Adopting a Patient-Centered Approach to Primary Outcome Analysis of Acute Stroke Trials Using a Utility-Weighted Modified Rankin Scale

Napasri Chaisinanunkul, MD; Opeolu Adeoye, MD, MS; Roger J. Lewis, MD, PhD; James C. Grotta, MD; Joseph Broderick, MD; Tudor G. Jovin, MD; Raul G. Nogueira, MD; Jordan J. Elm, PhD; Todd Graves, PhD; Scott Berry, PhD; Kennedy R. Lees, MD; Andrew D. Barreto, MD, MS; Jeffrey L. Saver, MD; for the DAWN Trial and MOST Trial Steering Committees*

**Background and Purpose**—Although the modified Rankin Scale (mRS) is the most commonly used primary end point in acute stroke trials, its power is limited when analyzed in dichotomized fashion and its indication of effect size challenging to interpret when analyzed ordinally. Weighting the 7 Rankin levels by utilities may improve scale interpretability while preserving statistical power.

**Methods**—A utility-weighted mRS (UW-mRS) was derived by averaging values from time-tradeoff (patient centered) and person-tradeoff (clinician centered) studies. The UW-mRS, standard ordinal mRS, and dichotomized mRS were applied to 11 trials or meta-analyses of acute stroke treatments, including lytic, endovascular reperfusion, blood pressure moderation, and hemicraniectomy interventions.

**Results**—Utility values were 1.0 for mRS level 0; 0.91 for mRS level 1; 0.76 for mRS level 2; 0.65 for mRS level 3; 0.33 for mRS level 4; 0.20 for mRS level 5; and 0 for mRS level 6. For trials with unidirectional treatment effects, the UW-mRS paralleled the ordinal mRS and outperformed dichotomous mRS analyses. Both the UW-mRS and the ordinal mRS were statistically significant in 6 of 8 unidirectional effect trials, whereas dichotomous analyses were statistically significant in 2 to 4 of 8. In bidirectional effect trials, both the UW-mRS and ordinal tests captured the divergent treatment effects by showing neutral results, whereas some dichotomized analyses showed positive results. Mean utility differences in trials with statistically significant positive results ranged from 0.026 to 0.249.

**Conclusions**—A UW-mRS performs similar to the standard ordinal mRS in detecting treatment effects in actual stroke trials and ensures the quantitative outcome is a valid reflection of patient-centered benefits. (*Stroke. 2015;46:00-00.* DOI: 10.1161/STROKEAHA.114.008547.)

**Key Words:** blood pressure ■ stroke ■ stroke management ■ thrombolysis

---

The modified Rankin Scale (mRS) is the most widely used measure of outcome after acute ischemic stroke in both research clinical trials and national and local quality improvement registries. However, there is much debate on how best statistically to analyze the mRS. Approaches include simple dichotomization, sliding dichotomy or responder analysis, and ordinal or shift analysis. The power of the mRS to detect treatment effects is often reduced when the scale is analyzed in dichotomized fashion, discarding substantial outcome information. In the simple dichotomous approach, the 7 possible mRS scores are collapsed into just 2 health states, and the optimal point for dichotomization depends on timing of the intervention and the anticipated distribution of severity of illness and prognosis of enrolled subjects. As data to guide selection of the most informative dichotomization are often incomplete, suboptimal selection may occur, missing a true treatment effect. Moreover, because they discard the preponderance of outcome information, dichotomized analyses...
always provide an incomplete delineation of treatment effects and may miss contrary harmful effects occurring at nonanalyzed health state transitions.

Analytic approaches that take into account all outcomes on the mRS provide a more complete depiction of treatment effect than collapsed analyses and will have greater statistical power than dichotomized analyses when treatment benefit accrues at several health state transitions rather than clustering at just one. Ordinal analysis approaches to the full distribution of outcomes may include the proportional odds model, the Mann–Whitney test, and the Cochran–Mantel–Haenszel test. However, all fail to reflect the varied width of transitions between different levels of the mRS, creating difficulty in interpreting treatment group differences, especially as patients’ valuation of each given mRS health state has been unclear.

Diverse organizations, including the Patient-centered Outcomes Research Institute and the National Institute for Health and Care Excellence and health economists, strongly advocate the use of outcome metrics that measure benefits of a given intervention to the patient. The most widely accepted patient-centered outcome measure is utility: the desirability of a specific health outcome to the patient. A promising approach to transforming the mRS into a patient-centered outcome measure is to weight the 7 levels of the mRS by their utilities. Utility weights would convert the spacing between ranks on the mRS from arbitrarily fixed intervals to distances that directly reflect patient and societal valuation of outcome disability states. Developing a utility-weighted (UW) version of the mRS has been recommended for acute stroke research by the Stroke Therapy Academic Industry Roundtable.

We aimed to derive a UW-mRS by averaging values from prior studies using time-tradeoff methodology in stroke survivors and person-tradeoff methodology in healthcare providers. To explore the feasibility and comparative statistical efficiency of the UW-mRS, we applied it, alongside standard dichotomized and ordinal mRS analytic approaches, to 11 clinical trials or meta-analyses of acute stroke treatments.

Methods

We derived utility weights for each level of the mRS by averaging utility values derived in 2 prior studies. In 1 study, the mRS and the European Quality of Life Scale (EQ-5D) were assessed among all stroke and transient ischemic attack survivors in a population-based study in Great Britain. The mRS scores were mapped to the EQ-5D in the same patients, and thence to utilities using utility values for the 243 possible EQ-5D health states derived using time-tradeoff methods in the general British population. In the other study, disability weights for mRS levels were derived using the methodology of the World Health Organization Global Burden of Disease Project. An international panel of neurovascular and cardiovascular physicians and nurses participated in the World Health Organization Global Burden of Disease Project person-tradeoff method. For the current study, these disability weights were converted to utility weights by taking their inverse.

For the analysis of completed trials, we selected 11 trials or meta-analyses for which group results in all 7 mRS categories were reported. Studies were selected to include all trials considered in the US national guidelines as providing supportive evidence for an acute ischemic stroke intervention (9 studies) and for illustrative purposes 1 trial providing supportive evidence for an intracerebral hemorrhage intervention and a trial showing neutral results. These trials included studies with unidirectional net benefits across every level of the mRS (8 studies), trials with bidirectional beneficial and harm effects across different mRS scale transitions (2 studies), and neutral unidirectional effects (1 study). In unidirectional benefit trials, all mRS cut points show better outcomes for treatment than control. In bidirectional treatment effect trials, some mRS cut points show better outcomes for treatment but others show better outcome for control, for example, in the Thrombectomy Revascularisation of Large Vessel Occlusions in Acute Ischemic Stroke (TREVO) 2 trial thrombectomy was associated with increased functional independence (mRS, 0–2), 39.9% versus 21.8%, but also with increased mortality, 34.1% versus 24.1%. To explore the informativeness of the UW-mRS in subgroup analysis, we also performed a subgroup analysis of the trial with neutral overall effects in 2 subgroups expected to have differential responses. Group mRS outcomes in each trial were reanalyzed using 5 statistical approaches: dichotomized at excellent outcome (0–1 versus 2–6), dichotomized at good or better outcome (0–2 versus 3–6), dichotomized at fair or better outcome (0–4 versus 5–6), ordinal analysis of the mRS, and analysis of the UW-mRS. The 3 particular dichotomizations of the mRS were selected because each had been used as the primary analysis for ≥1 of the analyzed trials or meta-analyses. Dichotomized analyses used Fisher exact test, ordinal analysis used the Mann–Whitney test, and utility analysis used the t test. In addition, with the UW-mRS, the mean utility differences between control and intervention arms and 95% confidence intervals were calculated.

In order for the t test to be valid for analyzing UW-mRS, the sampling distribution of the average utility score needs to be approximately normal. There are 6 possible utility values bounded from 0 to 1.0. Although the distribution of a single score is not normally distributed, it is bounded and well behaved. With a bounded set of 6 outcomes in as few as 15 observations, the distribution of the average utility is closely normal (Central Limit Theorem). In these studies there are typically much >100 per treatment arm. We have compared the results of the t test with modeling the 7 mRS outcomes exactly, using a Dirichlet distribution, and the results are extremely close.

Results

As shown in Figure 1, the utility values for each mRS level from the 2 source studies were close. Averaging the values produced the following utility weights: mRS 0 to 1.0, mRS 1 to 0.91, mRS 2 to 0.76, mRS 3 to 0.65, mRS 4 to 0.33, mRS 5 to 0, and mRS 6 to 0.

The P values for treatment arm differences associated with each mode of mRS analysis are shown in Figure 2. Among
the 8 trials with unidirectional beneficial effects, both the UW-mRS and the ordinal mRS were conventionally positive in 6, whereas dichotomization at 0 to 1 was positive in 4, dichotomization at 0 to 2 was positive in 4, and dichotomization at 0 to 4 was positive in 2. Among the 2 trials with bidirectional effects, both the UW-mRS and the ordinal mRS were nonpositive, reflecting the mixed overall treatment effects. In contrast, 3 of the 6 applied dichotomized analyses were positive, capturing benefit at the single interrogated health state transition but not also incorporating harm occurring at other points in the disability spectrum.

Table 1 shows the mRS values and UW-mRS values for all 11 trials and meta-analyses, ordered by nominal difference in mean utilities between the treatment groups. Utilities in the experimental arm of each study nominally exceeded the utilities in control arms in all 11 treatment comparison. Among the

Table 1. Measures of Effect Size Using Ordinal and Utility-Weight Analysis of the Modified Rankin Scale

<table>
<thead>
<tr>
<th>Trial</th>
<th>Median mRS</th>
<th>Median mRS Treatment</th>
<th>Delta mRS</th>
<th>Mean mRS</th>
<th>Mean mRS Treatment</th>
<th>Delta mRS Mean Values</th>
<th>Mean Utility</th>
<th>Mean Utility Treatment</th>
<th>Delta Utility</th>
<th>Mean Values 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemicraniectomy</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>5.20</td>
<td>3.91</td>
<td>1.29</td>
<td>0.150</td>
<td>0.399</td>
<td>0.249</td>
<td>0.134 to 0.364</td>
</tr>
<tr>
<td>SWIFT</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3.94</td>
<td>3.11</td>
<td>0.83</td>
<td>0.365</td>
<td>0.547</td>
<td>0.182</td>
<td>0.044 to 0.320</td>
</tr>
<tr>
<td>MR CLEAN</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3.95</td>
<td>3.49</td>
<td>0.46</td>
<td>0.362</td>
<td>0.462</td>
<td>0.100</td>
<td>0.043 to 0.157</td>
</tr>
<tr>
<td>NINDS tPA Trials</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>3.19</td>
<td>2.66</td>
<td>0.53</td>
<td>0.501</td>
<td>0.591</td>
<td>0.089</td>
<td>0.031 to 0.149</td>
</tr>
<tr>
<td>TREVO 2</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>3.74</td>
<td>3.40</td>
<td>0.34</td>
<td>0.395</td>
<td>0.464</td>
<td>0.069</td>
<td>−0.042 to 0.180</td>
</tr>
<tr>
<td>MELT</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3.07</td>
<td>2.59</td>
<td>0.48</td>
<td>0.516</td>
<td>0.583</td>
<td>0.067</td>
<td>−0.168 to 0.065</td>
</tr>
<tr>
<td>PROACT 2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>3.60</td>
<td>3.28</td>
<td>0.32</td>
<td>0.434</td>
<td>0.481</td>
<td>0.047</td>
<td>−0.168 to 0.074</td>
</tr>
<tr>
<td>IMS 3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3.20</td>
<td>3.01</td>
<td>0.19</td>
<td>0.505</td>
<td>0.541</td>
<td>0.036</td>
<td>−0.024 to 0.097</td>
</tr>
<tr>
<td>IST 3</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>3.83</td>
<td>3.67</td>
<td>0.16</td>
<td>0.418</td>
<td>0.448</td>
<td>0.030</td>
<td>0.001 to 0.058</td>
</tr>
<tr>
<td>ECASS 3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2.20</td>
<td>1.99</td>
<td>0.21</td>
<td>0.674</td>
<td>0.700</td>
<td>0.026</td>
<td>−0.070 to 0.018</td>
</tr>
<tr>
<td>INTERACT 2</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2.93</td>
<td>2.81</td>
<td>0.12</td>
<td>0.553</td>
<td>0.579</td>
<td>0.026</td>
<td>0.001 to 0.051</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; ECASS, The European Cooperative Acute Stroke Study; IMS, Interventional Management of Stroke; INTERACT, Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage trial; IST 3, International Stroke Trial 3; MELT, Middle Cerebral Artery Embolism Local Fibrinolytic Intervention Trial; MR CLEAN, Multicenter Randomized Clinical Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands; mRS, modified Rankin Scale; NINDS tPA, National Institute of Neurological Disorders and Stroke Tissue Plasminogen Activator; PROACT, Prolyse in Acute Cerebral Thromboembolism; SWIFT, Solitaire With the Intention for Thrombectomy; and TREVO, Thrombectomy Revascularisation of Large Vessel Occlusions in Acute Ischemic Stroke.
6 trials positive on the UW-mRS, group differences in mean utility ranged from 0.024 to 0.25. The greatest differences were seen with hemicraniectomy (mean utility delta, 0.25), endovascular recanalization (Solitaire With the Intention for Thrombectomy [SWIFT] mean utility delta, 0.18; Multicenter Randomized Clinical Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands [MR CLEAN] mean utility delta, 0.10), and early intravenous thrombolysis (National Institute of Neurological Disorders and Stroke Tissue Plasminogen Activator [NINDS tPA] study mean utility delta, 0.09).

As shown in Table 2, the UW-mRS did provide a signal of potentially different responsiveness in subgroup analysis. In the overall neutral Intervventional Management of Stroke 3 trial, nominal values for group utility differences were substantially larger in patients with severe (National Institutes of Health Stroke Scale ≥20) compared with moderate deficits (National Institutes of Health Stroke Scale 8–19) at entry, although these differences did not reach statistical significance.

### Discussion

We found that a utility approach to analyses of the mRS is feasible in acute stroke trials. The UW-mRS was able to be applied to a wide range of acute stroke trials, including studies enrolling diverse stroke subtypes (ischemic and hemorrhagic), in varying treatment time windows, and testing multiple therapeutic interventions. Analysis of the UW-mRS was computationally straightforward, using t tests. The UW-mRS showed similar statistical efficiency as ordinal analysis of the mRS and superior efficiency to dichotomized analyses in detecting treatment effects for interventions generally recognized as beneficial.

It is noteworthy that the 2 studies that were the sources for the utility values in the UW-mRS found extremely similar utility values for each of the 7 levels of the mRS, although they used different derivation techniques. In 1 study, utility values were derived by quality of life ratings by patients plus time-tradeoff judgments by laypeople; in the other, utility values were derived by person-tradeoff judgments by healthcare providers. The close correspondence of the utilities emerging from patients, laypeople, and healthcare providers perspectives reinforces the credibility of the averaged values used in the UW-mRS.

When treatment benefits were demonstrated at multiple health state transitions, the UW-mRS, like the ordinal mRS, showed stronger levels of statistical significance (smaller P values) when compared with dichotomized tests. When benefits were concentrated on single health state levels or there were signs of a bidirectional effect, the levels of significance for both the UW-mRS and the ordinal mRS were weaker than for the subset of dichotomized analyses focused on highly positive health state transitions (Middle Cerebral Artery Embolism Local Fibrinolytic Intervention Trial [MELT], Prolyse in Acute Cerebral Thromboembolism [PROACT] 2). For healthcare policy planners, economists, and methodologists, the results of UW-mRS analyses offered greater immediate interpretability than the ordinal mRS, with treatment group differences expressed directly in utility value differences standardly used in healthcare population-level decision making.

A particular important feature of generic utility measurement is the ability to generate quality-adjusted life years gained or lost by an intervention or treatment. Quality-adjusted life years allow healthcare decision makers the capability to understand the relative value of ostensibly different medical interventions for different diseases to maximize society-wide health-related quality of life. In many countries, particularly those with nationalized or socialized healthcare systems, decisions to fund certain treatments are closely linked with the cost per quality-adjusted life years gained.

A special challenge to trial interpretation occurs when treatments exert bidirectional effects, conferring benefit at some health state transitions but harm at others. For example, in The European Cooperative Acute Stroke Study (ECASS) 3 and TREVO trials, intervention patients had both more good outcomes and more poor outcomes than control patients. Dichotomized analyses of these trials focused only on the positive health state transitions indicated treatment benefit, failing to capture the harm conferred elsewhere in the disability spectrum. The UW-mRS integrated both positive and negative effects in a single metric and indicated no statistically significant overall beneficial effect on utility, although nominal differences in utility were favorable.

Although the UW-mRS and ordinal mRS analysis performed similarly in this set of 11 trials, it is possible they could provide divergent results in studies with particular outcome distributions. In the UW-mRS, the distances between mRS levels are not equal for each step, but rather varies substantially, reflecting patient and provider valuation of the worth of each health state transition. For example, in the UW-mRS, the mRS state of 5 (bedridden, 24-hour care, severely disabled) is
not valued as higher than the mRS state of 6 (dead). Moving patients from an mRS 6 to mRS 5 outcome will increase the likelihood of statistical significance on an ordinal mRS analysis but not on an UW-mRS analysis. This noncontribution is appropriate for a patient-centered outcome measure, given that on average patients and caregivers do not consider an mRS outcome of 5 as better (and many actually consider it worse) than an mRS outcome of 6. The greatest utility value steps on the UW-mRS are between the mRS health states of 5 to 4 and of 4 to 3. Consequently, treatments that improve outcomes preferentially in this more severe disability range will have greater statistical significance on the UW-mRS than treatments that improve outcomes preferentially in the milder disability range. For that reason, as well as the large proportion of patients it helps, hemicraniectomy for malignant infarction had the greatest utility gain among the treatments analyzed. Although excellent outcomes are the most highly desired among patients and providers, over the long arc of health states from normal to dead, transitions from dead to alive with at least some valuable function are valued even more highly than transitions between good and excellent health states.

An additional advantage of the UW-mRS is that it may allow more patients with prestroke disability to be informative when enrolled in acute stroke trials. Some degree of prestroke disability is present in up to 50% of patients with stroke in clinical practice, and many of these patients must be excluded from clinical trials with simple dichotomous mRS end points because their preexisting disability precludes them from crossing the single health state transition being examined. A utility approach to analyzing the mRS can allow inclusion of additional patients with stroke in trials despite varying degrees of baseline disability. However, it remains important to ensure balance in baseline mRS between treatment groups.

This study has limitations. Some of the available mRS distributions from trials were unadjusted and some were adjusted for major baseline prognostic factors. Cross-trial comparisons would be most fair using adjusted group comparisons from all trials. The utility values for the UW-mRS were derived by mapping patient quality of life ratings to utilities derived in a population-based sample from Great Britain and a healthcare provider sample from North America and Asia. Utility values derived directly from patients with stroke are desirable. Also, utility values derived from patients and providers in additional geographic settings may differ to a modest degree from those elicited in the current derivation studies. The minimally clinically important difference in utilities in patients with stroke has not been well studied. However, studies to define the minimally clinically important difference for utility have been performed in a wide range of other diseases and have found minimally clinically important differences generally ranging from 0.04 to 0.10.

We conclude that a utility approach to the analysis of the mRS is feasible, is statistically efficient, and provides a patient and societal-centered metric of the degree of benefit or harm of a tested intervention. Like ordinal analysis, the UW-mRS provides the advantages of comprehensiveness and of greater statistical power. Like dichotomized analysis, the UW-mRS provides the advantages of interpretability, stating the degree to which an intervention shows benefits valued by patients and clinicians, in the form of the utility values standard in health policy planning. The UW-mRS is potentially a useful outcome metric for future acute stroke trials.

Appendix

Additional contributors from DAWN Trial Steering Committee: Scott Berry, Anthony Furlan, Blaise Baxter, Helmi L. Lutsep, Marc Ribo, Olav Jansen, Rishi Gupta, Vitor Mendes Pereira.

Sources of Funding

This work was supported by the National Institute of Health National Institute of Neurological Disorders and Stroke Awards: U10 NS055892, P50 NS044263, and P50 NS044378.

Disclosures

Dr Berry is an owner of Berry Consultants, LLC, a statistical consulting company that provides clinical trial design services, including for trials of treatments for neurological conditions. Dr Broderick: Service on Executive Committee for PRISMS Trial, Genentech. Dr Grotta is a consultant, Specialists on Call. Dr Jovin received honoraria from Praxair (modest); ownership interest Silk Road Medical (modest). Dr Lees’ institution, the University of Glasgow, administers research grants relating to modified Rankin Scale as an outcome measure. Dr Lewis is a senior medical scientist at Berry Consultants, LLC, a statistical consulting company that provides clinical trial design services, including for trials of treatments for neurological conditions. Dr Nogueira: Stryker Neurovascular: Trevo 2 trial PI (modest), DAWN trial PI (no compensation); Covidien: SWIFT and SWIFT Prime Steering Committee (modest), STAIR trial Core Laboratory (significant); and Penumbra: Penumbra 3D trial Executive Committee (no compensation). Dr Saver is an employee of the University of California. The University of California, Regents receive funding for Dr Saver’s services as a scientific consultant regarding trial design and conduct to Medtronic/Covidien, Stryker, Brain’sGate, Pfizer, SiQuib, Boehringer Ingelheim (prevention only), ZI Biotech, and St. Jude Medical. Dr Saver has served as an unpaid site investigator in multicenter trials run by Lundbeck for which the UC Regents received payments on the basis of clinical trial contracts for the number of subjects enrolled. Dr Saver serves as an unpaid consultant to Genentech advising on the design and conduct of the PRISMS trial; neither the University of California nor Dr Saver received any payments for this voluntary service. The University of California has intellectual property rights in training vignettes for the Rankin Focused Assessment and in retrieval devices for stroke. The other authors report no conflicts.

References


Adopting a Patient-Centered Approach to Primary Outcome Analysis of Acute Stroke Trials Using a Utility-Weighted Modified Rankin Scale


Stroke. published online July 2, 2015;

Stroke is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 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/2015/07/02/STROKEAHA.114.008547

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/