Department of Veterans Affairs

JRRD Volume 47, Number 6, 2010
Pages 563–572
Journal of Rehabilitation Research & Development

Responsiveness and validity of three dexterous function measures in stroke rehabilitation

Keh-chung Lin, ScD, OTR;[1–2] Li-ling Chuang, PhD, PT;[1] Ching-yi Wu, ScD, OTR;[3*] Yu-wei Hsieh, MS;[1] Wan-ying Chang, BS[4]

[1]School of Occupational Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan; [2]Division of Occupational Therapy, Department of Physical Medicine and Rehabilitation, National Taiwan University Hospital, Taipei, Taiwan; [3]Graduate Institute of Clinical Behavioral Science and Department of Occupational Therapy, Chang Gung University, Taoyuan, Taiwan; [4]Rehabilitation Department, Taipei Hospital, Department of Health, Taipei, Taiwan

Abstract—In this study, we compared the responsiveness and validity of the Box and Block Test (BBT), the Nine-Hole Peg Test (NHPT), and the Action Research Arm Test (ARAT). We randomized 59 patients with stroke into one of three rehabilitation treatments for 3 weeks. We administered six outcome measures (BBT, NHPT, ARAT, Fugl-Meyer Assessment [FMA], Motor Activity Log [MAL], and Stroke Impact Scale [SIS] hand function domain) pretreatment and posttreatment. We used the standardized response mean (SRM) to examine responsiveness and the Spearman rank correlation coefficient (rho) to examine concurrent validity. The BBT, NHPT, and ARAT were moderately responsive to change and not significantly different (SRM = 0.64–0.79). The correlations within the BBT, NHPT, and ARAT were moderate to good at pretreatment (rho = –0.55 to –0.80) and posttreatment (rho = –0.57 to –0.71). The BBT and ARAT showed fair to moderate correlations with the FMA, MAL, and SIS hand function domain at pretreatment and posttreatment (rho = 0.31–0.59), whereas the NHPT demonstrated low to fair correlations with the FMA and MAL (rho = –0.16 to –0.33) and moderate correlations with the SIS hand function domain (rho = –0.58 to –0.66). Our results indicate that the BBT, NHPT, and ARAT are suitable to detect changes over time. While simultaneously considering the responsiveness and validity attributes, the BBT and ARAT can be considered more appropriate for evaluating dexterous function than the NHPT. Further studies with larger samples are needed to validate these findings.

Clinical Trial Registration: ClinicalTrials.gov; Relative Effects and Predictive Models of Contemporary Upper Limb Training Programs in Stroke Patients; NCT00778453; http://clinicaltrials.gov/ct2/show/NCT00778453?term=NCT00778453&rank=1.

Key words: bootstrapping, cerebrovascular accident, clinimetrics, dexterity, function, outcome, rehabilitation, responsiveness, upper limb, validity.

Abbreviations: ADL = activity of daily living, AOU = amount of use, ARAT = Action Research Arm Test, BAT = bilateral arm training, BBT = Box and Block Test, CIT = constraint-induced therapy, dCIT = distributed CIT, FMA = Fugl-Meyer Assessment, MAL = Motor Activity Log, NHPT = Nine-Hole Peg Test, SIS = Stroke Impact Scale, SRM = standardized response mean, UL = upper limb, WMFT = Wolf Motor Function Test.

[*]Address all correspondence to Ching-yi Wu, ScD, OTR; Department of Occupational Therapy, Chang Gung University, 259 Wen-Hwa 1st Road, Kwei-Shan, Tao-Yuan 333, Taiwan; 886-3-2118800, ext 5761; fax: 866-3-2118800, ext 3163. Email: cywu@mail.cgu.edu.tw

DOI:10.1682/JRRD.2009.09.0155

INTRODUCTION

The paretic hand is a common motor impairment after stroke. At 6 months poststroke, 38 percent of patients regained some dexterity in manipulative tasks but only 11.6 percent reached a complete functional recovery in dexterity of the paretic hand [1]. Dexterous hands are important for most activities of daily living (ADLs), such as food preparation and grooming. Loss of hand dexterity might limit ADL performance and social participation and thus reduce quality of life in patients with stroke [2]. The recovery of the use of a paretic hand is the main determinant of functional improvement of the affected arm [2]. Therefore, improvements in dexterous function to promote functional recovery are major goals of stroke rehabilitation.

Evidence suggests that constraint-induced therapy (CIT) [3] and bilateral arm training (BAT) [4] improve function of the upper limb (UL) after stroke, but clinicians and researchers need to identify appropriate measures that have sound clinimetric properties (i.e., reliability, validity, and responsiveness) to determine the effects of UL training on dexterity of the paretic hand. Use of appropriate measures for outcome evaluations would enhance the methodologic quality of controlled trials in stroke rehabilitation research [5]. Clinimetric research is needed to identify a sound measure of hand dexterity to facilitate interpretation and comparison of the results of controlled trials.

In the past 5 years, a number of UL function tests have been examined for their psychometric and clinimetric properties in people with stroke, including the Fugl-Meyer Assessment (FMA) [6–8], Wolf Motor Function Test (WMFT) [6,9–11], Box and Block Test (BBT) [7,12–14], Nine-Hole Peg Test (NHPT) [12,15–17], Action Research Arm Test (ARAT) [6–8,11,16–18], and Stroke Impact Scale (SIS) [16,19–20]. The comparisons of outcome measures of UL motor function have been investigated in patients with stroke [6–7,16,21]. The findings of Van der Lee et al. suggest that the ARAT is more responsive to improvement in UL function than the FMA in patients with chronic stroke undergoing forced-use treatment [21]. In contrast, Hsieh et al. reported that the FMA may be a better measure of motor function with respect to responsiveness and validity than the ARAT and WMFT [6]. Differences in the usage of responsiveness indices, training paradigms, and time after onset of stroke may contribute to the conflicting results among studies.

Among the UL function tests, the BBT [7,12–14,22–23], NHPT [5,12,14,17,24–26], and ARAT [1,5,7,16–17,27] are three commonly used clinical measures to assess hand dexterous function in patients with stroke. Limited research directly compared the clinimetric properties between pairs of the BBT, NHPT, and ARAT in patients with stroke [7,12,14,16].

- Higgins et al. evaluated the efficacy of a task-oriented intervention in enhancing arm function and concluded that the BBT was substantially more responsive to improvement in UL function than the NHPT for patients within 1 year of a first or recurrent stroke [14].
- Chen et al. investigated the test-retest reproducibility and smallest real difference of five hand function domain tests (BBT, NHPT, grip strength, palmar pinch strength, and lateral pinch strength) in 62 patients with stroke [12]. All five tests demonstrated satisfactory test-retest reproducibility, but levels of measurement error were higher for patients with hypertonicity of the affected hand.
- Desrosiers et al. validated the BBT as a measure of dexterity in 34 elderly patients with UL sensorimotor impairments from stroke and other conditions, including rheumatoid polyarthritis, osteoarthritis, multiple sclerosis, Parkinson disease, fracture, Friedreich ataxia, and carpal tunnel syndrome [23]. They reported that the BBT had high test-retest reliability and good construct validity, as shown by significant correlations between the BBT, ARAT, and functional independence measures.
- Platz et al. estimated the reliability and validity of the BBT, ARAT, and FMA UL subscale in 56 participants with UL paresis from stroke, multiple sclerosis, or traumatic brain injury [7]. The BBT, ARAT, and FMA had very high interrater and test-retest reliability, and high correlations were demonstrated among the three tests.
- Beebe and Lang examined the relationships and responsiveness of six UL function tests (ARAT, NHPT, grip strength test, pinch strength test, Jebsen-Taylor Hand Function Test, and SIS hand function domain) during the first 6 months poststroke [16]. All tests, including the ARAT and NHPT, were correlated with each other and moderately responsive to change over time.

In general, the relative capacity among the BBT, NHPT, and ARAT to detect a change in dexterous function in patients with chronic stroke after interventions remains unclear. No consensus exists on the optimal dexterous function measure for stroke intervention trials. Gaining knowledge about whether the BBT, NHPT, and ARAT are responsive to change after stroke interventions and how tests are related to each other and to other well developed measures over time would be crucial and beneficial to clinical practice and sound research. Therefore, the goal of our study was to investigate and compare the relative responsiveness and the concurrent validity of the BBT, NHPT, and ARAT for patients after stroke rehabilitation interventions.

METHODS

Participants

We randomized a cohort of 59 eligible participants into distributed CIT (dCIT), BAT, or control treatment for 2 hours per weekday for 3 weeks. The inclusion criteria comprised a first-ever stroke onset at least 6 months previously, demonstration of Brunnstrom stage IV to VI for the proximal and distal UL, no excessive UL spasticity (Modified Ashworth Scale score $\leq$2.5), and no cognitive impairment (Mini-Mental State Examination score $\geq$23).

Interventions and Procedures

Therapy in the dCIT group involved placing the unaffected hand in a mitt to restrict movement and intensive training of the affected UL in functional tasks, including reaching forward or upward to move a cup, picking up coins or a utensil to eat, and other functional movements simulating ADLs. To discourage use of the unaffected hand outside of therapy sessions, participants wore a restrictive mitt for a target of 6 hours per day during the treatment period. The BAT group concentrated on both ULs moving simultaneously in functional tasks by symmetrical or alternated patterns. The functional tasks emphasized both ULs moving synchronously, such as lifting two cups, picking up two pegs, reaching forward or upward to move two blocks, and grasping and releasing two towels. Control treatment applied neurodevelopmental techniques with emphasis on functional tasks. Therapy included stretching of the affected limb and training for strength, hand function, and coordination, as well as practicing functional tasks.

We provided the interventions at rehabilitation clinics under the supervision of three certified occupational therapists who were trained in the administration of the intervention protocols. We assessed the therapists and administered a written competency test before the therapists treated and assessed the participants. Three raters masked to the participant group and trained to properly administer the outcome measures administered outcome evaluations pre- and posttreatment.

Outcome Measures

Main

The three main outcome measures we used for evaluating hand dexterity were the BBT, NHPT, and ARAT. The BBT assesses gross manual dexterity by counting the number of blocks that can be transported individually from one compartment of a box to another within 1 minute. Higher scores are indicative of better manual dexterity [28]. The reliability, validity, and responsiveness of the BBT have been established in patients with stroke [7,12–13,23,28]. The NHPT is a timed test of fine manual dexterity [29]. Participants place nine pegs in nine holes and then remove them as quickly as possible. The time needed to complete the task is measured in seconds, and a lower score indicates better dexterity. The NHPT has been demonstrated to have high reliability, validity, and responsiveness in patients with stroke [12,24,29]. The ARAT assesses the ability to handle objects, with 19 items divided into 4 subscales of grasp, grip, pinch, and gross movement by using a 4-level ordinal scale [30] ranging from 0 (no movement) to 3 (normal movement). A total scale score maximum of 57 indicates normal performance. The ARAT has established reliability, validity, and responsiveness in patients with stroke [6–8,21].

Secondary

We administered three other well developed outcome measures—the FMA, the Motor Activity Log (MAL), and the SIS—to test the concurrent validity of the BBT, NHPT, and ARAT. Improving motor impairment, increasing daily function, and enhancing quality of life are the primary goals of stroke rehabilitation. The FMA assesses motor impairment, the MAL evaluates daily function, and the SIS determines quality of life; all have adequate reliability and validity and are widely used measures in stroke rehabilitation [6–7,31–33]. Thus, we considered the FMA, MAL, and SIS appropriate for validating the three main outcome measures.

We used the UL subscale of the FMA to assess motor impairment with a 3-point ordinal scale (0 = cannot perform, 1 = performs partially, 2 = performs fully) [31]. The 33 UL items measure the movement and reflexes of the shoulder, elbow, forearm, wrist, and hand and measure coordination and speed. The psychometric properties of the FMA have been established [6–8].

The MAL is a focal measure of self-perceived daily functions on tasks requiring UL use. It is a semistructured interview for assessing how much (amount of use [AOU]) and how well (quality of movement) patients use the affected UL. The MAL evaluates 30 tasks using a 6-point scale and scores each item ranging from 0 to 5 [32]. Adequate interrater reliability and internal consistency have been reported [32].

The SIS is a self-report questionnaire and a comprehensive measure of health-related quality of life in populations with stroke [33]. SIS version 3 includes 59 items that assess the effect of stroke in 8 functional domains by using a 5-point Likert scale. Test-retest reliability has been established for the hand function domain in patients with stroke [19]. We collected data for all domains of the SIS but only used the hand function domain subscale to investigate the validity of the three main outcome measures. Questions in the hand function domain assess the ability to carry heavy objects, turn a doorknob, open a can or jar, tie a shoelace, and pick up a dime.

Data Analysis

Examination of Responsiveness

We used two indices to compare the responsiveness of the three main outcome measures to change from pretreatment to posttreatment. The Wilcoxon matched-pairs signed rank test focuses on the statistical significance of the observed change in the measures. The standardized response mean (SRM), a variant of effect size, is the mean change in the score divided by the standard deviation of the change scores [34]. According to Cohen's criteria on effect size [34], 0.8 is large, 0.5 to 0.8 is moderate, and 0.2 to 0.5 is small. The 95 percent confidence intervals for the SRMs were estimated using 1,000 bootstrap samples with replacement [35]. To test for the differences of SRMs between two measures, we sorted the differences in SRMs for 1,000 paired bootstrap samples of two measures from lowest to highest. We then examined whether the value 0 was included between the 25th and 975th observations; if not, a significant difference exists between measures [36].

Examination of Validity

We studied concurrent validity to validate the three dexterous measures with each other and with the FMA, MAL, and SIS hand function domain obtained concurrently [37]. We used bivariate correlational analysis to examine the concurrent validity of the BBT, NHPT, and ARAT. We used the Spearman rank correlation coefficient (ρ) to examine the relation between the six outcome measures. We performed analyses separately pre- and posttreatment. We used the following criteria to interpret the magnitude of the correlation coefficients: <0.25 indicated low, 0.25 to 0.5 indicated fair, 0.5 to 0.75 indicated moderate to good, and >0.75 indicated good to excellent [37].

RESULTS

Table 1 lists the demographic and clinical characteristics of the 59 participants. Their mean age was 55.50 years and 79.7 percent were male. **Table 2** reports

Table 1.
Demographics and clinical characteristics of participants (n = 59).

Characteristic	Value
Age (years, mean ± SD)	55.50 ± 11.66
Sex, *n* (%)	
Female	12 (20.3)
Male	47 (79.7)
Side of Lesion, *n* (%)	
Left	28 (47.5)
Right	31 (52.5)
Time Since Stroke (months, mean ± SD)	16.14 ± 13.95
Brunnstrom Stage	
Proximal/Distal UL (median)	5/5
Mini-Mental State Examination (mean ± SD)	28.39 ± 1.52
Pretreatment Evaluations (mean ± SD)	
BBT	26.64 ± 11.76
NHPT	136.32 ± 99.17
ARAT	45.66 ± 10.40
FMA	50.93 ± 7.24
MAL-AOU	1.25 ± 0.94
MAL-QOM	1.38 ± 0.98
SIS Hand Function Domain	45.17 ± 27.75

ARAT = Action Research Arm Test, BBT = Box and Block Test, FMA = Fugl-Meyer Assessment, MAL-AOU = Motor Activity Log-Amount of Use, MAL-QOM = Motor Activity Log-Quality of Movement, NHPT = Nine-Hole Peg Test, SD = standard deviation, SIS = Stoke Impact Scale, UL = upper limb.

Table 2.
Responsiveness indices of three outcome measures.

Scale	SRM (95% CI)	Wilcoxon Test *Z*-value
BBT	0.74 (0.51–1.10)	5.12
NHPT	0.64 (0.41–0.86)	4.77
ARAT	0.79 (0.63–1.10)	5.76

Note: $p < 0.001$.
ARAT = Action Research Arm Test, BBT = Box and Block Test, CI = confidence interval, NHPT = Nine-Hole Peg Test, SRM = standardized response mean.

the responsiveness indices of the three outcome measures. The responsiveness of the BBT, NHPT, and ARAT was moderate from pretreatment to posttreatment (SRM = 0.64–0.79; Wilcoxon Z = 4.77–5.76; $p < 0.001$), and the difference between them was not significant (difference in SRM = 0.05–0.15; $p > 0.05$).

Table 3 reports concurrent validity of the six outcome measures at pretreatment and posttreatment. The correlations within the BBT, NHPT, and ARAT were moderate to good at pretreatment (ρ = –0.55 to –0.80) and posttreatment (ρ = –0.57 to –0.71). In addition, the BBT and ARAT had fair to moderate correlations with the FMA, MAL, and SIS hand function domain tests at pretreatment (ρ = 0.31–0.59) and posttreatment (ρ = 0.32–0.54). The NHPT showed relatively low to fair correlations with the FMA and MAL at pretreatment (ρ = –0.16 to –0.27) and posttreatment (ρ = –0.18 to –0.33). The correlation between NHPT and the SIS hand function domain was moderate at pretreatment (ρ = –0.58) and posttreatment (ρ = –0.66).

DISCUSSION

We examined and compared the responsiveness and validity of the BBT, NHPT, and ARAT in assessing hand dexterity after stroke rehabilitation. All three tests are moderately responsive to change over time. While simultaneously considering the responsiveness and validity attributes, the BBT and ARAT may be relatively sound dexterous outcome measures in stroke rehabilitation compared with the NHPT.

Given the rapid proliferation of randomized controlled trials in stroke rehabilitation, selecting outcome measures that are responsive, reliable, and valid is crucial. A responsive instrument used as an outcome measure in clinical trials should be able to detect change with improvement or deterioration and distinguish effective or useless treatments [38]. We found that the responsiveness of BBT, NHPT, and ARAT was not significantly different, as indicated by the three dexterous measures being moderately sensitive to detect changes after stroke interventions. The similar degree of responsiveness between the BBT and NHPT in this study differs from the findings of Higgins et al., who concluded that the BBT was substantially more responsive to improvement in UL function than the NHPT for patients with stroke [14]. This difference in findings may be because of the difference in participant age, number of strokes, time after stroke onset, and intervention protocols. We recruited younger patients with a first stroke and onset of at least 6 months for 3-week arm interventions in our study, whereas the patients in the Higgins et al. study were older, were within 1 year of having sustained a first or recurrent stroke, and received 6-week arm and walking interventions [14].

The result of sound responsiveness of the BBT is consistent with the previous study that indicated the BBT was responsive for changes in UL function and recovery >5 weeks for patients with acute stroke [13]. The finding of responsiveness of the NHPT is consistent with the results of previous studies [16,24–26]. The NHPT was moderately responsive to change of UL function during the first 6 months poststroke [16]. It was suggested that the NHPT detected further recovery after patients achieved maximal scores on the Frenchay Arm test [24]. The NHPT was more responsive than the UL Motricity Index for patients with stroke with useful UL function [25]. The good responsiveness of the ARAT is consistent with published responsiveness values that ranged from 0.51 [21] to 1.02 [18]. The result from this study (SRM = 0.79) is within this range, and differences in time points of testing and statistical methodologies could possibly account for variations in values. Previous studies showed that the ARAT had good sensitivity to detect change for patients with acute stroke [8,18], during the first 6 months poststroke [16], and for patients with chronic stroke [6,21].

Our findings that the correlations among the tests fluctuate only slightly at pretreatment and posttreatment suggest that the relationships among the tests are relatively stable over the period of 3-week interventions, which reflects constant and true relationships among the tests. The results of moderate to good correlations within the BBT, NHPT, and ARAT at pretreatment and posttreatment were similar to earlier studies demonstrating significant correlations between the BBT and the ARAT in

Table 3.
Concurrent validity (Spearman rank correlation coefficient) of dexterous function measures and criterion measures at pretreatment and posttreatment.

Scale	Pretreatment			Posttreatment		
	BBT (95% CI)	NHPT (95% CI)	ARAT (95% CI)	BBT (95% CI)	NHPT (95% CI)	ARAT (95% CI)
NHPT	–0.80* (–0.69 to –0.88)	—	—	–0.71* (–0.56 to –0.82)	—	—
ARAT	0.63* (0.45 to 0.76)	–0.55* (–0.34 to –0.71)	—	0.64* (0.46 to 0.77)	–0.57* (–0.37 to –0.72)	—
FMA	0.44* (0.21 to 0.63)	–0.27† (–0.02 to –0.49)	0.49* (0.27 to 0.66)	0.35* (0.10 to 0.56)	–0.18 (–0.42 to 0.08)	0.54* (0.33 to 0.70)
MAL-AOU	0.37* (0.13 to 0.57)	–0.16 (–0.40 to 0.10)	0.31† (0.06 to 0.52)	0.49* (0.27 to 0.66)	–0.23 (–0.46 to 0.03)	0.32† (0.07 to 0.53)
MAL-QOM	0.52* (0.31 to 0.68)	–0.26† (–0.48 to –0.01)	0.39* (0.15 to 0.59)	0.52* (0.31 to 0.68)	–0.33† (–0.54 to –0.08)	0.35* (0.10 to 0.56)
SIS Hand Function Domain	0.59* (0.39 to 0.74)	–0.58* (–0.73 to –0.38)	0.36* (0.12 to 0.56)	0.52* (0.31 to 0.68)	–0.66* (–0.78 to –0.49)	0.45* (0.22 to 0.63)

*$p < 0.01$.
†$p < 0.05$.
ARAT = Action Research Arm Test, BBT = Box and Block Test, CI = confidence interval, FMA = Fugl-Meyer Assessment, MAL-AOU = Motor Activity Log-Amount of Use, MAL-QOM = Motor Activity Log-Quality of Movement, NHPT = Nine-Hole Peg Test, SIS = Stroke Impact Scale.

elderly patients with UL impairments [7,23] and between the ARAT and NHPT during the first 6 months poststroke [16]. Despite the differences in sample characteristics and time since stroke, findings of our study together with results of prior research indicate that the three tests are related to each other with significant correlations.

Adequate validity has been established for the BBT in patients with acute stroke, multiple sclerosis, and traumatic brain injury and in elderly patients with UL impairment [7,13,23]. The ARAT has good construct validity in measuring UL motor function for patients with chronic stroke [6]. We found that the BBT and ARAT had better concurrent validity than the NHPT in patients with chronic stroke at pretreatment and posttreatment, with better correlations with the FMA and MAL. These findings attest to the relationships among the dexterous measures and motor impairment or daily functions. Understanding these relationships is particularly important in rehabilitation trials that are designed to improve impairments and real-world performance. It is of interest that the BBT and ARAT are more related to motor impairment and movement of the affected arm in daily functions than the NHPT. One possible reason for these differences might be the measurement criteria for the three tests. The BBT and ARAT both assess activity limitations for patients with UL paresis [7]. The BBT counts the number of blocks transported in 1 minute as a measure of gross manual dexterity [28]. The ARAT includes diverse items that involve gripping, pinching, and transport tasks relevant for daily life and, therefore, can be considered as an arm-specific measure of activity limitation [30]. The NHPT scale is a timed performance test to insert and remove nine pegs in nine holes as fast as possible [29]. Because it documents the speed of execution, it is possibly less sensitive to ADLs.

The magnitude of the correlations between the BBT and the FMA and between the ARAT and FMA at pretreatment and posttreatment was lower than the values reported by Platz et al. [7]. Differences in the populations studied may partly account for the discrepancies. Our sample included people with chronic stroke, whereas the sample of Platz et al. included individuals with various neurologic disorders, including stroke, multiple sclerosis, and traumatic brain injury [7]. Significant correlations between the BBT and MAL echo the previous studies [39–40] with regard to the importance of performance-based hand function tests and manual dexterity as correlates of functional dependence in the elderly. However, the results showing fair correlations between the ARAT and the MAL were different from the study by Dromerick et al., who demonstrated that patients with high scores on the ARAT still had residual disability on everyday affected arm use [11]. A difference in study sample and the ceiling effect might explain this discrepancy. Dromerick et al. chose patients who completed 3 months of an acute rehabilitation treatment trial as participants, and 41 percent (16/39) of their participants achieved maximum scores on the ARAT, which was a significant ceiling effect [11]. In our study, only 22 percent (13/59) of participants showed perfect scores on the ARAT.

Moreover, fair correlations between the ARAT and the SIS hand function domain at pretreatment and post-treatment were somewhat lower than the correlations reported in previous studies (0.57–0.83) [16–17]. Variations in the magnitude of the correlations between studies may arise from differences in time elapsed since the stroke and the initiation of intervention protocols. Our sample included people with chronic stroke, whereas the sample of Beebe and Lang included people with hemiparesis early after stroke [16]. Despite low correlations between the NHPT and the MAL, our data of moderate correlations between the NHPT and the SIS hand function domain were similar to the published data [16–17]. This likely reflects the fact that self-perceived actual use (MAL-AOU) captured information that was not assessed by the NHPT scale and that fine manual dexterity is an important component that determines the hand function aspect of quality of life after stroke.

The simplicity and feasibility of an outcome measure is also necessary to determine and weigh against the time, equipment, and training. The advantages of the BBT are its relative simplicity and the shorter time required for training and administration [28]. The advantages of the NHPT include simplicity, portability, brevity, and its relatively inexpensive cost. However, it is sensitive to changes at the upper level of performance, not when impairment is severe [26]. The NHPT should not be used to research the effectiveness of treatment to improve poor finger dexterity [29]. The main advantage of the ARAT is its ability to evaluate multiple tasks of varying complexity for a more comprehensive assessment of UL movement abilities. Its main limitations are that it is time-consuming to administer [24] and requires standardized equipment [8]. The ARAT takes an average of 8 minutes to complete and requires considerably more testing equipment [24]. This information should facilitate instrument selection in research and practice in accord with test purposes.

When the results of this study are interpreted, some potential limitations warrant consideration. First, the findings of this study were based on a sample of patients who received one of three treatments (dCIT, BAT, control treatment). Certain aspects of each form of rehabilitation might favor one instrument over the other, which could influence study results. Future studies should individually analyze the specific intervention to rule out this possibility. Second, we examined patients with chronic stroke with mild to moderate motor impairment, which might affect the generalization of results. Future studies evaluating patients with stroke in different stages are needed to determine which instrument is suitable. Third, a 3-week intervention might limit the recovery of dexterous function. A longer time to detect improvements in dexterity may be needed. Finally, the three dexterous measures involve unilateral tasks that are not representative of ADLs that often require bilateral use of the UL.

CONCLUSIONS

This article comprehensively examined and compared the responsiveness and validity of the BBT, NHPT, and ARAT. Our findings should inform clinicians and researchers in making decisions to choose appropriate tests for measuring hand dexterity in people receiving stroke interventions. All three tests are suitable to detect changes over the course of interventions. While simultaneously considering the responsiveness and validity attributes, the BBT and ARAT can be considered appropriate for evaluating dexterous function compared with the NHPT. Further research based on a larger sample is needed to validate the findings.

ACKNOWLEDGMENTS

Author Contributions:
Study concept and design: K. Lin, C. Wu.
Acquisition of data: C. Wu, K. Lin.
Analysis and interpretation of data: K. Lin, C. Wu, L. Chuang, Y. Hsieh.
Drafting and revision of manuscript: K. Lin, L. Chuang, W. Chang.
Critical revision of manuscript for important intellectual content: K. Lin, C. Wu.
Statistical analysis: L. Chuang, Y. Hsieh.
Obtained funding: K. Lin, L. Chuang, C. Wu.
Financial Disclosures: The authors have declared that no competing interests exist. No author had any paid consultancy or any other conflict of interest with this article.
Funding/Support: This material was based on work supported in part by the National Science Council (grants NSC-95-2314-B-002-225-MY2, NSC-96-2628-B-002-033-MY2, NSC-97-2314-B-002-008-MY3, NSC-97-2314-B-82-004-MY3, and NSC-98-2811-B-002-015) and the National Health Research Institutes (grants NHRI-EX98-9742PI and NHRI-EX99-9920PI) of Taiwan.

Additional Contributions: Dr. Wu is now with the Department of Occupational Therapy and Master's Degree Program of Behavioral Sciences, Chang Gung University, Taoyuan, Taiwan.
Institutional Review: We obtained institutional review board approval from the rehabilitation clinics and informed consent from all participants.
Participant Follow-Up: We do not plan to inform participants of the publication of this study. However, the participants have been encouraged to check the study Web site for updated information about the research.

REFERENCES

1. Kwakkel G, Kollen BJ, Van der Grond J, Prevo AJ. Probability of regaining dexterity in the flaccid upper limb: Impact of severity of paresis and time since onset in acute stroke. Stroke. 2003;34(9):2181–86. [PMID: 12907818] DOI:10.1161/01.STR.0000087172.16305.CD
2. Kwakkel G, Kollen B. Predicting improvement in the upper paretic limb after stroke: A longitudinal prospective study. Restor Neurol Neurosci. 2007;25(5–6):453–60. [PMID: 18334763]
3. Tuke A. Constraint-induced movement therapy: A narrative review. Physiotherapy. 2008;94(2):105–14. DOI:10.1016/j.physio.2007.07.007
4. Stewart KC, Cauraugh JH, Summers JJ. Bilateral movement training and stroke rehabilitation: A systematic review and meta-analysis. J Neurol Sci. 2006;244(1–2):89–95. [PMID: 16476449] DOI:10.1016/j.jns.2006.01.005
5. Salter KL, Teasell RW, Foley NC, Jutai JW. Outcome assessment in randomized controlled trials of stroke rehabilitation. Am J Phys Med Rehabil. 2007;86(12):1007–12. [PMID: 17912137] DOI:10.1097/PHM.0b013e3181587b3d
6. Hsieh YW, Wu CY, Lin KC, Chang YF, Chen CL, Liu JS. Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. Stroke. 2009; 40(4):1386–91. [PMID: 19228851] DOI:10.1161/STROKEAHA.108.530584
7. Platz T, Pinkowski C, Van Wijck F, Kim IH, Di Bella P, 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;19(4):404–11. [PMID: 15929509] DOI:10.1191/0269215505cr832oa
8. 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;87(7):962–66. [PMID: 16813784] DOI:10.1016/j.apmr.2006.02.036
9. Lin KC, Hsieh YW, Wu CY, Chen CL, Jang Y, Liu JS. Minimal detectable change and clinically important difference of the Wolf Motor Function Test in stroke patients. Neurorehabil Neural Repair. 2009;23(5):429–34. [PMID: 19289487] DOI:10.1177/1545968308331144
10. Bogard K, Wolf S, Zhang Q, Thompson P, Morris D, Nichols-Larsen D. Can the Wolf Motor Function Test be streamlined? Neurorehabil Neural Repair. 2009;23(5):422–28. [PMID: 19276293] DOI:10.1177/1545968308331141
11. Dromerick AW, Lang CE, Birkenmeier R, Hahn MG, Sahrmann SA, Edwards DF. Relationships between upper-limb functional limitation and self-reported disability 3 months after stroke. J Rehabil Res Dev. 2006;43(3):401–8. [PMID: 17041825] DOI:10.1682/JRRD.2005.04.0075
12. Chen HM, Chen CC, Hsueh IP, Huang SL, Hsieh CL. Test-retest reproducibility and smallest real difference of 5 hand function tests in patients with stroke. Neurorehabil Neural Repair. 2009;25(5):435–40. [PMID: 19261767] DOI:10.1177/1545968308331146
13. Higgins J, Mayo NE, Desrosiers J, Salbach NM, Ahmed S. Upper-limb function and recovery in the acute phase post-stroke. J Rehabil Res Dev. 2005;42(1):65–76. [PMID: 15742251] DOI:10.1682/JRRD.2003.10.0156
14. Higgins J, Salbach NM, Wood-Dauphinee S, Richards CL, Côté R, Mayo NE. The effect of a task-oriented intervention on arm function in people with stroke: A randomized controlled trial. Clin Rehabil. 2006;20(4):296–310. [PMID: 16719028] DOI:10.1191/0269215505cr943oa
15. Croarkin E, Danoff J, Barnes C. Evidence-based rating of upper-extremity motor function tests used for people following a stroke. Phys Ther. 2004;84(1):62–74. [PMID: 14992677]
16. Beebe JA, Lang CE. Relationships and responsiveness of six upper extremity function tests during the first six months of recovery after stroke. J Neurol Phys Ther. 2009; 33(2):96–103. [PMID: 19556918]
17. 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; 21(3):279–91. [PMID: 17353458] DOI:10.1177/1545968306296964
18. 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;87(12):1605–10. [PMID: 17141640] DOI:10.1016/j.apmr.2006.09.003
19. 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;43(5):507–15. [PMID: 15838417] DOI:10.1097/01.mlr.0000160421.42858.de

20. Lin KC, Fu T, Wu CY, Wang YH, Liu JS, Hsieh CJ, Lin SF. Minimal detectable change and clinically important difference of the stroke impact scale in stroke patients. Neurorehabil Neural Repair. 2010;24(5):486–93. [PMID: 20053950] DOI:10.1177/1545968309356295
21. 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;33(3):110–13. [PMID: 11482350] DOI:10.1080/165019701750165916
22. McEwen S, Mayo N, Wood-Dauphinee S. Inferring quality of life from performance-based assessments. Disabil Rehabil. 2000;22(10):456–63. [PMID: 10950499] DOI:10.1080/09638280050045938
23. Desrosiers J, Bravo G, Hébert R, Dutil E, Mercier L. Validation of the Box and Block Test as a measure of dexterity of elderly people: Reliability, validity, and norms studies. Arch Phys Med Rehabil. 1994;75(7):751–55. [PMID: 8024419]
24. Heller A, Wade DT, Wood VA, Sunderland A, Hewer RL, Ward E. Arm function after stroke: Measurement and recovery over the first three months. J Neurol Neurosurg Psychiatry. 1987;50(6):714–19. [PMID: 3612152] DOI:10.1136/jnnp.50.6.714
25. Jacob-Lloyd HA, Dunn OM, Brain ND, Lamb SE. Effective measurement of the functional progress of stroke clients. Br J Occup Ther. 2005;68(6):253–59.
26. Sunderland A, Tinson D, Bradley L, Hewer RL. Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. J Neurol Neurosurg Psychiatry. 1989;52(11):1267–72. [PMID: 2592969] DOI:10.1136/jnnp.52.11.1267
27. Kwakkel G, Wagenaar RC, Twisk JW, Lankhorst GJ, Koetsier JC. Intensity of leg and arm training after primary middle-cerebral-artery stroke: A randomised trial. Lancet. 1999; 354(9174):191–96. [PMID: 10421300] DOI:10.1016/S0140-6736(98)09477-X
28. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther. 1985;39(6):386–91. [PMID: 3160243]
29. Mathiowetz V, Weber K, Kashman N, Volland G. Adult norms for the Nine-Hole Peg test of finger dexterity. Occup Ther J Res. 1985;5(1):24–38.
30. Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. Int J Rehabil Res. 1981;4(4):483–92. [PMID: 7333761] DOI:10.1097/00004356-198112000-00001
31. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. Scand J Rehabil Med. 1975; 7(1):13–31. [PMID: 1135616]
32. Van der Lee JH, Beckerman H, Knol DL, De Vet HC, Bouter LM. Clinimetric properties of the motor activity log for the assessment of arm use in hemiparetic patients. Stroke. 2004;35(6):1410–14. [PMID: 15087552] DOI:10.1161/01.STR.0000126900.24964.7e
33. Duncan PW, Bode RK, Min Lai S, Perera S. Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: The Stroke Impact Scale. Arch Phys Med Rehabil. 2003; 84(7):950–63. [PMID: 12881816] DOI:10.1016/S0003-9993(03)00035-2
34. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): L. Erlbaum Associates; 1988.
35. Efron B, Tibshirani RJ. An introduction to the bootstrap. New York (NY): Chapman & Hall; 1993.
36. Stratford PW, Kennedy DM. Does parallel item content on WOMAC's pain and function subscales limit its ability to detect change in functional status? BMC Musculoskelet Disord. 2004;5:17. [PMID: 15189563] DOI:10.1186/1471-2474-5-17
37. Portney LG, Watkins MP. Foundations of clinical research: Applications to practice. 3rd ed. Upper Saddle River (NJ): Pearson/Prentice Hall; 2009.
38. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: A clarification. J Clin Epidemiol. 1989;42(5):403–8. [PMID: 2659745] DOI:10.1016/0895-4356(89)90128-5
39. Falconer J, Hughes SL, Naughton BJ, Singer R, Chang RW, Sinacore JM. Self report and performance-based hand function tests as correlates of dependency in the elderly. J Am Geriatr Soc. 1991;39(7):695–99. [PMID: 2061536]
40. Ostwald SK, Snowdon DA, Rysavy DM, Keenan NL, Kane RL. Manual dexterity as a correlate of dependency in the elderly. J Am Geriatr Soc. 1989;37(10):963–69. [PMID: 2507619]

Submitted for publication September 25, 2009. Accepted in revised form March 16, 2010.

This article and any supplementary material should be cited as follows:
Lin K, Chuang L, Wu C, Hsieh Y, Chang W. Responsiveness and validity of three dexterous function measures in stroke rehabilitation. J Rehabil Res Dev. 2010;47(6):563–72. DOI:10.1682/JRRD.2009.09.0155