Simplified Modified Rankin Scale Questionnaire
Reproducibility Over the Telephone and Validation With Quality of Life

Askiel Bruno, MD; Abiodun E. Akinwuntan, PhD; Chen Lin, BS; Brian Close, BS; Kristin Davis, MD; Vanessa Baute, MD; Tia Aryal, MD; Desiree Brooks, BS; David C. Hess, MD; Jeffrey A. Switzer, DO; Fenwick T. Nichols, MD

Background and Purpose—The simplified modified Rankin Scale questionnaire (smRSq) enables a reliable and rapid determination of the modified Rankin Scale score after stroke. We test the reliability and validity of a slightly revised smRSq.

Methods—Fifty consecutive outpatients 4.83±3.00 months after stroke were scored with a slightly revised smRSq by 3 raters selected consecutively from a list of 10: 4 stroke faculty, 3 neurology residents, 2 medial students, and 1 stroke research coordinator. Two ratings were in person within 20 minutes of each other and 1 was by telephone 1 to 3 days later. The telephone rating also included a quality of life scale, the Short-Form-12v2. Each rater was blinded to the other raters’ scores.

Results—The average estimated time to administer the smRSq was 1.29 minutes (range, 0.50 to 2.25 minutes). The in-person raters agreed 78% (κ=0.71; CI, 0.57 to 0.86 and weighted κ [κw]=0.86; CI, 0.79 to 0.94). The first in-person and telephone raters agreed 82% (κ=0.76; CI, 0.63 to 0.90 and κw=0.87; CI, 0.79 to 0.95). The second in-person and telephone rates agreed 82% (κ=0.77; CI, 0.63 to 0.90 and κw=0.89; CI, 0.82 to 0.96). The smRSq correlated with the physical (r=−0.50, P=0.005) than the mental (r=−0.36, P=0.048) components of the Short-Form-12v2.

Conclusions—The slightly revised smRSq appears to be useful in clinical stroke; it has excellent reliability in person and by telephone, can usually be administered in <1.5 minutes by a wide variety of raters, and correlates with quality of life. (Stroke. 2011;42:00-00.)

Key Words: clinimetrics ■ modified Rankin Scale ■ outcome assessment

R e liability (interrater agreement) and feasibility, including telephone assessment, of valid outcome measures in stroke are essential for reproducible and efficient determination of outcomes. The modified Rankin Scale (mRS) remains a popular validated functional outcome measure in acute stroke trials and other stroke studies. However, its usefulness has been limited by suboptimal reliability, and there are limited data from complete reports about its reliability over the telephone. We recently set out to improve the reliability of the mRS at the same time as keeping the assessments relatively simple and short and preserving the construct and validity of the original mRS. We created 5 relatively simple questions that address the key functional states assigned to each mRS score, the simplified mRS questionnaire (smRSq). This questionnaire requires yes or no answers by a patient or a caregiver, takes on average only 1.67 minutes to administer, and has very good reliability (agreement 78%, κ=0.72, and weighted κ [κw]=0.82). Because the agreement was weakest on the questions related to mRS scores 3 to 5 (63% to 68%), we then tried to improve the smRSq by making the 2 questions regarding mRS scores 3 to 5 more specific. In this study, we tested a slightly revised smRSq for interrater agreement in person and by telephone and for correlation with a quality of life scale, the Short-Form-12 version 2 (SF-12v2).

Patients and Methods
To improve the objectivity of the original smRSq, we made the questions related to scores 3 to 5 more specific (Figure; Table 1). We defined the ability to walk without assistance (mRS ≤3) as able to go from lying to sitting under own power.

We screened patients for this study consecutively in 4 weekly continuity clinics staffed by 4 stroke specialists and in 3 weekly resident continuity clinics at the Medical College of Georgia Hospital between July 2010 and October 2010. Included were patients with a diagnosis of ischemic or hemorrhagic stroke 1 to 12 months before enrolment who have been discharged from acute care. We confirmed the diagnosis of stroke if there was a sudden focal neurological deficit lasting >24 hours and cerebral imaging showed no cause.
The 10 mRS Web-certified raters included 4 faculty stroke specialists, 3 neurology residents, 2 medical students, and a stroke research coordinator. For each subject, 3 raters were identified consecutively as first, second, and third from an alphabetic list of all 10 raters. The first 2 raters administered the smRSq in person within 20 minutes of each other, and the third rater administered the questionnaire by telephone 1 to 3 days later. Patients were instructed to respond to each interview session independently of the others. The third rater also administered the SF-12v2 by telephone, when possible. The SF-12v2 is a quality-of-life scale derived from the longer SF-36 version, consists of physical and mental components, and has been validated by telephone. When a designated rater was unavailable, the next rater on the list was contacted until an available rater was found. For each new patient, Rater 1 followed directly Rater 3 for the preceding patient on the alphabetic list.

The in-person ratings were done with patients and their caregivers when possible by asking the smRSq questions from top to bottom (Figure). In cases of disagreement between patients and their caregivers, the caregivers’ answers were accepted as more accurate. The telephone smRSq was directed at the patients if their answers in person were not corrected by their caregivers. Otherwise, the telephone questionnaires were directed at the caregivers who assisted with the in-person interviews. Each rater’s scores were concealed in envelopes from the other raters until all subjects were rated.

Table 1. Percent Agreements Within Each Rater Pair on the 5 smRSq Questions*

<table>
<thead>
<tr>
<th>Revised smRSq Question</th>
<th>Rater Pair</th>
</tr>
</thead>
<tbody>
<tr>
<td>If you had to, could you live alone without any help from another person? This means</td>
<td>First to Second</td>
</tr>
<tr>
<td>being able to bathe, use the toilet, shop, prepare or get meals, and manage finances.</td>
<td>94 (50)</td>
</tr>
<tr>
<td>(score ≤2 or ≤3)</td>
<td>First to Third</td>
</tr>
<tr>
<td>96 (50)</td>
<td>Second to Third</td>
</tr>
<tr>
<td>96 (50)</td>
<td></td>
</tr>
<tr>
<td>Can you do everything that you were doing right before your stroke, even if slower and</td>
<td>86 (21)</td>
</tr>
<tr>
<td>not as much? (score 2 or ≤1)</td>
<td>100 (21)</td>
</tr>
<tr>
<td>91 (22)</td>
<td></td>
</tr>
<tr>
<td>Are you completely back to the way you were right before your stroke? (score 0 or 1)</td>
<td>87 (16)</td>
</tr>
<tr>
<td>82 (17)</td>
<td></td>
</tr>
<tr>
<td>81 (16)</td>
<td></td>
</tr>
<tr>
<td>Can you walk from one room to another without help from another person? (score 3 or</td>
<td>96 (26)</td>
</tr>
<tr>
<td>4); original question: Are you able to walk without help from another person?</td>
<td>96 (26)</td>
</tr>
<tr>
<td>92 (26)</td>
<td></td>
</tr>
<tr>
<td>Can you sit up in bed without any help? (score 4 or 5); original question: Are you</td>
<td>78 (9)</td>
</tr>
<tr>
<td>bedridden or needing constant supervision?</td>
<td>75 (8)</td>
</tr>
<tr>
<td>100 (8)</td>
<td></td>
</tr>
</tbody>
</table>

smRSq indicates simplified modified Rankin Scale questionnaire.

*Numbers in parentheses show patients who were asked the indicated question by each rater pair. For example, all 50 patients were asked the first question by all 3 rater pairs and so on. Ratings first and second were in person and the third was by telephone.
The SF-12v2 assessment was approved and added to this study after the first 11 patients were already scored and was administered only to the patients according to the scale instructions. Thus, some patients with cognitive deficits could not be scored with this instrument.

After rating all 50 patients in this study, the results were tabulated and analyzed using SAS statistical software (Cary, NC). We calculated the percentage agreement and determined $\kappa$ and $\kappa_w$ for agreement between each pair of raters. Biserial correlation analysis compared the physical and the mental component SF-12v2 scores with the smRSq scores. Because each patient had 3 smRSq scores, the majority score was taken in cases of disagreement as the final score for this correlation. In cases of 3 different smRSq scores, the middle score was taken. Also, each rater estimated the average time they took to score the smRSq. This study was approved by the Medical College of Georgia Institutional Review Board.

Results

The 50 patients in this study had a mean age of $61 \pm 14$ (SD) years, 24 (48%) were men, 27 (54%) were black, 22 (44%) were white, and 1 (2%) was Asian. Eight patients (16%) had intracerebral hemorrhage and 42 (84%) had ischemic stroke. Average time from stroke to assessment was 4.83 ± 3.00 (SD) months. The 4 faculty members performed a total of 62 ratings, the 3 residents 42, the 2 medical students 26, and the stroke research coordinator 20. The average time estimated by the 10 raters to administer the smRSq was 1.29 minutes (range, 0.5 to 2.25 minutes).

The in-person raters agreed 78% of the time with $\kappa=0.71$ (95% CI, 0.57 to 0.86) and $\kappa_w$ taking into account the extent of disagreement=0.86 (0.79 to 0.94). The first in-person and the telephone raters agreed 82% of the time with $\kappa=0.76$ (0.63 to 0.90) and $\kappa_w=0.87$ (0.79 to 0.95). The second in-person and the telephone raters agreed 82% of the time with $\kappa=0.77$ (0.63 to 0.90) and $\kappa_w=0.89$ (0.82 to 0.96). There were insufficient rater pairs with similar levels of clinical experience for a meaningful reliability analysis based on experience. However, greater clinical experience did not seem to increase smRSq reproducibility because the agreement was 65% among the 17 ratings where each rater was faculty.

Table 1 shows the agreements for the 3 rater pairs on each of the smRSq questions. Table 2 shows the crosstabulation of the smRSq scores for the 3 rating pairs. The SF-12v2 scores were available for correlation analysis in 31 patients and in this subgroup, the smRSq agreements were similar to that of the entire study sample ($\kappa_w$=0.79 to 0.86). Patients with higher smRSq scores were more likely to have a lower quality of life (lower SF-12v2 scores). The correlation between smRSq and the SF-12v2 was better for the physical component ($r=-0.50, P=0.005$) than the mental component ($r=-0.36, P=0.048$).

Table 2. Interrater Agreements on the smRSq for the 3 Rating Pairs*

<table>
<thead>
<tr>
<th>smRSq</th>
<th>Rating 1</th>
<th>Rating 2</th>
<th>Rating 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5</td>
<td>0 1 2 3 4 5</td>
<td>0 1 2 3 4 5</td>
</tr>
<tr>
<td>mRS 0</td>
<td>3 1</td>
<td>3 1</td>
<td>3 1</td>
</tr>
<tr>
<td>mRS 1</td>
<td>1 11 2</td>
<td>1 11 2</td>
<td>1 11 2</td>
</tr>
<tr>
<td>mRS 2</td>
<td>— 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>mRS 3</td>
<td>— — 2</td>
<td>16 1 —</td>
<td>2 16 1 —</td>
</tr>
<tr>
<td>mRS 4</td>
<td>— — — 3</td>
<td>2 4 —</td>
<td>2 4</td>
</tr>
<tr>
<td>mRS 5</td>
<td>— — — — 4</td>
<td>5 —</td>
<td>5</td>
</tr>
</tbody>
</table>

The percentage agreement and determined $\kappa$ and $\kappa_w$ for agreement between each pair of raters. Biserial correlation analysis compared the physical and the mental component SF-12v2 scores with the smRSq scores. Because each patient had 3 smRSq scores, the majority score was taken in cases of disagreement as the final score for this correlation. In cases of 3 different smRSq scores, the middle score was taken. Also, each rater estimated the average time they took to score the smRSq. This study was approved by the Medical College of Georgia Institutional Review Board.

Discussion

We tested a novel slightly revised simplified mRS scoring method, the smRSq (Figure), for reproducibility after stroke, both in person and by telephone. In a subgroup of patients, we also compared the smRSq with a telephone validated quality-of-life scale, the SF-12v2.

The SF-12v2 assessment was approved and added to this study after the first 11 patients were already scored and was administered only to the patients according to the scale instructions. Thus, some patients with cognitive deficits could not be scored with this instrument.

After rating all 50 patients in this study, the results were tabulated and analyzed using SAS statistical software (Cary, NC). We calculated the percentage agreement and determined $\kappa$ and $\kappa_w$ for agreement between each pair of raters. Biserial correlation analysis compared the physical and the mental component SF-12v2 scores with the smRSq scores. Because each patient had 3 smRSq scores, the majority score was taken in cases of disagreement as the final score for this correlation. In cases of 3 different smRSq scores, the middle score was taken. Also, each rater estimated the average time they took to score the smRSq. This study was approved by the Medical College of Georgia Institutional Review Board.
Although only a subgroup of 31 patients had the SF-12v2 done, the smRSq agreements in this subgroup ($k_{pa}$ 0.79 to 0.86) were similar to that of the entire study sample, suggesting similarity between these 2 populations. Other studies found significant correlations between the mRS and SF-36 (the parent instrument to the SF-12v2) after ischemic stroke,11–14 and 1 study found a significant correlation between the mRS and SF-12 after subarachnoid hemorrhage.15 In those studies, the mRS correlated better with the physical than the mental component of the quality-of-life scales, as might be expected from the physical construct of the mRS. In our study, the smRSq also correlated better with the physical than the mental component of the SF-12v2. The similar correlations with the quality-of-life scales in the previous studies and in our study suggest similarity between the smRSq and traditional mRS.

It is noteworthy that the agreements are $\geq 94\%$ on the first smRSq question (Table 1), which is a common mRS primary outcome cut point (mRS $\leq 2$) in clinical trials. Also, agreements are $\geq 86\%$ on the question distinguishing between scores 2 and $\leq 1$, which is another common primary outcome cut point. The agreements on the revised questions about walking ($\geq 92\%$) and about being bedridden ($\geq 75\%$) are better than in the original smRSq (68% and 63%, respectively).6

Repeat testing with the smRSq may have introduced recall bias. Such bias might be expected to be greater for the in-person ratings done within 20 minutes of each other than for the telephone ratings done 1 to 3 days later. We tried to limit such bias by instructing patients to treat each interview independently of the others. Also, some patients may have reconsidered and changed their answers between interviews, thus counteracting recall bias.

With the smRSq, we intended to create a simplified and more reliable method of scoring the mRS without altering its established validity. However, the smRSq method may yield somewhat different results than the traditional mRS,16 especially at the higher scores in which specific definitions about walking and being bedridden were introduced. However, because the ability to walk without assistance is a key distinction between mRS 3 and 4, the smRSq and the traditional mRS may yield similar results at this level. A direct comparison of the smRSq to the traditional or structured mRS, especially by other investigators, is needed and could identify differences between them, if present.

**Conclusions**

The slightly revised smRSq tested in this study shows excellent reliability that is similar to the original version and the structured interview mRS.4 In addition, this latest smRSq appears to have excellent reproducibility by telephone, correlates with a validated quality-of-life scale (SF-12v2),7 takes $<1.5$ minutes to administer, and can be used by a wide variety of raters. These features qualify it as a useful method of scoring the mRS.

**Disclosures**

None.

**References**


Simplified Modified Rankin Scale Questionnaire: Reproducibility Over the Telephone and Validation With Quality of Life
Askiel Bruno, Abiodun E. Akinwuntan, Chen Lin, Brian Close, Kristin Davis, Vanessa Baute, Tia Aryal, Desiree Brooks, David C. Hess, Jeffrey A. Switzer and Fenwick T. Nichols

Stroke, published online June 16, 2011; Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/early/2011/06/16/STROKEAHA.111.613273

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Stroke 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 Stroke is online at:
http://stroke.ahajournals.org//subscriptions/