Chronic Stroke Outcome Measures for Motor Function Intervention Trials

Expert Panel Recommendations

Cheryl Bushnell, MD, MHS; Janet Prvu Bettger, ScD; Kevin M. Cockroft, MD, MSc; Steven C. Cramer, MD, MMSc; Maria Orlando Edelen, PhD; Daniel Hanley, MD; Irene L. Katzian, MD, MS; Soeren Mattke, MD, DSc; Dawn M. Nilsen, EdD, OTR/L; Tepring Piquado, PhD; Elizabeth R. Skidmore, PhD, OTR/L; Kay Wing, PT, DPT, NCS, GCS; Gayane Yenokyan, MD, PhD, MHS, MPH, MPP

Background—About half of survivors with stroke experience severe and significant long-term disability. The purpose of this article is to review the state of the science and to make recommendations for measuring patient-centric outcomes in interventions for motor improvement in the chronic stroke phase.

Methods and Results—A 9-member expert panel reviewed evidence to identify measures of upper and lower extremity function used to date as outcomes in trials with patients who experienced a stroke ≥6 months before assessment. Outcome measures were screened using StrokEDGE consensus panel recommendations, and evaluated for availability of a published minimal clinically important difference. Measures meeting these criteria were further evaluated with regard to their level of measurement, psychometric properties, and ability of minimal clinically important difference to capture gains associated with improved function and clinical relevance to patients, to arrive at recommendations. A systematic literature review yielded 115 clinical trials of upper and lower extremity function in chronic stroke that used a total of 34 outcome measures. Seven of these had published minimal clinically important differences and were recommended or highly recommended by StrokEDGE. Those are the Fugl-Meyer Upper Extremity and Lower Extremity scales, Wolf Motor Function Test, Action Research Arm Test, Ten-Meter and Six-Minute Walk Tests, and the Stroke Impact Scale. All had evidence for their psychometric performance, although the strength of evidence for validity varied, especially in populations with chronic stroke Fugl-Meyer Upper and Lower Extremity scales showing the strongest evidence for validity.

Conclusions—The panel recommends that the Fugl-Meyer Upper and Lower Extremity scales be used as primary outcomes in intervention trials targeting motor function in populations with chronic stroke. The other 6 measures are recommended as secondary outcomes. (Circ Cardiovasc Qual Outcomes. 2015;8:S163-S169. DOI: 10.1161/CIRCOUTCOMES.115.002098.)

Key Words: chronic disease ■ lower extremity ■ stroke ■ survivors ■ upper extremity

In spite of great advances in acute treatment, many survivors with stroke experience debilitating sequelae.1–4 About 800000 people in the United States have a stroke each year and approximately two thirds of these require rehabilitation. Although evidence-based rehabilitation programs are able to improve functional abilities, the steepest slope of motor recovery occurs within the first 6 months.5 Residual and often disabling long-term deficits are common6,7 and frequently because of impaired motor function.8–10 Novel interventions to improve motor function, such as stem-cell therapy and brain–computer interfaces, are being developed to reduce disability for survivors with stroke in the chronic phase of recovery, ≥6 months after stroke onset. As with all proposed therapies, regulatory and clinical decision making is dependent on outcomes research with reliable and valid measures.

© 2015 American Heart Association, Inc.

Circ Cardiovasc Qual Outcomes is available at http://circoutcomes.ahajournals.org

DOI: 10.1161/CIRCOUTCOMES.115.002098

S163
WHAT IS KNOWN

• Stroke survivors often have residual and disabling deficits, frequently related to impaired motor function.
• In the chronic setting after stroke, although many measures of motor function are available, the best measure for evaluation of new therapies is not established.

WHAT THE STUDY ADDS

• No one measure met our evaluation criteria for: outcomes at the activity or participation level, solid psychometric evidence, and a minimal clinically important difference that was established based on patient rating.
• Given the strengths and weaknesses of multiple measures, the panel recommends the Fugl-Meyer upper extremity and lower extremity scales as primary outcomes for trials of changes in chronic stroke.
• This evidence-based review adds to the literature on chronic stroke outcome measures through its identification of gaps in research, which could enhance the use of additional measures that currently lack patient-based minimal clinically important differences (Wolf Motor Function Test and the Arm Research Action Test) and documentation of psychometric performance specifically for chronic stroke (Six-Minute Walk Test and Ten-Meter Walk Test).

makers must weigh their benefit against their risks. To do so requires evidence-based measures of motor function specific to the chronic stroke phase. Although numerous measures for poststroke motor function exist, it is not clear which most accurately measure meaningful change in the chronic phase.

The purpose of this article, therefore, is to review the state of the science of outcome measures for motor function in the chronic stroke phase and to make recommendations as to which measures should be used as outcomes for interventions aimed to reduce chronic motor deficits.

Methods

Expert Panel Selection

On the basis of nominations from national societies (American Heart Association, American Academy of Neurology, American Association of Neurological Surgeons, American Occupational Therapy Association, and the American Physical Therapy Association) and leading experts in the field, the research team (M.O.E., T.P., S.C.C., and S.M.) selected 9 individuals for an expert panel representing a diversity of practice settings, educational background, geographic regions, and research and clinical foci.

Procedures

Panel members met twice in person and twice by teleconference to guide the research team in 4 steps. First, the panel reviewed and provided input on the search strategy for the research team to identify publications of clinical trials of chronic stroke motor function (see Appendix I in the Data Supplement for description of search strategy). Second, the upper and lower extremity measures of motor function identified in the literature were screened to include only those that were either recommended or highly recommended by the StrokEDGE consensus group for use in chronic stroke rehabilitation. The details of StrokEDGE are reported in more detail elsewhere, but briefly, the Neurology Section of the American Physical Therapy Association completed a rigorous evaluation of outcome measure psychometric properties and clinical use for patients with stroke to provide recommendations on the appropriateness of use (highly recommended, recommended, unable to recommend at this time, and not recommended) by practice setting and patient acuity (subacute, acute, and chronic).

Third, the measures remaining were evaluated for the availability of a minimal clinically important difference (MCID) in the rehabilitation measures database. The database, developed under a grant from the Department of Education, was designed to help clinicians and researchers to identify optimal measures for assessing patient outcomes in all phases of rehabilitation. The panel further excluded measures without an MCID because this piece of information is critical for evaluating intervention benefit. Finally, the panel developed recommendations for primary and secondary measures of upper and lower extremity function based on analysis of each against 3 criteria:

1. Level of measurement after the World Health Organization’s International Classification of Functioning, Disability, and Health System framework, which characterizes measures as reflecting status in the domains of body function, activity, and participation. Measures of domains closer to patient experience (ie, activity or participation domain) were favored over measures of body function.
2. Strength of psychometric properties including reliability (inter-rater reliability, internal consistency reliability, and test–retest reliability), responsiveness (the measure’s ability to reflect change from pre to post-treatment), and validity (face validity, content validity, concurrent validity, predictive validity, and construct validity); and
3. Quality of MCID in that measures, which had their MCIDs specifically established for chronic stroke and derived using patient-based anchor methods, were favored over those with MCIDs only established for acute stroke or those based on distribution-based methods (ie, based on a statistical measure of spread) or expert judgment.

Results

Identifying Candidate Measures

We identified 115 clinical trials of interventions to improve motor function in patients with chronic stroke; 55 of those looked at lower extremity and 60 at upper extremity function (see Appendix I in the Data Supplement for list of references). Those trials yielded 34 unique measures, 14 of which were used in trials of upper extremity, 8 of lower extremity, and 12 of upper and lower extremity function. Table 1 indicates the frequency of use, the targeted area, the level of recommendation for use as a chronic stroke rehabilitation outcome according to the StrokEDGE panel for each measure; and for those recommended or highly recommended, the existence of an MCID. Notably, despite its frequent use as an assessment of functional mobility in the acute phase of recovery (onset to 6-month poststroke), the modified Rankin Scale, was not used as an outcome in any of the chronic stroke clinical trials. Furthermore, the modified Rankin Scale has no MCID defined for chronic stroke and only limited information exists for its responsiveness and prognostic value for recovery in chronic stroke.

The panel evaluated 7 measures in step 4 (indicated in dagger in Table 1) from which to develop recommendations for measuring upper and lower extremity motor function.
in the chronic stroke phase. Those were the Fugl-Meyer Upper Extremity subscale (FM-UE), the Fugl-Meyer Lower Extremity subscale (FM-LE), the Wolf Motor Function Test (WMFT), the Action Research Arm Test (ARAT), the Ten-Meter Walk Test (10MWT), and the Stroke Impact Scale (SIS). Six are performance-based measures (ie, they reflect clinician or therapist ratings of actions performed by the patient) and 1 measure, the SIS, is based on self-report (ie, patient or proxy answers questions about his/her ability to perform specific activities). These measures are described in more detail below and Table 2 summarizes the panel’s evaluation of the 3 criteria in step 4.

**Fugl-Meyer Upper Extremity**

The FM-UE\(^\text{17}\) assesses movement of the biceps, triceps, shoulder, elbow, forearm, hand, wrist, and finger with performance of 33 tasks. Clinicians rate patient performance on each task for quality of movement on a scale from 0 (no active motion) to 2 (motion seems to be normal). The maximum score on the FM-UE subscale is 66 points (possible range, 0–66).

The FM-UE measures performance at the body function domain. The FM-UE has excellent psychometric properties for assessment of motor function in chronic stroke, including high inter-rater reliability,\(^\text{18}\) and excellent test–retest reliability\(^\text{19}\) and responsiveness.\(^\text{15,20}\) There is also substantial evidence

Table 1. Properties of Measures Identified in 115 Chronic Stroke Rehabilitation Clinical Trials Reviewed

<table>
<thead>
<tr>
<th>Measure Name</th>
<th>Frequency Count, %</th>
<th>Target Extremity</th>
<th>StrokEDGE Recommendation*</th>
<th>MCID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fugl-Meyer upper extremity†</td>
<td>40 (34.8)</td>
<td>Upper</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Motor activity log</td>
<td>28 (24.3)</td>
<td>Upper</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>6-min Walk Test†</td>
<td>27 (23.5)</td>
<td>Lower</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Timed up and go</td>
<td>27 (23.5)</td>
<td>Both</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>10-Meter Walk Test†</td>
<td>26 (22.6)</td>
<td>Both</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Modified Ashworth Scale</td>
<td>26 (22.6)</td>
<td>Both</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Berg Balance Scale</td>
<td>23 (20.0)</td>
<td>Lower</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Wolf Motor Function Test†</td>
<td>22 (19.1)</td>
<td>Upper</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>Stroke Impact Scale†</td>
<td>15 (13.0)</td>
<td>Both</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Box and Block Test</td>
<td>12 (10.4)</td>
<td>Upper</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Fugl-Meyer lower extremity†</td>
<td>12 (10.4)</td>
<td>Lower</td>
<td>4</td>
<td>Yes</td>
</tr>
<tr>
<td>Action Research Arm Test†</td>
<td>10 (8.7)</td>
<td>Upper</td>
<td>3</td>
<td>Yes</td>
</tr>
<tr>
<td>Functional Independence Measure</td>
<td>7 (6.1)</td>
<td>Both</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>ABILHAND</td>
<td>6 (5.2)</td>
<td>Upper</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Functional Ambulation Category</td>
<td>6 (5.2)</td>
<td>Both</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Functional Reach</td>
<td>6 (5.2)</td>
<td>Upper</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Emory Functional Ambulation Profile</td>
<td>5 (4.3)</td>
<td>Both</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Activities-Specific Balance Confidence Scale</td>
<td>4 (3.5)</td>
<td>Lower</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Barthel Index</td>
<td>4 (3.5)</td>
<td>Both</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Dynamic Gait Index</td>
<td>4 (3.5)</td>
<td>Lower</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>Jebsen Hand Function Test</td>
<td>4 (3.5)</td>
<td>Upper</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Motricity Index</td>
<td>4 (3.5)</td>
<td>Both</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Rivermead Mobility</td>
<td>4 (3.5)</td>
<td>Lower</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>SF-36/12</td>
<td>4 (3.5)</td>
<td>Both</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>VO2 Max</td>
<td>4 (3.5)</td>
<td>Lower</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Active Range of Motion</td>
<td>3 (2.6)</td>
<td>Upper</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Arm Motor Ability Test</td>
<td>3 (2.6)</td>
<td>Upper</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Motor Assessment Scale</td>
<td>3 (2.6)</td>
<td>Both</td>
<td>n/a</td>
<td>No</td>
</tr>
<tr>
<td>Nine-Hole Peg Test</td>
<td>3 (2.6)</td>
<td>Upper</td>
<td>3</td>
<td>No</td>
</tr>
<tr>
<td>Canadian Occupational Performance Measure</td>
<td>2 (1.7)</td>
<td>Upper</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Stroke Specific Quality of Life Scale</td>
<td>2 (1.7)</td>
<td>Both</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Chedoke Arm and Hand Activity Index</td>
<td>1 (0.9)</td>
<td>Upper</td>
<td>1</td>
<td>No</td>
</tr>
<tr>
<td>Goal Attainment Scale</td>
<td>1 (0.9)</td>
<td>Upper</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Five Times Sit to Stand</td>
<td>1 (0.9)</td>
<td>Lower</td>
<td>2</td>
<td>No</td>
</tr>
<tr>
<td>Medical Research Council Scale</td>
<td>1 (0.9)</td>
<td>Lower</td>
<td>n/a</td>
<td>No</td>
</tr>
</tbody>
</table>

n/a indicates not applicable.

*StrokEDGE Recommendation: 4=highly recommended, 3=recommended, 2=unable to recommend at this time, 1=not recommended, and n/a=not considered.
†Indicates measure evaluated by panel.
Table 2. Properties of Short-Listed Measures

<table>
<thead>
<tr>
<th>Scale</th>
<th>ICF Domain</th>
<th>MCID*</th>
<th>Psychometric Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Reliability†</td>
</tr>
<tr>
<td>FM-UE</td>
<td>Body function</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>FM-LE</td>
<td>Body function</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>WMFT</td>
<td>Activity</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>ARAT</td>
<td>Activity</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>10-m walk</td>
<td>Activity</td>
<td>–</td>
<td>+++</td>
</tr>
<tr>
<td>6-min walk</td>
<td>Activity</td>
<td>–</td>
<td>++</td>
</tr>
<tr>
<td>SIS subscales</td>
<td>Activity, performance</td>
<td>++</td>
<td>++</td>
</tr>
</tbody>
</table>

ARAT indicates Action Research Arm Test; FM-LE, the Fugl-Meyer Lower Extremity subscale; FM-UE, Fugl-Meyer Upper Extremity subscale; ICF, International Classification of Functioning, Disability, and Health System; MCID, minimal clinically important difference; SIS, Stroke Impact Scale; and WMFT, Wolf Motor Function Test.

†+++; established for chronic stroke using patient-based anchor; +, established for chronic stroke using arbitrary anchor or distribution-based methods, –, not established for chronic stroke.

‡+++; large effect size; ++, moderate effect size; +, small effect size; and –, not established for chronic stroke.

§+++; substantial validity evidence; ++, moderate validity evidence; and +, minimal validity evidence.

for the validity of this measure among patients with chronic stroke. For example, Hsieh et al20 report correlations of the FM-UE with ARAT and WMFT scores ranging from 0.71 to 0.76 at pretreatment and 0.51 to 0.74 at post-treatment, reflecting high construct validity. Pretreatment scores on the FM-UE have also been shown to be significantly predictive of post-treatment scores on the Functional Independence Measure (FIM; ρ = 0.42 with both FIM-Total and FIM-Motor25). It has an established MCID for chronic stroke that was derived based on therapists’ evaluation of patients’ global rating of change after an intervention.21

Fugl-Meyer Lower Extremity

The FM-LE assesses movement of the Achilles, patellar, hip, knee, and ankle with performance of 16 tasks that are rated on a scale from 0 (no active motion) to 2 (motion seems to be normal). The maximum score on the FM-LE subscale is 34 points (possible range, 0–34).

The FM-LE is a body function domain measure. The FM-LE has high internal consistency and test–retest reliability, as well as excellent responsiveness.19 Furthermore, there is support in the literature for face and content validity22,23 of the FM-LE in the population with chronic stroke. Evidence for other types of validity for this measure is lacking. Although Gladstone et al22 suggest an MCID based on 10% of the total score, the FM-LE does not have an empirically established MCID for use with chronic stroke patients.

Wolf Motor Function Test

The WMFT24 is a measure of upper extremity motor ability that is quantified through 15 timed movement tasks and 2 strength-based tasks. The performance time of every WMFT item is measured to derive a time subscale (WOLF-Time) and a 6-point Functional Ability Scale (WOLF-Functional Ability Scale) rates the quality of movement for each of the tasks with values ranging from 0 (no attempt made to use the more affected upper extremity) to 5 (movement seems to be normal).

The WMFT reflects functioning at the activity domain. Like the FM-UE, both scales of the WMFT have excellent psychometric properties for assessment in chronic stroke, including high inter-rater reliability, excellent test–retest reliability,25 and evidence for responsiveness.20 There is substantial evidence for the validity of this measure among patients with chronic stroke. For example, as noted above, correlations of the WMFT scales with the FM-UE and ARAT are significant at pre and post-treatment, reflecting moderate construct validity.20 Hsieh et al20 also report strong predictive validity for the WMFT-Time (p of 0.47 and 0.43 with FIM-Total and Motor, respectively) but predictive validity coefficients for the WMFT-Functional Ability Scale were not significant (p of 0.17 and 0.19 with FIM-Total and Motor, respectively). The WMFT has an established anchor-based MCID for chronic stroke that was derived using a change on the FM-UE ranging from 6 to 10 as the anchor.26

Arm Research Action Test

The ARAT,27 an evaluative measure of upper extremity motor ability consisting of 19 movement tasks divided into 4 subtests (grasp, grip, pinch, and gross arm movement), assesses a patient’s ability to handle objects differing in size, weight, and shape.28 The ARAT contains a 3-point functional ability scale that rates the quality of movement for each of the tasks and has values ranging from 0 (not able to perform any part of the test) to 3 (movement seems to be normal). The maximum score on the ARAT is 57 points (possible range, 0–57).

The ARAT reflects measurement at the activity domain level. The ARAT has excellent inter-rater29 and test–retest reliability30 for assessment in chronic stroke, and strong evidence of responsiveness.29 Like the FM-UE and WMFT scales, there is evidence for the validity of the ARAT in patients with chronic stroke, as these 3 measures have been shown to correlate highly with one another at both pre and post-treatment assessments.20 However, like the WMFT-Functional Ability Scale, the predictive validity coefficients for the ARAT as reported by Hsieh et al20 were not significant (p of 0.22 and 0.26 with FIM-Total and Motor, respectively). The ARAT has an established anchor-based MCID for chronic stroke that was derived using a 10% change on the total scale as the anchor.29
Ten-Meter Walk Test (Gait Velocity)
The 10MWT is an evaluative measure of walking speed which requires a 20-m, indoor, flat straight hallway. The first and last 5 m are used to accelerate and decelerate while only the middle 10 m are recorded. The patient is instructed to walk at a self-selected speed, using whatever walking aids might be needed, such as a walker or cane. The velocity is calculated as distance divided by time.

The 10MWT is an activity level domain measure. Although there is evidence for excellent internal consistency and test–retest reliability of the 10MWT among patients with chronic stroke, the evidence for responsiveness is scarce for this population. Validity evidence is also limited, but can be inferred from its association with community ambulation. For example, a compilation of many studies has established walking speed thresholds appropriate for assessing and monitoring functional status in the home and community and overall health in a wide range of populations, including patients with stroke. However, there is no established MCID for the 10MWT in populations with chronic stroke, specifically.

Six-Minute Walk Test (Gait Endurance)
The 6MWT quantifies the distance a patient can walk at self-selected walking speed on a flat, hard surface in a period of 6 minutes.

The 6MWT is an activity level domain measure. Despite its frequent use in chronic stroke trials, there is little psychometric evidence for the 6MWT in populations with chronic stroke. With the exception of excellent test–retest reliability noted by Flansbjer et al, all of the available evidence is based on populations with acute stroke. Furthermore, much of the evidence for the psychometric performance of this measure, most notably that pertaining to responsiveness and MCID, pertains only to patients with acute stroke. However, the test has clear face validity among patients with chronic stroke. The 6MWT does not have an established MCID for chronic stroke.

Stroke Impact Scale
The SIS is a multidimensional self-reported measure of stroke outcomes. The SIS version 3.0 consists of 59 items, each of which is rated on a 5-point Likert scale ranging from 1 (unable to complete that item) to 5 (no difficulty experienced at all) and is divided into 8 subtests or domains, 4 of which are closely related to upper and lower extremity mobility (hand function, activities of daily living/instrumental activities of daily living, mobility, and strength). There is also a 16-item subtest from the SIS (the SIS-16) that measures physical functioning and focuses primarily on lower extremity function. However, there is no documentation of its performance in patients with chronic stroke, thus we do not consider it further here.

The SIS represents assessment across the activity and participation domains. The 4 SIS subscales have adequate internal consistency and test–retest reliability. However, 3 of the 4 subscales have low responsiveness. Aside from the validity evidence for the proxy version of the SIS scales in populations with chronic stroke, validity evidence for the subscales is limited; baseline measures of hand function are significantly correlated (r=0.51), with FM-UE scores at post-treatment in patients with chronic stroke providing some evidence for the predictive validity of this subscale. MCIDs have been established for the population with chronic stroke for each of the 4 subscales related to motor function.

Discussion
Recommendations of the multidisciplinary expert panel were aimed to identify which existing measures of motor status are best suited to capture the effects of emerging interventions targeting residual motor deficits in the chronic phase after stroke. Overall, no existing measure was recommended or highly recommended by StrokeEDGE, had a published MCID for chronic stroke, and met all 3 evaluation criteria of measuring outcomes at the activity or participation level, having solid psychometric evidence and having an MCID that was established based on patient rating. In weighing the strengths and weaknesses of all measures, this panel recommends using the FM-UE and FM-LE scales as primary outcomes. The panel further recommends using either the WMFT or ARAT for upper extremity and 10MWT or 6MWT for lower extremity functioning as secondary outcomes. As noted above, the modified Rankin Scale that is commonly used to assess changes in functional ability in the acute phase after stroke, is not regarded as suitable for the chronic phase, as it lacks an MCID, as well as sufficient information on its responsiveness and prognostic value for recovery.

The recommendation of the FM-UE and FM-LE scales is based on 3 considerations. First, their psychometric properties are strong and well documented. In particular, the excellent validity evidence that shows associations of FM-UE and FM-LE scores with activity measure scores, outweighs our concerns over the fact that they measure outcomes at the body function level, and facilitates their meaningful interpretation by clinicians and patients; concerns that these FM scales are body function level assessments are further mitigated by their close correlation with measures of activity limitation. Second, they are by far the most frequently used measures in relevant clinical trials that has resulted in a rich body of empirical evidence for their use and meaningfulness, as well as familiarity to the stroke rehabilitation field, which also strengthens ability to interpret and compare findings across studies. Third, a patient-centric MCID for chronic stroke has been established for the FM-UE.

Although conceptually attractive because of their inherent patient centeredness, none of the existing measures that are based on self-report can be recommended at this point. The panel’s main concerns were the limited evidence for validity in populations with chronic stroke and the lower reliability compared with other short-listed measures. Among the ramifications of this lower measurement precision is its negative impact on the responsiveness of the subscales. In addition, the self-report assessment format can be particularly problematic in situations where treatment assignment is not blinded to the patient, as the patient perspective on his or her own functioning can be biased in this case. In addition, frequent concomitants to motor deficits after stroke, such as aphasia and neglect, could significantly impact the use of self-report measures when applied to broad populations of patients with stroke.

The panel’s review of the evidence has highlighted several directions for future research. In light of the increased recognition of the value of patient-centered outcomes for
clinical and regulatory decision making, future studies are needed to more fully develop the psychometric evidence for the SIS, especially with respect to its validity in reflecting true improvement in functioning. Furthermore, if self-reported measures such as the SIS are to be appropriately used in clinical trials with chronic stroke patients, it will be important to use study designs that minimize the potential for bias that may be associated with a lack of blinding.

Although a body function score such as the Fugl-Meyer will work well where the goal is restoration of motor function such as stem-cell therapy or treatments targeted at brain recovery and reorganization, in the chronic phase many studies target disability or activity, which call for assessment of participation outcomes. Thus, studies are also needed to provide more evidence for the performance-based measures. In particular, the establishment of patient-based MCIDs would greatly enhance the value of the WMFT and ARAT measures. Future research is also needed to evaluate and document all aspects of the psychometric performance of the 10MWT and 6MWT in populations with chronic stroke. The use of these measures as secondary outcomes in future studies will facilitate the accumulation of such evidence.

One limitation of this work may be that the panelists for the StrokEdge and this consensus committee were all American. It is possible that country and continent differences in usage of outcome measures may make acceptability of the panel’s recommendations outside of the United States a challenge. However, among the 115 clinical trials we evaluated, 70% (n=80) were based on research conducted outside of the United States. Thus despite the exclusive US nationality of the panel’s recommendations outside of the United States a challenge. However, among the 115 clinical trials we evaluated, 70% (n=80) were based on research conducted outside of the United States. Thus despite the exclusive US nationality of the panelists, the evidence they considered was internationally participatory outcomes. Thus, studies are also needed to provide more evidence for the performance-based measures. In particular, the establishment of patient-based MCIDs would greatly enhance the value of the WMFT and ARAT measures. Future research is also needed to evaluate and document all aspects of the psychometric performance of the 10MWT and 6MWT in populations with chronic stroke. The use of these measures as secondary outcomes in future studies will facilitate the accumulation of such evidence.

One limitation of this work may be that the panelists for the StrokEdge and this consensus committee were all American. It is possible that country and continent differences in usage of outcome measures may make acceptability of the panel’s recommendations outside of the United States a challenge. However, among the 115 clinical trials we evaluated, 70% (n=80) were based on research conducted outside of the United States. Thus despite the exclusive US nationality of the panelists, the evidence they considered was internationally participatory outcomes. Thus, studies are also needed to provide more evidence for the performance-based measures. In particular, the establishment of patient-based MCIDs would greatly enhance the value of the WMFT and ARAT measures. Future research is also needed to evaluate and document all aspects of the psychometric performance of the 10MWT and 6MWT in populations with chronic stroke. The use of these measures as secondary outcomes in future studies will facilitate the accumulation of such evidence.


Chronic Stroke Outcome Measures for Motor Function Intervention Trials: Expert Panel Recommendations
Cheryl Bushnell, Janet Prvu Bettger, Kevin M. Cockroft, Steven C. Cramer, Maria Orlando Edelen, Daniel Hanley, Irene L. Katzan, Soeren Mattke, Dawn M. Nilsen, Tepring Piquado, Elizabeth R. Skidmore, Kay Wing and Gayane Yenokyan

_Circ Cardiovasc Qual Outcomes_. 2015;8:S163-S169
doi: 10.1161/CIRCOUTCOMES.115.002098

_Circulation: Cardiovascular Quality and Outcomes_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/8/6_suppl_3/S163

Data Supplement (unedited) at:
http://circoutcomes.ahajournals.org/content/suppl/2015/10/29/CIRCOUTCOMES.115.002098.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Quality and Outcomes_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation: Cardiovascular Quality and Outcomes_ is online at:
http://circoutcomes.ahajournals.org//subscriptions/
SUPPLEMENTAL MATERIAL

Supplemental Methods

Search strategy specifications

A search strategy of the past six years (2009-2014), using the following two sets of keywords: “chronic stroke” AND “motor function” OR "Stroke/rehabilitation" AND “chronic disease”; “stroke” AND “motor function” OR “motor skills” OR “upper extremity” OR “lower extremity” OR “gait” OR “grip strength” OR “hand strength” OR “mobility” OR “walking”, AND “clinical trial” OR “randomized control trial” AND “poststroke” OR post acute” OR “longer term” OR “chronic” OR “6 months” OR “12 months”, yielded 1608 references, which were evaluated for inclusion using the following criteria: stroke onset greater than or equal to six months prior; study included only stroke patients; target of intervention was upper or lower extremity function only. Exclusions were review papers, case studies, meta-analyses, single cohort pre-post designs, and studies with fewer than 10 subjects; also excluded were studies that used narrow outcome measures (e.g. only measured fMRI, knee flexion). We excluded 1493 articles and were left with 115 total; 55 lower extremity and 60 upper extremity.

Supplemental References


34. Embrey DG, Holtz SL, Alon G, Brandsma BA, McCoy SW. Functional electrical stimulation to dorsiflexors and plantar flexors during gait to improve walking in adults with chronic hemiplegia. *Archives of physical medicine and rehabilitation*. 2010; 91:687-696


111. Wu CY, Huang PC, Chen YT, Lin KC, Yang HW. Effects of mirror therapy on motor and sensory recovery in chronic stroke: A randomized controlled trial. *Archives of physical medicine and rehabilitation*. 2013; 94:1023-1030


