UCLA
UCLA Previously Published Works

Title

Development and validation of a simplified Stroke–Thrombolytic Predictive Instrument

Permalink

https://escholarship.org/uc/item/5qh7b15j

Journal

Neurology, 85(11)

ISSN

0028-3878

Authors

Kent, David M
Ruthazer, Robin
Decker, Carole
et al.

Publication Date

2015-09-15

DOI

10.1212/wnl.0000000000001925

Peer reviewed

Development and validation of a simplified Stroke–Thrombolytic Predictive Instrument

David M. Kent, MD, MSc
Robin Ruthazer, MPH
Carole Decker, RN, PhD
Philip G. Jones, MSc
Jeffrey L. Saver, MD
Erich Bluhmki, PhD
John A. Spertus, MD, MPH

Correspondence to
Dr. Kent:
dkent1@tuftsmedicalcenter.org

ABSTRACT

Objectives: The Stroke–Thrombolytic Predictive Instrument (Stroke-TPI) predicts the probability of good and bad outcomes with and without recombinant tissue plasminogen activator (rtPA). We sought to rebuild and externally validate a simpler Stroke-TPI to support implementation in routine clinical care.

Methods: Using the original derivation cohort of 1,983 patients from a combined database of randomized clinical trials (NINDS [National Institute of Neurological Disorders and Stroke] 1 and 2; ATLANTIS [Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke] A and B; and ECASS [European Cooperative Acute Stroke Study] II), we simplified the Stroke-TPI by reducing variables and interaction terms and by exploring simpler (3- and 8-item) stroke severity scores. External validation was performed in the ECASS III trial (n = 821).

Results: The following 6 variables were most predictive of good outcomes: age, systolic blood pressure, diabetes, stroke severity, symptom onset to treatment time, and rtPA therapy. Treatment effect modifiers included onset to treatment time and systolic blood pressure. For the models predicting a bad outcome (modified Rankin Scale [mRS] score ≥5), significant variables included age, stroke severity, and serum glucose. rtPA therapy did not change the risk of a poor outcome. Compared with models using the full NIH Stroke Scale, models using the 3-item severity score showed similar discrimination and excellent calibration. External validation on ECASS III showed similar performance (C statistics 0.75 [mRS score ≤1] and 0.80 [mRS score ≤2]).

Conclusion: A simpler model using a 3-item stroke severity score, instead of the 15-item NIH Stroke Scale, has similar prognostic value and may be easier to use in routine care. Future studies are needed to test whether it can improve process and clinical outcomes. *Neurology®* **2015;85:942–949**

GLOSSARY

AIS = acute ischemic stroke; **ATLANTIS** = Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke; **ECASS** = European Cooperative Acute Stroke Study; **IDI** = integrated discrimination improvement; **IST-3** = Third International Stroke Trial; **mRS** = modified Rankin Scale; **NIHSS** = NIH Stroke Scale; **NINDS** = National Institute of Neurological Disorders and Stroke; **OTT** = onset to treatment time; **PI** = prediction interval; **rtPA** = recombinant tissue plasminogen activator; **TPI** = Thrombolytic Predictive Instrument.

Thrombolytic therapy improves patients' functional outcomes after acute ischemic stroke (AIS),[1] and is a guideline-endorsed, Class 1A recommendation.[2,3] Despite this, more than 40% of AIS patients ideal for reperfusion therapy remain untreated.[4] Among the challenges in promptly treating patients is uncertainty regarding whether the risks of treatment outweigh the benefits in any given patient. While multiple prediction models have been developed to individualize prognosis,[5–11] these have not been routinely implemented in clinical care.

Several barriers have limited the use of predictive models in stroke care. For example, most prognostic models for AIS have been derived only on cohorts receiving thrombolytic therapy, and therefore are unable to provide estimates of treatment benefit. In addition, model complexity can pose a practical barrier, particularly when the model requires a detailed neurologic examination (e.g., the NIH Stroke Scale[12] [NIHSS]) in the emergency department.

Supplemental data at Neurology.org

The Stroke–Thrombolytic Predictive Instrument (Stroke-TPI) is the only model for predicting outcomes with and without rtPA developed using patient-level data from the first 5 randomized clinical trials using standard-dose rtPA.[6] It has been validated on an external cohort of patients receiving thrombolytic therapy,[13] but because it was derived on the totality of randomized data testing standard-dose rtPA available at the time, outcome predictions with and without thrombolysis could not be validated. Thus, we sought to develop a simpler model that could be more easily implemented in routine clinical care and to externally validate its predictions on more recent trial data.

METHODS **Conceptual overview for model development.** *Defining the outcomes.* The original Stroke-TPI predicted the probability of a normal/near-normal outcome (modified Rankin Scale [mRS] score = 0 or 1) and severe disability/death (mRS score = 5 or 6) both with and without thrombolytic therapy. Before this analysis, a qualitative research study was conducted with numerous patient, caregiver, and provider focus groups to evaluate these outcome measures for informing care decisions (unpublished, 2014). Based on these qualitative studies, we incorporated a new threshold at mRS score = 0–2 to capture the independent outcome, in addition to those thresholds previously included in the original Stroke-TPI.

Defining the optimal measure of stroke severity. The most common measure of stroke severity is the NIHSS, a 15-item neurologic assessment that reproducibly quantifies the extent of neurologic disability.[12] While excellent training programs exist,[14] the scale is often not used in routine clinical care; more than half (54.9%) of patients within the Get with the Guidelines–Stroke registry do not have an NIHSS score documented.[15] To simplify the implementation of the revised Stroke-TPI tool, we explored 2 simpler, previously developed stroke severity scores, an 8-item and a 3-item measure of stroke severity. The 8-item severity score[16] uses the following neurologic findings from the NIHSS: (1a) level of consciousness; (2) gaze; (3) visual fields; (4) facial paresis; (6a) motor–leg right; (6b) motor–leg left; (9) language; and (10) dysarthria. The 3-item score[17] uses disturbance of consciousness (none = 0, mild = 1, and severe = 2 points), gaze/head deviations (absent = 0, incomplete = 1, and forced gaze/head deviation = 2 points), and hemiparesis (absent = 0, moderate = 1, and severe = 2). These scores were mapped from individual items within the full NIHSS. We assumed that reduced-item severity scores validated in only the prehospital setting would be as reliable in the more controlled setting of the hospital, particularly since the items were drawn from the NIHSS, which is applied within this setting.

Minimizing the number of predictor variables in the model. For a model to be implemented in routine clinical care, it is important to minimize the required data elements. Model reduction was performed with a combination of clinical and statistical judgment. In addition, treatment interaction factors, the variables that modify the effects of therapy, were carefully evaluated. Treatment interactions are distinguished from prognostic factors, which affect the probability of outcome with or without

therapy, but not the proportional effects of therapy. In addition to longer onset to treatment time (OTT), which had been known to modify the effectiveness of therapy, the original Stroke-TPI project identified 3 other variables that appeared to modify the effect of therapy. Specifically, thrombolysis appeared less beneficial when initial systolic blood pressure was higher, the patient was male, and the patient had a prior stroke.

In this iteration of the Stroke-TPI, we assumed a more conservative approach to effect modification, eliminating those variables without strong external evidence.[18,19] We reviewed the literature to assess the consistency of observed interactions with existing evidence, including but not limited to the reported subgroup analyses from both the European Cooperative Acute Stroke Study (ECASS) III[20] and the Third International Stroke Trial (IST-3) reports.[21] Based on these comparisons, as well as statistical and clinical considerations, we included the 2 interaction terms evaluated as being most reliable (OTT and systolic blood pressure) and eliminated 2 others (sex and prior stroke). In addition, while the prior Stroke-TPI included models both with and without the ASPECT (Alberta Stroke Program Early CT) score, this simplified version did not include the variable, given the difficulty of obtaining this information in most settings and the fact that it does not appear to influence treatment effect for IV thrombolysis.[22,23]

Data source: The Stroke-TPI database. We used a combined, patient-level dataset from 5 clinical trials (National Institute of Neurological Disorders and Stroke [NINDS] 1 and 2, Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke [ATLANTIS] A[24] and B,[25] and ECASS II[26]) that had previously been used to develop the original Stroke-TPI. This dataset included common clinical characteristics and outcomes, including 90-day mRS scores. The only exception was that item-by-item detail for the NIHSS was not available for the NINDS trials, requiring multiple imputation in this dataset to estimate the scores for the reduced stroke severity scales. Of note, the range of OTT varied substantially across these randomized trials, with patients being treated between 58 and 360 minutes after symptom onset, enabling a broad range of treatment times to be modeled.

Validation dataset: ECASS III. ECASS III was a double-blind, parallel-group trial that enrolled 821 patients from 130 sites in 19 European countries. Criteria for study entry were similar to other studies in the derivation cohort of the Stroke-TPI, with some notable exceptions. In particular, because thrombolytic therapy was standard of care at the time of the trial when administered before 3 hours from symptom onset, patients in ECASS III were enrolled only if able to receive study drug in the window between 3 and 4.5 hours after the onset of symptoms. In addition, the combination of both prior stroke and diabetes was an exclusion from enrollment.

Statistical analysis. Unadjusted comparisons of the patients and outcomes across the trials were performed with t tests for continuous data and χ^2 for categorical data. As described above, 3 binary outcomes were modeled: 90-day mRS score ≤1 (normal/near-normal outcome), 90-day mRS score ≤2 (independent outcome), and 90-day mRS score ≥5 (severe disability/death outcome). Multivariable logistic regression models were used to predict each outcome. Variable selection relied on the previously published Stroke-TPI,[6] with no additional data exploration, incorporating the changes specified above: use of an alternative independent outcome threshold (mRS score ≤2); removal of 2 treatment interaction terms (tPA × sex and tPA × prior stroke); and testing reduced-item severity scores (including 3-item and

Table 1 Patient characteristics

Characteristic	All (N = 1,983)	ATLANTIS (n = 614)	ECASS (n = 778)	NINDS (n = 591)
Randomized to treatment, % (ratio)	49.9 (989/1,983)	49.3 (303/614)	50.9 (396/778)	49.1 (290/591)
Patient characteristics				
Age, y, mean ± SD (n)	66.0 ± 0.2 (1,983)	65.7 ± 0.2 (614)	65.4 ± 0.2 (778)	66.8 ± 0.2 (591)
Sex, male, % (ratio)	54.4 (1,078/1,983)	60.3 (370/614)	58.5 (455/778)	42.8 (253/591)
History of diabetes, % (ratio)	21.1 (418/1,983)	21.0 (129/614)	21.1 (164/778)	21.2 (125/591)
Prior stroke, % (ratio)	16.4 (326/1,983)	15.1 (93/614)	20.1 (156/778)	13.0 (77/591)
Baseline SBP,[a] mm Hg, mean ± SD (n)	152.5 ± 68.0 (1,983)	152.5 ± 68.0 (614)	152.3 ± 68.0 (778)	152.8 ± 68.0 (591)
Glucose, mmol/L, median [q1, q3] (n)[a]	7.0 [6.0, 9.0] (1,983)	7.0 [6.0, 8.0] (614)	7.0 [6.0, 8.0] (778)	7.0 [6.0, 9.0] (591)
Minutes from stroke to treatment, median [q1, q3] (n)	232 [155, 288] (1,983)	273 [240, 293] (614)	265 [205, 320] (778)	110 [89, 157] (591)
Stroke severity scales (at baseline)				
NIHSS, median [q1, q3] (n)	12 [8, 17] (1,983)	10 [7, 15] (614)	12 [8, 16] (778)	15 [9, 20] (591)
8-Variable scale, median [q1, q3] (n)	6 [4, 8] (1,392)	6 [4, 9] (614)	6 [4, 8] (778)	—[b]
3-Variable scale, median [q1, q3] (n)	2 [1, 3] (1,392)	2 [0, 3] (614)	2 [1, 3] (778)	—[b]
Outcomes, % (ratio)				
Functionally normal/near-normal: mRS score 0 or 1	37.9 (751/1,983)	40.4 (248/614)	38.4 (299/778)	34.5 (204/591)
Functionally independent: mRS score 0, 1, or 2	49.6 (984/1,983)	54.4 (334/614)	50.0 (389/778)	44.2 (261/591)
Severe disability/death: mRS score 5 or 6	18.5 (366/1,983)	16.0 (98/614)	14.9 (116/778)	25.7 (152/591)

Abbreviations: ATLANTIS = Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke; ECASS = European Cooperative Acute Stroke Study; mRS = modified Rankin Scale; NIHSS = NIH Stroke Scale; NINDS = National Institute of Neurological Disorders and Stroke; SBP = systolic blood pressure.
q1–q3 = interquartile range spanning 25th to 75th percentile.
[a] SBP truncated at 110 and 200 mm Hg and glucose truncated at 25 mmol/L.
[b] Reduced stroke severity scales not available in NINDS.

8-item neurologic assessments) against the full 15-item NIHSS on a subset of the data (ATLANTIS A and B and ECASS II). After these changes were made, we removed Stroke-TPI variables that were no longer significantly prognostic. A total of 9 separate models were developed using logistic regression, one for each of 3 outcomes using each of 3 stroke severity scales. The performance of each model was assessed by the C statistic and calibration plots, and was compared between the full and reduced models. Individual patient predictions with the different models were examined for clinical plausibility and compared across models. Model-to-model comparisons were made using Spearman rank correlation and integrated discrimination improvement (IDI)[27] indices.

For the 3-item severity score model, bootstrap sampling was used both to estimate 95% prediction intervals and to internally validate the model (i.e., correct for overoptimism). We drew 500 bootstrapped samples across the multiple imputed datasets of 1,983 patients. To compute 95% prediction intervals, new parameter estimates for the model were obtained, and a new C statistic was estimated, for each of the 500 bootstrapped samples. Optimism-corrected C statistics were calculated by applying each of the 500 models to the original sample. All analyses were conducted using SAS version 9.3 TS Level 1M2 in XP_PRO platform (SAS Institute Inc., Cary, NC); IDI was calculated using a program[28] run on R Studio (version 0.98.507, 2009–2013; RStudio, Inc., Boston, MA) via R software (version 3.1.0).[29]

External validation. Using fixed coefficients, we generated predictions on the ECASS III population for both the normal/near-normal outcome (mRS score 0 or 1) and the independent outcome (mRS score 0–2). Because baseline serum glucose was not available for ECASS III patients, a key patient characteristic associated with a bad outcome, predictions for the severe disability/death outcome (mRS score 5 or 6) could not be externally validated. Thus, a total of 6 models were evaluated (2 outcomes by 3 severity scores). For each model, performance on the validation cohort was assessed using the C statistic and by examining calibration plots, and calculating the mean (absolute) bias across quintiles. While model derivation was performed in Boston, model validation was performed independently in Boehringer, Germany.

Net benefit. To compare the potential benefits of using individualized predictions against a treatment recommendation based only on time window, we examined the distribution of the absolute benefit (predicted probability of a good outcome with rtPA *minus* predicted probability of a good outcome without rtPA) in the derivation dataset across each 30-minute OTT period. This analysis describes the proportion of patients expected to achieve benefit of varying magnitudes with treatment at a given time (i.e., to be harmed by treatment, to have an absolute risk reduction of 0% to 5% [i.e., number needed to treat >20] and an absolute risk reduction of >5% [number needed to treat <20]).

RESULTS There were 1,983 patients in the Stroke-TPI dataset (table 1). Given similar performance of the 3 stroke severity measures, table 2 describes the models predicting normal/near-normal outcomes (mRS score ≤1) and independent outcomes (mRS score ≤2) with the 3-item stroke severity measure. When the prior stroke and sex treatment interaction

Table 2 Models using 3-item stroke severity score

Measure	Model for functionally normal/near-normal outcome (mRS score ≤1) OR	p Value	Model for functionally independent outcome (mRS score ≤2) OR	p Value	Model for severe disability/death outcome (mRS score ≥5) OR	p Value
Intercept (exponentiated)	4.8158	—	48.4775	—	0.0013	—
Thrombolytic therapy (yes/no)	25.9936	0.0002	5.7047	0.0468	—	—
Age (y)	0.9996	0.9656	0.9793	0.0517	1.0482	<0.0001
Systolic blood pressure (mm Hg)	0.9968	0.3936	0.9963	0.3256	—	—
History of diabetes (yes/no)	0.4993	<0.0001	0.4499	<0.0001	—	—
Stroke severity score	0.8924	0.6576	0.6291	0.114	1.7485	<0.0001
Time to treatment (min)	0.9990	0.4401	0.9985	0.2594	—	—
Thrombolytic therapy × systolic blood pressure	0.9874	0.016	0.9948	0.3241	—	—
Thrombolytic therapy × time to treatment	0.9956	0.0008	0.9972	0.0292	—	—
Age × stroke severity score	0.9910	0.023	0.9954	0.2968	—	—
Serum glucose[a] (mmol/L)	—	—	—	—	1.0766	<0.0001

Abbreviations: mRS = modified Rankin Scale; OR = odds ratio.
[a] Truncated at 25 mmol/L.

terms were removed from the models, the main effects of these variables were no longer significant and they were also removed. Thus, only the following 6 variables were included in the models for normal/near-normal outcomes and independent outcomes: age, systolic blood pressure, history of diabetes, stroke severity, symptom onset to treatment time, and thrombolytic therapy. Models using both the full NIHSS and the 8-item severity scale are shown

in tables e-1 and e-2 on the *Neurology*® Web site at Neurology.org. Table 2 shows the models predicting severe disability/death outcome (mRS score ≥5). The 3 variables included in the model were the same as the original Stroke-TPI: age, stroke severity, and serum glucose. Thrombolytic therapy did not increase or decrease the risk of a poor outcome, indicating that the benefits of reperfusion and the harms of intracerebral hemorrhage are approximately balanced for this outcome.
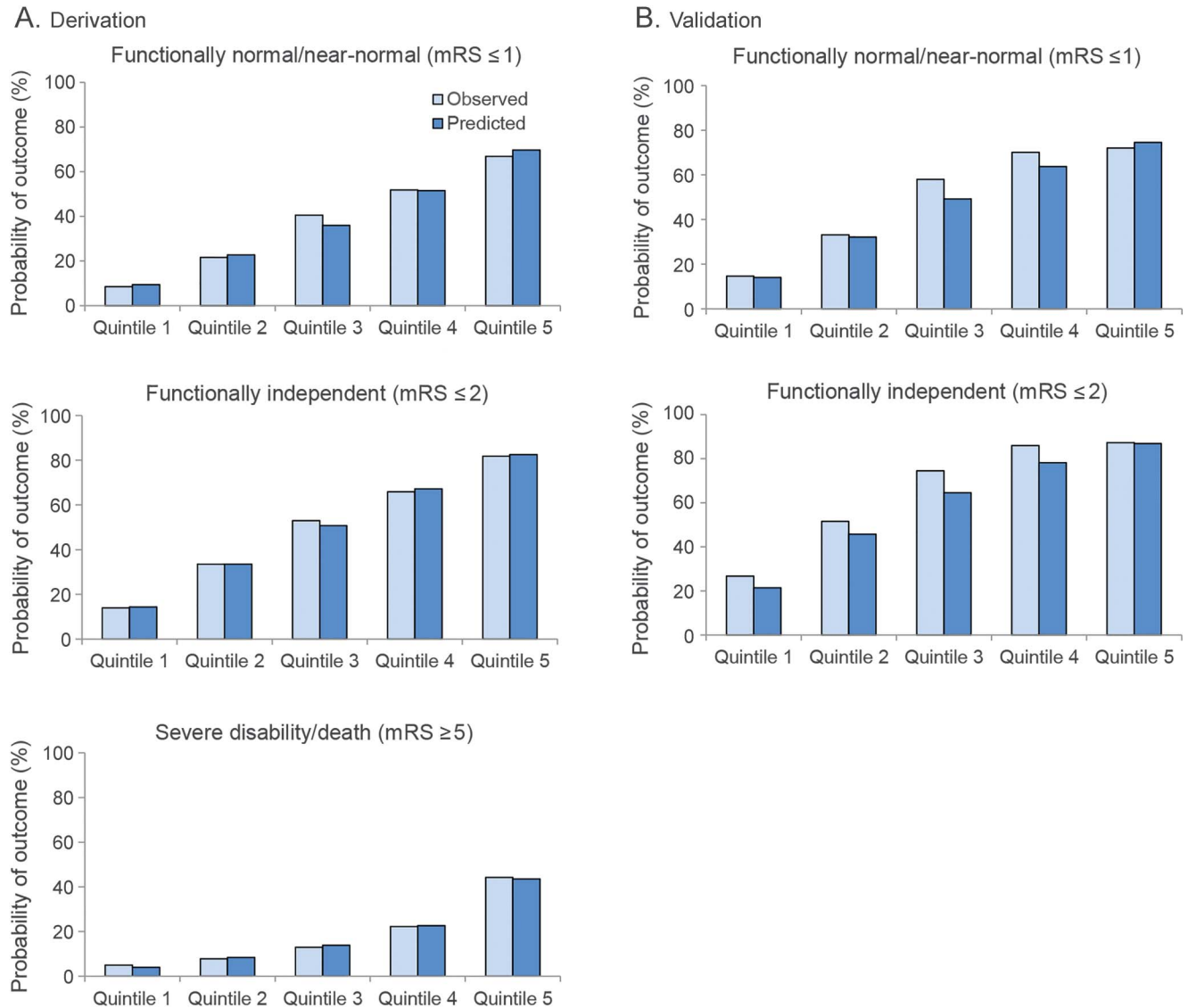
Model correlations, reclassification, and internal validation. We compared the predictions of individual patients across models using the 3 different severity scores. These predictions were found to be highly correlated. The Spearman rank correlations between the models using the full NIHSS and the 3-item stroke severity score were 0.80, 0.80, and 0.84 (all with $p < 0.0001$) for predicting mRS scores ≤1, ≤2, and ≥5, respectively. The IDI[27] indices of the models using the full NIHSS compared to that using the 3-item scale were 0.032 (−0.010 to 0.073), 0.013 (−0.025 to 0.051), and 0.026 (−0.019 to 0.071) for predicting mRS scores ≤1, ≤2, and ≥5, respectively, suggesting that the full NIHSS did not significantly improve the classification of patients with good outcomes.

The C statistics for all 9 models are shown in table 3, and all models had similar discrimination (C statistics between 0.75 and 0.80). On bootstrap validation, the models using the 3-item score had the following optimism-corrected C statistics: mRS score ≤1, 0.76 (95% prediction interval [PI]: 0.75–0.79); mRS score ≤2, 0.78 (95% PI: 0.77–0.80); and mRS

Table 3 C statistics for all 9 models

Model	C statistic (derivation pooled N = 1,983)	C statistic (validation: ECASS III)
Model for functionally normal/near-normal outcome (mRS score ≤1) using full NIHSS	0.784	0.760
Models for functionally independent outcome (mRS score ≤2) using full NIHSS	0.792	0.802
Model for severe disability/death outcome (mRS score ≥5) using full NIHSS	0.770	—
Model for functionally normal/near-normal outcome (mRS score <1) using 8-item score	0.756	0.752
Model for functionally independent outcome (mRS score ≤2) using 8-item score	0.771	0.794
Model for severe disability/death outcome (mRS score ≥5) using 8-item score	0.754	—
Model for functionally normal/near-normal outcome (mRS score <1) using 3-item score	0.764	0.749
Model for functionally independent outcome (mRS score ≤2) using 3-item score	0.781	0.790
Model for severe disability/death outcome (mRS score ≥5) using 3-item score	0.759	—

Abbreviations: ECASS = European Cooperative Acute Stroke Study; mRS = modified Rankin Scale; NIHSS = NIH Stroke Scale.

Figure 1   Calibration plots for 3-item severity score model

A. Derivation

Functionally normal/near-normal (mRS ≤1)

Functionally independent (mRS ≤2)

Severe disability/death (mRS ≥5)

B. Validation

Functionally normal/near-normal (mRS ≤1)

Functionally independent (mRS ≤2)

These plots show the proportion of patients with each outcome as both predicted by the model and as observed, for equal-sized quintiles in the derivation population (A) and validation population (B). Calibration is not shown for the severe disability/death outcome for the validation population, since ECASS III did not collect values for serum glucose at baseline. ECASS = European Cooperative Acute Stroke Study; mRS = modified Rankin Scale.

score ≥5, 0.76 (95% PI: 0.73–0.79). When calibration curves were examined across quintiles, bias of the model on the internal database was minimal, and did not appear to substantially depend on which stroke severity score was used. Figure 1A shows the calibration curve on the derivation dataset for the most parsimonious model using the 3-item severity score.

External validation. Discrimination and calibration were evaluated in ECASS III. All 6 models evaluated showed near identical discrimination as compared with that seen in the derivation dataset (table e-3). Calibration was excellent for the model predicting normal or near-normal outcome (mRS score ≤1) when the 15-item score was used. However, observed outcomes in ECASS III were slightly better than predicted for all other models. Chi-square goodness-of-fit tests across the quintiles showed no significant differences between observed or expected for any of the models using the full NIHSS or the 3-item score, although they were significant for the 8-item score (table e-3). Figure 1B shows the calibration curve on the validation dataset for the most parsimonious model using the 3-item severity score.

Net benefit. The distribution of predicted benefit over time for the 3-item severity model are shown in figure 2 for both the model predicting normal/near-normal outcomes (mRS score ≤1) and the model predicting independent outcomes (mRS score ≤2). As can be seen, there are a small number of patients that are

Functionally normal/near-normal (mRS ≤1): Estimated treatment benefit stratified by OTT

■ <0% benefit   ■ 0% - 5% benefit   □ >5% benefit

Functionally independent (mRS ≤2): Estimated treatment benefit stratified by OTT
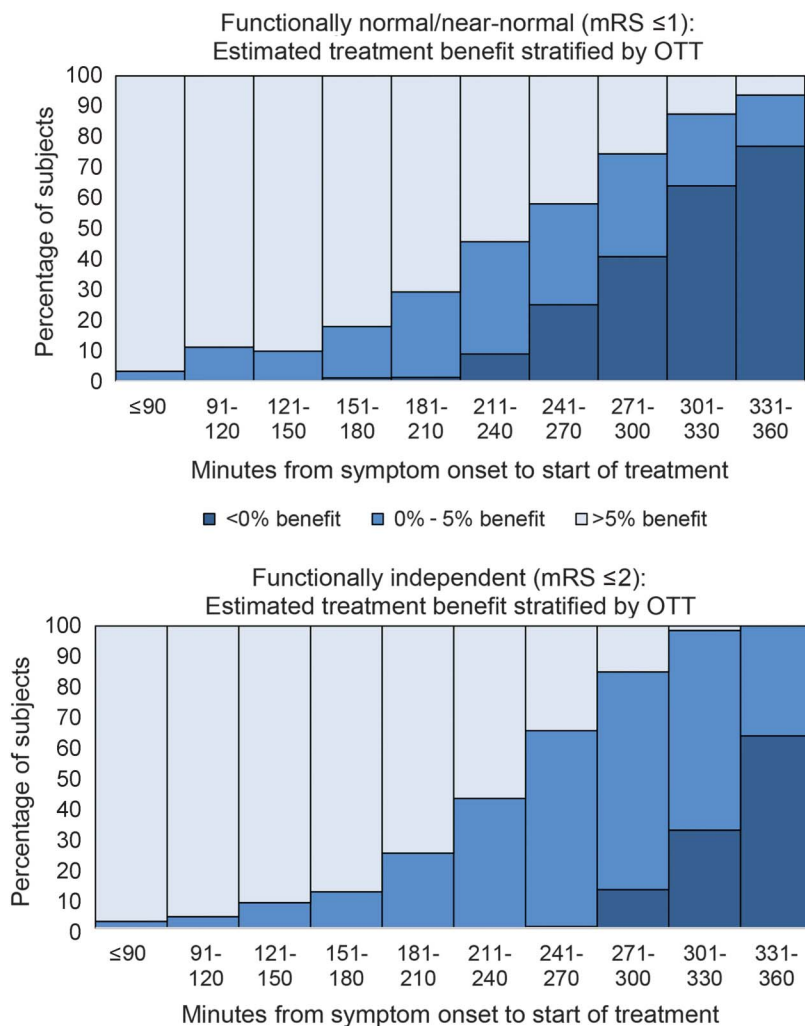
These figures illustrate the change in individual predicted net benefit (probability of a good outcome with *minus* without recombinant tissue plasminogen activator) in patients in the derivation database (N = 1,983) across different treatment onset to symptom times. As can be seen, there is variation in predicted net benefit at all time windows, with predicted harm in some patients treated before the 4.5-hour window expires and predicted benefit in some patients treated after this window. mRS = modified Rankin Scale; OTT = onset to treatment time.

predicted to be harmed (i.e., have treatment unfavorable characteristics), even with an OTT of less than 4.5 hours. Conversely, there are some patients predicted to have substantial benefit beyond this time window.

DISCUSSION To support the use of personalized estimates of the benefits of thrombolytic therapy in the treatment of AIS, we simplified a well-established prediction model, the Stroke-TPI. We found that removing several interaction terms and predictor variables, and using a reduced-item stroke severity measure, did not substantially degrade the discriminatory performance of the Stroke-TPI model. (For comparison, the C statistics in the

original Stroke-TPI were 0.788 and 0.775 for the models predicting mRS scores ≤1 and ≥5, respectively.) Furthermore, this study represents the first external validation of the Stroke-TPI on a large randomized dataset. These steps should increase confidence and ease of use of this modified Stroke-TPI, particularly in settings where the full NIHSS is not routinely assessed.

Multiple stroke predictive models have been developed to provide prognostic information, including the Stroke-TPI, iScore, ASTRAL, DRAGON, and others.[5–11] These models share several important features. In particular, prognosis in all models is largely predicted by stroke severity, age, and time to treatment. Additional variables provide further improvement in prognostic performance, yet with diminishing returns. While previous models include the full NIHSS, a simpler score, such as the 3-item version we use, may enable a broader range of clinicians to estimate treatment benefits and support broader adoption of personalized, evidence-based treatment. Also, unlike the other prognostic models, the Stroke-TPI was developed on a combined database of randomized clinical trials and has now been validated in an additional large randomized trial. This is important because the estimates of treatment benefit are unbiased due to the randomization process used in each trial.

A critical element needed to differentiate patients likely or unlikely to benefit is reliably identifying factors that modify the effects of tPA. Four factors have been shown with randomized data to produce statistically significant treatment modification.[6] The following characteristics have been shown to be associated with less benefit: longer symptom onset to treatment time, male sex, higher systolic blood pressure, and prior stroke. While the original Stroke-TPI included all these variables in treatment interaction terms, it should be recognized that inclusion of such terms has many of the same pitfalls as classic subgroup analysis. As yet, there is little rigorous guidance to help inform when such interaction terms are likely to be credible, although both clinical criteria (based on prior external evidence) and statistical criteria should be used.[30–32] In this iteration of the Stroke-TPI, we removed previously included interaction effects, pending further confirmation.

Our study should be interpreted in the context of several potential limitations. This database does not incorporate all randomized data currently available. After the development of the Stroke-TPI, several additional clinical trials were completed, most notably ECASS III,[20,33] which enrolled patients between 3 and 4.5 hours only, and IST-3,[21] which included patients in the 0- to 6-hour window. ECASS III was used as a validation dataset, but we did not have

access to IST-3 for these analyses. The Stroke-TPI is intended as a prognostic tool that may be of use for shared decision-making and informed consent. It is not intended as a diagnostic tool or to determine eligibility criteria for consideration of tPA. Selection of patients for lytic therapy should be based on the presence of a potentially disabling neurologic deficit evident on the general neurologic examination, and not on any particular score on the full NIHSS or the shorter scale used in this study.

As with all prediction models, prognostic outputs represent "evidential probabilities" conditional on the databases used and the mathematical assumptions of the model. Other models may yield different predictions at the individual patient level.[34] While the use of randomized controlled trials minimizes bias in estimating the benefits of rtPA, it is unknown whether similar results would be obtained in the broader population of patients presenting with AIS ineligible for the clinical trials. In addition, our validation dataset included patients treated only in the 3- to 4.5-hour time window. Finally, the clinical impact of this model on treatment rates and outcomes has not been tested.

Nevertheless, because estimates of absolute benefit vary substantially across the population, even among patients treated at the same symptom onset to treatment time, this tool can support decision-making. By having estimates of treatment benefit, clinicians can not only better clarify the benefits of treatment in their own minds, but can also share these estimates with patients as they jointly decide whether or not to treat. The next step is to test the implementation of this new risk model in clinical care to examine its effect on treatment rates, time to treatment, and, most importantly, patient outcomes.

REFERENCES

1. Tissue plasminogen activator for acute ischemic stroke. The National Institute of Neurological Disorders and Stroke rt-PA Stroke Study Group. N Engl J Med 1995; 333:1581–1587.
2. Jauch EC, Saver JL, Adams HP, et al. Guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2013;44:870–947.
3. American College of Emergency Physicians; American Academy of Neurology. Clinical policy: use of intravenous tPA for the management of acute ischemic stroke in the emergency department. Ann Emerg Med 2013;61: 225–243.
4. Fonarow GC, Reeves MJ, Smith EE, et al. Characteristics, performance measures, and in-hospital outcomes of the first one million stroke and transient ischemic attack admissions in Get with the Guidelines–Stroke. Circ Cardiovasc Qual Outcomes 2010;3:291–302.
5. Strbian D, Meretoja A, Ahlhelm FJ, et al. Predicting outcome of IV thrombolysis-treated ischemic stroke patients: the DRAGON score. Neurology 2012;78:427–432.
6. Kent DM, Selker HP, Ruthazer R, Bluhmki E, Hacke W. The Stroke-Thrombolytic Predictive Instrument: a predictive instrument for intravenous thrombolysis in acute ischemic stroke. Stroke 2006;37:2957–2962.
7. Konig IR, Ziegler A, Bluhmki E, et al. Predicting long-term outcome after acute ischemic stroke: a simple index works in patients from controlled clinical trials. Stroke 2008;39:1821–1826.
8. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional

outcome in acute ischemic stroke: the ASTRAL score. Neurology 2012;78:1916–1922.

9. Saposnik G, Kapral MK, Liu Y, et al. IScore: a risk score to predict death early after hospitalization for an acute ischemic stroke. Circulation 2011;123:739–749.

10. Weimar C, Ziegler A, Konig IR, Diener HC. Predicting functional outcome and survival after acute ischemic stroke. J Neurol 2002;249:888–895.

11. Saposnik G, Guzik AK, Reeves M, Ovbiagele B, Johnston SC. Stroke prognostication using age and NIH Stroke Scale: SPAN-100. Neurology 2013;80:21–28.

12. Brott T, Adams HP, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. Stroke 1989;20:864–870.

13. Uyttenboogaart M, Stewart RE, Vroomen PC, Luijckx GJ, De KJ. Utility of the Stroke-Thrombolytic Predictive Instrument. J Neurol Neurosurg Psychiatry 2008;79:1079–1081.

14. National Institute of Neurological Disorders and Stroke (NINDS). NIH Stroke Scale (NIHSS): English Program, Globally Distributed by Health Care Point. 2006. Available at: http://www.nihstrokescale.org. Accessed March 31, 2015.

15. Fonarow GC, Pan W, Saver JL, et al. Comparison of 30-day mortality models for profiling hospital performance in acute ischemic stroke with vs without adjustment for stroke severity. JAMA 2012;308:257–264.

16. Tirschwell DL, Longstreth WT Jr, Becker KJ, et al. Shortening the NIH Stroke Scale for use in the prehospital setting. Stroke 2002;33:2801–2806.

17. Singer OC, Dvorak F, du Mesnil de Rochemont R, Lanfermann H, Sitzer M, Neumann-Haefelin T. A simple 3-item stroke scale: comparison with the National Institutes of Health Stroke Scale and prediction of middle cerebral artery occlusion. Stroke 2005;36:773–776.

18. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2002;21:2917–2930.

19. Pocock SJ, Lubsen J. More on subgroup analyses in clinical trials. N Engl J Med 2008;358:2076–2077.

20. Hacke W, Kaste M, Bluhmki E, et al. Thrombolysis with alteplase 3 to 4.5 hours after acute ischemic stroke. N Engl J Med 2008;359:1317–1329.

21. Sandercock P, Wardlaw JM, Lindley RI, 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.

22. Dzialowski I, Hill MD, Coutts SB, et al. Extent of early ischemic changes on computed tomography (CT) before thrombolysis: prognostic value of the Alberta Stroke Program Early CT Score in ECASS II. Stroke 2006;37:973–978.

23. Kent DM, Hill MD, Ruthazer R, et al. "Clinical-CT mismatch" and the response to systemic thrombolytic therapy in acute ischemic stroke. Stroke 2005;36:1695–1699.

24. Clark WM, Albers GW, Madden KP, Hamilton S. The rtPA (alteplase) 0- to 6-hour acute stroke trial, part A (A0276g): results of a double-blind, placebo-controlled, multicenter study. Thrombolytic Therapy in Acute Ischemic Stroke Study Investigators. Stroke 2000;31:811–816.

25. Clark WM, Wissman S, Albers GW, Jhamandas JH, Madden KP, Hamilton S. Recombinant tissue-type plasminogen activator (alteplase) for ischemic stroke 3 to 5 hours after symptom onset. The ATLANTIS Study: a randomized controlled trial. Alteplase Thrombolysis for Acute Noninterventional Therapy in Ischemic Stroke. JAMA 1999;282:2019–2026.

26. Hacke W, Kaste M, Fieschi C, et al. Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators. Lancet 1998;352:1245–1251.

27. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008;27:157–172.

28. Steyerberg EW. Program for testicular cancer analyses: evaluate performance. 2009. Available at: http://download.lww.com/wolterskluwer_vitalstream_com/PermaLink/EDE/A/EDE_2009_10_02_STEYERBERG_200726_SDC1.pdf. Accessed March 31, 2015.

29. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2014. Available at: http://www.R-project.org. Accessed March 31, 2015.

30. Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. JAMA 2014;311:405–411.

31. Rothwell PM, Mehta Z, Howard SC, Gutnikov SA, Warlow CP. Treating individuals 3: from subgroups to individuals: general principles and the example of carotid endarterectomy. Lancet 2005;365:256–265.

32. van Klaveren D, Vergouwe Y, Farooq V, Serruys PW, Steyerberg EW. Estimates of absolute treatment benefit for individual patients required careful modelling of statistical interactions. J Clin Epidemiol Epub 2015 Feb 27.

33. Bluhmki E, Chamorro A, Davalos A, et al. Stroke treatment with alteplase given 3.0–4.5 h after onset of acute ischaemic stroke (ECASS III): additional outcomes and subgroup analysis of a randomised controlled trial. Lancet Neurol 2009;8:1095–1102.

34. Kent DM, Shah ND. Risk models and patient-centered evidence: should physicians expect one right answer? JAMA 2012;307:1585–1586.