Stroke

American Stroke Association.

A Division of American Heart Association

Improving Modified Rankin Scale Assessment With a Simplified Questionnaire

Askiel Bruno, Neel Shah, Chen Lin, Brian Close, David C. Hess, Kristin Davis, Vanessa Baute, Jeffrey A. Switzer, Jennifer L. Waller and Fenwick T. Nichols

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/41/5/1048

Improving Modified Rankin Scale Assessment With a Simplified Questionnaire

Askiel Bruno, MD; Neel Shah, BS; Chen Lin, BS; Brian Close, BS; David C. Hess, MD;
Kristin Davis, MD; Vanessa Baute, MD; Jeffrey A. Switzer, DO;
Jennifer L. Waller, PhD; Fenwick T. Nichols, MD

Background and Purpose—The modified Rankin Scale (mRS) is a popular primary stroke outcome measure, but its usefulness is limited by suboptimal reliability (inter-rater agreement).

Methods—We developed and tested the reliability of a simplified mRS questionnaire (smRSq) in 50 patients after stroke seen in outpatient clinics. Randomly chosen paired raters administered the smRSq within 20 minutes of each other and the ratings were blinded until the end of this study.

Results—Agreement among the raters was 78%, the κ statistic was 0.72 (95% CI, 0.58–0.86), and the weighted κ_w statistic taking into account the extent of disagreement was 0.82 (95% CI, 0.72–0.92). The average time to administer the smRSq was 1.67 minutes.

Conclusions—The smRSq appears to have very good reliability that is similar to that of a structured interview mRS and is considerably less time-consuming. (***Stroke***. 2010;41:1048-1050.)

Key Words: clinometrics ■ modified Rankin Scale ■ outcome assessment

Reliability (consistency) of measurements is of paramount importance in scientific research.[1] The modified Rankin Scale (mRS)[2] is a popular primary outcome measure in acute stroke trials, but its usefulness is limited by suboptimal reliability (inter-rater agreement). There is considerable variability in the reported reliability of the mRS.[3] A structured interview mRS that takes ≈15 minutes to administer was developed to help improve the mRS reliability.[4] In a recent systematic review the overall agreement between mRS raters without a standardized rating approach was 71%, the kappa (κ) statistic was 0.46 (95% CI, 0.41–0.51), and the weighted kappa (κ_w) statistic, taking into account the extent of all disagreements, was 0.90 (95% CI, 0.86–0.94).[3] Using the structured interview mRS, the overall agreement was ≈73%, κ was 0.62 (95% CI, 0.56–0.69), and κ_w was 0.87 (95% CI, 0.75–1.00).[3,5,6] Inter-rater agreement was significantly improved with the structured interview mRS among raters with varied professional backgrounds,[7] which simulates a multi-center clinical trial. In an effort to simplify, standardize, and increase further the reliability of the mRS, we developed a simplified mRS questionnaire (smRSq) and tested it among raters with varied professional experiences.

Materials and Methods

Four stroke faculty members with a total of 63 years experience applying the traditional unstructured mRS jointly created the smRSq

(Figure). Using the key issues distinguishing between consecutive mRS categories, we created relatively simple questions that could be answered "yes" or "no" by patients or caregivers with little or no explanation. The key mRS issues were having no residual symptoms (0), being able to resume all prestroke activities (≤1), being able to live independently (≤2), being able to walk without assistance (≤3), and not requiring constant supervision (≤4).

We screened patients for this study consecutively in 4 weekly clinics staffed by the stroke specialists and in 2 weekly resident continuity clinics at the Medical College of Georgia. Eligibility for this study required a diagnosis of ischemic or hemorrhagic stroke 1 to 12 months before the assessment and having been discharged from acute care. The diagnosis of stroke required documentation of at least a sudden focal neurological deficit lasting longer than 24 hours and cerebral neuroimaging showing no other cause.

The 9 mRS web-certified[8] raters in this study included 4 faculty stroke specialists, 2 neurology residents, 2 second-year medical students, and a stroke coordinator. We generated a random list of paired raters before this study began. The order of the paired raters was also random. The 2 raters administered the smRSq within 20 minutes of each other. If either of the designated paired raters was unavailable to administer the smRSq when a patient was enrolled, then the next pair of raters was called until a randomly selected pair was available. Ratings were performed face-to-face with patients and their caregivers by asking the prespecified questions (Figure). The 2 interviews of each patient were performed in an independent fashion. In cases of persistent disagreement between patients and their caregivers, the caregivers' answers were accepted as more accurate.[9] Each rater's scores were concealed in envelopes from the other raters until all subjects were rated.

After rating all 50 patients in this study, the results were tabulated and analyzed. The κ and κ_w statistic examined the degree of

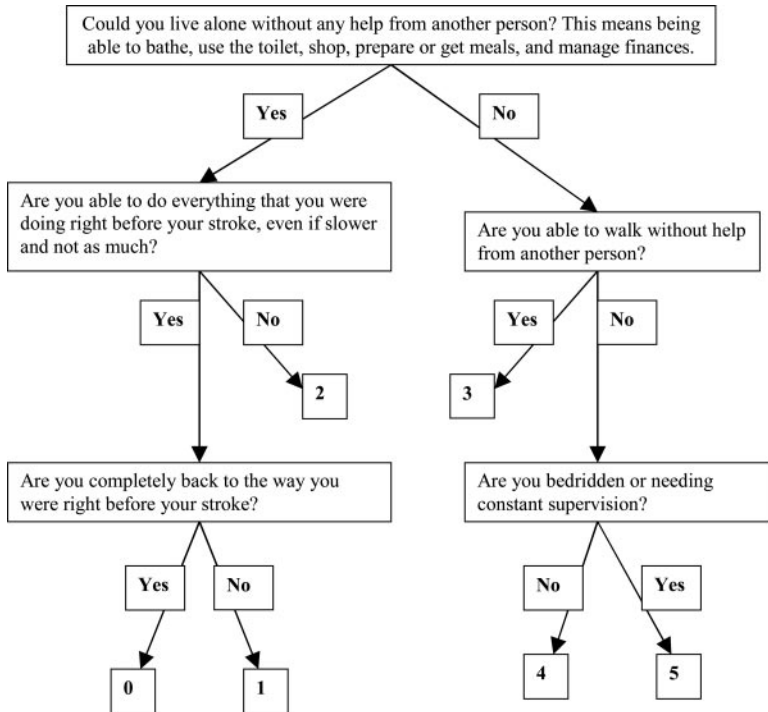*Stroke* **is available at http://stroke.ahajournals.org**

Figure. The simplified mRS questionnaire algorithm.

agreement between the raters in each pair. The Bowker test evaluated symmetry in cross-tabulation of the scores. This study was approved by the Medical College of Georgia Institutional Review Board.

Results

Of the 50 subjects in this study the mean age was 60 ± 13 years, 23 (46%) were men, 23 (46%) were black, and 26 (52%) were white. Six patients (12%) had intracerebral hemorrhage and the remainder had ischemic stroke. The 4 faculty members performed a total of 60 ratings, the 2 residents performed 17 ratings, the 2 medical students performed 11 ratings, and the stroke research coordinator performed 12 ratings. The average time estimated by the 9 raters to administer the smRSq was 1.67 minutes.

The overall agreement between the rater pairs was 78%. The Table shows the cross-tabulation of the smRSq scores by the first and second raters of the 50 rater pairs. All rater pairs asked all 50 patients about being able to live independently (first question) and the agreement was excellent (48/50; 96%). Subsequently, for scores 0 to 2, 13 pairs of

raters asked 13 patients about resumption of all prestroke activities and 3 pairs of raters asked 3 patients about complete recovery and the agreement on each of these questions was perfect. However, for scores 3 to 5, 22 paired raters asked 22 patients about walking unassisted and the agreement was 68% (15/22) and 8 paired raters asked 8 patients about being bedridden or constant supervision; the agreement was 63% (5/8).

Bowker test ($P=0.98$) indicated symmetry in the scores between raters. The κ statistic was 0.72 (95% CI, 0.58–0.86) and the weighted κ_w statistic, taking into account the extent of disagreement, was 0.82 (95% CI, 0.72–0.92). Our study was too small to test for smRSq reliability based on rater experience. However, all 3 disagreements by >1 point involved only the 4 faculty members.

Discussion

Our simplified questionnaire version of the mRS offers a new standardized approach to mRS rating that could prove advantageous in multicenter clinical stroke trials. The smRSq shows substantial[1] reliability in this study among raters with diverse professional experiences that is similar to the reliability of the structured interview mRS[3] and is considerably less time-consuming. The percent agreement between raters appears excellent for mRS scores 0 to 2, which are most relevant for the functional outcomes in acute stroke trials.

The smRSq questions can be understood by the majority of patients and caregivers with little or no explanation, and the assessment usually can be completed in <2 minutes. Because the smRSq is based on the key criteria that distinguish the mRS categories, we believe that the validity of the smRSq is similar to that of the traditional or structured interview mRS.

The paired ratings in this study performed within only 20 minutes of each other might have resulted in simple repetition of the answers from the first to the second interview, yielding

Table. Agreement Between the First and Second Raters in Scoring the smRSq

	Second Rater					
	0	1	2	3	4	5
First rater						
0	3	...	...	...	...	...
1	...	10	...	...	...	...
2	...	2	11	...	...	...
3	...	2	...	10	2	1
4	...	...	...	1	2	1
5	...	...	...	...	2	3

artificially high reliability. However, we believe that the patients and caregivers answered the study questions as accurately as they could during the first interview and that the second interview answers were at risk for being different because of the additional time for consideration.

As with the traditional or structured mRS rating, judgment is needed during the smRSq rating to decide who is providing more accurate answers in cases of disagreement between patients and their caregivers. Because stroke survivors tend to overstate their abilities,[9] it is probably best to accept the caregivers' answers. A rater's clinical experience does not appear to offer an advantage toward the reliability of smRSq in this study. Similarly, professional background did not affect the reliability of web-based mRS rating in the United Kingdom.[10] This suggests that these outcome assessments can be performed reliably by a wide variety of raters. We anticipate that the overall reliability of the smRSq and the agreement for scores 3 to 5 will improve by more specifically defining what qualifies as walking unassisted and what constitutes being bedridden and needing constant supervision.

Acknowledgments

Disclosures

None.

References

1. Feinstein AR. *The scientific importance of consistency*. *Clinimetrics*. New Haven, CT: Yale University Press; 1987:169–170.
2. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19:604–607.
3. Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin scale: A systematic review. *Stroke*. 2009;40:3393–3395.
4. Wilson JT, Hareendran A, Grant M, Baird T, Schulz UG, Muir KW, Bone I. Improving the assessment of outcomes in stroke: Use of a structured interview to assign grades on the modified Rankin scale. *Stroke*. 2002; 33:2243–2246.
5. Cincura C, Pontes-Neto OM, Neville IS, Mendes HF, Menezes DF, Mariano DC, Pereira IF, Teixeira LA, Jesus PA, de Queiroz DC, Pereira DF, Pinto E, Leite JP, Lopes AA, Oliveira-Filho J. Validation of the national institutes of health stroke scale, modified Rankin scale and Barthel Index in Brazil: The role of cultural adaptation and structured interviewing. *Cerebrovasc Dis*. 2009;27:119–122.
6. Shinohara Y, Minematsu K, Amano T, Ohashi Y. Modified Rankin scale with expanded guidance scheme and interview questionnaire: Interrater agreement and reproducibility of assessment. *Cerebrovasc Dis*. 2006;21: 271–278.
7. Wilson JT, Hareendran A, Hendry A, Potter J, Bone I, Muir KW. Reliability of the modified Rankin scale across multiple raters: Benefits of a structured interview. *Stroke*. 2005;36:777–781.
8. http://rankin-english.trainingcampus.net/uas/modules/trees/windex.aspx. Accessed December 14, 2009.
9. Knapp P, Hewison J. Disagreement in patient and carer assessment of functional abilities after stroke. *Stroke*. 1999;30:934–938.
10. Quinn TJ, Dawson J, Walters MR, Lees KR. Variability in modified Rankin scoring across a large cohort of international observers. *Stroke*. 2008;39:2975–2979.