The Reliability of the Wolf Motor Function Test for Assessing Upper Extremity Function After Stroke

David M. Morris, MSPT, Gitendra Uswatte, MA, Jean E. Crago, MSPT, Edwin W. Cook III, PhD, Edward Taub, PhD


Objective: To examine the reliability of the Wolf Motor Function Test (WMFT) for assessing upper extremity motor function in adults with hemiplegia.

Design: Interrater and test-retest reliability.

Setting: A clinical research laboratory at a university medical center.

Patients: A sample of convenience of 24 subjects with chronic hemiplegia (onset >1yr), showing moderate motor impairment.

Intervention: The WMFT includes 15 functional tasks. Performances were timed and rated by using a 6-point functional ability scale. The WMFT was administered to subjects twice with a 2-week interval between administrations. All test sessions were videotaped for scoring at a later time by blinded and trained experienced therapists.

Main Outcome Measure: Interrater reliability was examined by using intraclass correlation coefficients and internal consistency by using Cronbach’s alpha.

Results: Interrater reliability was .97 or greater for performance time and .88 or greater for functional ability. Internal consistency for test 1 was .92 for performance time and .92 for functional ability; for test 2, it was .86 for performance time and .92 for functional ability. Test-retest reliability was .90 for performance time and .95 for functional ability. Absolute scores for subjects were stable over the 2 test administrations.

Conclusion: The WMFT is an instrument with high interrater reliability, internal consistency, test-retest reliability, and adequate stability.

Key Words: Brain injuries; Cerebrovascular accident; Rehabilitation.

© 2001 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

ACcurate Assessment of motor function is important for documenting the outcome of rehabilitation programs. Assessment is also used (1) to match patients to appropriate interventions; (2) to compare the results of interventions across patients and among facilities; (3) to improve resources; and (4) to determine personnel needs. However, assessment practices vary greatly in rehabilitation research and practice and no single instrument is universally accepted.

A number of direct-observation performance tests have been described as useful for examining upper extremity (UE) ability in activities of daily living (ADL) tasks in the laboratory. The Frenchay Arm Test uses a binary pass-fail evaluation for each of 5 tasks, but it lacks assessment of quality of movement and performance time. The Jebsen-Taylor Hand Function Test uses performance time as an index of functional ability, but does not consider quality of movement and is of value only for higher functioning patients with adequate manual dexterity. The Action Research Arm Test uses a 4-step rating scale, but it does not measure performance time. The Test Evaluant les Membres superieurs des Personnes Agees (TEMPA) does assess both quality of movement and performance time; however, it was designed for use with geriatric subjects. As a result, most of the TEMPA tasks are too difficult for patients with moderate, and sometimes even mild, hemiparesis. The Fugl-Meyer Assessment (FMA) of Physical Performance examines the underlying impairments that influence functional use of both the UEs and lower extremities. Specifically, the FMA examines sensation, range of motion, joint pain, and the quality of movement of an extremity while a subject performs selected movement patterns. The FMA does not examine speed of movement or use of the extremities during functional tasks. It is useful for testing lower functioning patients, but is of less value for determining the full range of function in patients with mild to moderate impairments. The problems noted earlier limit the usefulness of these tests for assessing UE use after stroke.

The original version of the Wolf Motor Function Test (WMFT) was developed by Wolf et al to examine the effects of constraint-induced movement therapy (CIMT) for survivors of stroke and traumatic brain injury. The WMFT, as modified by this laboratory, along with other measures, has been used as an outcome measure in CIMT research projects. Because CIMT was originally used with patients with mild to moderate stroke, the WMFT was developed to be sensitive to the level of motor functioning characteristic of these types of patients. Recently, a graded WMFT has been developed for assessing lower functioning patients.

The original form of the test consisted of 21 simple tasks sequenced according to joints involved (shoulder to fingers) and level of difficulty (gross to fine motor skill). As recommended by Wolf, several of the original tasks were dropped from the testing protocol in the current version, which contains 17 tasks. Two of the tasks are simple measures of strength and, because performance on these items is not rated and is not included in the total performance time or functional ability score on the test, they were not included in this investigation.
Reliability for one of the strength measures has already been established. Interrater reliability for grip strength by using a Jaymar hand-held dynamometer is more than .80 and accuracy is $\pm 3\%$. In the original version of the test, performance time was the primary measure. Subsequently, rating scales for quality of movement and functional ability were added. Currently, only the Functional Ability Scale (FAS) is used because the 2 rating scales appeared to measure identical constructs ($r_{GOF,LA} = .98$). The WMFT is preferable to the commonly used UE performance tests because it tests a wide range of functional tasks (ie, from simple to complex) and explores both performance time and quality of movement.

Because over half of the items on the WMFT involved simple limb movements without functional endpoints, its relation to subjects’ ability to perform ADLs was uncertain. Taub and others developed the Arm Motor Ability Test (AMAT) to provide such information. The AMAT consists of 16 compound tasks, composed of 1 to 3 subtasks, each task representing a complete ADL commonly performed in the life situation (eg, donning a sweater, picking up a single dried bean on a spoon and bringing it to the mouth, unscrewing a jar cap). Subjects undergoing evaluation with the AMAT are rated for performance time, functional ability, and quality of movement. The AMAT was examined and found to be reliable, valid, internally consistent, and sensitive to change. Both the WMFT and AMAT have been used to examine the effectiveness of CIMT. However, the AMAT has been found to be less sensitive to subjects’ changes in motor abilities and more difficult to administer than the WMFT.

Therefore, the AMAT has been dropped from our CIMT research protocol in more recent studies and the WMFT is the sole laboratory motor test used. Given that the WMFT is used with regularity in CIMT research and that it might be considered a meaningful tool by other clinicians and researchers, it is important to assess its psychometric properties in more detail. In this study, we examined the interrater reliability, internal consistency, test-retest reliability, and stability of the WMFT for assessing UE motor function in adults with chronic hemiplegia.

**METHODS**

**Subjects**

Subjects were a sample of convenience of 15 men and 9 women with hemiplegia with onset longer than 1 year before testing (mean onset, 6yr; range, 2–17yr). They were recruited from a pool of suitable candidates identified in part through the use of our laboratory’s patient files. Subjects’ mean age was 61 years (range, 14–86yr); 72% were premorbidly right-handed; 68% had a right hemiparesis; and for 64%, the paresis was on the side of their premorbidly dominant UE. Twenty-four subjects were evaluated on the first testing occasion; 20 were tested on both occasions.

Inclusion criteria were: (1) the ability to extend the more affected wrist, thumb, and at least 2 additional digits 10\(^\circ\) at each joint (ie, the lower motor criterion for “second quartile” patients in our CIMT research), and (2) the ability to understand and follow verbal directions. An effort was made to recruit subjects with varying degrees of UE involvement to examine the use of the FAS for evaluating patients with hemiparesis at a variety of UE functional levels.

**Procedure**

We investigated the 15 items listed in table 1. A standardized protocol provided detailed information about conducting the test including: task descriptions, starting positions for subjects and equipment, timing procedures (including criteria for determining completion points), verbal instructions to be read to subjects, and scoring criteria. A template indicating the position of each test object was taped to the test table in a standard position to enable placement of test objects in the same location at each test administration. The material from the test manual for a single test item is given in table 2. Test performance was videotaped and later rated independently by 3 clinicians blinded to the order of test occasions. (In studies involving interventions, the clinicians are blinded to the pre- or posttreatment status of the subjects.) The FAS is presented in table 3. Ratings were allowed to rewind the videotape record of each item as many times as needed to confirm the rating they had assigned. Test performance was also timed off videotape. If a subject was unable to complete an item within 2 minutes, the attempt was stopped and a functional ability score of 1 and a performance time of 120 seconds were assigned.

The instructions for each task were read to the patients twice (3 times if there appeared to be some confusion). The general instructions read to the patient before test administration are as follows:

- Today we are going to take a look at how you are able to use your arm. Let me tell you how we are going to go about this. First, I will give you instructions on how to do the task, and then I will show you how to do it. I will describe and demonstrate each task 2 times. Do not practice the task while I’m describing and demonstrating it. However, I will be happy to clarify any points of confusion. Then I will say, “Ready, set, go” and you will do the task. Each of the activities you will be asked to do should be carried out as rapidly as possible. You can work on each task for up to 2 minutes. We ask that you attempt each part of the test even if you do not think that you can do it. If you are unable to carry out a task, then we will go on to the next one. Again, try to do each task as rapidly as possible. Do you have any questions?

Subjects were tested by using the WMFT on 2 occasions approximately 2 weeks apart. Subjects received no treatment during the 2-week interval. The first author (DM) conducted all test sessions according to a standardized written protocol. All task performances were videotaped from 1 of 3 task-specific positions. Three rehabilitation therapists (2 physical therapists,
1 occupational therapist) viewed the tapes independently and assigned scores for performance time and functional ability ratings. Each panel member possessed between 14 and 20 years of clinical experience treating adults with neurologic deficits. Members of the rating panel were trained by a physical therapist experienced in administering and scoring the WMFT. They were also given a scoring manual and directions for administering the test.

To establish a common frame of reference for scoring, the panel first rated as a group 6 patients tested with the WMFT. The videotapes of 14 additional practice subjects were rated independently and group review sessions were held until interrater agreement of at least .90 was reached for the mean functional ability and median performance time scores. None of these 20 subjects was part of the current sample. Each individual’s testing sessions were placed on different videotapes to reduce the chances of recall of a previously viewed session. Approximately 5 sessions of different subjects were on each of the videotapes.

The summary score for functional ability is the mean of the item scores. The summary score for the performance time is the median because it is less sensitive to outliers than the mean score. In addition, the mean score for individual subjects would include arbitrary truncated scores of 120 seconds, which are assigned to tasks that are not completed; means are therefore not veridical. In calculating the median, a low rank is assigned to tasks that are not completed. Thus, scores truncation does not affect the median, which is therefore the preferable measure.

Data Analysis

The interrater reliability of the functional ability ratings and performance time scores were estimated by intraclass correlation coefficients (ICCs), types (3,1) and (2,1), separately for tests 1 and 2. ICC3,1 evaluates the consistency of the ratings in terms of relative ranking (ie, if a rater gives a subject a high score, the extent to which other raters also assign a high score). ICC2,1 evaluates the agreement of the ratings in terms of absolute scores (ie, if a rater gives a functional ability score of 4, the extent to which the other raters also assign a 4). Interrater reliability was calculated on all subjects tested on each occasion.

The internal consistencies of the functional ability ratings and performance time scores were assessed by Cronbach’s alpha. The test-retest reliabilities of the mean functional ability ratings and median performance time scores were estimated by Pearson’s correlation coefficient. The stability of the test was examined by testing for significant change in func-

---

Table 2: Testing Protocol for Task Item 1: Forearm to Table*

<table>
<thead>
<tr>
<th>Set Up</th>
<th>Task</th>
<th>Verbal Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Starting Position:</strong></td>
<td><strong>Task Description:</strong></td>
<td><strong>Verbal Instructions:</strong></td>
</tr>
<tr>
<td>Chair position A (Chair placed sideways and 8.5cm from the front table edge. Side to be tested adjacent to the table. Back of chair is 6.5cm beyond edge of side end of template.)</td>
<td>Patient attempts to place forearm on the table (adjacent and parallel to front edge) by abduction at the shoulder. (Some shoulder flexion will probably also be necessary to get arm past the edge of table.) “Forearm” is defined as the wrist and elbow. The palmar surface of the hand need not be flat. Timing ends when both the forearm and hand touch the table. <strong>Timing Procedure:</strong> Starts on word “Go” and ends when patient's forearm and hand touch the table in the required position. <strong>Measure:</strong> The time elapsed from the starting point to the moment the forearm and hand touch the table in the required fashion.</td>
<td>“Place your forearm on the table as quickly as you can. Do it just like this (examiner demonstrates). At the end of the movement, your forearm and hand should be touching the surface of the table” (repeat instructions). “Do you have any questions?” “Ready, set, go.” <strong>Scoring:</strong> Functional ability scoring should take into account the extent to which the head and trunk are maintained in normal alignment and the speed, fluidity, and precision with which movements are performed.</td>
</tr>
<tr>
<td>Hips against back of chair. Hands in lap. Both feet on floor. Filming position A (Profile of the whole body: the front edge of the camera tripod should be placed 3 feet to the side of the patient being tested and directly in line with the back edge of the desk. The camera view should include the patient’s entire body.)</td>
<td>“Ready, set, go.” “Do you have any questions?”</td>
<td>“Place your forearm on the table as quickly as you can. Do it just like this (examiner demonstrates). At the end of the movement, your forearm and hand should be touching the surface of the table” (repeat instructions). “Do you have any questions?” “Ready, set, go.” <strong>Scoring:</strong> Functional ability scoring should take into account the extent to which the head and trunk are maintained in normal alignment and the speed, fluidity, and precision with which movements are performed.</td>
</tr>
</tbody>
</table>

*Functional ability of the shoulder (shoulder abduction); tasks performed to the side of the patient (ie, away from the midsagittal plane of the patient).

---

Table 3: Functional Ability Scale

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Does not attempt with involved arm.</td>
</tr>
<tr>
<td>1</td>
<td>Involved arm does not participate functionally; however, an attempt is made to use the arm. In unilateral tasks the uninvolved extremity may be used to move the involved extremity.</td>
</tr>
<tr>
<td>2</td>
<td>Arm does participate, but requires assistance of uninvolved extremity for minor adjustments or change of position, or requires more than 2 attempts to complete, or accomplishes very slowly. In bilateral tasks the involved extremity may serve only as a helper or stabilizer.</td>
</tr>
<tr>
<td>3</td>
<td>Arm does participate, but movement is influenced to some degree by synergy or is performed slowly and/or with effort.</td>
</tr>
<tr>
<td>4</td>
<td>Arm does participate; movement is close to normal,* but slightly slower; may lack precision, fine coordination or fluidity.</td>
</tr>
<tr>
<td>5</td>
<td>Arm does participate; movement appears to be normal.*</td>
</tr>
</tbody>
</table>

* For the determination of normal, the uninvolved limb can be used as an available index for comparison, with premorbid limb dominance taken into consideration.
tional ability ratings and performance time scores between tests 1 and 2 by using a repeated-measures analysis of variance. Analysis of individual WMFT test items concerning interrater reliability, item-total correspondence, and test-retest reliability was also conducted in the same manner described earlier.

RESULTS

The interrater reliabilities (ICC3,1) of the functional ability ratings and performance time scores were high (both ≥.93) (table 4). The agreement (ICC2,1) of the raters was also high, (both ≥.88). These results suggest that the raters not only ranked the functional ability of the subjects in a similar manner, but also tended to give individual subjects the same absolute scores. For performance time, the consistency of the raters was also high for both tests (both ≥.97), as was the agreement of the raters (both ≥.97).

The consistency and agreement of the functional ability ratings for individual items was generally adequate for test 1 (median ICC3,1 = .76; range, .50–.93; median ICC2,1 = .71; range, .47–.94) and test 2 (median ICC3,1 = .75; range, .41–.93; median ICC2,1 = .75; range, .36–.93). Nine items, however, displayed less than adequate interrater reliability (ICC3,1 < .75)20 for functional ability on at least 1 of the test administrations (table 5). The interrater reliability of performance time scores for individual items was high; all of the items had an ICC3,1 of at least .90.

Internal consistency for test 1 (coefficient alpha) for both functional ability and performance time was .92; for test 2, coefficient alpha for functional ability was .92, whereas for performance time it was .86 (table 4). The item-total correlations were generally high. They were somewhat lower for test 2 performance time than for that of test 1 or functional ability on either test. Items 1, 4, 7, 10, and 11 (table 1) had consistently low item-total correlations across administrations and scales (<.50).21

The relative ranks of subjects’ scores on functional ability and performance time were highly stable between the 2 testing occasions (test-retest reliabilities = .95, .90, respectively). Individual item scores were less reliable across administrations, though they were adequate (> .60)21 for all but 3 items (5, 7, 11) on functional ability and 3 items (9, 12, 14) on performance time.

In addition, the absolute scores of the patients were stable between testing occasions. Given that the participants in this study were people with chronic stroke who were not receiving treatment, no substantial changes in performance were expected between the 2 test administrations. This was the case for median performance times (mean change, −.30s, not significant). The change in the mean functional ability ratings was also small (mean change, 0.1 scale points; by comparison, CIMT patients with similar initial motor deficits improve on average 0.4 points from pre- to posttreatment).10 The increase, however, was statistically significant because it was consistent across subjects (F1,19 = 5.05, p < .05).

DISCUSSION

The WMFT was modified at the University of Alabama at Birmingham to measure motor impairment through the use of a quantitative physical variable, performance time, while at the same time assessing through a rating scale the fine coordination, fluidity, and other general, clinically relevant characteristics of movement (FAS). Although more difficult to quantify objectively, these characteristics are nevertheless critically important parameters of movement. The WMFT includes items that cover a range of movements that can be evaluated within a realistic time period for both clinical and research purposes. It has been recognized that there is a need for more and better functional assessment tools to examine UE function after a stroke.2 The WMFT represents an attempt to address this need. The AMAT12,16,17 was also developed at this institution to assess UE function, but, as noted, has been found to be less sensitive for measuring the effect of UE rehabilitation than the WMFT.

Both WMFT functional ability ratings and performance time scores had high interrater reliabilities and high internal consistency for each of the 2 test sessions. The high internal consistency obtained indicates that the WMFT measures 1 underlying construct or ability; the similar interrater reliabilities and internal consistency values obtained for both testing occasions suggest that the properties of the test, with regard to these 2 characteristics, do not vary with time of administration. The test-retest reliabilities of the functional ability ratings and the performance time scores were also high. It should be noted that the psychometric data for the FAS were essentially as good as for the presumably more objective performance time measure, which should help to address clinicians’ concerns about the legitimacy of using standardized, formal functional assessment scales.22

### Table 4: Interrater Consistency, Interrater Agreement, Internal Consistency, and Test-Retest Reliability for the WMFT

<table>
<thead>
<tr>
<th>Functional ability</th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC3,1</td>
<td>.93</td>
<td>.95</td>
</tr>
<tr>
<td>ICC2,1</td>
<td>.88</td>
<td>.92</td>
</tr>
<tr>
<td>Internal Consistency (α)</td>
<td>.92 (n = 15)</td>
<td>.92 (n = 15)</td>
</tr>
<tr>
<td>Test-Retest Reliability (r)</td>
<td>.95 (test 2, n = 20)</td>
<td>NA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance time</th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC3,1</td>
<td>.99</td>
<td>.97</td>
</tr>
<tr>
<td>ICC2,1</td>
<td>.99</td>
<td>.86</td>
</tr>
<tr>
<td>Internal Consistency (α)</td>
<td>.92 (n = 20)</td>
<td>.96 (n = 14)</td>
</tr>
<tr>
<td>Test-Retest Reliability (r)</td>
<td>.90 (test 2, n = 20)</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviation: NA, not applicable.

### Table 5: Items With Inadequate Interrater Reliability* on the FAS for Both Tests

<table>
<thead>
<tr>
<th>Item</th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Forearm to table (side)</td>
<td>.52</td>
<td>.75</td>
</tr>
<tr>
<td>2. Forearm from table to 25.4-cm box (side)</td>
<td>.78</td>
<td>.57</td>
</tr>
<tr>
<td>3. Extend elbow 28cm on table top (side)</td>
<td>.68</td>
<td>.66</td>
</tr>
<tr>
<td>4. Extend elbow 28cm on table top (.45-kg weight)</td>
<td>.59</td>
<td>.65</td>
</tr>
<tr>
<td>5. Hand to table (front)</td>
<td>.50</td>
<td>.46</td>
</tr>
<tr>
<td>6. Hand to box (front)</td>
<td>.66</td>
<td>.41</td>
</tr>
<tr>
<td>7. Retrieve .45-kg weight from 28-cm line on table top by elbow flexion</td>
<td>.80</td>
<td>.61</td>
</tr>
<tr>
<td>9. Lift pencil from table</td>
<td>.88</td>
<td>.72</td>
</tr>
<tr>
<td>13. Turn key in lock: clockwise to 180°, counterclockwise to 180°, and back to the starting position</td>
<td>.67</td>
<td>.82</td>
</tr>
</tbody>
</table>

* ICC3,1 < .75.
Item interrater and test-retest reliabilities were generally adequate, yet certain test items showed low reliability. These results suggest that individual items on the WMFT should not be used as indicators of motor function. Instead, the score for the entire test (mean item score for functional ability; median item score for performance time) should be used when interpreting results. The 3 items exhibiting consistently poor reliability were items 4, 5, and 7 (table 5). However, when these items were dropped and the remaining data reexamined, the results for interrater reliability, internal consistency, test-retest reliability, and stability as a whole did not change to a meaningful degree. Thus, if the test was to be shortened for administration in a clinical setting, these items would be the most appropriate items to drop. For research purposes, the 17-item WMFT assessed here would be most useful because it includes a wide range of tasks (ie, simple to complex) and examines multiple joints of the UE.

As noted, the median time, rather than the mean time, is used as the test score for performance time. The results of this investigation support this approach. The reliability of the mean performance time score ($r = .60, p < .01$) was substantially lower than the median performance time score ($r = .90, p < .001$). Also, analysis of test stability indicated that the mean performance time scores experienced a change from test 1 to test 2 that approached statistical significance (ie, mean decrease of 16.8%), whereas the change in median performance time scores from test 1 to test 2 was much more modest (ie, mean decrease of 5.2%). The 2 results suggest that the median performance time is a more reliable and stable measure than the mean performance time and appropriately preserves the rank and absolute level of the participants over the 2-week test-retest period. The mean performance time may be inferior because it is more sensitive to outliers (scores of 120s) than the median performance time.14

The detailed WMFT written protocol promotes standardized administration and probably contributes to its psychometric stability. As noted, the written instructions include detailed task descriptions, starting positions for subjects and test objects, timing procedures, verbal instructions for subjects, and scoring criteria. Testers using the WMFT are advised to adhere closely to the written instructions to assure reliability. Moreover, a test object template taped to the testing table and placement of the subject’s chair on adhesive dots at measured distances from the table ensures the same position on different test administrations for test objects and subjects, respectively.

The WMFT should be useful to clinicians in that it examines many different aspects of UE function (eg, single and multijoint coordinated movements). In addition, the test has both physical (performance time score) and qualitative (functional ability rating) measures of performance. Although specific equipment is needed for test administration, most of the items are common, inexpensive, and easily obtained. This fact combined with its reliability, internal consistency, and stability makes the WMFT of value for research purposes. One potential obstacle for clinicians may be the time needed to conduct the test (~30min). However, in situations in which it is important to document change over time because of treatment, this may be time well spent.

The data collected in this study were derived from the functional ability and performance time ratings from experienced, trained clinicians. Moreover, these raters were allowed to rewind and view the videotaped performance of each task as many times as needed before assigning a score. The reliability of assigning scores in “real time” while simultaneously administering the WMFT was not assessed in this study. However, for 26 patients with chronic hemiplegia who were treated in another clinical research project, the agreement between a WMFT-trained therapist who rated potential patients “live” and an independent clinician who rated the patients from videotape was adequate ($ICC_{1,1} = .85, ICC_{2,1} = .79$).11 Clinicians and researchers are still advised to videotape WMFT testing sessions and assign scores in the manner described here because the interrater reliability is higher when both raters are working from videotape. Also, it is important that WMFT raters receive adequate training before assigning scores independently.

The reason for testing patients with chronic stroke (>1yr postevent) in this study was to reduce the possibility of spontaneous recovery affecting the results by improving the performance on test 2. However, there is no reason why the WMFT could not be used with acute or subacute patients with some recovered or spared motor ability.

In this laboratory, the WMFT showed considerable sensitivity to treatment effects in subjects undergoing CIMT. Higher functioning patients with chronic stroke in terms of the severity of their initial motor deficit have exhibited significant reductions in median performance time scores from 6.0 seconds before treatment to 3.2 seconds after treatment, and significant increases in mean functional ability from 3.5 to 4.0 rating points. Attention-placebo controls did not show significant changes in either measure.15 The mean effect size for this test in experimental subjects was .90. In the meta-analysis literature, effect sizes of .80 are deemed large.23

CONCLUSION

The WMFT was designed to evaluate the benefits of CIMT for survivors of stroke and traumatic brain injury. This study shows that the test has high interrater reliability, internal consistency, and test-retest reliability, and adequate stability when used with chronic hemiplegic subjects. These psychometric properties support the use of the WMFT in both the research and clinical settings.

Acknowledgment: The authors thank Dr. Steven Wolf for a critical reading of the manuscript and for providing access to 3 experimental subjects.

References

Supplier
a. Jaymar Hand Evaluation/Preston, PO Box 89, Jackson, MS 49204-0089.