HHS Public Access

Author manuscript

Stroke. Author manuscript; available in PMC 2024 July 22.

Published in final edited form as:

Stroke. 2021 August ; 52(8): 2547–2553. doi:10.1161/STROKEAHA.120.032487.

Predicting 90-Day Outcome After Thrombectomy:

Baseline-Adjusted 24-Hour NIHSS Is More Powerful Than NIHSS Score Change

Eva A. Mistry, MBBS,
Sharon Yeatts, PhD,
Adam de Havenon, MD,
Tapan Mehta, MBBS, MPH,
Niraj Arora, MBBS,
Felipe De Los Rios La Rosa, MD,
Amy K. Starosciak, PhD,
James E. Siegler III, MD,
Akshitkumar M. Mistry, MD,
Shadi Yaghi, MD,
Pooja Khatri, MD, MSc

Department of Neurology (E.A.M.) and Department of Neurosurgery (A.M.M.), Vanderbilt University Medical Center, Nashville, TN. Public Health Sciences, Medical University of South Carolina, Charleston (S.Y.). Department of Neurology, University of Missouri, Columbia (N.A.). Ayer Neuroscience Institute, Hartford HealthCare, CT (T.M.). Baptist Health Neuroscience Center, Miami, FL (F.D.L.R.L.R., A.K.S.). Department of Neurology, Cooper University Hospital, Camden, NJ (J.E.S.). Department of Neurology, New York Langone Health (S.Y.). Department of Neurology and Rehabilitation Medicine, University of Cincinnati, OH (P.K.).

Abstract

BACKGROUND AND PURPOSE: The National Institutes of Health Stroke Scale (NIHSS) measured at an early time point is an appealing surrogate marker for long-term functional outcome of stroke patients treated with endovascular therapy. However, definitions and analytical methods for an early NIHSS-based outcome measure that optimize power and precision in clinical studies are not well-established.

METHODS: In this post-hoc analysis of our prospective observational study that enrolled endovascular therapy-treated patients at 12 comprehensive stroke centers across the US, we compared the ability of 24-hour NIHSS, ΔNIHSS (baseline minus 24-hour NIHSS), and percentage change (NIHSS×100/baseline NIHSS), analyzed as continuous and dichotomous measures, to predict 90-day modified Rankin Scale (mRS) using logistic regression (adjusted

Correspondence to: Eva A. Mistry, MBBS, Department of Neurology, Vanderbilt University Medical Center, 2525 W End Ave, St 612, Nashville, TN 37232. eva.a.mistry@vumc.org.

This manuscript was sent to Ajay K. Wakhloo, Guest Editor, for review by expert referees, editorial decision, and final disposition.
The Data Supplement is available with this article at https://www.ahajournals.org/doi/suppl/10.1161/STROKEAHA.120.032487.
The other authors report no conflicts.

for age, baseline NIHSS, glucose, hypertension, Alberta Stroke Program Early CT Score, time to recanalization, recanalization status, and intravenous thrombolysis) and Spearman **ρ**.

RESULTS: Of 485 patients in the BEST (Blood Pressure After Endovascular Stroke Therapy) cohort, 446 (92%) with 90-day follow-up data were included. An absolute 24-hour NIHSS, adjusted for baseline in multivariable modeling, had the highest predictive power of all definitions evaluated (aR^2 0.368 and adjusted odds ratio 0.79 [0.75–0.84], $P<0.001$ for mRS score 0–2; aR^2 0.444 and adjusted odds ratio 0.84 [0.8–0.86] for ordinal mRS). For predicting mRS score of 0–2 with a cut point, the second most efficient approach, the optimal threshold for 24-hour NIHSS score was ≤7 (sensitivity 80.1%, specificity 80.4%; adjusted odds ratio 12.5 [7.14–20], $P<0.001$), followed by percent change in NIHSS (sensitivity 79%, specificity 58.5%; adjusted odds ratio 4.55 [2.85–7.69], $P<0.001$).

CONCLUSIONS: Twenty-four–hour NIHSS, adjusted for baseline, was the strongest predictor of both dichotomous and ordinal 90-day mRS outcomes for endovascular therapy-treated patients. A dichotomous 24-hour NIHSS score of ≤7 was the second-best predictor. Although ΔNIHSS, continuous and dichotomized at ≥4, predicted 90-day outcomes, absolute 24-hour NIHSS definitions performed better.

Keywords

glucose; hypertension; National Institutes of Health; odds ratio; thrombectomy

As research surrounding outcomes after endovascular therapy (EVT) for acute ischemic stroke progresses, strategies to improve research efficiency become imperative.[1] One important component of the design of efficient EVT clinical trials is to define an early and optimal clinical surrogate end point that can differentiate efficacious and nonefficacious interventions with a degree of precision. Furthermore, a baseline and early follow-up (usually 24 hours) evaluation of neurological status, most frequently using the National Institutes of Health Stroke Scale (NIHSS) score, is routinely and easily performed. Early neurological status measured on the NIHSS has been strongly associated with patient-centered long-term functional outcomes.[2–5] Thus, early measurement of NIHSS following acute stroke intervention is an attractive early surrogate clinical end point for clinical trials. However, the definition and statistical analysis of this end point has widely varied in the acute stroke literature. Previous trials have considered an absolute decrease in NIHSS score by 4,[6] 8,[7,8] or 10 points, or 24-hour NIHSS score of 0–1,[9,10] to reflect early neurological recovery. Thus, 24-hour NIHSS is frequently analyzed as a binary outcome, often without an appropriate accounting of the baseline measurement to allow for increased efficiency. We aimed to identify the early, NIHSS-based outcome measures that best predicted 90-day functional outcome in patients treated with EVT using a large, prospective, modern, multicenter, and real-world data set and validate our findings using an external data set. We also provide data-driven arguments on the ideal analytical method for these early NIHSS-based surrogate outcome measures.

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request. We conducted a post hoc analysis of our prospective, multicenter cohort study, BEST (Blood Pressure After Endovascular Stroke Therapy), that enrolled adult (≥18 years) participants treated with EVT for an anterior cerebral circulation acute ischemic stroke (internal carotid artery or M1 or M2 segment of the middle cerebral artery) in the routine practice of medicine from November 2017 to September 2018 at 12 comprehensive stroke centers across the United States. Patients with a preexisting modified Rankin Scale score (mRS) ≥2, known terminal medical condition, left ventricular assist device, and an incident stroke in the perioperative or inpatient setting were excluded. Detailed methods of the BEST study have been previously published.[11] The BEST study was approved by the Institutional Review Boards of all participating institutions.

We included patients with a completed 90-day follow-up in this analysis. The following NIHSS-based early clinical outcomes were explored: (1) a threshold of NIHSS at 24 hours (dichotomous measure); (2) a threshold of change in NIHSS within 24 hours of EVT (baseline minus 24-hour NIHSS [ΔNIHSS]; dichotomous measure); (3) NIHSS at 24 hours (continuous measure) adjusted for baseline NIHSS in multivariable modeling; (4) ΔNIHSS (continuous measure); (5) percent change from baseline (calculated as ΔNIHSS/baseline NIHSS × 100). The primary outcome was mRS at 90 days. Youden index, a summary measure of the performance of a diagnostic test, was used to identify the optimal thresholds of 24-hour NIHSS and ΔNIHSS that best predicted 90-day mRS score of 0–2. This index was used to identify the cut point of a measure that maximally differentiates good versus bad outcomes by taking the specificity and sensitivity into account.[12] The 24-hour NIHSS and ΔNIHSS were dichotomized at these thresholds as well as other commonly used thresholds.[6–10] Continuous values of 24-hour NIHSS and ΔNIHSS were also considered. The strength of association with 90-day mRS was assessed using logistic regression (mRS score 0–2) and ordinal regression model (7-level ordinal mRS), adjusted for age, admission glucose, history of hypertension, baseline NIHSS (not included in the model with ΔNIHSS), Alberta Stroke Program Early CT Score, time from last known well to recanalization, successful recanalization (modified Thrombolysis in Cerebral Ischemia, ≥2b), and intravenous thrombolysis administration. The covariates for adjustment were selected a priori. To account for patient clustering by institution, as a sensitivity analysis, we fit mixed-effects logistic regression models to our outcome with institution as random effect. Furthermore, Spearman ρ was used to determine the potential predictive power of each covariate for all models for 90-day mRS. Spearman ρ is a measure of correlation between a set of predictors and an outcome. Higher value denotes that a higher correlation of a variable for that outcome. The proportional odds assumption of each ordinal regression model was tested using the Brant test. To test the linearity assumption for continuous 24-hour NIHSS, the general linear *F* test was applied to compare models with 24-hour NIHSS included with linear restrictions versus non-linear terms. The nonlinear 24-hour NIHSS was constructed using the restricted cubic spline function in STATA. A significant F-statistic denoted violation of the linearity assumption. The 24-hour NIHSS was used as a nonlinear term in the model if linearity assumption was violated.[13] Last, the Akaike

information criterion (AIC), which evaluates the in-sample fit of the model to estimate the likelihood of a model to predict the future values, commonly used for model selection, was calculated for each model (lower AIC values reflect a better model).[14] The workflow of the statistical analyses is outlined in Figure I in the Data Supplement. Missing data were not imputed. The models for the optimal definition were externally validated using the publicly available IMS-III trial (Interventional Management of Stroke III) data.[15] All statistical analyses were performed using the STATA/IC 16.0 (College Station, TX) and GraphPad Prism (GraphPad Software, San Diego, CA). The level of significance was set at 0.05 for all statistical analyses. All *P* values are 2-sided. All effect sizes are reported with 95% confidence intervals in addition to *P* values. To improve clinical utility of our results, we provide a tool for automated calculation of predicted probability of 90-day mRS score of 0–2 based on 24-hour NIHSS and baseline NIHSS using the formula:

$$Z\ \text{score} = \text{intercept} + (\beta 1 \times 24 - \text{hour}\ \ \text{NIHSS}) + (\beta 2 \times \text{baseline NIHSS})$$

$$\text{Probability of}\ \ 90 - \text{day mRS score}\ \ 0 - 2 = (1/(1 + 2.718281^{(-z\ \ \text{score})})) \times 100$$

RESULTS

Of 485 patients enrolled in the BEST cohort, 446 (92%) complete cases with 90-day follow-up data were included in this study (228 [51%] females; mean age 68±15 years; median baseline NIHSS 16 [interquartile range, 11–20]). Additional baseline characteristics of this cohort are outlined in Table 1.

Twenty-Four–Hour NIHSS (Dichotomous Measure)

For 24-hour NIHSS, the optimal Youden index to predict 90-day mRS score of 0–2 was achieved at ≤7 (sensitivity 80.1%, specificity 80.4%, area under the curve [AUC] 0.855 [0.819–0.887], *P*<0.001, Figure 1A, Table 2). Of 446 patients, 200 (45%) had a 24-hour NIHSS score of ≤7. Patients with 24-hour NIHSS score of ≤7 had an increased odds of having mRS score of 0–2 at 90 days (odds ratio [OR], 16.67 [95% CI, 10–25], *P*<0.001; adjusted OR [aOR], 12.5 [95% CI, 7.14–20], *P*<0.001, AIC 401.7; Table 2). NIHSS score of ≤7 was also associated with a more favorable mRS distribution (common OR [cOR], 14.9 [95% CI, 9.8–22.7]; acOR 9.8 [95% CI, 6.25–15.4]). The receiver operating characteristics for a 24-hour NIHSS cut point of ≤2 are provided in Table 2 for reference.

ΔNIHSS (Dichotomous Measure)

The optimal cut point for ΔNIHSS for predicting 90-day mRS score of 0–2 was ≥4 (sensitivity 79%, specificity 58.5%, AUC 0.73 [0.685–0.77], *P*<0.001, Figure 1B, Table 2). Of 446 patients, 255 (57%) had a ΔNIHSS score of ≥4. Patients with ΔNIHSS score of ≥4 had an increased odds of having mRS score of 0–2 at 90 days (OR, 5.27 [95% CI, 3.44–8.33], *P*<0.001; aOR, 4.55 [95% CI, 2.85–7.69], *P*<0.001, AIC 478.6; Table 2). ΔNIHSS score of ≥4 was also associated with a more favorable mRS distribution (cOR, 4.54 [95% CI, 3.18–6.45]; acOR, 3.5 [95% CI, 2.4–5.2]). The receiver operating characteristics for the

ΔNIHSS cut point of ≥8 is provided in Table 2. The difference between the AUC for the 24-hour NIHSS and ΔNIHSS was 0.126, P<0.001(Figure 1C), favoring the former.

Twenty-Four–Hour NIHSS (Continuous Measure) Adjusted for the Baseline

The 24-hour NIHSS adjusted for baseline, along with all prespecified covariates, in multivariable modeling was associated with 90-day mRS score of 0–2 with an OR of 0.79 (95% CI, 0.75–0.82), P<0.001 and aOR 0.79 (95% CI, 0.75–0.84), P<0.001, AIC 390.9. The algebraic equation incorporating the regression coefficients for calculation of probability of 90-day mRS score of 0–2 is as follows:

$$Z \text{ score} = 1.764354 + (-0.237319 \times 24 - \text{hour NIHSS}) + (-0.0043874 \times \text{baseline NIHSS})$$

$$\text{Probability of } 90-\text{day mRS score } 0-2 = (1/(1 + 2.718281^{(-Z \text{ score})})) \times 100$$

The OR for having a favorable shift in 90-day mRS was 0.82 (95% CI, 0.8–0.84; acOR 0.84 [95% CI, 0.8–0.86]); however, the linearity assumption of 24-hour NIHSS was violated in both models and limited the validity of these results. Treating the 24-hour NIHSS as a nonlinear term (restricted cubic spline with 3 knots) improved the fit of the ordinal regression model (χ^2 321 versus 312), but not the logistic model. The nonlinear prediction of 90-day mRS based on 24-hour NIHSS (unadjusted for the baseline) is outlined in Figure 2.

ΔNIHSS (Continuous Measure)

The ΔNIHSS, when treated as a continuous measure, was associated with 90-day mRS score of 0–2 with an OR of 1.11 (95% CI, 1.07–1.13, P<0.001 and aOR 1.11 [95% CI, 1.07–1.14], P<0.001), AIC 481. The OR for favorable shift in 90-day mRS was 1.09 (95% CI, 1.08–1.12; acOR1 1.08 [95% CI, 1.05–1.11]).

Percent Change in NIHSS (Continuous Measure)

The optimal cut point for percent change in NIHSS for predicting 90-day mRS score of 0–2 was >41.2% (sensitivity 79.5%, specificity 73.5%, AUC 0.819 [0.78–0.854], P<0.001, Figure 1C). Percent change in NIHSS from the baseline was associated with 90-day mRS score of 0–2 with an OR of 1.016 (95% CI, 1.01–1.02), P<0.001 and aOR of 1.016 (95% CI, 1.01–1.02), P<0.001, AIC 465.9. The OR for favorable shift in 90-day mRS was 1.013 (95% CI, 1.01–1.016), P<0.001 and acOR 1.014 (95% CI, 1.01–1.017), P<0.001.

Of all potential surrogates considered, 24-hour NIHSS adjusted for baseline (continuous) had the highest predictive power with an adjusted Spearman ρ^2 of 0.368 for mRS score of 0–2 and 0.444 for ordinal mRS. Dichotomous 24-hour NIHSS score of ≤7 had the next best predictive power. The values of unadjusted and adjusted Spearman ρ^2 for these definitions and other model covariates are outlined in Table I in the Data Supplement. We validated the models for this optimal definition using the IMS-III data set. The 24-hour NIHSS (nonlinear term with 2 knots) adjusted for baseline strongly predicted 90-day mRS with comparable

β coefficients, that is, −0.2809 for 24-hour NIHSS and −0.00149 for the baseline NIHSS (aOR, 0.75 [95% CI, 0.71–0.80], $P<0.001$, AUC 0.913 for mRS score 0–2 and acOR 0.79 [95% CI, 0.76–0.82], $P<0.001$). The results of sensitivity analyses were unchanged in the mixed-effects models to account for institutional clustering of the patients.

DISCUSSION

Among various definitions of early NIHSS-based surrogate outcome measures, we determined and externally validated that a 24-hour NIHSS adjusted for baseline was the strongest predictor of both dichotomous and ordinal 90-day mRS outcomes for acute ischemic stroke patients treated with EVT. Of the models containing different definitions, the one with absolute 24-hour NIHSS adjusted for the baseline was the best model (with the lowest AIC value). A dichotomous 24-hour NIHSS of ≤7 was the second-best predictor. Although ΔNIHSS, continuous and dichotomized at ≥4, predicted 90-day outcomes, it performed less well than the continuous 24-hour NIHSS definitions.

Consistent with prior studies, early neurological status or recovery, defined in any number of ways, was associated with functional outcomes in our study.[2,16–18] However, the early neurological status-based outcome measures have been heterogeneously defined and analyzed in the literature. Prior studies evaluating early neurological status as a surrogate end point for 90-day outcome have largely used a binary measure of an improvement (change) in NIHSS of greater than a certain arbitrary value.[16–18] Furthermore, the analytical methods used in prior studies are also constrained by restrictive linear assumptions for the NIHSS. Our study provides valuable insight into the comparative predictive ability of early, NIHSS-based outcome measures defined as various continuous or binary measures in EVT-treated cohort of patients with external validation. It also provides guidance on the optimal analytical methods, such as allowing flexibility for nonlinear prediction, to maximize the power of this surrogate end point for future clinical trials and observational studies.

It is not entirely surprising that a baseline-adjusted, 24-hour NIHSS best predicted 90-day outcomes, and that this association was the strongest when mRS was treated as an ordinal outcome. This is likely due to the fact that, by not dichotomizing both of these variables, we maximize the statistical power.[19–21] It is well known that dichotomization of continuous and ordinal variables at a certain, often arbitrary, threshold leads to loss of critical clinical and statistical information,[22,23] such as, and most importantly, a nonlinear relationship between NIHSS and mRS as highlighted in our study. Moreover, dichotomization can lead to false-positive results.[24] While a major movement within the field exists to use 90-day mRS as a nondichotomous outcome, most studies continue to define early neurological status or recovery in a dichotomous fashion. This may be, in part, due to continuous baseline-adjusted 24-hour NIHSS being less intuitive to the clinician for understanding patient outcomes. To address this limitation to the approach, we suggest using the algebraic equation of our model provided in above, which will calculate the probability of 90-day mRS score of 0–2 based on the BEST data set upon inputting baseline and 24-hour NIHSS values (automated output of this equation can be found available at https://redcap.vanderbilt.edu/surveys/?s=734P978D9E; which is now validated using the IMS-III trial data set).

It is important to note that despite a relatively minor contribution of baseline NIHSS in the overall prediction of 90-day outcomes, it is imperative to adjust for it in multivariable modeling when the 24-hour NIHSS is used as surrogate outcome in clinical trials and prospective studies to improve the power and reduce the bias. The 24-hour NIHSS is often dependent on the baseline NIHSS and accounting for these baseline differences in analysis, especially when baseline NIHSS is not used for stratification of randomization, is imperative to reduce biased findings.[25] Accounting for these individual differences in the baseline with regression modeling can improve the precision of the 24-hour NIHSS outcome measure and thus improve the power to detect treatment effect.

In addition to performing least well of all early NIHSS-based outcome definitions, ΔNIHSS is a suboptimal outcome for the following reasons. First, in our study, this end point was associated with unacceptably low specificity as a surrogate end point. Second, ΔNIHSS does not appropriately account for the baseline NIHSS or clinical assumptions surrounding the NIHSS. For example, a ΔNIHSS of 4 has vastly different implications for a patient with baseline NIHSS of 6 versus 22. Additionally, a certain ΔNIHSS is likely to have different implications, according to the affected hemisphere of the stroke. Third, using ΔNIHSS as an outcome requires statistical assumptions of perfect correlation between 24-hour and baseline NIHSS and constant variance to be met.[25,26] Due to these deficiencies, 24-hour NIHSS, adjusted for the baseline using regression modeling, provides added power and precision over ΔNIHSS. In fact, in a post hoc analysis of the National Institute of Neurological Disorders and Stroke Recombinant Tissue Plasminogen Activator Stroke Study, intravenous thrombolysis had a significant effect on early neurological recovery when measured as a relative change from baseline as opposed to an absolute change.[6,27] However, we found that 24-hour NIHSS was more strongly associated with 90-day mRS compared to percent change from baseline as demonstrated by higher correlation (Table I in the Data Supplement). Thus, our study provides evidence that regression modeling of the 24-hour NIHSS may be a superior analytic approach to relative change.[26]

Our results should be interpreted in light of several limitations. First, our study only included patients treated with EVT and only for anterior circulation strokes. The optimal early NIHSS-based outcome measures in patients treated exclusively with thrombolytics, not treated with endovascular reperfusion therapies, or having a posterior circulation stroke may be different. Second, BEST is a prospective observational study that collected data from routine clinical care of EVT-treated patients. The adjudication of 24-hour NIHSS and 90-day mRS may not occur with high accuracy in the routine clinical practice; arguably, this can also be a strength of our study in that it provides results generalizable to real-world clinical practice. Although a limited number of covariates were included in our prediction model, we were able to account for the most established predictors of outcome in this patient population.[28,29] Our findings should be confirmed in post hoc analyses of additional large clinical trials.

CONCLUSIONS

We determined that absolute 24-hour NIHSS, treated as a continuous variable and adjusted for baseline NIHSS, best predicts 90-day functional outcomes, as commonly defined in

research and clinical practice for patients treated with EVT for anterior circulation acute ischemic stroke. Investigators should consider this early surrogate end point to maximize the power and precision of their research studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Hayden LaFever for his assistance with creating the publicly available survey to calculate predicted probability of 90-day modified Rankin Scale score of 0–2.

Sources of Funding

Dr Mistry reports funding from the National Institutes of Health/National Institute of Neurological Disorders and Stroke (NIH/NINDS) K23NS113585. Dr Khatri reports funding from NIH/NINDS U24NS107241.

For Sources of Funding and Disclosures, see page 2552.

Disclosures

Dr Yeatts reports grants from the National Institute of Neurological Disorders and Stroke (NINDS) outside the submitted work.

Nonstandard Abbreviations and Acronyms

ΔNIHSS	baseline minus 24-hour NIHSS
AIC	Akaike information criterion
aOR	adjusted odds ratio
AUC	area under the curve
BEST	Blood Pressure After Endovascular Stroke Therapy
EVT	endovascular therapy
IMS-III	Interventional Management of Stroke III
mRS	modified Rankin Scale
NIHSS	National Institutes of Health Stroke Scale

REFERENCES

1. Liebeskind DS, Derdeyn CP, Wechsler LR; STAIR X Consortium. STAIR X: emerging considerations in developing and evaluating new stroke therapies. Stroke. 2018;49:2241–2247. doi: 10.1161/STROKEAHA.118.021424 [PubMed: 30355006]
2. Rangaraju S, Frankel M, Jovin TG. Prognostic value of the 24-hour neurological examination in anterior circulation ischemic stroke: a post hoc analysis of two randomized controlled stroke trials. Interv Neurol. 2016;4:120–129. doi: 10.1159/000443801 [PubMed: 27051408]
3. Powers WJ, Rabinstein AA, Ackerson T, Adeoye OM, Bambakidis NC, Becker K, Biller J, Brown M, Demaerschalk BM, Hoh B, et al. ; American Heart Association Stroke Council. 2018 guidelines for the early management of patients with acute ischemic aroke: a guideline

for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2018;49:e46–e110. doi: 10.1161/STR.0000000000000158 [PubMed: 29367334]

4. Saver JL, Altman H. Relationship between neurologic deficit severity and final functional outcome shifts and strengthens during first hours after onset. Stroke. 2012;43:1537–1541. doi: 10.1161/STROKEAHA.111.636928 [PubMed: 22492517]
5. Irvine HJ, Battey TW, Ostwaldt AC, Campbell BC, Davis SM, Donnan GA, Sheth KN, Kimberly WT. Early neurological stability predicts adverse outcome after acute ischemic stroke. Int J Stroke. 2016;11:882–889. doi: 10.1177/1747493016654484 [PubMed: 27334760]
6. National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. Tissue plasminogen activator for acute ischemic stroke. N Engl J Med. 1995;333:1581–1587. doi: 10.1056/NEJM199512143332401 [PubMed: 7477192]
7. Campbell BC, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, Yan B, Dowling RJ, Parsons MW, Oxley TJ, et al. ; EXTEND-IA Investigators. Endovascular therapy for ischemic stroke with perfusion-imaging selection. N Engl J Med. 2015;372:1009–1018. doi: 10.1056/NEJMoa1414792 [PubMed: 25671797]
8. Goyal M, Menon BK, van Zwam WH, Dippel DW, Mitchell PJ, Demchuk AM, Dávalos A, Majoie CB, van der Lugt A, de Miquel MA, et al. ; HERMES Collaborators. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. Lancet. 2016;387:1723–1731. doi: 10.1016/S0140-6736(16)00163-X [PubMed: 26898852]
9. Nogueira RG, Jadhav AP, Haussen DC, Bonafe A, Budzik RF, Bhuva P, Yavagal DR, Ribo M, Cognard C, Hanel RA, et al. ; DAWN Trial Investigators. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. N Engl J Med. 2018;378:11–21. doi: 10.1056/NEJMoa1706442 [PubMed: 29129157]
10. Mocco J, Zaidat OO, von Kummer R, Yoo AJ, Gupta R, Lopes D, Frei D, Shownkeen H, Budzik R, Ajani ZA, et al. ; THERAPY Trial Investigators*. Aspiration thrombectomy after intravenous alteplase versus intravenous alteplase alone. Stroke. 2016;47:2331–2338. doi: 10.1161/STROKEAHA.116.013372 [PubMed: 27486173]
11. Mistry EA, Sucharew H, Mistry AM, Mehta T, Arora N, Starosciak AK, De Los Rios La Rosa F, Siegler JE 3rd, Barnhill NR, Patel K, et al. Blood pressure after endovascular therapy for ischemic stroke (BEST): a Multicenter Prospective Cohort Study. Stroke. 2019;50:3449–3455. doi: 10.1161/STROKEAHA.119.026889 [PubMed: 31587660]
12. Youden WJ. Index for rating diagnostic tests. Cancer. 1950;3:32–35. doi: 10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3 [PubMed: 15405679]
13. Stat 501 Regression Methods. 6.2 - The General Linear F-Test. Accessed August 25, 2020. https://online.stat.psu.edu/stat501/lesson/6/6.2.2020.
14. Portet S. A primer on model selection using the Akaike Information Criterion. Infect Dis Model. 2020;5:111–128. doi: 10.1016/j.idm.2019.12.010 [PubMed: 31956740]
15. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, Jauch EC, Jovin TG, Yan B, Silver FL, et al. ; Interventional Management of Stroke (IMS) III Investigators. Endovascular therapy after intravenous t-PA versus t-PA alone for stroke. N Engl J Med. 2013;368:893–903. doi: 10.1056/NEJMoa1214300 [PubMed: 23390923]
16. Heit JJ, Mlynash M, Kemp SM, Lansberg MG, Christensen S, Marks MP, Ortega-Gutierrez S, Albers GW. Rapid neurologic improvement predicts favorable outcome 90 days after thrombectomy in the DEFUSE 3 Study. Stroke. 2019;50:1172–1177. doi: 10.1161/STROKEAHA.119.024928 [PubMed: 30932783]
17. Rudilosso S, Urra X, Amaro S, Llull L, Renú A, Laredo C, Obach V, Chamorro Á. Timing and relevance of clinical improvement after mechanical thrombectomy in patients with acute ischemic stroke. Stroke. 2019;50:1467–1472. doi: 10.1161/STROKEAHA.118.024067 [PubMed: 31113338]
18. Soize S, Fabre G, Gawlitza M, Serre I, Bakchine S, Manceau PF, Pierot L. Can early neurological improvement after mechanical thrombectomy be used as a surrogate for final stroke outcome? J Neurointerv Surg. 2019;11:450–454. doi: 10.1136/neurintsurg-2018-014332
19. Broderick JP, Adeoye O, Elm J. Evolution of the modified Rankin Scale and its use in future stroke trials. Stroke. 2017;48:2007–2012. doi: 10.1161/STROKEAHA.117.017866 [PubMed: 28626052]

20. 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. doi: 10.1161/STROKEAHA.111.619122 [PubMed: 21719772]
21. Saver JL, Gornbein J. Treatment effects for which shift or binary analyses are advantageous in acute stroke trials. Neurology. 2009;72:1310–1315. doi: 10.1212/01.wnl.0000341308.73506.b7 [PubMed: 19092107]
22. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predic-tors in multiple regression: a bad idea. Stat Med. 2006;25:127–141. doi: 10.1002/sim.2331 [PubMed: 16217841]
23. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ. 2006;332:1080. doi: 10.1136/bmj.332.7549.1080 [PubMed: 16675816]
24. Austin PC, Brunner LJ. Inflation of the type I error rate when a continuous confounding variable is categorized in logistic regression analyses. Stat Med. 2004;23:1159–1178. doi: 10.1002/sim.1687 [PubMed: 15057884]
25. Harrell FE, Slaughter JC. Biostatistics for Biomedical Research, Section 14.4, 2019. Accessed August 25, 2020. https://hbiostat.Org/doc/bbr.Pdf
26. Frank Harell P. How Should Change be Measured? Newer Material, Section 14.4, 2020. Accessed August 25, 2020. http://biostat.Mc.Vanderbilt.Edu/wiki/main/measurechange.
27. Agarwal S, Scher E, Lord A, Frontera J, Ishida K, Torres J, Rostanski S, Mistry E, Mac Grory B, Cutting S, et al. Redefined measure of early neurological improvement shows treatment benefit of alteplase over placebo. Stroke. 2020;51:1226–1230. doi: 10.1161/STROKEAHA.119.027476 [PubMed: 32102629]
28. Ozdemir O, Giray S, Arlier Z, Baş DF, Inanc Y, Colak E. Predictors of a good outcome after endovascular stroke treatment with stent retrievers. ScientificWorldJournal. 2015;2015:403726. doi: 10.1155/2015/403726 [PubMed: 26137591]
29. Yoon W, Kim SK, Park MS, Baek BH, Lee YY. Predictive factors for good outcome and mortality after stent-retriever thrombectomy in patients with acute anterior circulation stroke. J Stroke. 2017;19:97–103. doi: 10.5853/jos.2016.00675 [PubMed: 28178407]

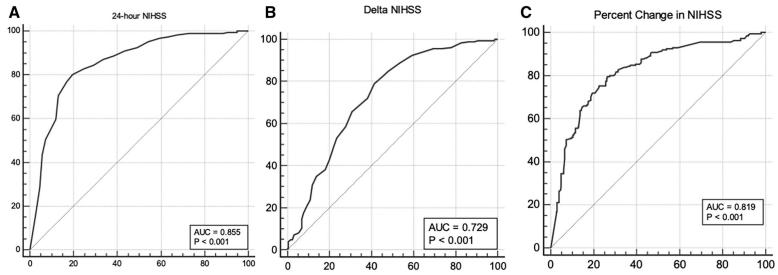

Figure 1. Receiver operating characteristics of National Institutes of Health Stroke Scale (NIHSS)-based outcome measures at predicting 90-d modified Rankin Score (mRS) 0–2.

A, 24-h NIHSS, (**B**) NIHSS (baseline mines 24-h NIHSS), and (**C**) percent change in NIHSS (NIHSS/baseline NIHSS×100). The sensitivity (*y* axis) and 100-specificity (*x* axis) and the area under the curve (AUC) for 24-h NIHSS, ΔNIHSS, and percent change in NIHSS at differentiating 90-d mRS score 0–2 vs 3–6 is shown.

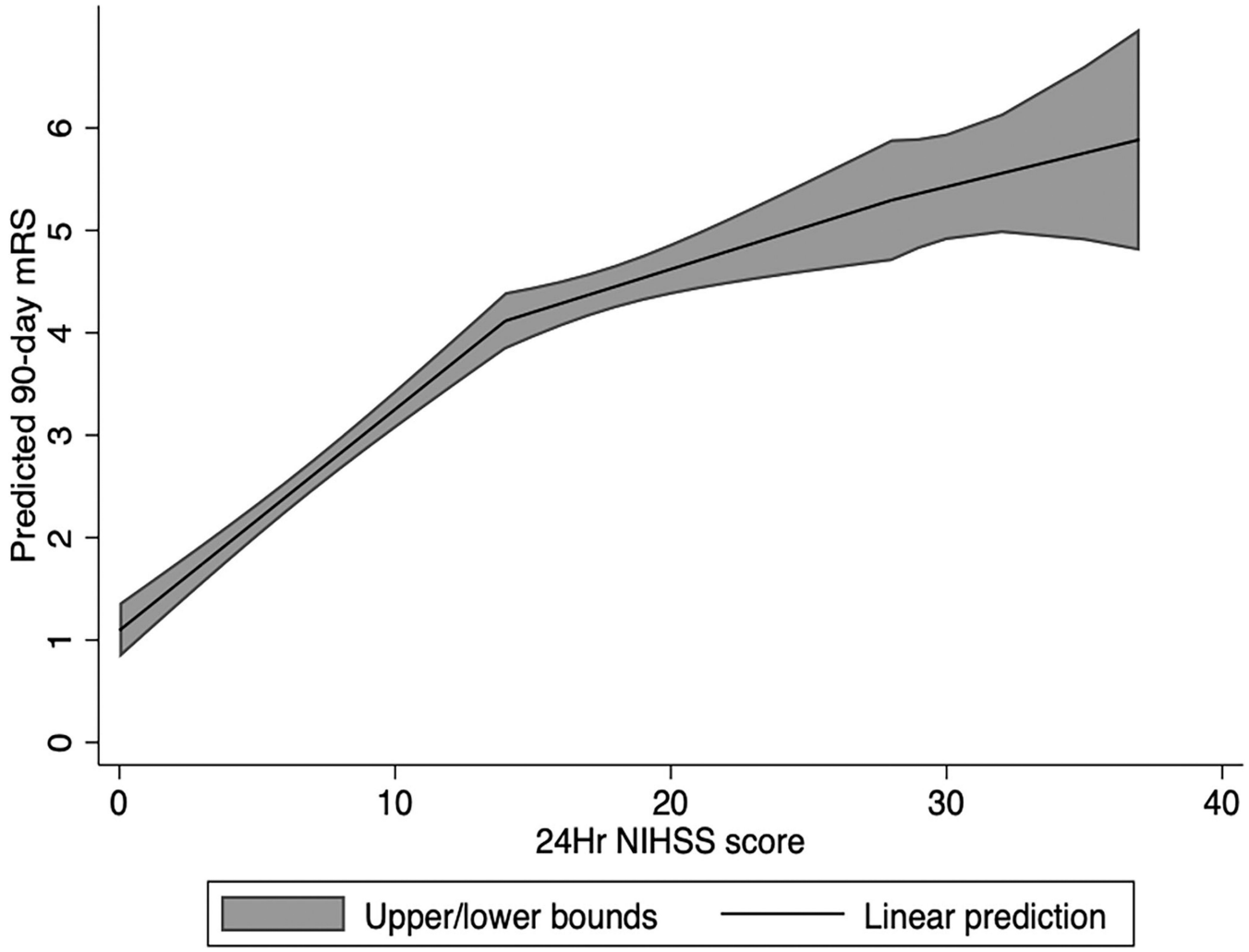

Figure 2. Nonlinear prediction of 90-d modified Rankin Scale score based on the 24-h National Institutes of Health Stroke Scale (NIHSS) score.

In this linear prediction model, 24-h NIHSS is treated as a nonlinear variable (restricted cubic spline with 3 knots). Upper and lower bounds of this prediction are shown with the gray area plot.

Table 1.

Baseline Characteristics

Age (mean±SD)	68 (±15)
Females, n (%)	228 (51)
Hypertension, n (%)	332 (74)
Diabetes, n (%)	125 (28)
Atrial fibrillation, n (%)	161 (36)
Glucose, median [IQR]	124 [105–150]
Baseline NIHSS, median [IQR]	16 [11–20]
ASPECTS, median [IQR]	8 [7–9]
Time to recanalization, median [IQR], min	293 [190–563]
Successful recanalization, n (%)	395 (89)
Intravenous thrombolysis, n (%)	217 (49)
24-h NIHSS, median [IQR]	9 [3, 16]
Intracerebral hemorrhage, n (%)	
Asymptomatic	92 (21)
None	336 (75)
Symptomatic	18 (4)
24-h NIHSS score ≤7	200 (45)
ΔNIHSS score ≥4 (%)	255 (57)

ΔNIHSS indicates baseline minus 24-h NIHSS; ASPECTS, Alberta Stroke Program Early CT Score; IQR, interquartile range; and NIHSS, National Institutes of Health Stroke Scale.

Table 2.

Association of Various Thresholds of 24-Hour NIHSS and ΔNIHSS With 90-Day Modified Rankin Scale Score 0–2

Definitions	Unadjusted odds ratio	*P* value	Adjusted odds ratio	*P* value	Sensitivity	Specificity
24-h NIHSS score ≤7	16.67 [10–25]	<0.001	12.5 [7.14–20] *	<0.001	80.1%	80.4%
24-h NIHSS score ≤2	12.5 [7.14–25]	<0.001	7.69 [4.16–16.67] *	<0.001	43.55%	94.23%
ΔNIHSS score ≥4	5.27 [3.44–8.33]	<0.001	4.55 [2.85–7.69] †	<0.001	79%	58.5%
ΔNIHSS score ≥8	3.7 [2.5–5.56]	<0.001	3.33 [2.12–5.26] †	<0.001	53.2%	76.5%
24-h NIHSS as continuous	0.79 [0.75–0.82] ‡	<0.001	0.79 [0.75–0.84] *	<0.001	NA	NA
ΔNIHSS as continuous	1.11 [1.07–1.13]	<0.001	1.11 [1.07–1.14] †	<0.001	NA	NA
Percent change in NIHSS as continuous	1.016 [1.01–1.02]	<0.001	1.016 [1.01–1.02] †	<0.001	NA	NA

ΔNIHSS indicates baseline minus 24-h NIHSS; ASPECTS, Alberta Stroke Program Early CT Score; NIHSS, National Institutes of Health Stroke Scale; and tPA, tissue-type plasminogen activator.

*Adjusted for baseline NIHSS, age, hypertension, baseline glucose, ASPECTS, time to recanalization, successful recanalization, and tPA administration.

†Adjusted for age, hypertension, ASPECTS, time to recanalization, successful recanalization, and tPA administration.

‡Adjusted for baseline NIHSS.