

Association of Modified Rankin Scale With Recovery Phenotypes in Patients With Upper Extremity Weakness After Stroke

Kimberly S. Erler, PhD, Rui Wu, BA, Julie A. DiCarlo, MS, Marina F. Petrilli, BA, Perman Gochyev, PhD, Leigh R. Hochberg, MD, PhD, Steven A. Kautz, PhD, Lee H. Schwamm, MD, Steven C. Cramer, MD, Seth P. Finklestein, MD, and David J. Lin, MD

Correspondence

Dr. Lin
dlin7@mgh.harvard.edu

Neurology® 2022;98:e1877-e1885. doi:10.1212/WNL.0000000000200154

Abstract

Background and Objectives

Precise measurement of outcomes is essential for stroke trials and clinical care. Prior research has highlighted conceptual differences between global outcome measures such as the modified Rankin Scale (mRS) and domain-specific measures (e.g., motor, sensory, language or cognitive function). This study related motor phenotypes to the mRS, specifically aiming to determine whether mRS levels distinguish motor impairment and function phenotypes, and to compare mRS outcomes to meaningful changes in impairment and function from acute to subacute recovery after stroke.

Methods

Patients with upper extremity weakness after ischemic stroke were assessed with a battery of impairment and functional measures within the first week and at 90 days after stroke. Impairment and functional outcomes were examined in relation to 90-day mRS scores. Clinically meaningful changes in motor impairment, activities of daily living, and mobility were examined in relation to 90-day mRS score.

Results

In this cohort of 73 patients with stroke, impairment and functional outcomes were associated with 90-day mRS scores but showed substantial variability within individual mRS levels: within mRS level 2, upper extremity impairment ranged from near hemiplegia (with an upper extremity Fugl-Meyer score 8) to no deficits (upper extremity Fugl-Meyer score 66). Overall, there were few differences in impairment and functional outcomes between adjacent mRS levels. While some outcome measures were significantly different between mRS levels 3 and 4 (Nine-Hole Peg, Leg Motor, gait velocity, Timed Up and Go, NIH Stroke Scale, and Barthel Index), none of the outcome measures differed between mRS levels 1 and 2. Fugl-Meyer and grip strength were not different between any adjacent mRS levels. A substantial number of patients experienced clinically meaningful changes in impairment and function in the first 90 days after stroke but did not achieve good mRS outcome (mRS score ≤ 2).

Discussion

The mRS broadly relates to domain-specific outcomes after stroke, confirming its established value in stroke trials, but it does not precisely distinguish differences in impairment and function, nor does it sufficiently capture meaningful clinical changes across impairment, activities of daily living status, and mobility. These findings underscore the potential utility of incorporating detailed phenotypic measures along with the mRS in future stroke trials.

From the Center for Neurotechnology and Neurorecovery (K.S.E., R.W., J.A.D., L.R.H., D.J.L.), Division of Neurocritical Care (L.R.H., D.J.L.), Stroke Service (L.R.H., L.H.S., S.P.F., D.J.L.), Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston; Department of Occupational Therapy (K.S.E.), Massachusetts General Hospital, Boston; School of Health and Rehabilitation Sciences (K.S.E., M.F.P., P.G., D.J.L.), MGH Institute of Health Professions, Boston, MA; VA RR&D Center for Neurorestoration and Neurotechnology (L.R.H., D.J.L.), VA Medical Center, Providence, RI; School of Health Sciences and Research (S.A.K.), Medical University of South Carolina, Charleston; Ralph H. Johnson VA Medical Center (S.A.K.), Charleston, SC; Department of Neurology (S.C.C.), University of California, Los Angeles; California Rehabilitation Institute (S.C.C.), Los Angeles.

Go to [Neurology.org/N](https://www.neurology.org/N) for full disclosures. Funding information and disclosures deemed relevant by the authors, if any, are provided at the end of the article.

Glossary

ADL = activities of daily living; **BBT** = Box and Blocks Test; **BI** = Barthel Index; **FMA-UE** = Fugl Meyer Assessment of Upper Extremity; **GV** = gait velocity; **MCID** = minimal clinically important difference; **mRS** = modified Rankin Scale; **NIHSS** = NIH Stroke Scale; **9-HP** = Nine Hole Peg; **PCA-1** = first principal component; **SIS** = Stroke Impact Scale-16; **TUG** = Timed Up and Go; **UE** = upper extremity.

Precise measurement of outcomes is essential for stroke trials and clinical care,¹⁻⁴ yet there are different approaches to measuring recovery after stroke. Global approaches capture general function and disability, while modality-specific approaches provide details on the behavioral output of specific neural systems (e.g., motor, language, cognitive).^{5,6} Stroke clinical trials most commonly use the modified Rankin Scale (mRS), a 1-item, 7-level, clinician-rated, ordinal measure of global disability, as the primary endpoint.⁷⁻¹² The mRS is efficient to administer, practical, and reliable and encompasses the full range of clinically relevant outcomes after stroke, with levels ranging from no deficits to death.^{13,14} While the mRS is widely accepted as the gold standard and a meaningful endpoint in stroke trials, it has also been challenged for its restricted responsiveness, lack of granularity, and inconsistent ability to detect meaningful change.^{15,16} Ordinal mRS levels are unequally and somewhat arbitrarily spaced; it can be challenging to interpret the differences in disability between each level.^{17,18} These weaknesses become more pronounced in examinations of longer-term recovery after stroke. The mRS has limited utility for detecting change for trials spanning the first months after stroke because the accuracy of mRS levels measured during hospitalization in the acute stroke period is questionable.^{19,20} The mRS is a valuable outcome measure after stroke, yet understanding the limits of measuring stroke outcomes with the mRS is key to defining its most effective and appropriate use in stroke trials.

Upper extremity (UE) deficits are the most common source of disability and a contributing factor to overall quality of life and subjective well-being after stroke.²¹⁻²³ Therefore, to further understand precisely what the mRS measures when it classifies stroke outcomes, we examined differences in impairment and function across levels of the mRS among a sample of patients with UE weakness 90 days after ischemic stroke. We specifically aimed to determine whether mRS levels distinguish distinct detailed phenotypes of impairment and function and to compare mRS outcomes to meaningful changes in impairment and function from the acute to subacute stage of stroke recovery among a cohort of patients with UE weakness. We hypothesized that the mRS would be highly correlated with commonly used impairment and function outcome measures for patients with stroke but that adjacent mRS levels would not statistically capture distinct phenotypes of impairment or function.

Methods

This study analyzed data from a prospective single-center study that enrolled participants with UE weakness after stroke

(Stroke Motor Rehabilitation and Recovery Study; ClinicalTrials.gov identifier: NCT03485040). For this substudy, patients admitted to Massachusetts General Hospital between May 21, 2017, and November 11, 2019, with stroke were eligible to participate when they met the following criteria at the time of consent: (1) between 18 and 90 years of age; (2) evidence of an ischemic stroke resulting in unilateral UE motor weakness as defined by a score of ≥ 1 on the NIH Stroke Scale (NIHSS) arm motor drift questions (5A or 5B); and (3) ability to follow simple commands in English. Patients with a history of developmental, neurologic, or major psychiatric disorders and those with visual or auditory disorders limiting their ability to participate in testing procedures were excluded.

During the acute stroke hospitalization, demographic and stroke characteristics, including age, sex, handedness,²⁴ affected UE, stroke risk factors, status with respect to tissue plasminogen activator or endovascular therapy treatment, and infarct location, were documented. The following battery of outcomes was performed: (1) Fugl Meyer Assessment of Upper Extremity (FMA-UE), (2) dynamometer-assessed grip strength (Grip), (3) Box and Blocks Test (BBT), (4) Nine Hole Peg (9-HP), (5) Leg Motor (Leg), (6) gait velocity (GV), (7) Timed Up and Go (TUG), (8) NIHSS, and (9) Barthel Index (BI). These 9 impairment and function-based outcomes were selected because of their common use in poststroke clinical care and research trials. The battery was repeated at 90 days after stroke with the addition of the Stroke Impact Scale-16 (SIS) and the mRS. For 9-HP and TUG, participants who were unable to complete the assessments as a direct consequence of their level of impairment were assigned the poorest possible score. Leg was characterized by the NIHSS Leg Motor Items 6a and 6b. Table 1 provides an overview of each outcome measure. Outcome measures were administered by study staff who underwent standardized training on the study protocol and assessment standard operating procedures. Testing during acute hospitalization was completed by licensed occupational therapists who were not involved in the clinical care of the patients they were assessing. Follow-up testing in the outpatient setting was performed by trained research assistants. While raters typically administered multiple outcome assessments during 1 time period, they followed strict standard operating procedures to avoid results on 1 assessment influencing scores on other assessments. The raters in the outpatient setting were blinded to scores from the prior time point until after they completed the full assessment battery. Strong interrater reliability has previously been established for raters in this study.²⁵

Table 1 Battery of Impairment and Functional Outcome Measures

Outcome measure	Construct assessed
Fugl-Meyer Assessment of Upper Extremity	Upper extremity motor impairment
Grip strength	Isometric strength of hand and forearm
Box and Blocks Test	Gross upper extremity manual dexterity
Nine Hole Peg	Finger dexterity
Leg Motor (NIH Stroke Scale Items 6a and 6b)	Lower extremity motor impairment
Gait velocity	Walking speed
NIH Stroke Scale	Global stroke severity
Timed Up and Go	Functional mobility
Barthel Index	Global activities of daily living
Stroke Impact Scale-16	Global activities of daily living

Standard Protocol Approvals, Registrations, and Patient Consents

All participants in the study provided written informed consent. The Institutional Review Board at Mass General Brigham approved the study.

Statistical Analyses

Baseline demographic and stroke characteristics were examined descriptively with mean and SD, frequency and percentage, or median and interquartile range to characterize the cohort. Only participants who had complete data (i.e., all outcome measures in battery captured) at 90 days were included in the analyses.

To determine the univariate relationships between the mRS and UE and lower extremity functional and impairment outcomes, we performed Bonferroni-corrected Spearman ρ correlations pairwise among all variables. Given the high correlations found among all variables (eFigure 1, links.lww.com/WNL/B871), we performed a principal components analysis across outcome variables to generate an overall outcome score (first principal component [PCA-1]).

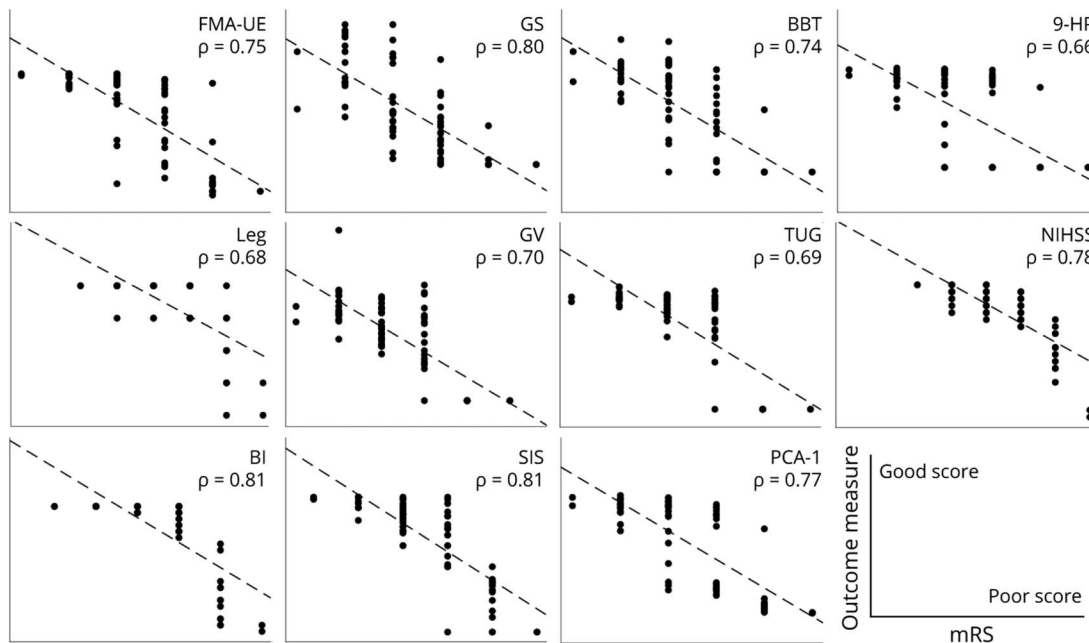
To determine whether individual levels of the mRS capture distinct profiles of function and impairment at 90 days after stroke, we performed separate Kruskal-Wallis tests with mRS levels 1, 2, 3, and 4 as the independent variable and each outcome measure (plus PCA-1) as the dependent variable. We hypothesized that adjacent mRS levels would not differ on impairment-based outcomes; therefore, we performed Bonferroni post hoc comparisons between levels to determine which mRS levels differed from each other. We excluded levels 0 ($n = 2$) and 5 ($n = 2$) from these analyses given the small number of participants in these levels. The α value was set at $p < 0.005$ to correct for multiple comparisons.

Table 2 Participant Demographics and Characteristics (n = 73)

Age, y	61.90 (12.91)
Sex, n (%)	
Male	38 (52.1)
Race/ethnicity, n (%)	
White	64 (87.7)
Black	7 (9.6)
Hispanic	2 (2.7)
Stroke hemisphere, n (%)	
Left	30 (41.1)
Infarct, n (%)	
Right middle cerebral artery	22 (30.1)
Left middle cerebral artery	35 (47.9)
Brainstem	8 (11.0)
Multiple/other	7 (9.6)
Dominant upper extremity, n (%)	
Right	64 (87.7)
Dominant upper extremity affected	27 (36.9)
Acute stroke therapy, n (%)	
tPA administered	18 (24.7)
Endovascular intervention	16 (21.9)
Premorbid mRS category	
No disability (mRS score 0)	53 (72.6)
No significant disability (mRS score 1)	10 (13.7)
Slight Disability (mRS score 2)	10 (13.7)
Baseline Fugl-Meyer Assessment of Upper Extremity score	30.00 (4.50–54.00)
Baseline grip strength score	3.70 (0–17.35)
Baseline Box and Blocks Test score	0.00 (0.00–24.50)
Baseline Nine Hole Peg score	180.00 (60.70–180.00)
Baseline gait velocity score	0.00 (0.00–0.45)
Baseline Timed Up and Go score	45.00 (27.20–45.00)
Baseline Barthel Index score	45.00 (25.00–63.75)
Baseline NIH Stroke Scale score	7.00 (4.00–10.00)
Days from stroke to baseline testing	3.00 (2.00–4.00)
Days from stroke to 90-d testing	89.00 (86.00–94.50)

Abbreviations: mRS = modified Rankin Scale; tPA = tissue plasminogen activator. Values are mean (SD), median (interquartile range), or number (percent).

Figure 1 Relationship Between the mRS and Impairment- and Function-Based Outcomes 90 Days After Stroke



Each panel shows pairwise scatterplots and the Spearman ρ correlation between the modified Rankin Scale (mRS) and each impairment and function outcome measure: Fugl-Meyer Assessment of Upper Extremity (FMA-UE), grip strength (Grip), Box and Blocks Test (BBT), Nine Hole Peg (9-HP), Leg Motor (Leg), gait velocity (GV), Timed Up and Go (TUG), NIH Stroke Scale (NIHSS), Barthel Index (BI), Stroke Impact Scale-16 (SIS), and the first principal component of all variables (PCA-1). Best possible score appears on left side of each panel; worst possible score appears on the right side for all outcome measures. All pairwise relationships were significant at $p < 0.001$. mRS score is strongly correlated to all impairment and function outcomes at the group level.

As secondary analyses, we examined the relationships that 90-day mRS score had with changes from baseline to 90 days after stroke in UE impairment, mobility, and independence with activities of daily living (ADL). We first determined the number of participants who achieved the minimal clinically important difference (MCID) on the FMA-UE (MCID = 10 points),²⁶ the BI (MCID = 10 points),²⁷ and GV (MCID = 0.16 m/s).²⁸ Next, we examined the relationships between clinically relevant cutoffs (FMA-UE²⁹: ≤ 28 severe-moderate, ≤ 42 moderate, > 42 mild; BI³⁰: ≤ 20 severe, ≤ 60 moderate, > 60 mild; GV³¹: < 0.4 severe, 0.4–0.8 moderate, > 0.8 mild) of these outcome measures, initially and at 90 days, and the mRS at 90 days. The Cohen κ was used to determine whether there was agreement between good outcome on the mRS (mRS level ≤ 2) and clinically meaningful changes (MCID change and transitions in clinical severity) on the FMA-UE, BI, and GV.¹⁰

Data Availability

Anonymized data not published within this article will be made available by request from any qualified investigator.

Results

One hundred nineteen participants provided consent for this study. Seventy-three participants completed the 90-day research follow-up and were included in the final analysis. Testing during the acute stroke hospitalization occurred a median of 3.00 (interquartile range 2.00–4.00) days after

stroke, and 90-day follow-up occurred a median of 89.00 (interquartile range 86.00–94.50) days after stroke. Patients were not available for follow-up because of the following reasons: deceased (7), withdrew from the study (9), lost to follow-up (18), unable to tolerate testing (10), or time constraints of study visit (2). The patients not included in the analysis did not differ in acute NIHSS score ($p = 0.30$), age ($p = 0.64$), race ($p = 0.65$), or sex ($p = 0.44$) from those included in the analysis. Patients who experienced UE impairment on their dominant side did not differ from those who had UE impairment on their nondominant side on any of the outcome measures at baseline. At 90 days, these 2 groups did not differ on the FMA-UE (0.149), grip strength (0.068), and BBT (0.056), but patients with dominant hand affected performed better on the NIHSS (0.045), Barthel (0.031), SIS (0.041), GV (0.007), TUG (0.017), Leg Motor (0.012), and 9-HP (0.027). Baseline demographic and clinical characteristics of the analyzed cohort are summarized in Table 2.

The mRS was highly correlated with all impairment and function outcomes (all $p < 0.001$) at 90 days after stroke across both upper and lower extremities. Figure 1 shows the strength of these associations. The PCA-1 of the battery of impairment and function outcomes accounted for 86% of variance and was highly correlated with the 90-day mRS score (Spearman $\rho = 0.77$, $p < 0.001$). Taken together, the 90-day mRS is closely associated with outcomes that represent impairment and function in both the upper and the lower extremity.

Table 3 Impairment and Activity Scores Within Individual Levels of the mRS at 90 Days After Stroke

	mRS score at 90 d after stroke					
	0 (n = 2)	1 (n = 16)	2 (n = 21)	3 (n = 20)	4 (n = 12)	5 (n = 2)
FMA-UE	65.50	64.00 (60.00–66.00)	61.00 (52.50–64.50)	46.50 (27.75 to 59.50)	5.50 (4.00 to 10.50)	4.00
Grip	28.85	32.35 (27.78–44.15)	17.70 (12.50–35.70)	10.85 (4.18 to 16.23)	0.00 (0.00 to 0.00)	0.00
BBT, # blocks	52.50	48.00 (40.50–52.75)	45.00 (29.00–53.50)	23.50 (23.50 to 37.00)	0.00 (0.00 to 0.00)	0.00
9-HP, s	27.91	33.55 (26.42–42.20)	30.72 (27.07–86.30)	180.00 (37.38 to 180.00)	180.00 (180.00 to 180.00)	180.00
Leg	0.00	0.00 (0.00–0.00)	0.00 (0.00–0.00)	0.00 (0.00 to 1.00)	2.00 (1.00 to 3.00)	3.50
GV, m/s	1.06	1.13 (1.04–1.32)	0.89 (0.74–1.10)	0.88 (0.48 to 1.04)	0.00 (0.00 to 0.00)	0.00
TUG, s	9.55	9.03 (7.66–9.86)	11.56 (9.26–13.94)	11.44 (9.60 to 19.59)	45.00 (45.00 to 45.00)	45.00
NIHSS	0.00	1.00 (1.00–1.75)	2.00 (1.00–3.50)	3.50 (2.00 to 5.00)	9.00 (7.00 to 11.75)	18.50
BI	100.00	100.00 (100.00–100.00)	100.00 (97.50–100.00)	95.00 (81.25 to 100.00)	25.00 (10.00 to 37.75)	2.50
SIS	79.50	80.00 (75.50–80.00)	71.00 (64.00–75.50)	65.00 (53.25 to 71.00)	35.50 (23.75 to 39.75)	16.00
PCA-1	76.09	69.48 (58.31–80.19)	67.13 (10.17–75.82)	–74.79 (–85.74 to 52.83)	–119.75 (–126.81 to –111.87)	–129.83

Abbreviations: BBT = Box and Blocks Test; BI = Barthel Index; FMA-UE = Fugl Meyer Assessment of Upper Extremity; GV = gait velocity; mRS = modified Rankin Scale; NIHSS = NIH Stroke Scale; 9-HP = Nine Hole Peg; PCA-1 = first principal component; SIS = Stroke Impact Scale-16; TUG = Timed Up and Go. Values are median (interquartile range).

Within each mRS level, however, median and interquartile ranges of individual impairment and function revealed substantial variation (Table 3). Figure 2 shows the distribution of scores on outcome measures within each mRS level. For example, for the clinically critical mRS level 2, there was a wide range of neurologic phenotypes, from near hemiplegia (FMA-UE score 8) to no UE motor impairment (FMA-UE score 66).

We examined the ability of mRS levels to discriminate differences in impairment and function at 90 days after stroke. While there were expected significant differences on the Kruskal-Wallis tests for each outcome measure across the full spectrum of the mRS ($p < 0.001$ for all), post hoc comparisons revealed that adjacent mRS levels 1 and 2 did not differ on any of the outcome measures; mRS levels 2 and 3 differed only on BBT, 9-HP, and SIS; and mRS levels 3 and 4 differed on the 9-HP, Leg, GV, TUG, NIHSS, BI, and PCA-1. links.lww.com/WNL/B871, shows the full results of the post hoc statistics comparing mRS levels pairwise for each outcome measure. Strikingly, there were no differences in UE impairment–focused measures (i.e., FMA-UE and grip) between any adjacent levels of the mRS. Taken together, adjacent levels of the mRS do not sufficiently capture distinct UE and lower extremity impairment and function phenotypes at 90 days after stroke.

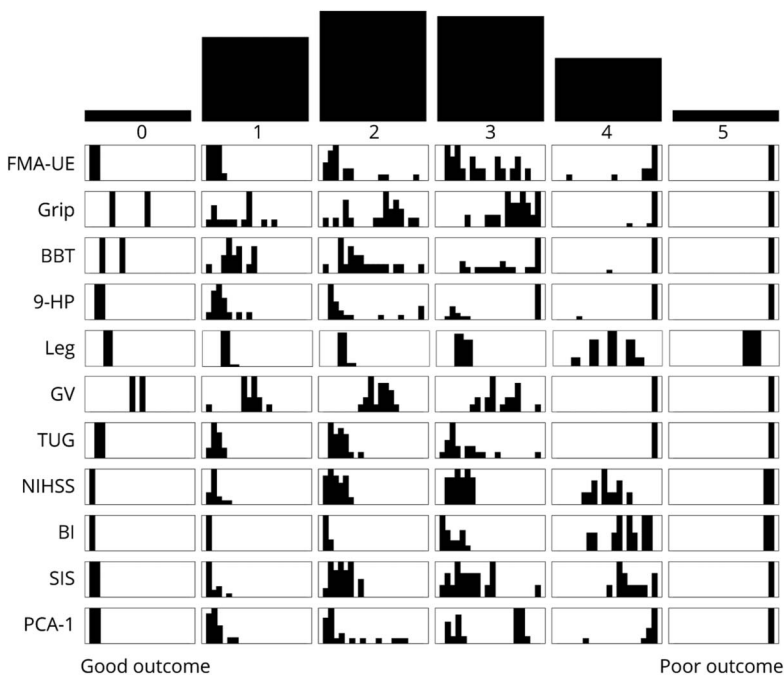
Last, we examined whether 90-day mRS levels reliably reflect clinically important changes in UE impairment, ADLs, and mobility during the period of acute to subacute stroke recovery. Figure 3 shows relationships between clinically relevant changes in each measure and the 90-day mRS score. Good outcome on the mRS is considered an mRS score ≤ 2 .

Of the patients who achieved at least MCID change on FMA-UE, BI, and GV (Figure 3, all markers left of gray diagonal bars), 48%, 46%, and 41% did not achieve a good outcome on the mRS, respectively (Figure 3). There was poor agreement between achieving mRS good outcome and MCID change on FMA-UE ($\kappa = -0.031$), BI ($\kappa = 0.016$), and GV ($\kappa = 0.14$). Furthermore, only 4 of the 13 patients who started with initially severe arm motor impairment and improved by 90 days to mild or moderate arm motor impairment had good outcome on the mRS. For BI, only 1 of the 9 patients who started with initial severe disability and improved to no longer be fully dependent at 90 days achieved good outcome on the mRS. For GV, 17 of 30 patients who started with severe walking disability and improved to moderate or full walking ability achieved good outcome on the mRS. There was poor agreement between achieving good mRS outcome and transitions from severe to nonsevere clinical categories on the FMA-UE ($\kappa = -0.15$), BI ($\kappa = -0.19$), and GV ($\kappa = 0.05$). The relationship between meaningful change in leg from baseline to 90 days with 90-day mRS score was similarly poor (eFigure 2, links.lww.com/WNL/B871). Taken together, 90-day mRS score does not sufficiently capture meaningful clinical changes across impairment, ADL status, and mobility from baseline to 90 days after stroke.

Discussion

In a cohort of stroke survivors with UE weakness, we examined relationships between the 90-day mRS score and a battery of impairment and functional outcomes after stroke. Our results indicate that while the mRS is broadly associated with

Figure 2 Distribution of Impairment and Function Outcome Scores Within Each Level of the mRS



Top histogram represents the number of participants within each level of the modified Rankin Scale (mRS) (mRS 0 n = 2, mRS 1 n = 16, mRS 2 n = 21, mRS 3 n = 20, mRS 4 n = 12, and mRS 5 n = 2). Each of the following rows shows the distribution of impairment and outcome scores within each level of the mRS for Fugl-Meyer Assessment of Upper Extremity (FMA-UE) (x-axis range 0–66), grip strength (Grip) (range 0–50), Box and Blocks Test (BBT) (range 0–70), Nine Hole Peg (9-HP) (range 15–180), Leg Motor (Leg) (range 0–4), gait velocity (GV) (0–2.5), Timed Up and Go (TUG) (range 5–45), NIH Stroke Scale (NIHSS) (range 0–25), Barthel Index (BI) (range 0–100), Stroke Impact Scale-16 (SIS) (range 16–80), and the first principal component of all variables (PCA-1) (range 0–1). Best possible score appears on the left side of each panel; worst possible score appears on the right side for all outcome measures. There was a wide range of neurologic phenotypes within mRS levels 2 to 4.

impairment and function, levels of the mRS do not discriminate detailed UE or lower extremity impairment and function phenotypes. Poor and good outcomes on the 90-day mRS are not consistently associated with meaningful clinical improvements in UE impairment, mobility, and ADL independence.

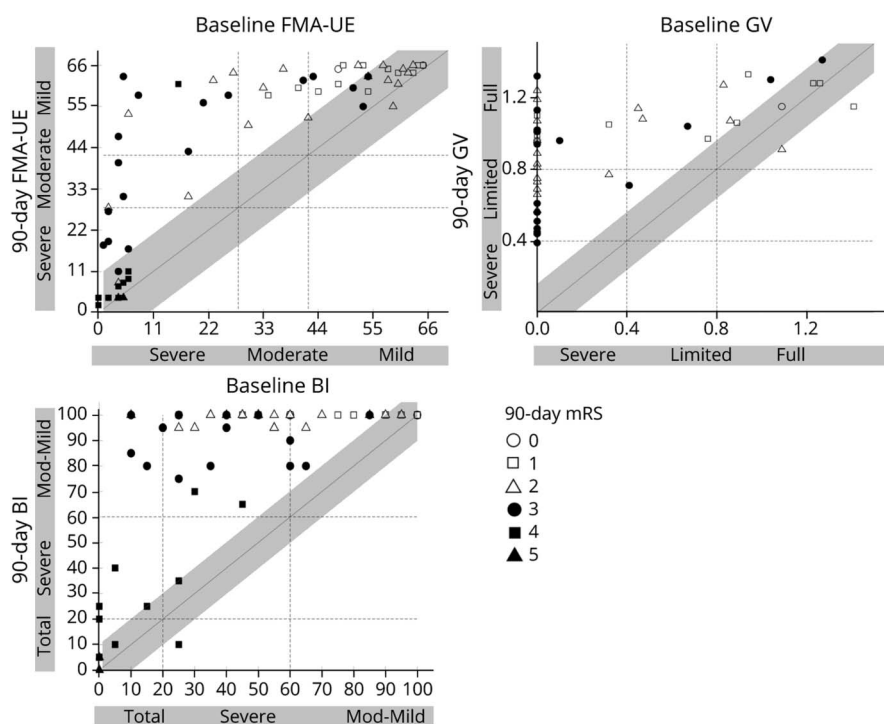
Our findings provide insights into what the mRS captures well and what it does not. At the group level, the 90-day mRS score is associated with cross-sectional measurements of both UE and lower extremity motor impairment and function. Furthermore, the 90-day mRS score was closely related to the PCA-1 of the 90-day outcome measure battery, which accounted for the majority of variance across outcome measures. Thus, via a relatively straightforward protocol and short administration period, the mRS categorizes patients into a global disability level, which relates to an array of both domain-specific UE and lower extremity impairment and global measures of function. Our findings thus support the use of the mRS as a global outcome measure in clinical studies testing treatments with large effect sizes translating to broad improvements in global disability. The mRS is likely to be particularly useful in large, multisite stroke clinical trials in which standardization and ease of outcome administration is key.³²

However, our findings also reveal a number of key limitations with relying solely on the 90-day mRS score in stroke trials. First, adjacent mRS levels could not consistently discriminate impairment- and function-based outcomes at 90 days after stroke. While mRS levels 3 and 4 separated the most outcome

measures (9-HP, GV, TUG, NIHSS, and BI), mRS levels 1 and 2 did not separate any outcome measures. Failing to detect a statistically significant difference between mRS levels may be a function of our study sample size; however, the usefulness of the measurement tool for individual-level decisions should not depend on the sample characteristics (i.e., sample size or variance in the sample). Thus, the mRS may be useful for stroke trials in which the experimental intervention is hypothesized to produce large changes in global function (i.e., tissue plasminogen activator or endovascular therapy) but likely will have significant limitations for other types of trials targeting specific domains.⁶ For example, a trial targeting arm impairment would need to have substantial effect sizes or a large sample size in order to be reflected in 90-day mRS, yet our sample size (n = 73) is representative of phase I/II stroke recovery studies.^{33,34}

Second, scores on mRS are not specific to impairment or function at 90 days after stroke; similar impairment and functional outcomes were seen across many different mRS levels. For example, patients with severe arm impairment (FMA-UE score ≤ 28) were represented across 4 mRS categories (mRS 2–5). For the FMA-UE, this implies that the mRS is not able to distinguish true restitution from compensation: a patient may have no improvement in UE impairment but learn to compensate to achieve a better mRS level. This was true for other non-impairment-based outcomes as well. For example, for the SIS, a self-report quality of life scale, patients reporting low scores in all categories (strength, function, mobility, and ADL; SIS score <40) were

Figure 3 Relationships Between mRS Outcome at 90 Days and Clinically Relevant Changes in Impairment and Function From Baseline to 90 Days



Top panel shows change in Fugl-Meyer Assessment of Upper Extremity (FMA-UE) defined by minimal clinically important difference (MCID) = 10 (gray-shaded diagonal) and severity category shift (FMA-UE: ≤ 28 severe-moderate, ≤ 42 moderate, > 42 mild upper extremity motor impairment). Middle panel shows change in gait velocity (GV) defined by MCID = 0.1 (gray-shaded diagonal) and severity category shift (GV: ≤ 0.46 severe, ≤ 0.93 limited, > 0.93 full ambulation). Bottom panel shows change in Barthel Index (BI) defined by MCID = 10 (gray-shaded diagonal) and severity category shift (BI: ≤ 20 severe, ≤ 60 moderate, > 60 mild disability). Shapes represent mRS levels at 90 days as indicated in legend. Filled markers represent poor outcome (mRS score ≥ 3); open markers represent good outcome (mRS score ≤ 2). Of the patients who achieved at least MCID change on FMA-UE, BI, and GV (all markers left of gray diagonal bars), 48%, 46%, and 41% did not achieve a good outcome on the mRS, respectively. There was poor agreement between achieving mRS good outcome and meaningful change in arm impairment, gait speed, and activities of daily living.

seen across 3 mRS categories (mRS levels 3–5). Moreover, the mRS was not precise in categorizing motor impairment, function, or quality of life outcomes in our data. Prior research revealed similar limitations of the mRS related to categorizing cognition and depression.³⁵ Together, these findings indicate that the patients within the same mRS level are markedly heterogeneous; such differences may be more important to some stroke trials than others. Future work should explore whether incorporating additional endpoints along with the mRS would make stroke outcome categorization more precise.

Third, the 90-day mRS score incompletely accounts for change in impairment and function from poststroke baseline. A substantial number of patients in our cohort achieved clinically meaningful gains in UE impairment, mobility, and ADL status in the first 90 days after stroke but were not categorized as good outcome by the 90-day mRS. Trialists should be aware that 90-day mRS score, common in acute stroke trials, and clinically meaningful gains from acute stage to 90 days, common in stroke recovery trials, are different therapeutic targets. Use of a cross-sectional mRS score at 90 days is insensitive to the amount of recovery a patient made to reach the score, but such differences may be of central importance to some types of stroke trials. This concept is akin to the importance of controlling for baseline in regression models. Treatments that target neural repair processes after stroke may be more likely to succeed if trial designs

incorporate outcomes that can be serially measured starting immediately after stroke and calculate clinically meaningful gains rather than measuring 90-day mRS only.³⁶

There are several limitations to this study. The sample of patients included only those for whom we could capture our entire battery of outcome measures at poststroke baseline (during the acute stroke hospitalization) and at poststroke day 90. The sample size also may limit the detection of the full spectrum of possible phenotypes across mRS levels. A larger sample size may also permit a more detailed understanding of how endpoints change over time. However, the sample size in this study is representative of phase I/II stroke recovery studies. The patients were also mostly White and younger than the average age of patients with stroke in the United States, reflecting the patient population at the single-site academic medical center at which the study was conducted. Therefore, there may be limitations in generalizing findings to the entire stroke population.³⁷ Furthermore, our outcome measure battery was weighted to UE and lower extremity motor impairment and function; it did not include detailed measurements of poststroke language, cognition, mood, or social factors. Future studies that assess relationships between mRS levels and nonmotor factors are needed. A notable finding in this study was that, while patients with dominant-hand (and thus hemisphere) injury did not differ on baseline assessments, they, in general, had better outcomes at 90 days. Given that nearly all patients in this study were right-hand

dominant, this suggests the need to examine hemisphere-specific effects on functional recovery.

To advance novel treatments for recovery after stroke, we should rethink and expand how we measure and implement outcome measures after stroke.^{6,12,16,35,36} Further attention to the development of sensitive measures that reflect biological processes should be prioritized.³⁸ The mRS is useful for measuring a patient's global disability, but there is a need to further understand what components of impairment improve function and result disability.³⁹ In addition to the traditional 90-day mRS, we recommend establishing a comprehensive minimum dataset that includes high-quality, responsive measures that capture various types of modality-specific impairment and function-based outcomes, measured serially from the acute stroke period. It will be essential that we develop disability measures that have greater levels of granularity and that capture improvements across different modalities: motor, language, cognitive.⁶ For feasibility in clinical settings, item-response theory and computer adaptive testing could be used to optimize data collection. Progress in stroke recovery requires the optimal selection and deployment of outcome measures that reflect neurologic and functional status and are meaningful to patients and caregivers.

Our findings suggest that the mRS, the current gold standard stroke clinical trial endpoint, has excellent correlation with domain-specific impairment and global functional outcomes but may not be precise enough to capture distinct phenotypes of impairment and function at 90 days after stroke. The mRS is a globally recognizable tool that is easy to implement and interpret that allows clinicians and researchers to measure the impact of stroke on disability. While there are many strengths of this tool, similar impairment and function-outcomes were represented across many different mRS levels, and the mRS at 90 days incompletely accounted for meaningful changes in UE impairment, ADL independence, and mobility from poststroke baseline. Taken together, these findings emphasize the urgent need to improve measurement in stroke clinical care and research.

Study Funding

This research was supported by the MGH Institute of Health Professions (Faculty Development Grant, Dr. Erler), the Andrew David Heitman Foundation (Young Investigator Award, Dr. Lin), and the American Society of Neuro-rehabilitation (Seed Funding Collaborative Clinical Research Project, Dr. Lin). Dr Kautz was funded by VA Rehabilitation R&D 1IK6RX003075 and NIH P20 GM109040. Dr. Finklestein was supported in part by NIH 2R44NS095381-02.

Disclosure

K.S. Erler, R. Wu, J.A. DiCarlo, M.F. Petrilli, and P. Gochyyev report no conflicts relevant to this work. L.R. Hochberg reports that the MGH Translational Research Center has a clinical research support agreement with Neuralink, Paradromics, and

Synchron for which Dr. Hochberg provides consultative input. S.A. Kautz reports no conflicts relevant to this work. L.H. Schwamm: reports the following relationships relevant to research grants or companies that manufacture products for stroke treatment: scientific consultant regarding trial design and conduct for Genentech (late-window thrombolysis) and member of Steering Committee (TIMELESS [Tenecteplase in Stroke Patients Between 4.5 and 24 Hours], NCT03785678); consultant on user interface design and usability for LifeImage; member of a Data Safety Monitoring Boards for Penumbra (Artemis in the Removal of Intracerebral Hemorrhage [MIND] NCT03342664) and for Diffusion Pharma (Efficacy and Safety of Trans Sodium Crocetin [TSC] for Treatment of Suspected Stroke [PHAST-TSC], NCT03763929); serving as national principal investigator (PI) for Medtronic (Rate of Atrial Fibrillation Through 12 Months in Patients With Recent Ischemic Stroke of Presumed Known Origin [Stroke AF], NCT02700945); national co-PI, late-window thrombolysis trial, National Institute of Neurological Disorders and Stroke (P50NS051343, A Study of Intravenous Thrombolysis With Alteplase in MRI-Selected Patients [MR WITNESS], NCT01282242); and alteplase provided free of charge to Massachusetts General Hospital, as well as supplemental per-patient payments to participating sites by Genentech. S.C. Cramer serves as a consultant for Abbvie, Constant Therapeutics, MicroTransponder, Neuroolutions, SanBio, Fujifilm Toyama Chemical Co, NeuExcell, Elevian, Medtronic, and TRCare. S.P. Finklestein is a consultant to Constant Therapeutics, AZTherapies, Elevian, Alkermes, Eleusis, and Cycle-rion. He is a principal in Stemetix and Recovery Therapeutics. D.J. Lin reports that the MGH Translational Research Center has a clinical research support agreement with Elevian for which Dr. Lin provides consultative input. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* July 25, 2021. Accepted in final form January 18, 2022.

Appendix Authors

Name	Location	Contribution
Kimberly S. Erler, PhD	Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Rui Wu, BA	Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Julie A. DiCarlo, MS	Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design

Appendix (continued)

Name	Location	Contribution
Marina F. Petrilli, BA	School of Health and Rehabilitation Sciences, MGH Institute of Health Professions, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data
Perman Gochyyev, PhD	School of Health and Rehabilitation Sciences, MGH Institute of Health Professions, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Leigh R. Hochberg, MD, PhD	Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Steven A. Kautz, PhD	Department of Health Sciences and Research, Medical University of South Carolina, Charleston	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Lee H. Schwamm, MD	Stroke Service, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Steven C. Cramer, MD	Department of Neurology University of California, Los Angeles	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
Seth P. Finklestein, MD	Stroke Service, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design
David J. Lin, MD	Center for Neurotechnology and Neurorecovery, Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

References

- Hachinski V, Donnan GA, Gorelick PB, et al. Stroke: working toward a prioritized world agenda. *Cerebrovasc Dis*. 2010;30(2):127-147.
- Kwakkel G, van Wegen EEH, Burridge JH, et al. Standardized measurement of quality of upper limb movement after stroke: consensus-based core recommendations from the Second Stroke Recovery and Rehabilitation Roundtable. *Neurorehabil Neural Repair*. 2019;33(11):951-958.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008;337:a1655.
- Quinn TJ, Singh S, Lees KR, Bath PM, Myint PK. Validating and comparing stroke prognosis scales. *Neurology*. 2017;89(10):997-1002.
- Cramer SC, Koroshetz Walter J, Finklestein Seth P. The case for modality-specific outcome measures in clinical trials of stroke recovery-promoting agents. *Stroke*. 2007;38(4):1393-1395.
- Braun RG, Heitsch L, Cole JW, et al. What the modified Rankin isn't ranking: domain-specific outcomes for stroke clinical trials. *Neurology*. 2021;97(8):367-377.
- van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJA, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19(5):604-607.
- Banks JL, Marotta CA. Outcomes validity and reliability of the modified Rankin Scale: implications for stroke clinical trials. *Stroke*. 2007;38(3):1091-1096.
- Bath PMW, Lees KR, Schellinger PD, et al. Statistical analysis of the primary outcome in acute stroke trials. *Stroke*. 2012;43(4):1171-1178.
- Broderick JP, Adeoye O, Elm J. Evolution of the modified Rankin Scale and its use in future stroke trials. *Stroke*. 2017;48(7):2007-2012.
- Quinn TJ, Dawson J, Walters MR, Lees KR. Functional outcome measures in contemporary stroke trials. *Int J Stroke*. 2009;4(3):200-205.
- Saver JL. Optimal end points for acute stroke therapy trials: best ways to measure treatment effects of drugs and devices. *Stroke*. 2011;42(8):2356-2362.
- Quinn TJ, Dawson J, Walters MR, Lees KR. Exploring the reliability of the modified Rankin Scale. *Stroke*. 2009;40(3):762-766.
- Quinn TJ, Dawson J, Walters MR, Lees KR. Reliability of the modified Rankin Scale: a systematic review. *Stroke*. 2009;40(10):3393-3395.
- Ward NS, Carmichael ST. Blowing up neural repair for stroke recovery. *Stroke*. 2020;54(10):3169-3173.
- Cramer SC, Wolf S, Saver JL, et al. The utility of domain-specific endpoints in acute stroke trials. *Stroke*. 2021;52(3):1154-1161.
- Hong KS, Saver JL. Quantifying the value of stroke disability outcomes: WHO Global Burden of Disease Project disability weights for each level of the modified Rankin Scale. *Stroke J Cereb Circ*. 2009;40(12):3828-3833.
- Dijkland Simone A, Voormolen Daphne C, Venema E, et al. Utility-weighted modified Rankin Scale as primary outcome in stroke trials. *Stroke*. 2018;49(4):965-971.
- Zhao H, Collier JM, Quah DM, Purvis T, Bernhardt J. The modified Rankin Scale in acute stroke has good inter-rater-reliability but questionable validity. *Cerebrovasc Dis*. 2010;29(2):188-193.
- Bugnicourt JM, Godefroy O. Day-7 or day-90 modified Rankin Scale score: what is the best measure of outcome after thrombolysis in ischemic stroke? *J Thromb Thrombolysis*. 2013;36(3):316.
- Winstein CJ, Stein J, Arena R, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*. 2016;47(6):e98-e169.
- Stewart JC, Cramer SC. Patient-reported measures provide unique insights into motor function after stroke. *Stroke*. 2013;44(4):1111-1116.
- Wyller TB, Sveen U, Södring KM, Pettersen AM, Bautz-Holter E. Subjective well-being one year after stroke. *Clin Rehabil*. 1997;11(2):139-145.
- Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*. 1971;9(1):97-113.
- Ranford J, Asiello J, Cloutier A, et al. Interdisciplinary stroke recovery research: the perspective of occupational therapists in acute care. *Front Neurol*. 2019;10:1327.
- Shelton FD, Volpe BT, Reding M. Motor impairment as a predictor of functional recovery and guide to rehabilitation treatment after stroke. *Neurorehabil Neural Repair*. 2001;15(3):229-237.
- Hsieh YW, Wang CH, Wu SC, Chen PC, Sheu CF, Hsieh CL. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. *Neurorehabil Neural Repair*. 2007;21(3):233-238.
- Tilson JK, Sullivan KJ, Cen SY, et al. Meaningful gait speed improvement during the first 60 days poststroke: minimal clinically important difference. *Phys Ther*. 2010;90(2):196-208.
- Woytowicz EJ, Rietschel JC, Goodman RN, et al. Determining levels of upper extremity movement impairment by applying cluster analysis to upper extremity Fugl-Meyer assessment in chronic stroke. *Arch Phys Med Rehabil*. 2017;98(3):456-462.
- Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol*. 1989;42(8):703-709.
- Bowden MG, Balasubramanian CK, Behrman AL, Kautz SA. Validation of a speed-based classification system using quantitative measures of walking performance poststroke. *Neurorehabil Neural Repair*. 2008;22(6):672-675.
- Lees KR, Bath PMW, Schellinger PD, et al. Contemporary outcome measures in acute stroke research. *Stroke*. 2012;43(4):1163-1170.
- Scheidtmann K, Fries W, Müller F, Koenig E. Effect of levodopa in combination with physiotherapy on functional motor recovery after stroke: a prospective, randomised, double-blind study. *Lancet*. 2001;358(9284):787-790.
- Chollet F, Tardy J, Albucher JF, et al. Fluoxetine for motor recovery after acute ischaemic stroke (FLAME): a randomised placebo-controlled trial. *Lancet Neurol*. 2011;10(2):123-130.
- Polding LC, Tate WJ, Mlynash M, et al. Quality of life in physical, social, and cognitive domains improves with endovascular therapy in the DEFUSE 3 trial. *Stroke*. 2021;52(4):1185-1191.
- Hermann DM, Bassetti CL, Marx U, Audoli-Inthavong ML, Chabriat H. Refining endpoints for stroke recovery trials. *Lancet Neurol*. 2020;19(5):381-382.
- Virani SS, Alonso A, Aparicio HJ, et al. Heart disease and stroke statistics—2021 update: a report from the American Heart Association. *Circulation*. 2021;143(8):e254-e743.
- Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the Stroke Recovery and Rehabilitation Roundtable taskforce. *Int J Stroke*. 2017;12(5):444-450.
- Cramer SC, Le V, Saver JL, et al. Intense arm rehabilitation therapy improves the modified Rankin Scale score: association between gains in impairment and function. *Neurology*. 2021;96(14):e1812-e1822.