ARTICLE   OPEN ACCESS

Ordinal vs dichotomous analyses of modified Rankin Scale, 5-year outcome, and cost of stroke

Aravind Ganesh, MD, Ramon Luengo-Fernandez, DPhil, Rose M. Wharton, MSc,
and Peter M. Rothwell, FMedSci, on behalf of the Oxford Vascular Study

Correspondence
Prof. Rothwell
peter.rothwell@
clneuro.ox.ac.uk

Abstract

Objective
To compare how 3 common representations (ordinal vs dichotomized as 0–1/2–6 or 0–2/3–6) of the modified Rankin Scale (mRS)—a commonly used trial outcome measure—relate to long-term outcomes, and quantify trial ineligibility rates based on premorbid mRS.

Methods
In consecutive patients with ischemic stroke in a population-based, prospective, cohort study (Oxford Vascular Study; 2002–2014), we related 3-month mRS to 1-year and 5-year disability and death (logistic regressions), and health/social care costs (generalized linear model), adjusted for age/sex, and compared goodness-of-fit values (C statistic, mean absolute error). We also calculated the proportion of patients in whom premorbid mRS score >1 or >2 would result in exclusion from trials using dichotomous analysis.

Results
Among 1,607 patients, the ordinal mRS was more strongly related to 5-year mortality than both the 0–1/2–6 and 0–2/3–6 dichotomies (all $p < 0.0001$). Results were similar for 5-year disability, and 5-year care costs were also best captured by the ordinal model (change in mean absolute error vs age/sex: −\$3,059 for ordinal, −\$2,805 for 0–2/3–6, −\$1,647 for 0–1/2–6). Two hundred forty-four (17.1%) 3-month survivors had premorbid mRS score >2 and 434 (30.5%) had mRS score >1; both proportions increased with female sex, socioeconomic deprivation, and age (all $p < 0.0001$).

Conclusion
The ordinal form of the 3-month mRS relates better to long-term outcomes and costs in survivors of ischemic stroke than either dichotomy. This finding favors using ordinal approaches in trials analyzing the mRS. Exclusion of patients with higher premorbid disability by use of dichotomous primary outcomes will also result in unrepresentative samples.

RELATED ARTICLE

Editorial
Which version of the modified Rankin Scale should we use for stroke trials? Lump or split?

Glossary

The modified Rankin Scale (mRS) has been widely used as an outcome measure in stroke trials[1] and in more than 30 other neurologic conditions including Parkinson disease and autoimmune encephalitis (data available from Dryad, table e-1, doi.org/10.5061/dryad.609bp7m). However, there is lack of consensus on how the mRS should be analyzed, with more than half of stroke trials since 2007 still using dichotomous analysis, splitting the mRS into favorable/unfavorable outcomes, while about a quarter used ordinal (shift) analysis, assessing all changes across the range.[2] There is also disagreement about optimal dichotomy, with many recent trials using 0–1/2–6 instead of 0–2/3–6.[3,4]

Dichotomous analyses provide results that are easily explained using the absolute risk reduction in outcome between treatment groups but they may often be statistically inefficient, requiring larger sample sizes, and risk overvaluing a beneficial transition at one end of the distribution while disregarding harm at the other.[5] The ordinal approach, first advocated for stroke trials by Bath et al.[6] and Saver,[7] reduces this risk and allows trials to reduce sample size by 14% to 53%[8]—particularly appealing for trials in rarer conditions—but might also reduce statistical power if benefit is clustered at a few predictable state transitions.[9] When seeking a clinically interpretable parameter (common odds ratio), ordinal analysis is also limited by the proportional odds assumption, although alternatives have been proposed.[10] The dilemma is complicated by interrater variability, which is greatest for grades 1 and 2, around which the mRS is dichotomized.[11]

Aside from the issue of which approach is most statistically robust for short-term outcome analysis, the long-term trajectories of patients in different mRS grades can also inform the choice of primary outcome analysis, by testing the validity of the ordinal vs dichotomous forms of the mRS as they relate to long-term clinical or health-economic outcomes.[12] If meaningful differences in outcomes are seen across the range of mRS scores, resulting in a significantly better fit in regression models using the ordinal mRS, this would favor ordinal approaches, but if meaningful differences are seen only between certain groupings of mRS levels (e.g., mRS 0–2 vs 3–6), this would favor dichotomization.

Moreover, a dichotomous approach usually results in exclusion of patients with premorbid disability, since patients with mRS score >1 or >2 before their stroke cannot achieve a "favorable" poststroke outcome if defined as a 3-month mRS of 0–1 or 0–2, respectively. For example, recent trials of thrombectomy in acute stroke generally excluded patients with premorbid mRS score ≥2.[13] However, such patients could still contribute in ordinal analysis. It is unclear to what extent dichotomies risk excluding key segments of the population, potentially limiting the external validity of results; quantifying this risk may help trialists decide whether to dichotomize the mRS.

We therefore compared how the 3 commonly used mRS representations (ordinal/full-range vs dichotomized as 0–1/2–6 or 0–2/3–6) relate to 5-year disability, survival, and health/social care costs in 3-month survivors of ischemic stroke in a population-based study. We also examined the proportion of patients who had premorbid mRS score >1 or >2—who would risk exclusion with dichotomous approaches—and examined whether the relationship of the mRS representations to long-term outcomes would differ on excluding these patients.

Methods

The Oxford Vascular Study population comprises all individuals registered with about 100 general practitioners (GPs) in 9 practices across Oxfordshire, UK (midpoint population between 2002 and 2012 was 92,728; corresponding estimated midpoint population in Oxfordshire was 639,900).[14] The published characteristics of the study population, such as age and sex structure, closely resemble those of the general population of the United Kingdom.[14] Recruitment began in April 2002 and is ongoing. Near-complete ascertainment[15] of suspected stroke or TIA cases is achieved using multiple overlapping methods of "hot" and "cold" pursuit: (1) a daily, rapid access clinic to which participating GPs and the local emergency department refer all individuals with suspected TIA/stroke but are not hospitalized; (2) daily searches of admissions to medical, cardiology, stroke, neurology, and other relevant wards; (3) daily searches of the local emergency attendance register; (4) daily searches of in-hospital death records via the Bereavement Office; (5) monthly searches of all death certificates and coroners' reports for out-of-hospital deaths; (6) monthly searches of GP diagnostic coding and hospital discharge codes; and (7) monthly searches of all brain and vascular imaging referrals.

Consenting patients with ischemic stroke recruited from April 2002 to March 2014 were included. Patients were assessed urgently by study clinicians and considered for inclusion. Stroke was diagnosed per the World Health Organization definition.[16] Assessments of neurologic impairment, clinical presentation, medical/social history, and risk factors were made. All cases were reviewed by a senior neurologist (P. M.R.) daily, and imaging results were assessed by the study

neuroradiologist. Patients were followed up face to face by a study nurse or physician either in a hospital clinic or at home at 1, 3, and 6 months, 1 year, and 5 and 10 years. Recurrent vascular events and disability (mRS) were recorded at each follow-up. Raters were all trained in the use of the mRS using an instructional DVD with accompanying written materials produced by the University of Glasgow that has been used in large-scale clinical trials.[17]

Patients who moved out of the study area were followed up by telephone. Additional information was obtained from a carer in patients with impaired cognition or speech. All deaths were recorded via death certificates, coroners' reports, and the Office for National Statistics Central Register. Health and social care resource use was obtained from the date of the first stroke in the study period ("index" stroke) until 5 years poststroke or May 15, 2017, whichever was first. The methods on how resource use and costs were collected have been reported previously.[18,19] Briefly, patients' medical records from the Oxford University Hospitals Trust were reviewed for any emergency visit/transport, outpatient-care visit, day case, or hospitalization. For each stay in hospital, the dates of admission, discharge, and transfers between specialty wards were recorded. We estimated the number of institutionalized days as the difference between either date of 5-year follow-up or death, whichever was earliest, and the date of admission into the institution. Hospital resource use was valued using unit costs from the National Health Service's schedule of reference costs.[20] Institutionalization cost was recorded as the cost per week in a private nursing home, £795 ($1,145) in 2016.[21] All costs were presented in 2016 prices and converted from UK pounds sterling (£) to US dollars ($) using the 2016 rate of purchasing power parities ($ = £0.694, stats.oecd.org/).

Standard protocol approvals, registrations, and patient consents

The study was approved by the Oxfordshire Research Ethics Committee. Informed consent was obtained from patients whenever possible; otherwise, assent was obtained from caregivers if patients were unable to consent.

Statistical analyses

Since we sought to relate the 3-month mRS (as used in trials) to 5-year clinical outcomes and costs, and 3-month mRS = 6 (death) would be perfectly correlated with 5-year death/disability, we focused on patients surviving 3 months past their index stroke. However, since 3-month deaths would contribute to the primary outcome analysis of a trial and might improve the associations of each mRS representation with long-term outcomes to different extents, we ran regressions for 1- and 5-year death both including and excluding these patients.

The proportions of 3-month survivors who had premorbid mRS score >2 or >1 were calculated both for the overall cohort and in relation to age (<45, 45–54, 55–64, 65–74,

75–84, ≥85 years), sex, and index of socioeconomic deprivation. The proportions disabled, dead/disabled, or dead at 1 year and 5 years were also calculated and stratified by 3-month mRS. Disability at 5 years was defined as mRS score >2, but analysis was repeated using mRS score >1. Logistic regression was used to adjust associations of 3-month mRS and long-term outcomes for age/sex. Proportions of interest were compared using χ^2 tests.

Analyses were censored at May 15, 2017. Since we did not yet have full 5-year data for patients recruited after May 15, 2012, we examined the effect of censoring on costs,[22] partitioning the study period into smaller time periods (by day) within each of which the total cost incurred for all patients alive at the beginning of the period was calculated. Estimated costs of patients with complete data for each time period were weighted by the Kaplan-Meier sample average estimator and summed over all periods to estimate mean censor-adjusted costs. Costs were stratified by 3-month mRS and reported as means with 95% confidence intervals from 1,000-bootstrap estimates. To assess whether health/social care costs varied over time by 3-month mRS, we constructed generalized gamma linear models (GLMs) assuming a log identity, adjusted for age and sex. Statistical significance was set at $p < 0.050$.

We compared how each mRS representation related to disability and/or death using the area under the receiver operating characteristic (ROC) curve (AUC or C statistic) from age- and sex-adjusted logistic regressions, with each of the following forms of the 3-month mRS: (1) dichotomized as 0–2/3–6, (2) dichotomized as 0–1/2–6, and (3) including the full range of the mRS in the models. Therefore, the 3-month mRS was modeled in logistic regressions as either a binary variable (0–1 vs 2–5 or 0–2 vs 3–5) or a categorical variable with 6 levels (0, 1, 2, 3, 4, 5). Given that increases in goodness of fit in models using the ordinal mRS may relate to the increased number of parameters estimated, we used k-fold cross-validation to generate the AUCs for each logistic regression model (used for disability, death/disability, or death). This involves randomly splitting the dataset into k equally sized groups (we used k = 10). Data on one group is excluded, and the model is fitted to the data on the other k – 1 groups. The resulting model is then applied to the excluded group, and the AUC is calculated. This process is repeated k times, excluding each of the groups in turn. The resulting k AUCs are averaged to produce a single, overall, optimism-corrected estimate of the AUC.[23] We compared these AUCs using the standard nonparametric approach and calculated the ΔAUC by subtracting the AUC obtained from a logistic model including only age and sex.[24]

We compared the goodness of fit for each mRS representation as it related to 5-year health and social care costs using k-fold cross-validation to estimate the mean absolute error (MAE) for age- and sex-adjusted GLMs containing each representation. We also calculated the ΔMAE vs a GLM including only

age/sex (more negative values indicating better goodness of fit).

These analyses were repeated after excluding patients with premorbid mRS score >2, then premorbid mRS score >1. Analyses were performed using Stata 13.1 software (StataCorp LP, College Station, TX).

Data availability

Requests for access to the data used in this report will be considered by the corresponding author.

Results

Of 1,607 patients with an index ischemic stroke between 2002 and 2014, 181 (11.3%) died within 3 months. Complete baseline data were available for 1,421 (99.6%) 3-month survivors (table 1) and follow-up data for 1,403 (98.4%) survivors. Of the 23 excluded survivors, 19 refused inclusion and follow-up (refusal rate of 1.3%) and 4 had mRS assessments only beyond 3 months. Two hundred forty-three (17.1%)
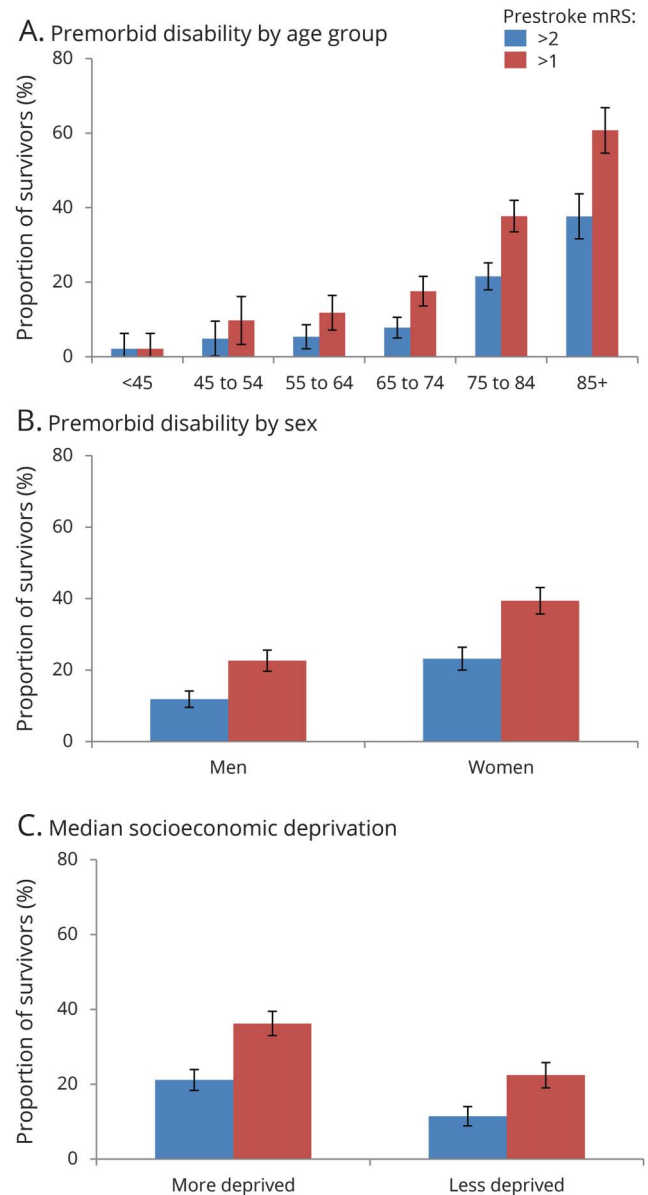
3-month survivors had a premorbid mRS score >2; this proportion increased with age from 5 patients (3.9%) younger than 55 years, to 228 (20.6%) older than 65 years, and 92 (37.9%) older than 85 years (figure 1, A). Four hundred thirty-three patients (30.5%) had a premorbid mRS score >1, including 9 patients (7.0%) younger than 55 years, 402 (36.3%) older than 65, and 150 (60.7%) older than 85. Women were more likely than men to have a premorbid mRS score >2 (155/670 [23.1%] vs 89/752 [11.8%]) or >1 (264/670 [39.4%] vs 170/752 [22.6%], $p < 0.0001$; figure 1, B). Patients at or worse than the median socioeconomic

Table 1 Patient sample and characteristics for 3-month survivors of ischemic stroke (n = 1,421; 3-month outcome available for 1,403 patients)

Characteristic	
Age, y, mean (SD)	73.2 (12.7)
Sex, male	753 (52.9)
History	
Myocardial infarction	177 (12.5)
Angina	239 (16.8)
Atrial fibrillation	260 (18.3)
Hypertension	889 (62.6)
Dyslipidemia	469 (33.0)
Diabetes	205 (14.4)
PVD	108 (7.6)
Stroke	158 (11.1)
TIA	205 (14.4)
Smoking	836 (58.8)
Cancer	22 (15.5)
Prior disability: mRS >2	243 (17.1)
Prior disability: mRS >1	433 (30.5)
3-mo mRS 0–2	869 (61.9)
3-mo mRS 0–1	564 (40.2)

Abbreviations: mRS = modified Rankin Scale; PVD = peripheral vascular disease.
Data represent n (%) unless otherwise indicated.

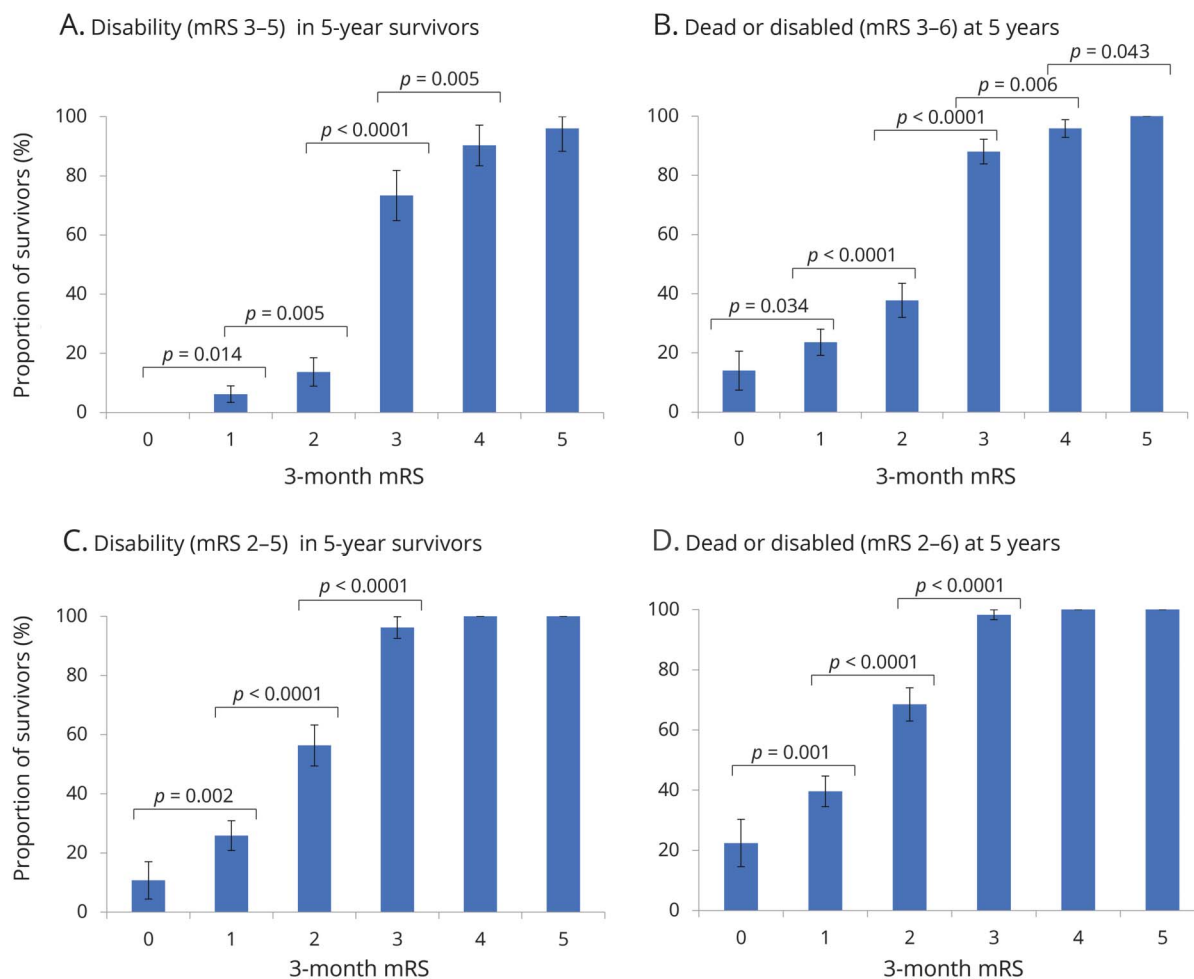**Figure 1** Proportion of 3-month survivors of ischemic stroke with premorbid disability

Premorbid disability was defined as a prestroke modified Rankin Scale (mRS) score of >2 (blue) or >1 (red) and is presented by (A) age, (B) sex, and (C) socioeconomic deprivation index. Bars represent 95% confidence intervals.

deprivation index were more likely than those less deprived to have a premorbid mRS score >2 (175/828 [21.1%] vs 68/593 [11.5%]) or >1 (300/828 [36.2%] vs 133/593 [22.5%], $p <$ 0.0001; figure 1, C). Even when considering only the 1,268 3-month survivors with first-in-lifetime strokes, 188 (14.8%) had a premorbid mRS score >2 and 347 (27.4%) had a premorbid mRS score >1. Age, female sex, and deprivation independently predicted premorbid disability on multivariable logistic regression even upon adjusting for significant comorbidities, including prior stroke (data available from Dryad, table e-2, doi.org/10.5061/dryad.609bp7m).

At 5 years, 465 (32.6%) 3-month survivors were dead and 214 (27.3%) of those alive had a 5-year mRS score >2 while 395 (50.3%) had an mRS score >1. One hundred seventy-three patients (12.1%) had not yet reached 5-year follow-up, but 1-year follow-up data were available for 169 of these, including everyone with a 3-month mRS. When disability was defined as mRS score >2, the proportion of survivors who were disabled at

5 years increased with 3-month mRS, with a bigger step change from mRS 2 to 3 vs mRS 1 to 2 (figure 2, A). Similar trends were observed with the composite outcome of 5-year death/disability (figure 2, B). On age- and sex-adjusted logistic regression, 3-month mRS was strongly related to 5-year disability (data available from Dryad, tables e-3 and e-4, doi.org/10.5061/dryad.609bp7m) and 5-year death/disability (data available from Dryad, tables e-5 and e-6), with AUC ≥0.8 for all 3 models. Both when including/excluding patients with premorbid mRS score >2, the AUCs for the ordinal model were consistently superior to the 0–2/3–5 and 0–1/2–5 dichotomies (e.g., for 5-year death/disability in all, ΔAUC [vs age/sex] = 0.132 for ordinal vs 0.124 for 0–2/3–5, $p = 0.016$; 0.080 for 0–1/2–5, $p < 0.0001$; ROC curves data available from Dryad, figures e-1 and e-2). When disability was defined as mRS score >1 (figure 2, C and D, and data available from Dryad, tables e-7 to e-10), the 2 dichotomies became similar in their relationship to 5-year disability and death/disability but the ordinal model remained superior when including/

Figure 2 Five-year disability and death/disability outcomes in survivors of ischemic stroke, stratified by 3-month mRS scores
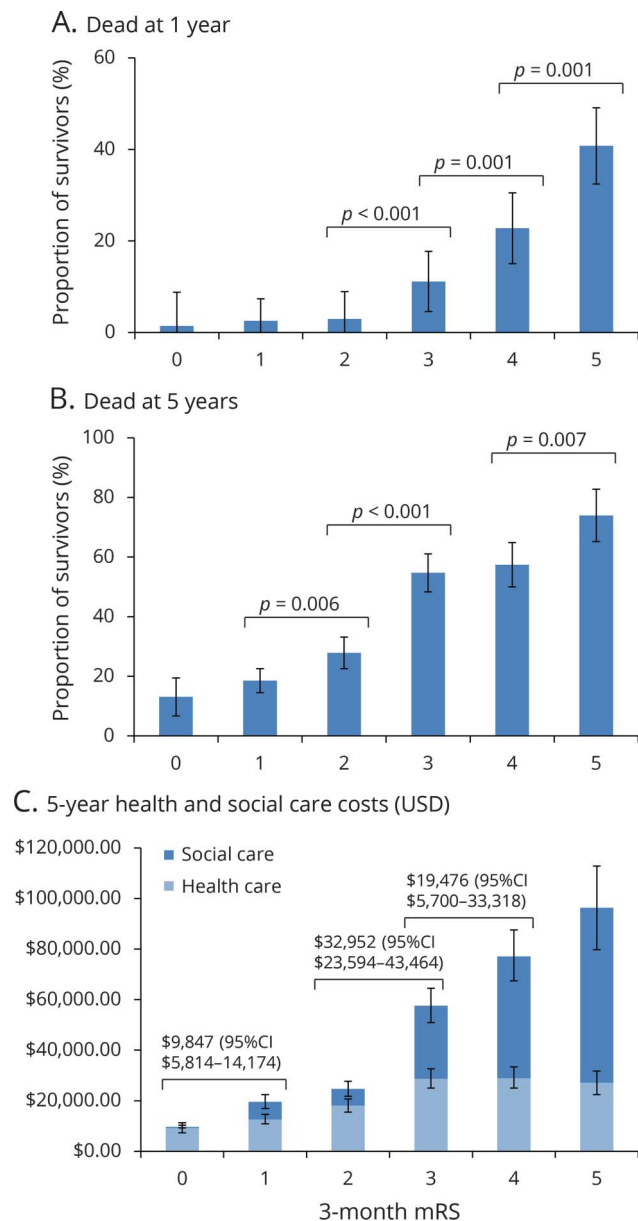

The graphs show the proportion of 3-month survivors, also alive at 5 years, who were disabled at 5 years (A–C), and the proportion of 3-month survivors who were dead/disabled at 5 years (B–D), with disability defined as 5-year mRS score >2 (A and B) or mRS score >1 (C and D). Significant differences between mRS grades are indicated using p values from χ² analysis. Bars represent 95% confidence intervals. mRS = modified Rankin Scale.

excluding those with premorbid mRS score >1 (ROC curves data available from Dryad, figures e-3 and e-4); e.g., for 5-year death/disability in all, ΔAUC = 0.071 for 0–2/3–5 vs 0.079 for 0–1/2–5, $p = 0.22$; 0.102 for ordinal, $p = 0.0001$.

Three-month mRS score was strongly related to 1-year mortality (figure 3, A), with 22 of 869 patients (2.5%) with 3-month mRS score of 0–2 dead at 1 year (no significant difference among 0, 1, or 2) vs 28 of 251 (11.2%) with mRS score 3, 41 of 180 (22.8%)

with mRS score 4, and 42 of 103 (40.8%) with mRS score 5 ($p \le 0.001$ for each comparison). Consequently, on age- and sex-adjusted logistic regression, the 0–2/3–5 dichotomy better related to 1-year death than 0–1/2–5 while the ordinal scale was superior to both (table 2). On examining 5-year risk of death, mortality differences emerged between mRS 1 (66/356, 18.5%) and 2 (76/273, 27.8%, $p = 0.006$) while mRS 3 and 4 were similar (figure 3, B). The ordinal model remained the best fit for 5-year death (table 3, ROC curves data available from Dryad (figure e-5, doi.org/10.5061/dryad.609bp7m) while the 0–2/3–5 dichotomy performed better than 0–1/2–5 ($p = 0.03$). In particular, jumps in odds of 5-year death were seen from mRS 2 to 3 and 3–4 to 5. When patients with 3-month mRS of 6 were included in the regressions for 1- and 5-year death, the ordinal model remained superior to both dichotomies (data available from Dryad, tables e-11 and e-12).

When patients with premorbid mRS score >2 were excluded (data available from Dryad, tables e-13 and e-14, doi.org/10.5061/dryad.609bp7m), the ordinal model no longer significantly outperformed the 0–2/3–5 dichotomy but was still consistently superior to 0–1/2–5 (e.g., for 5-year death, ΔAUC = 0.037 for ordinal vs 0.026 for 0–2/3–5, $p = 0.05$; 0.019 for 0–1/2–5, $p = 0.009$; ROC curves data available from Dryad, figure e-6). However, the 0–2/3–5 dichotomy no longer had a significantly better fit than 0–1/2–5 for 1-year deaths (ΔAUC = 0.072 for 0–2/3–5 vs 0.046 for 0–1/2–5, $p = 0.086$) and 5-year deaths ($p = 0.30$). When we further excluded those with premorbid mRS score >1 (data available from Dryad, tables e-15 and e-16), the 0–1/2–5 dichotomy became less useful with an AUC similar to age/sex alone for 1-year death (ΔAUC = 0.027, $p = 0.17$). The ordinal scale remained consistently superior, whereas the 0–2/3–5 dichotomy again was superior to 0–1/2–5 only in relation to 1-year death (ROC curves data available from Dryad, figure e-7).

Three-month mRS was also correlated with 5-year health and social care costs, which increased with each mRS grade (figure 3, C). Patients with 3-month mRS score of 5 accrued the highest costs despite high mortality, and there were significant differences in costs between patients with 3-month mRS score of 1 vs 2, 2 vs 3, and 3 vs 4 (e.g., mean difference between mRS 2 and 3: $32,952, 95% confidence interval $23,594–$43,464) vs the insignificant difference between mRS 1 and 2 ($5,104, −$1,090 to $10,614). Five-year costs were best captured by the ordinal model, which had the lowest MAE, i.e., minimized error (table 4). Results were similar when censored cases were included (data available from Dryad, table e-17, doi.org/10.5061/dryad.609bp7m); even upon excluding patients with premorbid mRS >2 and mRS >1 (data available from Dryad, tables e-18 and e-19), ordinal analysis was superior and 0–1/2–5 inferior to the 0–2/3–5 dichotomy.

Discussion

We have demonstrated that the ordinal mRS relates better to 5-year disability (whether defined as mRS >1 or >2), death/

Figure 3 Deaths and health and social care costs in survivors of ischemic stroke, stratified by 3-month mRS scores

The graphs show the proportion of 3-month survivors with full 5 years of follow-up who were dead at 1 year (A), dead at 5 years (B), and the cumulative 5-year health care costs for all survivors (C). Significant differences between mRS grades are indicated using p values from χ^2 analysis (A and B) and mean cost differences with 95% CIs (C). CI = confidence interval; mRS = modified Rankin Scale; USD = US dollars.

Table 2 Association of 3-month mRS, dichotomized as 0–1/2–5 or 0–2/3–5 or included as the full ordinal scale, with 1-year death in 3-month ischemic stroke survivors (n = 1,403)

	Age/sex alone	mRS dichotomized as 0–1 vs 2–5	mRS dichotomized as 0–2 vs 3–5	Full range of mRS
3-mo mRS				
0	NA	0–1 = reference	0–2 = reference	0 = reference
1				1.55 (0.34–7.14)
2		4.84 (2.67–8.79)		1.52 (0.32–7.22)
3			7.33 (4.50–12.0)	4.81 (1.10–21.0)
4				12.6 (2.91–54.1)
5				31.8 (7.31–139)
Age	1.09 (1.07–1.12)	1.07 (1.05–1.10)	1.06 (1.04–1.09)	1.07 (1.04–1.09)
Male	1.36 (0.93–1.98)	1.46 (0.99–2.14)	1.58 (1.06–2.34)	1.92 (1.26–2.91)
AUC (95% CI)[a]	0.743 (0.704–0.781)	0.790 (0.756–0.824)	0.823 (0.789–0.857)	0.843 (0.810–0.876)
ΔAUC (vs age/sex)[a]	NA	0.047	0.080	0.100
$p > |\chi^2|$ (vs age/sex)[a]	NA	<0.0001	<0.0001	<0.0001

Abbreviations: AUC = area under the receiver operating characteristic curve; CI = confidence interval; mRS = modified Rankin Scale; NA = not applicable.
Data represent adjusted odds ratio (95% CI). p = 0.0009 for comparison of AUC for 0–1/2–5 dichotomy vs 0–2/3–5 dichotomy (favoring the latter); p = 0.0001 for comparison of 0–1/2–5 dichotomy vs ordinal mRS (favoring the latter); p = 0.03 for comparison of AUC for 0–2/3–5 dichotomy vs ordinal mRS (favoring the latter).
[a] Generated by k-fold cross-validation.

disability, mortality, and health/social care costs, than both the 0–1/2–6 and 0–2/3–6 dichotomies. In addition to further validating the mRS as an outcome measure, these findings have implications for trials using the mRS as the primary outcome measure.

First, our findings favor retaining the full ordinal range of the mRS over dichotomizing it for analysis, by adding clinical data to prior arguments about statistical efficiency. Efficiency depends on expected treatment effects and does not always favor ordinal analysis.[9] However, a clinically robust primary outcome measure should capture differences in long-term disability, care needs (costs), and survival—objective treatment goals in neurologic disease—and our analyses demonstrate that the ordinal mRS best reflects these outcomes, with differences seen across the range of the scale. These results imply that any shifts in mRS can be meaningful, and that it may be misleading to consider certain states as being equivalent with respect to long-term outcomes. It is also inappropriate to view patients with premorbid disability as having no potential to benefit from disability-lowering treatments, given significant differences in long-term outcome among patients with 3-month mRS grades 3, 4, and 5, and certainly between mRS 2 and 3.

Second, our results highlight the substantial proportion of patients with premorbid disability who risk exclusion with dichotomous approaches. The 0–1/2–6 dichotomy would have excluded 30.5% of our cohort from attaining a favorable

outcome even prestroke and was more likely to exclude older, female, and socioeconomically deprived patients. Defining a favorable outcome as mRS score 0–2 is less exclusionary, but would still have prevented 17.2% of our patients from demonstrating any treatment benefit. Previous inpatient studies have reported higher rates of premorbid disability.[25] Even when trials use block randomization or actively enroll older patients, such as the Third International Stroke Trial, their sample will still be unrepresentative if they exclude patients with premorbid disability, as did the Third International Stroke Trial.[26] Ageism has been noted in interventional stroke studies,[27] and older patients, women, and socioeconomically deprived patients are less likely to receive appropriate acute stroke care.[28,29] By permitting patients with premorbid disability to demonstrate treatment benefit, ordinal analysis can encourage trials to enroll such patients and better represent the stroke population. However, clinical trials have generally enrolled healthier patients to optimize the detection of treatment effects, and enrolling patients with various levels of premorbid disability can complicate sample-size calculations and add another factor for adjustment when comparing treatment and control arms.

Third, for acute stroke trials that still wish to dichotomize the mRS, our data suggest that in addition to promoting exclusion of more patients, trials using the 0–1/2–6 dichotomy risk ultimately comparing more similar patient groups with respect to long-term outcomes than those using a 0–2/3–6 dichotomy, since patients with a 3-month mRS score of 2

Table 3 Association of 3-month mRS, dichotomized as 0–1/2–5 or 0–2/3–5 or included as the full ordinal scale, with 5-year death in 3-month ischemic stroke survivors with full 5 years of follow-up or death between 3 months and 5 years (n = 1,235)

	Age and sex alone	mRS dichotomized as 1–1 vs 2–5	mRS dichotomized as 0–2 vs 3–5	Full range of mRS
3-mo mRS				
0	NA	0–1 = reference	0–2 = reference	0 = reference
1				1.20 (0.62–2.31)
2		2.89 (2.13–3.92)		1.64 (0.85–3.17)
3			3.61 (2.74–4.75)	3.81 (1.97–7.37)
4				4.59 (2.31–9.10)
5				10.4 (4.83–22.6)
Age	1.12 (1.10–1.13)	1.10 (1.09–1.12)	1.10 (1.08–1.12)	1.10 (1.08–1.12)
Male	1.13 (0.87–1.48)	1.20 (0.91–1.57)	1.25 (0.94–1.65)	1.29 (0.97–1.71)
AUC (95% CI)[a]	0.796 (0.771–0.821)	0.817 (0.793–0.840)	0.829 (0.806–0.852)	0.839 (0.817–0.862)
ΔAUC (vs age/sex)[a]	NA	0.021	0.033	0.043
$p > \lvert \chi^2 \rvert$ (vs age/sex)[a]		0.0006	<0.0001	<0.0001

Abbreviations: AUC = area under the receiver operating characteristic curve; CI = confidence interval; mRS = modified Rankin Scale; NA = not applicable.
Data represent adjusted odds ratio (95% CI). $p = 0.031$ for comparison of AUC for 0–1/2–5 dichotomy vs 0–2/3–5 dichotomy (favoring the latter); $p = 0.0005$ for comparison of AUC for 0–1/2–5 dichotomy vs ordinal mRS (favoring the latter); $p = 0.026$ for comparison of AUC for 0–2/3–5 dichotomy vs ordinal mRS (favoring the latter).
[a] Generated by k-fold cross-validation where k = 10.

seem likely to fare similarly to those with an mRS of 1—and much more favorably than those with mRS of 3 ("step change"), considering long-term outcomes. This is a concern given the 0–1/2–6 dichotomy's popularity in recent trials.[3,4] However, 0–2/3–6 may not always be the most optimal dichotomy for trials based on other factors such as expected case

Table 4 Association of 3-month mRS, dichotomized as 0–1/2–5 or 0–2/3–5 or included as the full ordinal scale, with 5-year health care costs in 3-month ischemic stroke survivors with full 5 years of follow-up, or death between 3 months and 5 years (n = 1,235)

	Age and sex alone	mRS dichotomized as 1–1 vs 2–5	mRS dichotomized as 1–2 vs 3–5	Full range of mRS
3-mo mRS				
0	NA	0–1 = reference	0–2 = reference	0 = reference
1				9,808 (4,399 to 15,218)
2		39,723 (28,972 to 50,474)		12,032 (5,878 to 18,186)
3			48,299 (36,295 to 60,304)	39,851 (28,130 to 51,573)
4				61,622 (42,105 to 79,139)
5				80,052 (49,563 to 110,542)
Age	1,465 (1,028 to 1,902)	1,185 (782 to 1,588)	1,169 (770 to 1,569)	1,171 (773 to 1,570)
Male	−10,417 (−16,487 to −4,347)	−9,160 (−17,570 to −751)	−10,189 (−18,583 to −1,796)	−8,257 (−16,454 to −60)
MAE (95% CI)[a]	26,684 (25,022 to 28,345)	25,037 (23,031 to 27,042)	23,879 (22,727 to 25,031)	23,625 (22,796-24,454)
ΔMAE (vs age/sex)	NA	−1,647	−2,805	−3,059

Abbreviation: CI = confidence interval; MAE = mean absolute error; mRS = modified Rankin Scale; NA = not applicable.
Data represent margins, $ (95% CI).
[a] Generated by k-fold cross-validation where k = 10.

mix and stroke severity. For example, trials of hemi-craniectomy in malignant middle cerebral artery infarction used a dichotomy of 0–4/5–6.[30]

Fourth, our findings demonstrate nonlinear trends in the long-term outcomes predicted by different mRS grades, suggesting that not all shifts should be weighted equally in an ideal ordinal analysis. The shift from 0–2 to 3–6 appears to merit the highest weighting, but there are also jumps in mortality and care costs from 3 to 4 to 5. This is in contrast to conventional approaches to ordinal analysis, such as proportional odds logistic regression, which assume that all step increases in the mRS are equal. The utility-weighted mRS may be one method to address this issue, although it most values transitions from 3 to 4 to 5.[31] However, as recently suggested in a simulation study, differently weighted approaches such as the utility-weighted mRS may reduce the statistical power of randomized trials vs conventional ordinal analysis, although in this case, it was still more efficient than dichotomous approaches.[32] Using a sliding dichotomy, in which a good outcome—and thereby the state transition of interest for a given patient—is defined based on initial severity or other characteristics is more informative than a fixed dichotomy, but still fails to utilize information from the entire range of mRS outcomes.[5]

Our analysis has several strengths, including a robust population-based design, high rates of ascertainment of all incident strokes, completeness of follow-up, replication of findings for several outcomes, and generalizability, with similar 5-year mortality and disability rates as prior population-based studies.[33,34] However, there are some potential shortcomings. First, a randomized-controlled trial is required to prove a causal association between shifting patients to lower mRS scores and reducing long-term disability, mortality, and costs, and to validate the differential value of various state transitions suggested by our analyses. Our findings would also benefit from further validation in other large cohort studies. Second, since assessors were not blinded to prior mRS scores, 1- and 5-year disability assessments could have been unfavorably biased in patients with higher 3-month mRS scores. However, ordinal analysis remained superior with different disability definitions, and trends were similar for 5-year death and costs, which are untainted by disability assessments. Third, we only studied patients with ischemic stroke; distinctions among mRS grades may differ in other diseases. Fourth, given that our aim was to examine how the mRS relates to long-term outcomes, we did not adjust for all the potential factors that could also influence long-term disability, mortality, and costs, including stroke subtype, impairment on the NIH Stroke Scale, various cardiovascular and noncardiovascular comorbidities (such as cancer), recurrent strokes, socioeconomic status, or adherence to rehabilitation or secondary prevention. Ethnicity, potentially a relevant factor for such long-term outcomes, was also not examined; the Oxford Vascular Study population is 95% white. Such factors may need to be accounted for by studies seeking to more accurately estimate long-term outcomes for purposes like resource allocation. Fifth, given our wide recruitment period, secular trends in society and the health care system may have shifted the mRS distribution of our population over time. For instance, increasing longevity could have resulted in a higher proportion of patients with increased premorbid mRS over time; therefore, our estimate of the burden of premorbid disability may be an underestimate compared with present-day populations.

Furthermore, examining relationship to long-term outcomes such as mortality and costs is only one way to compare different 3-month mRS representations. Other measures that have been previously related to 3-month mRS include simultaneously assessed metrics such as quality of life (concurrent validity),[35] or retrospective metrics such as home time within the first 90 days after hospitalization.[36] Such metrics could also be used to examine differences between the different 3-month mRS representations in future studies, and it is possible that their findings may differ from ours.

Our findings demonstrate that ordinal analysis of the 3-month mRS better relates to long-term outcomes of ischemic stroke than either dichotomy. Exclusion of patients with higher premorbid disability by use of dichotomous primary outcomes will also result in unrepresentative samples. Similar considerations may apply to other comparable ordinal scales used for trial outcomes.

Author contributions
A.G. collected data, performed statistical analysis and interpretation, and wrote and revised the manuscript. R.L.-F. acquired data, performed statistical analysis, and revised the manuscript. R.M.W. performed statistical analysis. P.M.R. conceived and designed the study, provided study supervision and funding, analyzed and interpreted data, and wrote and revised the manuscript.

References

1. Saver JL, Goyal M, van der Lugt A, et al. Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis. JAMA 2016;316: 1279–1288.

2. Nunn A, Bath PM, Gray LJ. Analysis of the modified Rankin Scale in randomised controlled trials of acute ischaemic stroke: a systematic review. Stroke Res Treat 2016; 2016:9482876.

3. Anderson CS, Robinson T, Lindley RI, et al. Low-dose versus standard-dose intravenous alteplase in acute ischemic stroke. New Engl J Med 2016;374:2313–2323.

4. Thomalla G, Fiebach JB, Ostergaard L, et al. A multicenter, randomized, double-blind, placebo-controlled trial to test efficacy and safety of magnetic resonance imaging-based thrombolysis in wake-up stroke (WAKE-UP). Int J Stroke 2014;9:829–836.

5. Bath PM, Lees KR, Schellinger PD, et al. Statistical analysis of the primary outcome in acute stroke trials. Stroke 2012;43:1171–1178.

6. Bath PM, Geeganage C, Gray LJ, Collier T, Pocock S. Use of ordinal outcomes in vascular prevention trials: comparison with binary outcomes in published trials. Stroke 2008;39:2817–2823.

7. Saver JL. Novel end point analytic techniques and interpreting shifts across the entire range of outcome scales in acute stroke trials. Stroke 2007;38:3055–3062.

8. Optimising Analysis of Stroke Trials Collaboration. Calculation of sample size for stroke trials assessing functional outcome: comparison of binary and ordinal approaches. Int J Stroke 2008;3:78–84.

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

10. Howard G, Waller JL, Voeks JH, et al. A simple, assumption-free, and clinically interpretable approach for analysis of modified Rankin outcomes. Stroke 2012;43:664–669.

11. Quinn TJ, Dawson J, Walters MR, Lees KR. Variability in modified Rankin scoring across a large cohort of international observers. Stroke 2008;39:2975–2979.

12. Wade DT. Choosing a measure. In: Wade DT, editor. Measurement in Neurological Rehabilitation. Oxford, UK: Oxford University Press; 1992:35–43.

13. Goyal M, Menon BK, van Zwam WH, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. Lancet 2016;387:1723–1731.

14. Rothwell PM, Coull AJ, Silver LE, et al. Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (Oxford Vascular Study). Lancet 2005;366:1773–1783.

15. Coull AJ, Silver LE, Bull LM, Giles MF, Rothwell PM; Oxford Vascular (OXVASC) Study. Direct assessment of completeness of ascertainment in a stroke incidence study. Stroke 2004;35:2041–2045.

16. Hatano S. Experience from a multicentre stroke register: a preliminary report. Bull World Health Organ 1976;54:541–553.

17. Quinn TJ, Lees KR, Hardemark HG, Dawson J, Walters MR. Initial experience of a digital training resource for modified Rankin Scale assessment in clinical trials. Stroke 2007;38:2257–2261.

18. Luengo-Fernandez R, Paul NL, Gray AM, et al. Population-based study of disability and institutionalization after transient ischemic attack and stroke: 10-year results of the Oxford Vascular Study. Stroke 2013;44:2854–2861.

19. Luengo-Fernandez R, Gray AM, Rothwell PM; Oxford Vascular Study. A population-based study of hospital care costs during 5 years after transient ischemic attack and stroke. Stroke 2012;43:3343–3351.

20. NHS reference costs 2012 to 2013 [online]. Available at: gov.uk/government/publications/nhs-reference-costs-2012-to-2013. Accessed October 15, 2015.

21. Personal Social Services Research Unit. Unit costs of health and social care 2016 [online]. Available at: pssru.ac.uk/project-pages/unit-costs/2016/. Accessed February 10, 2017.

22. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics 1997;53:419–434.

23. Smith GC, Seaman SR, Wood AM, Royston P, White IR. Correcting for optimistic prediction in small data sets. Am J Epidemiol 2014;180:318–324.

24. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845.

25. Kwok CS, Potter JF, Dalton G, et al. The SOAR stroke score predicts inpatient and 7-day mortality in acute stroke. Stroke 2013;44:2010–2012.

26. Sandercock P, Lindley R, Wardlaw J, et al. Third International Stroke Trial (IST-3) of thrombolysis for acute ischaemic stroke. Trials 2008;9:37.

27. Hadbavna A, O'Neill D. Ageism in interventional stroke studies. J Am Geriatr Soc 2013;61:2054–2055.

28. Asdaghi N, Romano JG, Wang K, et al. Sex disparities in ischemic stroke care: FL-PR CReSD study (Florida–Puerto Rico Collaboration to Reduce Stroke Disparities). Stroke 2016;47:2618–2626.

29. Kimball MM, Neal D, Waters MF, Hoh BL. Race and income disparity in ischemic stroke care: nationwide inpatient sample database, 2002 to 2008. J Stroke Cerebrovasc Dis 2014;23:17–24.

30. Vahedi K, Hofmeijer J, Juettler E, et al. Early decompressive surgery in malignant infarction of the middle cerebral artery: a pooled analysis of three randomised controlled trials. Lancet Neurol 2007;6:215–222.

31. Chaisinanunkul N, Adeoye O, Lewis RJ, et al. Adopting a patient-centered approach to primary outcome analysis of acute stroke trials using a utility-weighted modified Rankin Scale. Stroke 2015;46:2238–2243.

32. Dijkland SA, Voormolen DC, Venema E, et al. Utility-weighted modified Rankin Scale as primary outcome in stroke trials: a simulation study. Stroke 2018;49:965–971.

33. Feigin VL, Barker-Collo S, Parag V, et al. Auckland Stroke Outcomes Study. Part 1: gender, stroke types, ethnicity, and functional outcomes 5 years poststroke. Neurology 2010;75:1597–1607.

34. Hankey GJ, Jamrozik K, Broadhurst RJ, Forbes S, Anderson CS. Long-term disability after first-ever stroke and related prognostic factors in the Perth Community Stroke Study, 1989–1990. Stroke 2002;33:1034–1040.

35. 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. Med Decis Making 2010;30:341–354.

36. Yu AYX, Rogers E, Wang M, et al. Population-based study of home-time by stroke type and correlation with modified Rankin score. Neurology 2017;89:1970–1976.