke prognosis.

Dataset	Sensitivity	Specificity	Accuracy
All 8Target-8Trials	84.05	90.82	86.34
8Target-6Trials	83.91	89.71	85.91
8Target-4Trials	85.32	89.94	86.87
4Target-Diagonal	81.84	86.68	83.46
4Target-Straight	86.34	90.36	87.11
8Target-2Trials	85.51	88.26	86.12
p-value	**5.91e-12**	**6.33e-24**	**4.34e-17**

Table 4: Sensitivity, specificity, and accuracy analysis for reduced trials/targets.

REFERENCES

[1] American Stroke Association. Web site. http://www.strokeassociation.org. Accessed March 2014.

[2] R. P. Donahue, R. D. Abbott, D. M. Reed, and K. Yano, "Alcohol and Hemorrhagic Stroke: The Honolulu Heart Program," JAMA, vol. 255, no. 17, pp. 2311–2314, 1986.

[3] P. W. Duncan, L. B. Goldstein, R. D. Horner, P. B. Landsman, G. P. Samsa, and D. B. Matchar, "Similar motor recovery of upper and lower extremities after stroke". Stroke, 25(6):1181-1188. 1994.

[4] J. Biernaskie, G. Chernenko, and D. Corbett. "Efficacy of rehabilitative experience declines with time after focal ischemic brain injury". *The Journal of neuroscience*, *24*(5), 1245-1254. 2004.

[5] S.H. Scott, S.P. Dukelow. "Potential of robots as next-generation technology for clinical assessment of neurological disorders and upper-limb therapy". *Journal of Rehabilitation Research & Development*, *48*(4). 2011.

[6] A.M. Coderre, A.A. Zeid, S.P. Dukelow, M.J. Demmer, K.D. Moore, M.J. Demers, H. Bretzk, T.M. Herter, J.I. Glasgow, K.E. Norman, S.D. Bagg, S.H. Scott. "Assessment of upper-limb sensorimotor function of subacute stroke patients using visually guided reaching". Neurorehabil Neural Repair. 2010; 24(6):528-41.

[7] S.H. Scott. "Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching". J Neurosci Methods. 1999 Jul 15; 89(2):119-27.

[8] B.S. Everitt. The Cambridge Dictionary of Statistics, CUP. ISBN 0-521-81099-X. 2003.