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

Napasri Chaisinanunkul, MD[1,2], **Opeolu Adeoye, MD MS**[3], **Roger J. Lewis, MD, PhD**[4], **James C. Grotta, MD**[5], **Joseph Broderick, MD**[6], **Tudor G. Jovin, MD**[7], **Raul G. Nogueira, MD**[8], **Jordan Elm, PhD**[9], **Todd Graves, PhD**[10], **Scott Berry, PhD**[10], **Kennedy R. Lees, MD**[11], **Andrew D. Barreto, MD MS**[12], **Jeffrey L. Saver, MD**[1], and **the DAWN Trial and MOST Trial Steering Committees**

[1] Department of Neurology and Comprehensive Stroke Center, University of California, Los Angeles, Los Angeles, California [2] Phyathai Stroke Center, Phyathai 1 Hospital, Bangkok, Thailand [3] Departments of Emergency Medicine and Neurosurgery, and Neuroscience Institute, University of Cincinnati, Cincinnati, OH [4] Department of Emergency Medicine at Harbor-UCLA Medical Center and Berry Consultants, LLC [5] Memorial Hermann Hospital, Houston, TX [6] Department of Neurology and Neuroscience Institute, University of Cincinnati, Cincinnati, OH [7] Stroke Institute, University of Pittsburgh Medical Center, Pittsburgh, PA [8] Marcus Stroke and Neuroscience Center, Grady Memorial Hospital, Emory University, Atlanta, GA [9] Division of Biostatistics, Medical University of South Carolina, Charleston, SC [10] Berry Consultants, LLC [11] Department of Stroke Research, University of Glasgow, Glasgow, United Kingdom [12] University of Texas Health Science Center at Houston, Houston, TX

Abstract

Background and Purpose—Although the modified Rankin Scale (mRS) is the most commonly employed primary endpoint 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 seven 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: mRS 0–1.0; mRS 1 - 0.91; mRS 2 - 0.76; mRS 3 - 0.65; mRS 4 - 0.33; mRS 5 & 6 - 0. 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 six of eight unidirectional effect trials, while dichotomous analyses were statistically significant in two to four of eight. 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.

Conclusion—A utility-weighted mRS performs similarly 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.

The modified Rankin score (mRS) is the most widely used as a measure of outcome after acute ischemic stroke (AIS) in both research clinical trials and national and local quality improvement registries. However, there is much debate regarding how best statistically to analyze the mRS.[1, 2] Approaches include simple dichotomization, sliding dichotomy or responder analysis, and ordinal or "shift" analysis.[2] 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 seven possible mRS scores are collapsed into just two 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.[1] As data to guide selection of the most informative dichotomization is 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 non-analyzed 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 worth 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 (PCORI) and the National Institute for Health and Care Excellence (NICE) and health economists, strongly advocate the use of outcome metrics that measure benefits of a given intervention to the patient.[3-5] The most widely-accepted patient-centered outcome measure is utility – the desirability of a specific health outcome to the patient.[6] A promising approach to transforming the modified Rankin Scale into a patient-centered outcome measure is to weight the seven level of the modified Rankin Scale 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 version of the mRS has been recommended for acute stroke research by the Stroke Therapy Academic Industry Roundtable (STAIR).[7]

We aimed to derive a utility-weighted mRS (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 modified Rankin Scale by averaging utility values derived in two prior studies. In one 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.[8] 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 employing 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 (WHO-GBD).[9] An international panel of neurovascular and cardiovascular physicians and nurses participated in the WHO-GBD 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 seven mRS categories were reported.[10-20] Studies were selected to include all trials considered in United States national guidelines as providing supportive evidence for an acute ischemic stroke intervention (9), and for illustrative purposes one 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), trials with bidirectional beneficial and harm effects across different mRS scale transitions (2), and neutral unidirectional effects (1). In unidirectional benefit trials, all mRS cutpoints show better outcomes for treatment than control. In bidirectional treatment effect

trials, some mRS cutpoints show better outcomes for treatment but others show better outcome for control, e.g. in the TREVO 2 trial thrombectomy was associated with increased functional independence (mRS 0-2), 39.9% vs 21.8%, but also with increased mortality, 34.1% vs 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 two subgroups expected to have differential responses. Group modified Rankin Scale outcomes in each trial were re-analyzed using 5 statistical approaches: dichotomized at excellent outcome (0-1 vs 2-6); dichotomized at good or better outcome (0-2 vs 3-6); dichotomized at fair or better outcome (0-4 vs 5-6); ordinal analysis of the mRS; and analysis of the UW-mRS. The three particular dichotomizations of the mRS were selected because each had been used as the primary analysis for one or more of the analyzed trials or meta-analyses. Dichotomized analyses used Fisher's Exact test, ordinal analysis used the Mann Whitney test, and utility analysis used the t test. Additionally, 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. While 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 very closely normal (Central Limit Theorem). In these studies there are typically much greater than 100 per treatment arm. We have compared the results of the t-test to 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 two source studies were quite close. Averaging the values produced the following utility weights: mRS 0 - 1.0; mRS 1 - 0.91; mRS 2 - 0.76; mRS 3 - 0.65; mRS 4 - 0.33; mRS 5 - 0; mRS6 - 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, while dichotomization at 0-1 was positive in 4, dichotomization at 0-2 was positive in 4, and dichotomization at 0-4 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 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 (SWIFT mean utility

delta 0.18, MR CLEAN mean utility delta 0.10), and early intravenous thrombolysis (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 IMS 3 trial, nominal values for group utility differences were substantially larger in patients with severe (NIHSS 20) compared with moderate deficits (NIHSS 19) at entry, though these differences did not reach statistical significance.

Discussion

We found that a utility approach to analyses of the modified Rankin Scale 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 two studies that were the sources for the utility values in the UW-mRS found extremely similar utility values for each of the seven levels of the mRS, even though they employed very different derivation techniques. In one 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) compared to dichotomized tests.[21] When benefits were concentrated on single health state levels or there were signs of a bi-directional 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 (MELT, PROACT 2). For healthcare policy planners, economists, and methodololgists, the results of UW-mRS analyses offered greater immediate interpretability than the ordinal mRS, with treatment groups 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 (QALYs) gained or lost by an intervention or treatment. QALYs allow healthcare decision-makers the capability to understand the relative value of ostensibly different medical interventions for different diseases in order to maximize a societies' 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 QALY 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 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 is 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 non-contribution 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 stroke patients in clinical practice,[22] and many of these patients must be excluded from clinical trials with simple dichotomous mRS endpoints because their pre-existing disability precludes them from crossing the single health state transition being examined. A utility approach to analyzing the mRS can allow inclusion of additional stroke patients 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

Author Manuscript

derived in a population-based sample from Great Britain and a healthcare provider sample from North America and Asia. Utility values derived directly from stroke patients 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 (MCID) in utilities in stroke patients has not been well-studied. However, studies to define the MCID for utility have been performed in a wide range of other diseases, and have found MCIDs generally ranging from 0.04-0.10.[23-26]

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.

Acknowledgments

References

1. Bath PM, Lees KR, Schellinger PD, Altman H, Bland M, Hogg C, et al. Statistical analysis of the primary outcome in acute stroke trials. Stroke. 2012; 43:1171–1178. [PubMed: 22426314]

2. Saver JL. Optimal end points for acute stroke therapy trials: Best ways to measure treatment effects of drugs and devices. Stroke. 2011; 42:2356–2362. [PubMed: 21719772]

3. Hallan S, Asberg A, Indredavik B, Wideroe TE. Quality of life after cerebrovascular stroke: A systematic study of patients' preferences for different functional outcomes. J Intern Med. 1999; 246:309–316. [PubMed: 10475999]

4. Gage BF, Cardinalli AB, Albers GW, Owens DK. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. JAMA : the journal of the American Medical Association. 1995; 274:1839–1845. [PubMed: 7500532]

5. Solomon NA, Glick HA, Russo CJ, Lee J, Schulman KA. Patient preferences for stroke outcomes. Stroke; a journal of cerebral circulation. 1994; 25:1721–1725.

6. Feeny D. A utility approach to the assessment of health-related quality of life. Med Care. 2000; 38:II151–154. [PubMed: 10982100]

7. Albers GW, Goldstein LB, Hess DC, Wechsler LR, Furie KL, Gorelick PB, et al. Stroke treatment academic industry roundtable (stair) recommendations for maximizing the use of intravenous thrombolytics and expanding treatment options with intra-arterial and neuroprotective therapies. Stroke. 2011; 42:2645–2650. [PubMed: 21852620]

8. Rivero-Arias O, Ouellet M, Gray A, Wolstenholme J, Rothwell PM, Luengo-Fernandez R. Mapping the modified rankin scale (mrs) measurement into the generic euroqol (eq-5d) health outcome. Medical Decision Making. 2010; 30:341–354. [PubMed: 19858500]

9. 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. 2009; 40:3828–3833. [PubMed: 19797698]

10. NINDS rt-PA Stroke Group. Tissue plasminogen activator for acute ischemic stroke. New England Journal of Medicine. 1995; 333:1581–1587. [PubMed: 7477192]

11. Sandercock P, Wardlaw JM, Lindley RI, Dennis M, Cohen G, Murray G, et al. The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [ist-3]): A randomised controlled trial. Lancet. 2012; 379:2352–2363. [PubMed: 22632908]

12. Furlan A, Higashida R, Wechsler L, Gent M, Rowley H, Kase C, et al. Intra-arterial prourokinase for acute ischemic stroke. The proact ii study: A randomized controlled trial. Prolyse in acute cerebral thromboembolism. JAMA. 1999; 282:2003–2011. [PubMed: 10591382]

13. Ogawa A, Mori E, Minematsu K, Taki W, Takahashi A, Nemoto S, et al. Randomized trial of intraarterial infusion of urokinase within 6 hours of middle cerebral artery stroke: The middle cerebral artery embolism local fibrinolytic intervention trial (melt) japan. Stroke. 2007; 38:2633–2639. [PubMed: 17702958]

14. Saver JL, Jahan R, Levy EI, Jovin TG, Baxter B, Nogueira RG, et al. Solitaire flow restoration device versus the merci retriever in patients with acute ischaemic stroke (swift): A randomised, parallel-group, non-inferiority trial. Lancet. 2012; 380:1241–1249. [PubMed: 22932715]

15. Fransen PS, Beumer D, Berkhemer OA, van den Berg LA, Lingsma H, van der Lugt A, et al. Mr clean, a multicenter randomized clinical trial of endovascular treatment for acute ischemic stroke in the netherlands: Study protocol for a randomized controlled trial. Trials. 2014; 15:343. [PubMed: 25179366]

16. Anderson CS, Heeley E, Huang Y, Wang J, Stapf C, Delcourt C, et al. Rapid blood-pressure lowering in patients with acute intracerebral hemorrhage. N Engl J Med. 2013; 368:2355–2365. [PubMed: 23713578]

17. Hofmeijer J, Kappelle LJ, Algra A, Amelink GJ, van Gijn J, van der Worp HB. Surgical decompression for space-occupying cerebral infarction (the hemicraniectomy after middle cerebral artery infarction with life-threatening edema trial [hamlet]): A multicentre, open, randomised trial. Lancet Neurol. 2009; 8:326–333. [PubMed: 19269254]

18. Hacke W, Kaste M, Bluhmki E, Brozman M, Davalos A, Guidetti D, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med. 2008; 359:1317–1329. [PubMed: 18815396]

19. Nogueira RG, Lutsep HL, Gupta R, Jovin TG, Albers GW, Walker GA, et al. Trevo versus merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (trevo 2): A randomised trial. Lancet. 2012; 380:1231–1240. [PubMed: 22932714]

20. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, et al. Endovascular therapy after intravenous t-pa versus t-pa alone for stroke. N Engl J Med. 2013; 368:893–903. [PubMed: 23390923]

21. Saver JL, Gornbein J. Treatment effects for which shift or binary analyses are advantageous in acute stroke trials. Neurology. 2009; 72:1310–1315. [PubMed: 19092107]

22. Kwok CS, Potter JF, Dalton G, George A, Metcalf AK, Ngeh J, et al. The soar stroke score predicts inpatient and 7-day mortality in acute stroke. Stroke. 2013; 44:2010–2012. [PubMed: 23780953]

23. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in eq-5d utility and vas scores in cancer. Health Qual Life Outcomes. 2007; 5:70. [PubMed: 18154669]

24. Luo N, Johnson J, Coons SJ. Using instrument-defined health state transitions to estimate minimally important differences for four preference-based health-related quality of life instruments. Med Care. 2010; 48:365–371. [PubMed: 20355266]

25. Le QA, Doctor JN, Zoellner LA, Feeny NC. Minimal clinically important differences for the eq-5d and qwb-sa in post-traumatic stress disorder (ptsd): Results from a doubly randomized preference trial (drpt). Health Qual Life Outcomes. 2013; 11:59. [PubMed: 23587015]

26. McDonough CM, Tosteson TD, Tosteson AN, Jette AM, Grove MR, Weinstein JN. A longitudinal comparison of 5 preference-weighted health state classification systems in persons with intervertebral disk herniation. Med Decis Making. 2011; 31:270–280. [PubMed: 21098419]
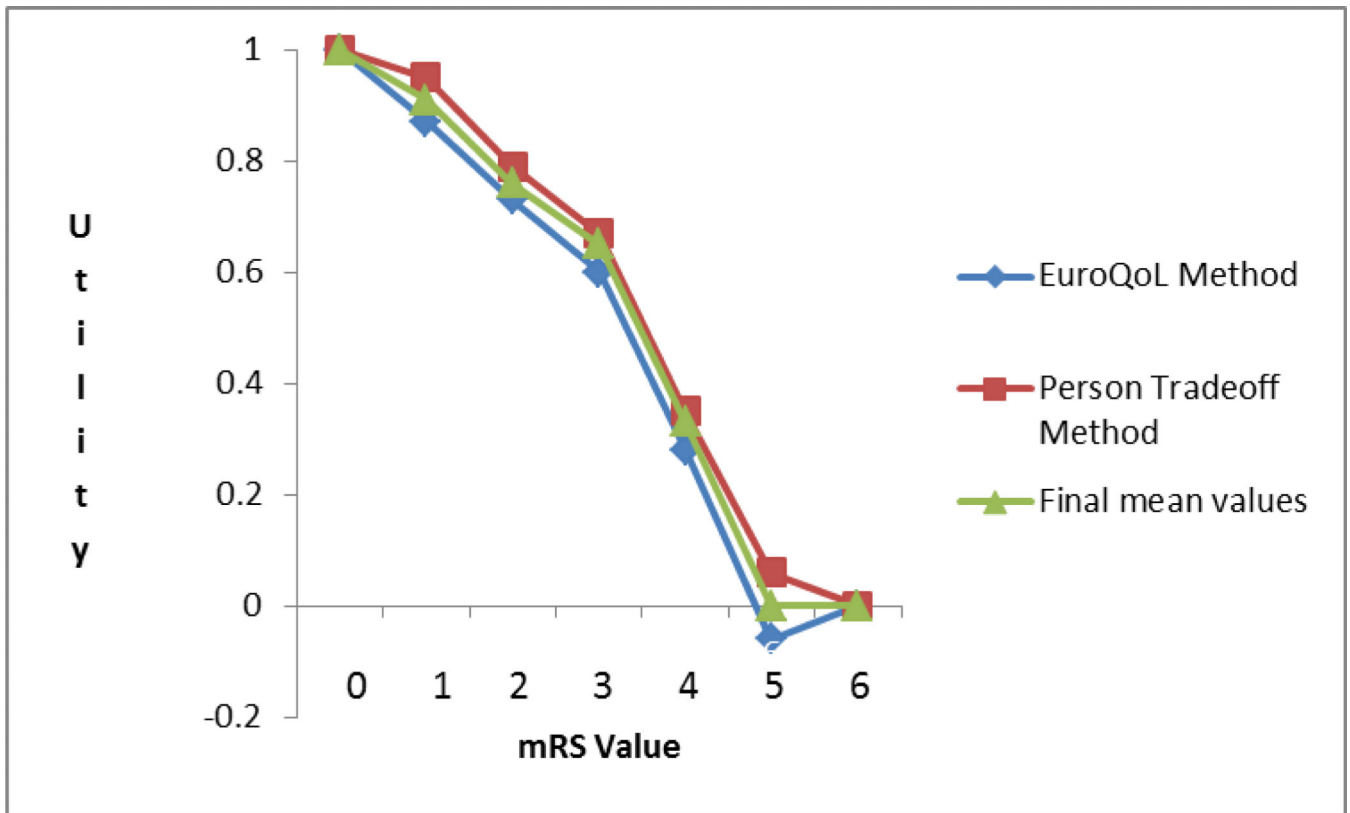
Figure 1. Comparability of Utility Values for Modified Rankin Scale Levels
The utility weights derived from patient informants by the time tradeoff method (blue) and
from healthcare provider informants by the person tradeoff method (red) are nearly identical,
as consequently are their averaged values comprising the UW-mRS (green). Transitions
from mRS health states 0-1, 1-2, and 2-3 are all valued moderately, from 3-4 and 4-5 valued
more substantially, and from 5-6 not valued at all.

Effect Direction	Intervention	Trial	Utility weighted mRS (t test)	Ordinal mRS (Mann Whitney)	Dichotomized 0-1 vs 2-6	Dichotomized 0-2 vs 3-6	Dichotomized 0-4 vs 5-6
Unidirectional	IV TPA	NINDS TPA Trials	0.0031	0.001	0.00	0.003	0.276
Unidirectional	IV TPA	IST 3	0.041	0.021	0.056	0.43	0.15
Unidirectional	Endovascular 1st Generation	PROACT 2	0.44	0.33	0.177	0.047	1.0
Unidirectional	Endovascular 1st Generation	MELT	0.33	0.19	0.03	0.28	0.89
Unidirectional	Endovascular 2nd Generation	SWIFT	0.011	0.012	0.308	0.308	0.004
Unidirectional	Endovascular 2nd Generation	MR CLEAN	0.0006	0.0006	0.037	0.0007	0.08
Unidirectional	BP Lowering	INTERACT 2	0.045	0.044	0.033	0.063	0.201
Unidirectional	Surgery	Hemicraniectomy Meta-Analysis	0.0001	0.00002	1.00	0.039	0.00
Bidirectional	IV TPA	ECASS 3	0.28	0.059	0.037	0.136	0.535
Bidirectional	Endovascular 2nd Generation	TREVO 2	0.22	0.35	0.047	0.009	0.535
Neutral	Endovascular 1st Generation	IMS 3	0.24	0.67	0.58	0.56	0.19

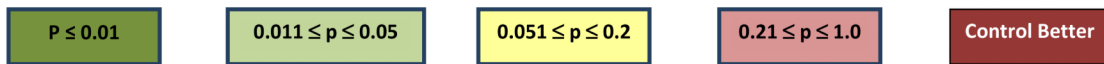| P ≤ 0.01	0.011 ≤ p ≤ 0.05	0.051 ≤ p ≤ 0.2	0.21 ≤ p ≤ 1.0	Control Better

Figure 2.
Trial Hypothesis Testing Using 5 Modes of Statistical Analysis of the Modified Rankin Scale P values are shown for trials using utility-weighting of the mRS, ordinal analysis of the mRS, and three dichotomized analyses of the mRS.

Table 1

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

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

Table 2

Subgroup Analysis of IMS 3 with UW-mRS

Study	Treatment Arm	N	Proportion with Primary Outcome	Mean Utility	Utility Difference	95% CI
IMS 3 Overall	TPA alone	222	0.39	0.505	0.036	−0.024 - 0.097
	TPA + Endovascular	434	0.41	0.541		
IMS 3, NIHSS 8-19	TPA alone	143	0.52	0.614	0.002	−0.069 - 0.072
	TPA + Endovascular	285	0.51	0.612		
IMS 3, NIHSS 20	TPA alone	71	0.17	0.287	0.101	−0.004 - 0.21
	TPA + Endovascular	130	0.24	0.388		