Published in final edited form as:

J Neurol Phys Ther. 2009 June ; 33(2): 96–103. doi:10.1097/NPT.0b013e3181a33638.

Relationships and responsiveness of six upper extremity function tests during the first 6 months of recovery after stroke

Justin A. Beebe¹ and Catherine E. Lang^{1,2,3}

¹Program in Physical Therapy, Washington University, St. Louis, MO

²Program in Occupational Therapy, Washington University, St. Louis, MO

³Department of Neurology, Washington University, St. Louis, MO

Abstract

Background—Knowing how clinical tests are related to each other, if tests are responsive to change and what constitutes change is critical to evidence-based practice and sound research.

Objective—To determine 1) relationships between six clinical tests of upper extremity function, and 2) responsiveness for each test during the first 6 months after stroke.

Methods—Grip strength, Pinch strength, Action Research Arm Test (ARAT), Jebsen Taylor test of hand function, 9-Hole Peg Test (9HPT), and the Stroke Impact Scale-Hand (SIS-Hand) domain were administered to 33 subjects within 1 month, 3 and 6 months after stroke. Spearman's correlations were used to analyze relationships between tests. Responsiveness was calculated using the single population effect size method.

Results—All tests were correlated with each other with absolute values ranging from 0.54 to 0.92 at 1 month, 0.57 to 0.97 at 3 months, and 0.41 to 0.97 at 6 months. All tests were at least moderately responsive to change, with effect sizes ranging from 0.50 to 1.02 from 1–3 months, and 0.56 to 0.86 from 1–6 months.

Conclusions—Our data can assist clinicians and researchers in making decisions to use specific tests for measuring upper extremity function in people with hemiparesis in the first 6 months of recovery.

Introduction

Of the almost 600,000 newly diagnosed strokes per year in the United States¹, nearly 80% of these individuals have acute paresis affecting the upper extremity². Research suggests that most recovery occurs within the first few months after stroke^{2–7} and not coincidentally rehabilitation begins as early as possible, often within the first week. To evaluate changes in upper extremity function early after stroke, clinicians can choose from a number of standardized clinical tests. Recent research from our lab has shown that many of these tests measure the same underlying construct of upper extremity function^{8, 9}. Using principal components analysis, we found, in two separate samples, that six standardized clinical tests all loaded onto a single principal component. This suggests that each of these tests was measuring the same thing, and that standardized upper extremity function tests are interchangeable, i.e. one is as good as another. Knowing how these same six tests are related to each other and whether or not the relationships change over time would be of great benefit to clinicians as

they choose tests to assess upper extremity function of individuals with acute hemiparesis and throughout the first few months of recovery.

The starting point for selecting upper extremity tests is to identify the ones with established reliability and validity¹⁰. Beyond the typical types of validity, clinicians and researchers need to know the responsiveness of the test, i.e. can it accurately detect change over time^{11–13}. Responsiveness of a test, as with other psychometric properties, needs to be assessed for each population in which the test is to be used¹⁴. For people with stroke, a few upper extremity function tests have been examined for responsiveness. The Action Research Arm Test (ARAT)^{15–17}, the Motor Assessment Scale¹⁵, Fugl-Meyer Assessment¹⁶, the Stroke Impact Scale (SIS)¹⁸, and the Chedoke Arm and Hand activity inventory¹⁹ have been shown to be responsive to change early after stroke. Initial testing in all of these studies was conducted within 1 month of stroke onset. Two of these studies involving the ARAT¹⁷ and the SIS¹⁸, examined responsiveness over a greater period with follow up testing at 3 months^{17,18} or 6 months¹⁸. In rehabilitation settings, other upper extremity function tests are often administered post stroke, such as the Nine-Hole Peg Test (9HPT), the Jebsen Taylor Test of hand function, and Grip and Pinch strength. Grip and Pinch strength, while impairment level measures, have been to be proposed to be surrogate measures of upper extremity functional outcomes²⁰ and will be treated as such in this manuscript. The responsiveness of these other tests has not been examined in people during the first few weeks and months after stroke.

The purpose of this investigation was twofold: 1) to determine how six clinical tests of upper extremity function were related to each other in the first weeks and months after stroke, and 2) to determine how responsive the tests were to change over the first 6 months after stroke. This information is needed by clinicians and researchers because it is critical for evidence based practice and sound research.

Methods

Subjects

Thirty-three subjects with hemiparesis due to stroke participated in this study. Subjects were recruited from the Cognitive Rehabilitation Research Group Stroke Registry based on the presence of hemiparesis. Subjects were included if they 1) had a diagnosis of ischemic or hemorrhagic stroke by a stroke neurologist within one month of onset of symptoms, 2) had CT or MRI imaging data consistent with clinical presentation, 3) had persistent hemiparesis with a score of 1 to 4 on the Motor Arm item of the National Institutes of Health Stroke Scale (NIHSS), 4) had evidence of preserved cognition as indicated by a score of 0 or 1 on the Consciousness and Communication item of the NIHSS, and 5) had the ability to follow 2-step commands. Patients were excluded from the study if they 1) had orthopaedic or other medical conditions that limited the more affected upper extremity prior to the stroke, 2) had a prior history of hemiparesis or stroke, 3) had hemispatial neglect as evidenced by a score of 2 on the Extinction and Inattention item of the NIHSS, 4) had severe aphasia as evidenced by a score of 2 or 3 on the Language item of the NIHSS, 5) had complete hemianopsia as evidenced by a score of 2 or 3 on the Visual item of the NIHSS, or 6) the subject was unable to give informed consent. Characteristics of the group are provided in Table 1. Subjects were recruited from a tertiary care center where a large portion of the people with stroke either lived too far away for follow-up research testing, had multiple co-morbidities that excluded them from our study, or were discharged to a skilled nursing facility where we did not have access to continue testing them. The 33 subjects that were tested accounted for approximately 10% of the total subjects screened for this study. This study was approved by the Washington University Human Research Protection Office, and all participants provided informed consent prior to participation.

Over the duration of the study, subjects received standard stroke rehabilitation services as directed by their neurology and rehabilitation physicians. At the first visit, most subjects were in inpatient rehabilitation and were therefore receiving daily occupational and physical therapy services. The subjects that were living at home at the time of the first visit typically had outpatient services a few days per week. By the 3 and 6 month visits, subjects had progressively decreasing services or no services at all as determined by their progress and by their rehabilitation team. We did not record subject specific information on the type and duration of therapy received.

Testing paradigm

Subjects were tested for their ability to use their upper extremity for functional activities at three time points: within 1 month, 3 months, and 6 months after stroke. At each time point, testing was completed during a single session lasting approximately one hour. The tests were administered randomly with the Stroke Impact Scale questionnaire always being administered last. The authors, both physical therapists, administered all of the tests. The least affected side was tested first. Physical fatigue did not appear to be a concern as subjects were permitted rest time as needed between tests.

Measurement of upper extremity function

All subjects underwent a battery of six standardized clinical tests of upper extremity function. The battery of tests were evaluative measures used to measure upper extremity function in multiple ways such as criterion-rated, timed performance, and self report. We were careful as to the number of tests administered so as not to place too great a testing burden on our subjects, particularly at the 1 month time point. The tests in this study were selected based on published data at the time of study initiation regarding reliability, validity, normative values, and appropriateness for use with people with stroke. Several tests were excluded for the following reasons. The Motor Assessment Scale was not included because it contains items examining other body segments and is not specific to the more affected upper extremity. The Fugl-Meyer was not included because it has already been shown to be highly correlated with the Action Research Arm Test (r values = 0.91 – 0.94)²¹. The Chedoke Arm and Hand activity inventory was not included because information about it was not available at the time of study initiation. It has since been shown to be highly correlated with the Action Research Arm Test (r = 0.93)¹⁹. Two of the tests, Grip and Pinch strength, were included because they are quick and easy to administer, and they are common in most clinics. All clinical tests were performed bilaterally such that the less affected upper extremity served as the matched control for the more affected side. The following tests were used:

*Grip Strength*²² is a dynamometer measurement of the maximum amount of force produced during a five-finger grip. Test-retest reliability (ICC = 0.91) has been reported for the more affected side of subjects with chronic stroke²⁰. Three trials were performed on each side with the subject sitting, arm at their side, the elbow at 90 degrees, forearm neutral, and the wrist in slight extension. If subjects were not able to hold the test position, the dynamometer was supported, however care was taken to ensure the support did not influence force production. The average of three trials²³ was computed, and scores were expressed in kilograms. Normal grip strength for similarly aged individuals has been reported to be 32.9 (8.5) kg²⁴ (mean (SD)).

*Pinch Strength*²⁴ is a dynamometer measurement of the maximum amount of force produced during a three-fingered key pinch. Test-retest reliability of an average between left and right hands (r = 0.85) has been reported for young, healthy subjects²⁵. Three trials were performed on each side with the subject sitting, arm at their side, the elbow at 90 degrees, forearm neutral, and the wrist in slight extension. If subjects were not able to hold the test position, the

dynamometer was supported, however extreme care was taken to ensure the support did not influence force production. The average of three trials²³ was computed, and scores were expressed in kilograms. Normal pinch strength for similarly aged individuals has been reported to be 8.8 (1.5) kg²⁴.

Action Research Arm Test^{17, 26–30} is a test for upper extremity function with four subscales: grasping, gripping, pinching, and gross movement. It uses ordinal scoring on 19 items, where 0 indicates no movement and 3 indicates normal movement. Items in each subscale are summed for grasping (18-point maximum), gripping (12-point maximum), pinching (18-point maximum), and gross movement (9-point maximum), with a total scale score of 57, indicating normal. Test-retest reliability has been reported (ICC = 0.97) for the more affected side of individuals with stroke of varying duration³¹.

Jebsen Taylor Test of Hand Function^{32, 33} is a functional assessment scored by the summed times to complete seven common tasks. The tasks are: writing a sentence, simulated page turning, picking up small objects, simulated feeding, stacking checkers, picking up large light objects, and picking up large heavy objects. The first task, writing a sentence, was not used because it is dependent on hand dominance and education level³³. The maximal amount of time allotted for each subtest was 120 seconds³⁴. Test-retest reliability has been reported ($r = 0.92$) for patients with stable hand disorders resulting from stroke, brain injury, or rheumatoid arthritis³². Normal time to complete the Jebsen for similarly aged individuals has been reported to be 30.4 (1.11) seconds^{35, 36}.

*Nine-Hole Peg Test*³⁷ is a finger coordination measure involving timed performance to insert and remove nine pegs from a wooden block. Test-retest reliability has been reported ($r = 0.98$) for patients with acquired neurological disorders resulting from stroke, multiple sclerosis, brain injury, or tumor³⁸. Scores were expressed in seconds, with the maximal amount of time allotted being 120 seconds. Normal time to complete the 9HPT for similarly aged individuals has been reported to be 19.4 (2.68) seconds³⁷.

Stroke Impact Scale^{18, 39, 40} is a self report questionnaire to measure the impact of stroke in multiple domains. Test-retest reliability has been reported (ICC = 0.92) for the hand function domain in individuals with stroke of varying duration⁴⁰. The subject answers the questions based on a 5-point Likert scale, and the scores are computed to obtain a number between 0 and 100, where 100 indicates normal³⁹. Data on all domains of the Stroke Impact Scale were collected, but only the hand function domain (asking about the most affected upper extremity) was part of our upper extremity function battery. Questions in the hand function domain pertain to carrying heavy objects, turning a doorknob, opening a can or jar, tying a shoelace, and picking up a dime.

Additional testing

Additional tests were conducted to provide a more thorough description of the sample (Table 1). This testing occurred randomly during each testing session. To quantify upper extremity strength, a hand-held dynamometer was used to assess both flexion and extension of the shoulder, elbow, wrist, and index finger using standard manual muscle test positions⁴¹. Both more affected and less affected upper extremities were tested. The values obtained were highly correlated within each arm and thus, the values were combined into one variable and were expressed as a percentage of less affected side (Table 1, Composite Strength)⁴². More affected side shoulder pain was assessed using a standard 11-point numeric rating scale, where 0 = no pain (Table 1, Shoulder Pain). Joint position sense was evaluated on both sides at the index finger using standard clinical techniques where normal = correct on $\geq 3/5$ trials (Table 1, Index finger joint position sense). Lastly, spasticity was evaluated on the more affected side using the Modified Ashworth Scale⁴³ for flexion/extension movements at four joints:

metacarpophalangeal joints of the hand, the wrist, the elbow, and the shoulder (Table 1, Modified Ashworth Scale).

Data analysis

SPSS version 13 (SPSS Inc., Chicago, IL) was used for analyses. Spearman's correlations were used to evaluate the relationships between test scores at each time point. For ease of interpretation, all correlations were expressed as absolute values. This was done because lower times reflect better performance on timed tests (i.e. Jebsen, and 9HPT), and higher scores reflect better performance on the other tests (i.e. Grip, Pinch, ARAT, and SIS-Hand). Bonferroni correction for statistical significance was set to $p < 0.0033$ to allow for 15 comparisons within each time point. Interpretation of the magnitude of the correlation coefficients were as follows: $r > 0.25$ were considered fair; $r > 0.50$ were considered moderate; and $r > 0.75$ were considered excellent/strong⁴⁴.

For each test, we calculated responsiveness between testing at 1 month and 3 months, and 1 month and 6 months using the single population effect size method^{11, 12}. The effect size from 1 to 3 months was calculated as the mean absolute value change from 1 month to 3 months divided by the standard deviation at 1 month. The effect size from 1 to 6 months was calculated as the mean absolute value change from 1 month and 6 months divided by the standard deviation at 1 month. Similar to the interpretation of the correlation coefficients, responsiveness values > 0.20 were considered fair; values > 0.50 were considered moderate; and values > 0.80 were considered highly responsive to change⁴⁵. It is possible for values to exceed 1.0, however for values that are less than one, those closer to 1.0 indicate better responsiveness.

Results

Thirty-three subjects with acute hemiparesis were included in the study. By the 3 and 6 month follow up visits, 28 and 19 subjects remained, respectively. Subject attrition by the 3 month time point was attributed to 1 deceased, 2 with additional medical complications that prohibited them from continuing, and 2 who could not be reached for follow-up. Subject attrition by the 6 month time point was attributed to 4 with additional medical complications that prohibited them from continuing, 4 whose original testing paradigm only included 1 and 3 month visits and did not include a 6 month follow-up visit, and 1 who could not be reached for follow-up.

Subject demographic information is included in Table 1. The average age of the individuals was 56.9 (10.2) (mean (SD)) years old, and average time since stroke was 18.6 (5.6) days for the first visit, 98.3 (14.9) days for the second visit, and 186.7 (12.3) days for the third visit. Within this sample, there were varying degrees of hemiparesis from nearly complete plegia to just barely detectable paresis as indicated by the range of values for composite strength (Table 1). The subjects had minimal shoulder pain, and minimal spasticity throughout their upper extremity. Only six subjects had impaired joint position sense at the index finger and/or wrist. The type of lesion was ischemic in 97% of the subjects, with the dominant hand affected 52% of the time. Paired t-tests were conducted to determine if differences existed between the subjects who were lost to attrition and those that remained in the study, between 1 month and 3 months, and between 3 months and 6 months. These analyses did not show any differences between the subjects who were lost and those that remained on any of the six clinical tests at either time point.

Clinical Test Scores

The clinical tests of upper extremity function were administered to each subject within 1 month, and at 3 and 6 months post stroke. The group mean, standard deviation, and range for each time point are shown in Table 2. Results of the tests show that the sample had decreased upper

extremity function early after stroke which generally improved over time, as expected. For example, grip strength values represented approximately 30%, 45%, and 47% of the less affected side for each time point, respectively. Subjects improved on all other tests with time with the exception of the SIS-Hand, which showed a decrease in perceived function between the 3 and 6 month time points. By and large however, the scores indicate an overall improvement in upper extremity function, most notably during the first 3 months post stroke.

Clinical Test Correlations

To determine how the tests were related to each other, we conducted correlational analyses at each time point. Our results showed that the clinical tests were moderately to strongly correlated with each other at each time point (Table 3). At the 1 month time point, all correlations were significant at the $p < 0.0033$ level. The highest correlation was between Grip and Pinch (0.92), and the lowest correlation between Pinch and the SIS-Hand (0.54). At the 3 month time point, all correlations were again significant at the $p < 0.0033$ level. The highest correlation at the three month time point was between the Jebsen and the 9HPT (0.97), and the lowest between the ARAT and the SIS-Hand (0.57). At the 6 month time point, 9 correlations were significant at the $p < 0.0033$ level, 3 correlations had values of $p < 0.01$ level, 1 had a value of $p < 0.05$, and 1 had a value of $p > 0.05$ (see Table 3, footnotes). The highest correlation at 6 months continued to be between the Jebsen and the 9HPT (0.97), and the lowest was between Pinch and the SIS-Hand (0.41). Overall, the correlation coefficients between pairs of tests were generally similar at each time point, indicating that the relationships between tests remain stable over the first 6 months post stroke.

Responsiveness

To determine how responsive the tests were to change over time, we used the single population effect size method. A higher effect size (closer to 1.0) is considered more responsive to change. Values for responsiveness between the 1 and 3 month time points are shown in Table 4 (first column), with the highest being the SIS-Hand (1.02), and the lowest being Grip (0.50). Values for responsiveness between the 1 and 6 month time points are shown in Table 4 (second column), with the highest being the SIS-Hand (0.86), and the lowest being Pinch (0.56). The values between each time point are similar indicating that these upper extremity function tests are at least moderately responsive to change in this sample.

Discussion

Correlations between Clinical Tests

We found that the tests were correlated with each other at each time point. Based on the interpretation of a strong correlation being $r > 0.75$, all correlations except that between the Jebsen and Pinch, and most of the pairs which included the SIS-Hand were considered strong or excellent at the time of initial testing. The correlation between the Jebsen and Pinch, and those that included the SIS-Hand were considered to be fair to moderate. At the 3 month time point, all of the tests were considered at least moderately related to each other. And at 6 months, the relationships became somewhat more variable with pinch strength and SIS-Hand having weaker relationships with the other tests. These data extend the available information regarding relationships between upper extremity function tests by investigating how six commonly used measures are related to each other early after stroke and during the first 6 months of recovery. Below, we discuss how our results relate to previously published data.

Some of our correlational data were similar to those published by others. The relationship between Grip strength and the 9HPT has been reported to be 0.71 and 0.79 at 1 and 6 months respectively⁴⁶. Relationships between a modified version of the Jebsen test with the 9HPT

have been reported to be between 0.86 and 0.88³⁸. Finally, in previously published data from our lab with a sample of individuals with chronic hemiparesis, using the same tests, we reported similar correlations between all tests except for the SIS-Hand⁸. Despite differences in samples and time elapsed since stroke, the converging evidence of our data with other published results confirms that we are getting closer to the true correlational values.

Our correlational data were different from that of Bovend'Eerd et al. (2004). They reported the relationships between Grip strength and the 9HPT to be between 0.39 and 0.43, and the modified Jebsen and Grip strength to be between 0.53 and 0.44³⁸. Also the values reported by Lang and Beebe (2007) were somewhat higher between the SIS-Hand and all the other clinical tests⁸. A few methodological differences exist between our study and those listed above which may account for some of the discrepancies. One obvious difference is that we used 6 Jebsen subtests, Bovend'Eerd et al. (2004) used 3: flipping index cards, stacking checkers, and simulated feeding. Another difference is the populations studied^{8, 38}. Our sample included people with hemiparesis early after stroke. Bovend'Eerd et al.'s sample included various neurological disorders ranging from stroke to multiple sclerosis to head injury to tumor³⁸, and our other study was done with people with chronic (greater than 6 months) stroke⁸. Also, it is interesting to note in the present study that the SIS-Hand was the only one of the test battery to show a decline in the average scores from 3 to 6 months. We speculate that a longer time period post stroke without full return led to lower scores on the self report SIS-Hand while the more objective measures showed improvements. This may account for the lower correlations between tests in our study. These substantial differences between studies could possibly account for the variations in the strength of the relationships.

Our finding that the correlations fluctuate only slightly at each time point suggests that the relationships between tests are relatively stable over the first 6 months of recovery. While other studies have investigated relationships between 1 or 2 tests over time⁴⁶, our results show that the relationships between a larger number of standardized upper extremity tests is stable over the course of recovery. Our data suggests that if a person with stroke attains a high score/better times for one test, they could expect to also attain high scores/better times on the other tests.

Responsiveness of the Clinical Tests

We found that all tests were responsive to change during the first 6 months after stroke. Between 1 and 3 months the SIS-Hand was considered highly responsive to change and all of the other tests were moderately responsive to change. Though the values changed a little, this same pattern existed between 1 and 6 months. As mentioned in the Introduction, the investigation of the responsiveness of these tests is in its infancy, especially in this population. Our data add to this literature by examining multiple tests simultaneously, and permitting comparison of test responsiveness within the same sample.

Of the six clinical tests examined in this study, responsiveness values for people post stroke have been published for only the ARAT and the SIS-Hand^{15–17, 29}. For the ARAT, published responsiveness values ranged from 0.51²⁹ to 1.02¹⁷. The results from the present study are within this range, and differences between values may arise from differences in time points of testing. For the SIS-Hand, the previously published data for responsiveness used different statistical methodologies but was determined to be responsive to change¹⁸. Unfortunately, due to the methodological differences no direct comparison of values with the present study is possible. We found it particularly interesting that the SIS-Hand was the most responsive to change in our sample. This 5-question instrument was more sensitive to change than other, longer tests that required performance of upper extremity movements. Because it is a patient-focused functional assessment, quick to administer, and sensitive to change, the SIS-Hand is an excellent tool for evaluating upper extremity clinical outcomes^{19, 47}.

Grip and Pinch strength were the least responsive to change at each time point respectively. This may be because, despite our choice to consider these tests as surrogate measures for functional outcomes²⁰, these are impairment level measures versus activity level measures. Responsiveness of these tests however, was not largely different from the other more time consuming, and/or better constructed measures of upper extremity activity. Thus, they continue to have high clinical utility.

Limitations of our study

There are two primary limitations of our study. First, our sample size was small ($n = 33$) for the 1 month time point and decreased at the 3 month ($n = 28$), and 6 month ($n = 19$) time points. While we determined that the lost subjects did not differ statistically from the ones that remained, we cannot rule out the possibility that results could have been different had we been able to maintain the original sample size. The reduced sample size also reduced statistical power in the correlational analyses. Despite this, most of the correlations continued to be significant even after strict Bonferroni corrections for multiple comparisons. Still, further investigation with a larger sample size is needed to confirm these results. Second, our inclusion/exclusion criteria restricted our sample to those primarily with motor deficits. Like most other studies of motor function after stroke, we did not recruit individuals who had aphasia, neglect, or who could not give informed consent. Thus our results should be interpreted with caution with regard to individuals with deficits in multiple, non-motor domains.

Implications

Our results are important for clinicians and researchers working with individuals with a recent onset of stroke. Our data can be used to help make decisions about which tests to choose when measuring upper extremity function in groups or in individual patients. Studying six tests with published reliability and validity values, we have shown that these tests yield scores that are highly correlated, and that the tests are moderately to strongly responsive to change over the course of recovery. Because of these findings, we suggest that when selecting upper extremity function tests for individual patients, the circumstances can dictate the test choice, instead of any one test being considered the “gold standard”. In other words, the choice of which test to use could be made based on the individual patient and what is available in the therapist’s practice setting.

Below we discuss some potential situations and/or limitations we have identified for using each test. If there is a need to make frequent assessments, choosing Grip or Pinch strength may be a good option. There is no need to perform both tests because of strong relationships between them ($r = 0.83$ to 0.92). From our experience, most clinics have grip or pinch dynamometers available for use. These tests are quick and easy to administer. The limitation to using Grip or Pinch is that they are impairment level measures which may be useful for more severely affected patients, but can only provide indirect information on the broader range of upper extremity movement capabilities.

If the goal is to obtain a more comprehensive assessment of upper extremity movement capabilities, then choosing the ARAT or the Jebsen test may be a preferred option. Both of these tests evaluate multiple tasks of varying complexity and require controlled movement of the entire upper extremity. Some potential drawbacks about the ARAT and the Jebsen are that they take more time to administer and require specific testing equipment that may not be available in a particular clinic. The Jebsen test is available commercially, but the ARAT must be created/manufactured. Building an ARAT is feasible as the specifications have recently been published⁴⁸. There is no need to perform both tests because of strong relationships between them ($r = 0.87$ to 0.95).

Finally, the 9HPT and SIS-Hand are good choices because they do not take much time to administer and are easy to obtain. The SIS-Hand can be procured from the internet¹⁸, and could be given to a patient to fill out while they sit in the waiting room⁴⁰. It is useful to note that this test is valid as an interview, filled out independently, and from caregiver report⁴⁹. The 9HPT does not take long to administer, can be purchased as a kit, and may be a good choice for use with a patient where additional information about fine motor dexterity is needed. Whichever test the clinician or researcher decides to use, it is important that the same measure is used at each follow-up time point because only through consistent use of the same test, within the same patient, can change be measured over time.

Conclusions

This study provides data to assist clinicians and researchers in making decisions to use specific tests for measuring upper extremity function in people with hemiparesis in the first 6 months of recovery. Our data establish the relationships between tests, and their responsive to change when used in people with hemiparesis early after stroke. From this information, choices can be made based on individual patient, clinic, or laboratory specific needs.

Acknowledgments

We thank Lucy Morris, MD, MPH, for her assistance with lesion data analysis, and Dustin Hardwick, DPT, for his assistance with data collection. Subject recruitment was facilitated by the Stroke Registry supported by the James S. McDonnell Foundation 21002032. Supported by National Institutes of Health, and the National Institute of Child Health and Human Development HD047669, HD007434, and Foundation for Physical Therapy Promotion of Doctoral Studies II Scholarship.

Reference List

1. Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, Hailpern SM, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell C, Roger V, Sorlie P, Steinberger J, Thom T, Wilson M, Hong Y. Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 2008 January 29;117(4):e25–e146. [PubMed: 18086926]
2. Parker VM, Wade DT, Langton HR. Loss of arm function after stroke: measurement, frequency, and recovery. *Int Rehabil Med* 1986;8(2):69–73. [PubMed: 3804600]
3. Bard G, Hirschberg GG. Recovery of voluntary motion in upper extremity following hemiplegia. *Arch Phys Med Rehabil* 1965 August;46:567–572. [PubMed: 14340366]
4. Feys H, De WW, Nuyens G, van de WA, Selz B, Kiekens C. Predicting motor recovery of the upper limb after stroke rehabilitation: value of a clinical examination. *Physiother Res Int* 2000;5(1):1–18. [PubMed: 10785907]
5. Olsen TS. Arm and leg paresis as outcome predictors in stroke rehabilitation. *Stroke* 1990 February; 21(2):247–251. [PubMed: 2305400]
6. Twitchell TE. The restoration of motor function following hemiplegia in man. *Brain* 1951 December; 74(4):443–480. [PubMed: 14895765]
7. Duncan PW, Goldstein LB, Horner RD, Landsman PB, Samsa GP, Matchar DB. Similar motor recovery of upper and lower extremities after stroke. *Stroke* 1994 June;25(6):1181–1188. [PubMed: 8202977]
8. Lang CE, Beebe JA. Relating movement control at 9 upper extremity segments to loss of hand function in people with chronic hemiparesis. *Neurorehabil Neural Repair* 2007 May;21(3):279–291. [PubMed: 17353458]
9. Beebe JA, Lang CE. Absence of a proximal to distal gradient of motor deficits in the upper extremity early after stroke. *Clin Neurophysiol* 2008;119(9):2074–2085. [PubMed: 18571981]
10. Guyatt G, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40(2):171–178. [PubMed: 3818871]

11. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989 March;27(3 Suppl):S178–S189. [PubMed: 2646488]
12. Wallace D, Duncan PW, Lai SM. Comparison of the responsiveness of the Barthel Index and the motor component of the Functional Independence Measure in stroke: the impact of using different methods for measuring responsiveness. *J Clin Epidemiol* 2002 September;55(9):922–928. [PubMed: 12393081]
13. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil* 2008 January;89(1):69–74. [PubMed: 18164333]
14. Liang MH, Lew RA, Stucki G, Fortin PR, Daltroy L. Measuring clinically important changes with patient-oriented questionnaires. *Med Care* 2002 April;40(4 Suppl):II45–II51. [PubMed: 12064581]
15. Hsueh IP, Hsieh CL. Responsiveness of two upper extremity function instruments for stroke inpatients receiving rehabilitation. *Clin Rehabil* 2002 September;16(6):617–624. [PubMed: 12392336]
16. Rabadi MH, Rabadi FM. Comparison of the action research arm test and the Fugl-Meyer assessment as measures of upper-extremity motor weakness after stroke. *Arch Phys Med Rehabil* 2006 July;87(7):962–966. [PubMed: 16813784]
17. Lang CE, Wagner JM, Dromerick AW, Edwards DF. Measurement of upper-extremity function early after stroke: properties of the action research arm test. *Arch Phys Med Rehabil* 2006 December;87(12):1605–1610. [PubMed: 17141640]
18. Duncan PW, Wallace D, Lai SM, Johnson D, Embretson S, Laster LJ. The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke* 1999 October;30(10):2131–2140. [PubMed: 10512918]
19. Barreca SR, Stratford PW, Lambert CL, Masters LM, Streiner DL. Test-retest reliability, validity, and sensitivity of the Chedoke arm and hand activity inventory: a new measure of upper-limb function for survivors of stroke. *Arch Phys Med Rehabil* 2005 August;86(8):1616–1622. [PubMed: 16084816]
20. Boissy P, Bourbonnais D, Carlotti MM, Gravel D, Arsenault BA. Maximal grip force in chronic stroke subjects and its relationship to global upper extremity function. *Clin Rehabil* 1999 August;13(4):354–362. [PubMed: 10460123]
21. De Weerd W, Harrison MA. Measuring recovery of arm-hand function in stroke patients: a comparison of the Brunnstrom-Fugl-Meyer test and the action research arm test. *Physiotherapy Canada* 1985;37(2):65–70.
22. Schmidt RT, Toews JV. Grip strength as measured by the Jamar dynamometer. *Arch Phys Med Rehabil* 1970 June;51(6):321–327. [PubMed: 5423802]
23. Bertrand AM, Mercier C, Bourbonnais D, Desrosiers J, Gravel D. Reliability of maximal static strength measurements of the arms in subjects with hemiparesis. *Clin Rehabil* 2007 March;21(3):248–257. [PubMed: 17329282]
24. Mathiowetz V, Kashman N, Volland G, Weber K, Dowe M, Rogers S. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil* 1985 February;66(2):69–74. [PubMed: 3970660]
25. Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg* 1984 March;9(2):222–226.
26. Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. *Int J Rehabil Res* 1981;4(4):483–492. [PubMed: 7333761]
27. Hsieh CL, Hsueh IP, Chiang FM, Lin PH. Inter-rater reliability and validity of the action research arm test in stroke patients. *Age Ageing* 1998 March;27(2):107–113. [PubMed: 16296669]
28. Van der Lee JH, De G, V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. *Arch Phys Med Rehabil* 2001 January;82(1):14–19. [PubMed: 11239280]
29. Van der Lee JH, Beckerman H, Lankhorst GJ, Bouter LM. The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. *J Rehabil Med* 2001 March;33(3):110–113. [PubMed: 11482350]
30. Lang CE, Edwards DF, Birkenmeier RL, Dromerick AW. Estimating minimal clinically important differences of upper extremity measures early after stroke. *Arch Phys Med Rehabil*. 2008 In Press.
31. Platz T, Pinkowski C, van WF, Kim IH, di BP, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and

- Box and Block Test: a multicentre study. *Clin Rehabil* 2005 June;19(4):404–411. [PubMed: 15929509]
32. Jebsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. *Arch Phys Med Rehabil* 1969 June;50(6):311–319. [PubMed: 5788487]
 33. Hummel F, Celnik P, Giraux P, Floel A, Wu WH, Gerloff C, Cohen LG. Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke. *Brain* 2005 March;128(Pt 3):490–499. [PubMed: 15634731]
 34. Duncan PW, Richards L, Wallace D, Stoker-Yates J, Pohl P, Luchies C, Ogle A, Studenski S. A randomized, controlled pilot study of a home-based exercise program for individuals with mild and moderate stroke. *Stroke* 1998;29:2055–2060. [PubMed: 9756581]
 35. Agnew PJ, Maas F. Hand function related to age and sex. *Arch Phys Med Rehabil* 1982 June;63(6):269–271. [PubMed: 7082153]
 36. Hackel ME, Wolfe GA, Bang SM, Canfield JS. Changes in hand function in the aging adult as determined by the Jebsen Test of Hand Function. *Phys Ther* 1992 May;72(5):373–377. [PubMed: 1631206]
 37. Mathiowetz V, Weber K, Kashman N, Volland G. Adult norms for the nine hole peg test of finger dexterity. *OTJR* 1985;5(1):24–38.
 38. Bovend'Eerdt TJ, Dawes H, Johansen-Berg H, Wade DT. Evaluation of the Modified Jebsen Test of Hand Function and the University of Maryland Arm Questionnaire for Stroke. *Clin Rehabil* 2004 March;18(2):195–202. [PubMed: 15053129]
 39. Duncan PW, Wallace D, Studenski S, Lai SM, Johnson D. Conceptualization of a new stroke-specific outcome measure: the stroke impact scale. *Top Stroke Rehabil* 2001;8(2):19–33. [PubMed: 14523743]
 40. Duncan P, Reker D, Kwon S, Lai SM, Studenski S, Perera S, Alfrey C, Marquez J. Measuring stroke impact with the stroke impact scale: telephone versus mail administration in veterans with stroke. *Med Care* 2005 May;43(5):507–515. [PubMed: 15838417]
 41. Andrews AW, Thomas MW, Bohannon RW. Normative values for isometric muscle force measurements obtained with hand-held dynamometers. *Phys Ther* 1996 March;76(3):248–259. [PubMed: 8602410]
 42. Wagner JM, Lang CE, Sahrman SA, Edwards DF, Dromerick AW. Sensorimotor impairments and reaching performance in subjects with poststroke hemiparesis during the first few months of recovery. *Phys Ther* 2007 June;87(6):751–765. [PubMed: 17442839]
 43. Bohannon RW, Smith MB. Interrater reliability of a modified Ashworth scale of muscle spasticity. *Physical Therapy* 1987 February;67(2):206–207. [PubMed: 3809245]
 44. Portney, LG.; Watkins, MP. *Foundations of Clinical Research: Applications to Clinical Practice*. 1st ed.. Norwalk, CT: Appleton & Lange; 1993.
 45. Wanders AJ, Gorman JD, Davis JC, Landewe RB, van der Heijde DM. Responsiveness and discriminative capacity of the assessments in ankylosing spondylitis disease-controlling antirheumatic therapy core set and other outcome measures in a trial of etanercept in ankylosing spondylitis. *Arthritis and Rheum* 2004 February 15;51(1):1–8. [PubMed: 14872448]
 46. Sunderland A, Tinson D, Bradley L, Hower RL. Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. *J Neurol Neurosurg Psychiatry* 1989 November;52(11):1267–1272. [PubMed: 2592969]
 47. Rudman D, Hannah S. An instrument evaluation framework: description and application to assessments of hand function. *J Hand Ther* 1998 October;11(4):266–277. [PubMed: 9862265]
 48. Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the action research arm test. *Neurorehabil Neural Repair* 2008 January;22(1):78–90. [PubMed: 17704352]
 49. Duncan PW, Lai SM, Tyler D, Perera S, Reker DM, Studenski S. Evaluation of proxy responses to the Stroke Impact Scale. *Stroke* 2002 November;33(11):2593–2599. [PubMed: 12411648]

Table 1

Subject Characteristics

	mean (SD)	range/percent
Age (years)	56.9 (10.2)	31–77
Gender	19 males 14 females	58% 42%
Time since stroke 1 st visit (days)	18.6 (5.6)	8–30
Time since stroke 2 nd visit (days)	98.3 (14.9)	78–138
Time since stroke 3 rd visit (days)	186.7 (12.3)	159–208
Hand affected	14 left 19 right	42% 58%
Type of stroke	32 ischemic 1 hemorrhagic	97% 3%
Dominant hand affected	17	52%
Composite strength ^a	0.35 (0.36)	0 – 1
Shoulder pain	0.76 (1.64)	0 – 6
Index finger joint position sense	27 normal 6 impaired	90% 10%
Modified Ashworth Scale ^b	0.58 (0.67)	0 – 3

Values are means $\pm$ SD, and (range) or number (%).

^aExpressed as a ratio of more affected to less affected side strength.

^bModified Ashworth Scale at 4 joints, where score = average of 4 sites and normal = 0.

Table 2

Clinical Test Raw Scores

	mean (SD) range	mean (SD) range	mean (SD) range
	1 month (n = 33)	3 months (n = 28)	6 months (n = 19)
Grip Strength (kg)	9.2 (9.6) 0 – 27	14.0 (10.3) 0 – 33	15.4 (11.4) 0 – 37
Pinch Strength (kg)	3.5 (3.5) 0 – 10	5.3 (3.3) 0 – 12	5.5 (3.5) 0 – 12
ARAT	26.4 (23.9) 0 – 57	39.5 (19.7) 0 – 57	41.3 (20.8) 0 – 57
Jebsen (sec)	418 (316) 720 – 38	200 (250) 661 – 28	189 (254) 441 – 30
9HPT (sec)	88.8 (40.2) 120 – 23	67.8 (41.7) 120 – 20	60.8 (39.7) 120 – 19
SIS: Hand function	19.9 (28.0) 0 – 85	48.4 (32.7) 0 – 100	43.9 (34.2) 0 – 100

Values are means $\pm$ SD, and range of scores for the more affected upper extremity.

Range of scores is listed where poorer performance is first, and better performance is second. The Jebsen and 9HPT are timed tests, therefore a lower number is considered better performance.

Table 3

Spearman's (r_s) Clinical Test Correlations

	Grip	Pinch	ARAT	Jebson	9HPT	Time point
Pinch	0.92					1 month
	0.89					3 months
	0.83					6 months
ARAT	0.86	0.79				“
	0.73	0.80				
	0.80	0.65				
Jebson	0.79	0.72	0.87			“
	0.79	0.79	0.95			
	0.81	0.60*	0.90			
9HPT	0.80	0.77	0.87	0.84		“
	0.78	0.78	0.93	0.97		
	0.82	0.61*	0.85	0.97		
SIS-Hand	0.61	0.54	0.73	0.83	0.66	“
	0.65	0.69	0.57	0.66	0.62	
	0.53 [‡]	0.41 [‡]	0.69	0.61*	0.53 [‡]	

Correlation coefficients are expressed as absolute values.

All correlation coefficients at 1, and 3 months are statistically significant at the Bonferroni corrected level of $p < 0.0033$.

* $p < 0.01$

[‡] $p < 0.05$

[‡] $p > 0.05$

Table 4

Responsiveness

	Responsiveness 1–3 months	Responsiveness 1–6 months
Grip	0.50	0.65
Pinch	0.52	0.56
ARAT	0.55	0.63
Jebsen	0.69	0.73
9HPT	0.52	0.66
SIS-Hand	1.02	0.86

Responsiveness was calculated using the single population effect size method. Values closer to 1.00 = more responsive to change. Low responsiveness < 0.20; moderate responsiveness < 0.50, and high responsiveness < 0.80⁴⁵.