Validating and comparing stroke prognosis scales

Terence J. Quinn, MD*
Sarjit Singh, BSc
    (Med Sci)*
Kennedy R. Lees, MD
Philip M. Bath, PhD
Phyo K. Myint, MD
On behalf of the VISTA
    Collaborators

Correspondence to
Dr. Quinn:
Terry.Quinn@glasgow.ac.uk

ABSTRACT

Objective: To compare the prognostic accuracy of various acute stroke prognostic scales using a large, independent, clinical trials dataset.

Methods: We directly compared 8 stroke prognostic scales, chosen based on focused literature review (Acute Stroke Registry and Analysis of Lausanne [ASTRAL]; iSCORE; iSCORE-revised; pre-admission comorbidities, level of consciousness, age, and neurologic deficit [PLAN]; stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale [mRS] [SOAR]; modified SOAR; Stroke Prognosis Instrument 2 [SPI2]; and Totaled Health Risks in Vascular Events [THRIVE]) using individual patient-level data from a clinical trials archive (Virtual International Stroke Trials Archive [VISTA]). We calculated area under receiver operating characteristic curves (AUROC) for each scale against 90-day outcomes of mRS (dichotomized at mRS >2), Barthel Index (>85), and mortality. We performed 2 complementary analyses: the first limited to patients with complete data for all components of all scales (simultaneous) and the second using as many patients as possible for each individual scale (separate). We compared AUROCs and performed sensitivity analyses substituting extreme outcome values for missing data.

Results: In total, 10,777 patients contributed to the analyses. Our simultaneous analyses suggested that ASTRAL had greatest prognostic accuracy for mRS, AUROC 0.78 (95% confidence interval [CI] 0.75–0.82), and SPI2 had poorest AUROC, 0.61 (95% CI 0.57–0.66). Our separate analyses confirmed these results: ASTRAL AUROC 0.79 (95% CI 0.78–0.80 and SPI2 AUROC 0.60 (95% CI 0.59–0.61). On formal comparative testing, there was a significant difference in modified Rankin Scale AUROC between ASTRAL and all other scales. Sensitivity analysis identified no evidence of systematic bias from missing data.

Conclusions: Our comparative analyses confirm differences in the prognostic accuracy of stroke scales. However, even the best performing scale had prognostic accuracy that may not be sufficient as a basis for clinical decision-making. *Neurology*® **2017;89:997–1002**

GLOSSARY

ASTRAL = Acute Stroke Registry and Analysis of Lausanne; **AUROC** = area under the receiver operating characteristic curve; **BI** = Barthel Index; **iSCORE-r** = revised iSCORE; **mRS** = modified Rankin Scale score; **NIHSS** = NIH Stroke Scale; **OCSP** = Oxfordshire Community Stroke Project; **PLAN** = preadmission comorbidities, level of consciousness, age, and neurologic deficit; **ROC** = receiver operating characteristic; **SOAR** = stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; **SOARm** = modified stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; **SPI2** = Stroke Prognosis Instrument 2; **THRIVE** = Totaled Health Risks in Vascular Events; **VISTA** = Virtual International Stroke Trials Archive.

Various tools exist to assist clinicians to quantify risk of good or poor outcome. For example, the ABCD2 score uses clinical features to predict risk of stroke following TIA.[1] Although there are criticisms of ABCD2, it is widely used and included in stroke guidelines.[2] Numerous scales purport to predict acute stroke outcomes from baseline features.[3–5] In general, these prognosis scales have had limited traction and have not been incorporated into routine clinical practice or guidelines.[3]
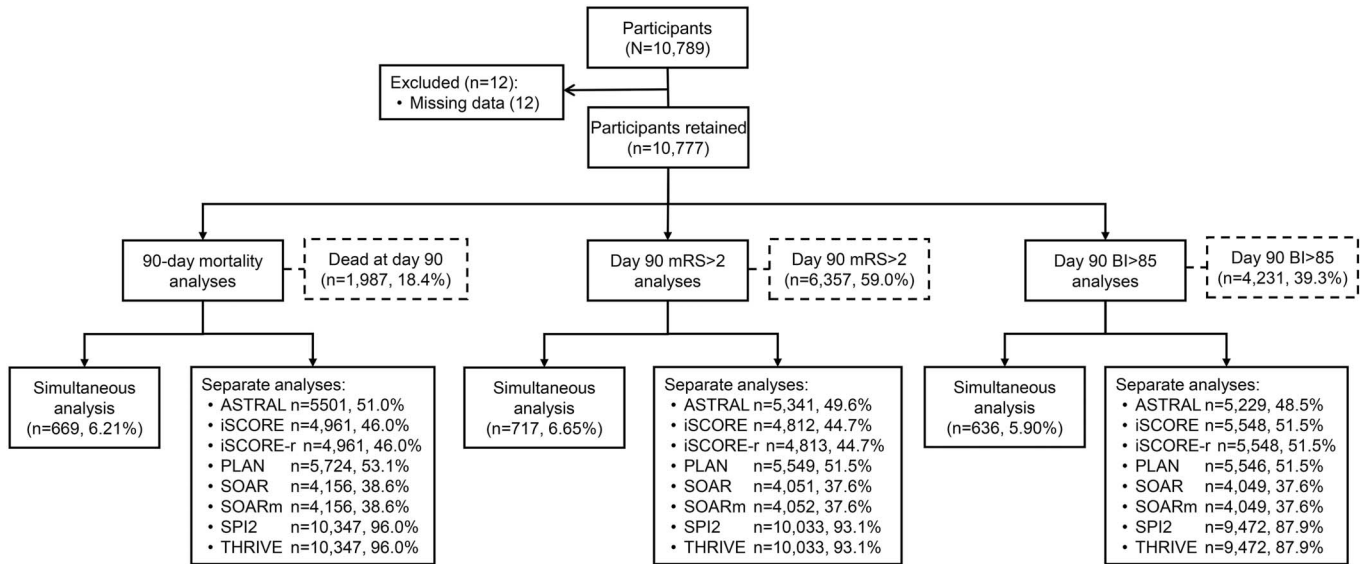
**Supplemental data
at Neurology.org**

Figure 1   Patient flow diagram

Flow diagram shows numbers and percentage of patients with each outcome (dashed boxes) and number and percentage of patients used in each analysis (complete boxes). ASTRAL = Acute Stroke Registry and Analysis of Lausanne; BI = Barthel Index; iSCORE-r = revised iSCORE; mRS = modified Rankin Scale score; PLAN = preadmission comorbidities, level of consciousness, age, and neurologic deficit; SOAR = stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SOARm = modified stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SPI2 = Stroke Prognosis Instrument 2; THRIVE = Totaled Health Risks in Vascular Events.

Table 1   Baseline characteristics of the Virtual International Stroke Trials Archive (VISTA) dataset

| Characteristic	Overall	Simultaneous analyses
Age, y	69.5 ± 12.5	70.9 ± 12.1
Sex, male (%)	5,909 (54.8)	327 (52.2)
Baseline NIHSS	13.4 ± 5.6	14.72 ± 4.3
Time to rtPA, min	65.1 ± 9.1	80.62 ± 82.9
Baseline glucose, mmol/L	7.6 ± 3.1	7.5 ± 2.6
Prestroke mRS		
0 or 1	9,131 (84.7)	623 (99.5)
2 to 5	6 (0.0)	3 (0.5)
Not recorded	1,640 (15.2)	
Stroke type		
Hemorrhagic	571 (5.3)	626 (100.0)
Ischemic	4,621 (42.9)	0 (0.0)
Not recorded	5,585 (51.8)	
OCSP		
TACS	5,049 (46.8)	379 (60.5)
PACS	3,458 (32.1)	235 (37.5)
LACS	1,699 (15.8)	8 (1.3)
POCS	571 (5.3)	4 (0.6)

Abbreviations: mRS = modified Rankin Scale score; NIHSS = NIH Stroke Scale; OCSP = Oxfordshire Community Stroke Project; rtPA = recombinant tissue plasminogen activator. Baseline characteristics of the included VISTA data given as mean ± SD or number (%). For reference, in the complete VISTA acute dataset mean age is 71; proportion male is 54.9%; median baseline NIHSS is 11. LACS = lacunar anterior circulation syndrome; PACS = partial anterior circulation syndrome; POCS = posterior circulation syndrome; TACS = total anterior circulation syndrome.

There are various plausible reasons why these scales have not been adopted by the stroke community. Scales may be perceived as too complex to use or may require information that is not routinely available in the acute setting (for example, sophisticated neuroimaging).[3] Often these scales report mortality; however, given the disabling nature of stroke, scales predicting longer-term disability may be more useful.[6]

Before adopting a scale for clinical use, clinicians and policy makers will want to see robust evidence of external validity, i.e., evidence that the scale performs well in cohorts independent of the population used to derive the scale. Attempts to validate scales are emerging but such data remain lacking for many stroke prognosis scales.[3,5] Where more than one scale is available, clinicians will want to compare and contrast the properties of the various scales in a common dataset. Again, such analyses are infrequent and we have to rely on indirect comparison across heterogeneous studies.

We aimed to validate and compare the accuracy of stroke prognosis scales for predicting death and disability in an independent dataset using patient-level data.

Figure 2   Receiver operating characteristic (ROC) curves generated for modified Rankin Scale (mRS) (simultaneous analysis)
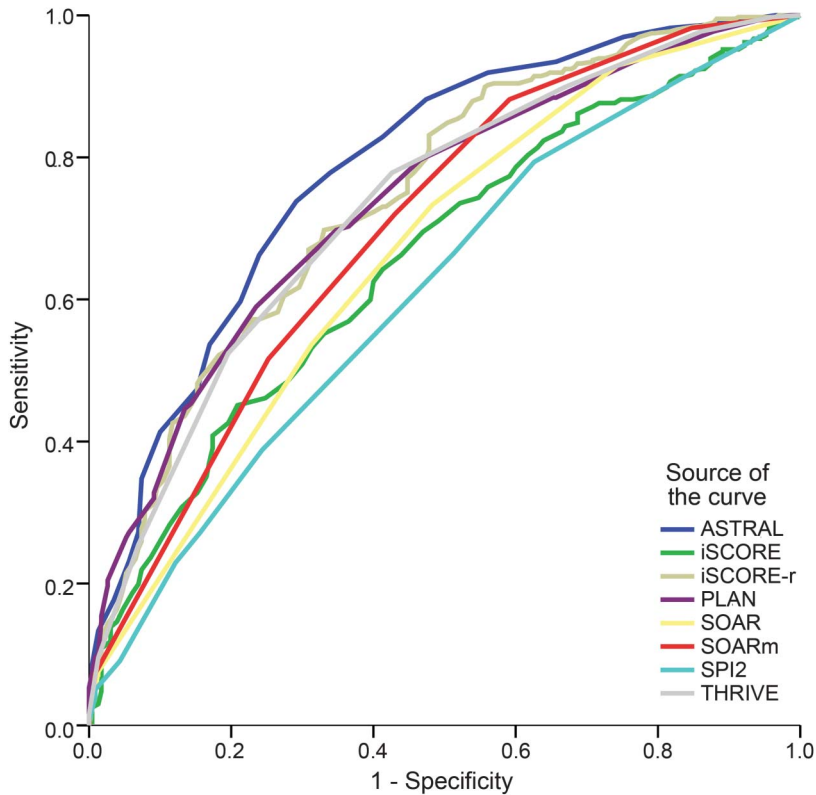
ROC curve generated for 90-day mRS >2. ASTRAL = Acute Stroke Registry and Analysis of Lausanne; iSCORE-r = revised iSCORE; PLAN = preadmission comorbidities, level of consciousness, age, and neurologic deficit; SOAR = stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SOARm = modified stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SPI2 = Stroke Prognosis Instrument 2; THRIVE = Totaled Health Risks in Vascular Events.

METHODS Data sources. We used the acute section of the Virtual International Stroke Trials Archive (VISTA; vista.gla.ac. uk). VISTA is a not-for-profit, anonymized data repository containing individual patient-level data from completed stroke trials. All included trials were performed under appropriate institutional review board and regulatory approvals, and only fully anonymized data are held by VISTA.[7]

Choice of prognostic scales. We selected multi-item scales (more than 2 items) based on focused literature review. The search was not a systematic review; we identified scales according to the following criteria: intended for use in acute stroke; comprises baseline clinical (not neuroimaging) and demographic data; predicts outcome at 30 days (or more) poststroke; published since 2000; published in a leading topic specific journal (table e-1 at Neurology.org).

From the chosen scales, we created a list of variables needed to derive each and requested these data from VISTA. VISTA data are anonymized and it is a condition of use that shared VISTA data are not mapped back to identify the source trial. To ensure our VISTA dataset was truly independent of the datasets used to originally derive the scores, we collated the primary article describing each of our included scales and shared these with the VISTA data management team, who confirmed that the VISTA data were independent (table e-2).

Outcomes. Our primary outcome was 90-day modified Rankin Scale score (mRS). We dichotomized 90-day mRS with poor outcome defined as mRS >2. Secondary outcomes were 90-day all-cause mortality and 90-day Barthel Index (BI), dichotomized with poor outcome defined as BI ≤85.

Data management. We worked with the dataset to maximize available data. Where mortality data were not directly described, we calculated 90-day mortality from day of death or 90-day mRS (taking mRS 6 as death). Where Oxford stroke classification was not available, we derived this based on NIH Stroke Scale (NIHSS) using a validated algorithm.[8] Where Canadian Neurological Scale score data were not reported, we calculated this from NIHSS data. Certain scales included specific clinical features and we derived the presence of these (visual field defects, loss of consciousness, limb paresis, neglect, and aphasia) from NIHSS. Where scales used comorbid disease, we interrogated data on comorbidity for conditions of relevance using both categorized and free text searches.

Statistical analyses. Our primary analyses of interest were based around the discriminative ability of the scales. We generated receiver operating characteristic (ROC) curves for each scale. We calculated the area under the receiver operating characteristic curves (AUROC) with 95% confidence intervals. We created 2 sets of ROC curves. The first included analysis of all scales for those patients who had sufficient data, i.e., we only included patients who had data sufficient to derive a score for all of our chosen scales (simultaneous analyses). For the second analysis, each scale was assessed individually using as many patients as possible (separate analyses). The simultaneous analysis was designed to facilitate valid comparisons; the separate analysis was designed to give a precise assessment of prognostic accuracy using a dataset independent from the derivation dataset.

In addition to ROC analyses, we described individual values for sensitivity and specificity limiting to values corresponding to point of optimal performance on ROC curve for our primary outcome of 90-day mRS. To describe calibration, we performed Hosmer-Lemeshow tests of each scale's performance against primary and secondary outcomes.
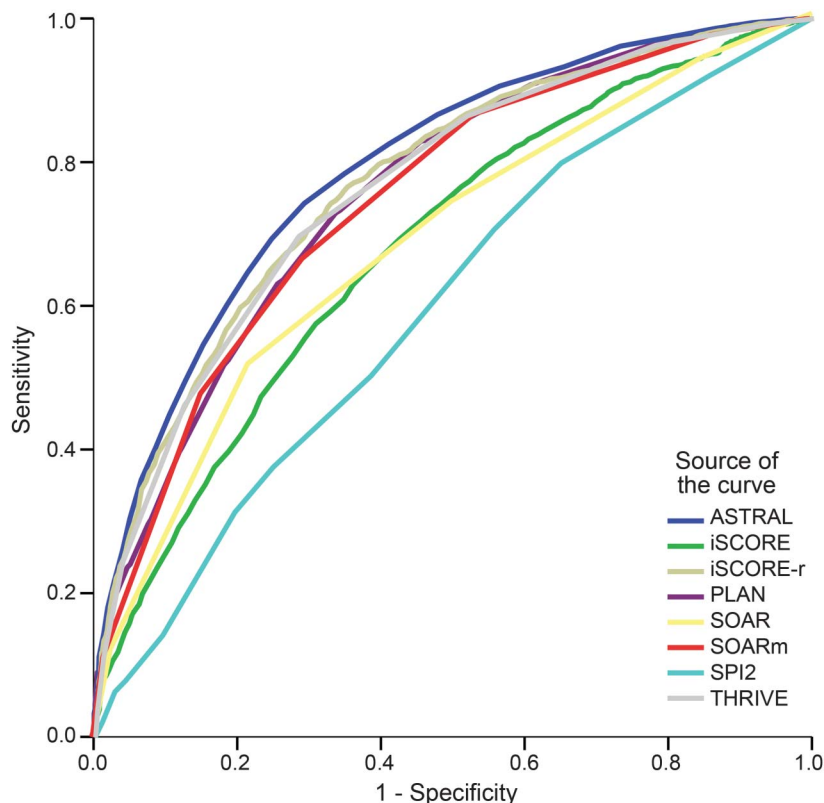
Certain variables of interest had missing data at the individual patient level. We performed sensitivity analyses to explore the effect of missing data by replacing missing data with extreme high and low values for each scale.

We performed all discrimination and calibration analyses using SPSS version 22.0 (IBM Corporation, Armonk, NY).

We formally tested the difference in AUROC between scales, accounting for multiplicity of analyses[9] (MedCalc version 16.0, MedCalc Software, Ostend, Belgium), comparing the 6 scales with the highest AUROC using the results of our simultaneous and separate analyses.

RESULTS Patients. We had access to individual patient-level data for 10,789 patients. Twelve patients were excluded as they lacked any relevant data. Figure 1 describes included patients and their outcomes. Table 1 contains baseline characteristics of included patients. We derived Oxfordshire Community Stroke Project (OCSP) data for 7,533 (69.9%) patients based on their baseline NIHSS data. All Canadian Neurological Scale scores were derived from baseline NIHSS.

Figure 3 Receiver operating characteristic (ROC) curves generated for modified Rankin Scale (mRS) (separate analyses)

ROC curve generated for 90-day mRS >2. ASTRAL = Acute Stroke Registry and Analysis of Lausanne; iSCORE-r = revised iSCORE; PLAN = preadmission comorbidities, level of consciousness, age, and neurologic deficit; SOAR = stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SOARm = modified stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SPI2 = Stroke Prognosis Instrument 2; THRIVE = Totaled Health Risks in Vascular Events.

Included scales. We assessed 8 prognostic scales: Acute Stroke Registry and Analysis of Lausanne (ASTRAL); iSCORE; iSCORE-r (a revised iSCORE); preadmission comorbidities, level of consciousness, age, and neurologic deficit (PLAN); stroke subtype, OCSP, age, and prestroke mRS (SOAR); SOARm (a modified SOAR); SPI2 (Stroke Prognosis Instrument 2); and THRIVE (Totaled Health Risks in Vascular Events).[10–17]

Validation. For our primary outcome, prediction of 90-day mRS, ASTRAL had the largest and SPI2 the smallest AUROC in all analyses (figures 2 and 3). This pattern of results was similar in our separate analyses (table 2). Sensitivity analysis where missing data were imputed with a best and worst case scenario for 90-day mRS (i.e., mRS 0 and mRS 6) yielded differing AUROC for 2 of our included scales (iSCORE-r and SOARm). In each case, the AUROC was less than in the original analysis (table e-3).

For prediction of 90-day mortality, ASTRAL had the largest and SPI2 the smallest AUROC in both analyses (figure e-1). Sensitivity analysis yielded differing AUROC for 2 scales (iSCORE-r and SOARm). In each case, the AUROC was less than in the original analyses (tables e-4 and e-5).

For prediction of 90-day BI, ASTRAL had the largest and SPI2 the smallest AUROC in both analyses (figure e-2). Sensitivity analysis for 90-day BI yielded differing AUROC for 4 scales (both iSCORE and both SOAR variants). In each case, the AUROC was less than in the original analyses (tables e-6 and e-7).

On formal significance tested, we observed differences in the AUROC of included scales (table e-8). ASTRAL had significantly higher AUROC than all other scales for 90-day mRS <2 and significantly higher AUROC for BI than most other scales. Differing tests had differing patterns of sensitivity vs specificity at point of optimal performance (table e-9).

Our analysis of calibration suggested that all scales other than THRIVE had good calibration for 90-day mRS. Scales of iSCORE, iSCORE-r, PLAN, and SPI2 demonstrated good calibration for each of our chosen outcomes (table e-10).

DISCUSSION Validation and quantitative comparison can assist in choice of prognostic scale, yet there have been few articles offering such analyses in stroke.[3–5] We provide these data for recognized acute stroke prognostic scales, with ASTRAL showing superior prognostic utility and SPI2 the lowest prognostic accuracy.

Our data highlight the value of independent, external validation of prognostic scales. The original ASTRAL study reported an AUROC of 0.85, with subsequent AUROC of 0.90 in a validation study by the ASTRAL team.[10] Our data suggest a lower prognostic accuracy (AUROC 0.82 and 0.81); indeed, most of our analyses gave prognostic values that are less than described in the original derivation articles. This pattern of independent analyses reporting less favorable scale properties has been described in other settings.[18] Although best practice statements for prognostic research highlight the importance of external validation, it remains true that most prognosis research focuses on model development with relatively limited further validation.[18] We have also shown the value of comparative analyses using a common dataset. Indirect comparisons using results from derivation studies may have given the erroneous impression that all scales perform equally well, potentially because of differences in the sampling frames in the original studies. Our formal testing suggests that scales have differing discriminative properties.

Strengths of our approach include access to a large dataset of trial quality data from a variety of sources. We explored the possibility of missing data introducing bias through our sensitivity analyses. We offer both comparative and individual scale level analyses

Table 2 Area under receiver operating characteristic curve (AUROC) for primary analysis (modified Rankin Scale [mRS])

	Separate analyses		Simultaneous analysis		Derivation study
	AUROC (95% CI)	No. (%)	AUROC (95% CI)	No. (%)	AUROC (95% CI)
ASTRAL	0.79 (0.78–0.80)	5344 (49.6)	0.78 (0.75–0.82)	627 (5.8)	0.85
iSCORE	0.68 (0.67–0.70)	5551 (51.5)	0.66 (0.61–0.70)		0.82 (0.81–0.83)
iSCORE-r	0.77 (0.76–0.78)	5551 (51.5)	0.74 (0.70–0.78)		0.80
PLAN	0.76 (0.75–0.77)	5549 (51.5)	0.74 (0.70–0.78)		0.84 (0.82–0.85)[a]
SOAR	0.68 (0.66–0.70)	4051 (37.6)	0.66 (0.62–0.71)		0.82 (0.79–0.84)[b]
SOARm	0.75 (0.73–0.76)	4051 (37.6)	0.70 (0.66–0.74)		0.83 (0.79–0.86)[c]
SPI2	0.60 (0.59–0.61)	10033 (93.1)	0.61 (0.57–0.66)		0.63 (0.62–0.65)[d]
THRIVE	0.77 (0.76–0.78)	10033 (93.1)	0.73 (0.69–0.77)		0.71

Abbreviations: ASTRAL = Acute Stroke Registry and Analysis of Lausanne; CI = confidence interval; iSCORE-r = revised iSCORE; PLAN = preadmission comorbidities, level of consciousness, age, and neurologic deficit; SOAR = stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SOARm = modified stroke subtype, Oxfordshire Community Stroke Project, age, and prestroke modified Rankin Scale; SPI2 = Stroke Prognosis Instrument 2; THRIVE = Totaled Health Risks in Vascular Events.
AUROC (95% CI) and n (%) of patients in our analyses for 90-day mRS >2.
[a] Primary outcomes for derivation studies: 1-year mortality.
[b] Primary outcomes for derivation studies: 7-day mortality.
[c] Primary outcomes for derivation studies: discharge mortality.
[d] Primary outcomes for derivation studies: stroke or death within 2 years.
International Stroke Trials Archive.

and describe ROC and individual sensitivity, specificity, and calibration values. We hope these data will help clinicians choose whether to use a prognosis scale and which is most appropriate.

There were limitations to our analyses. We maximized data available for analyses by deriving certain input variables; we recognize this may introduce bias. To facilitate analyses, we dichotomized mRS and BI data. Our selected thresholds were in keeping with those commonly used in trials and prognostic work.[19] However, we appreciate that this approach fails to capture the range of functional outcomes that can follow stroke. We limited our analysis to those multi-item scales, described in selected journals that use variables potentially available in VISTA. Thus, we not have included certain prognostic scales.[5,20] A comprehensive, systematic review could be a future direction for this work. For any analysis that uses trial data there is always a concern regarding generalizability to a real-world population. Patients included in our analysis were younger and less impaired than may be seen in an unselected cohort. We would encourage similar analyses using large-scale observational cohort data.

Our data would support the use of the ASTRAL scale, as this scale had the greatest predictive accuracy and already has published external validation studies. However, choice of prognostic scale, or indeed the decision to use these scales at all, is dependent on many factors in addition to prognostic accuracy. Ease of use at the bedside is an important consideration,

particularly in the hyperacute stroke setting. The scale with the fewest included variables was SOAR and this showed moderate predictive accuracy.

We deliberately chose not to have a threshold value that signified "acceptable" accuracy. The value of a prognostic or diagnostic instrument is relative to the context of the assessment and absolute values are unhelpful in determining the utility of a scale in practice.[18,21] Scales can serve an assistive or decisive function.[22,23] The appropriateness of basing clinical decisions solely or partly on prognostic scales is a factor of both the accuracy of the scale and the cost of a misclassification. In the acute setting, if the scales are being used to inform decisions regarding life-saving treatment, it is debatable whether accuracy values perceived as statistically acceptable are also clinically acceptable. For example, with a specificity of around 0.8, 1 in 5 may be wrongly assigned a poor prognosis and treated accordingly.

Potential solutions to improving scale properties include adding the most discriminating factors from other scales to create a new scale or using scales in series. Our analysis of individual sensitivity and specificity scores suggest that some scales favor sensitivity and vice versa. Thus there may be utility in using, for example, a highly sensitive scale followed by a highly specific scale. Such approaches may have value, but will add complexity to assessment and ultimately would represent new scales that require further validation. It may also be useful for future research to focus on a better understanding of the characteristics of

those patients who are most likely to be misclassified by scales. Our work highlights recommendations by groups such as MRC PROGRESS[18] that development of future scales include comparison against preexisting scales and inclusion of systems to allow for independent, external validation of the scale.

There are many scales designed to provide prognostic information following acute stroke. Our comparative analyses suggest that certain scales (e.g., ASTRAL) may have better prognostic accuracy than others, although this accuracy was substantially lower than described in the original derivation articles. Based on our data, even the scales with the best properties may not have sufficient prognostic power to be used as the sole basis for important treatment decisions.

DISCLOSURE

The authors report no disclosures relevant to the manuscript. Go to Neurology.org for full disclosures.

REFERENCES

1. Johnston KC, Wagner DP, Wang XQ, et al. Validation of an acute ischemic stroke model: does diffusion-weighted imaging lesion volume offer a clinically significant improvement in prediction of outcome? Stroke 2007;38:1820–1825.
2. National Institute for Health and Clinical Excellence (NICE). Stroke and Transient Ischaemic Attack in Over 16s: Diagnosis and Initial Management [online]. Nice Guideline. 2008. Available at: nice.org.uk/guidance/CG68. Accessed May 3, 2017.
3. Counsell C, Dennis M. Systematic review of prognostic models in patients with acute stroke. Cerebrovasc Dis 2001;12:159–170.
4. Mattishent K, Kwok CS, Mahtani A, Pelpola K, Myint PK, Loke YK. Prognostic indices for early mortality in ischaemic stroke: meta-analysis. Acta Neurol Scand 2016;133:41–48.
5. Teale EA, Forster A, Munyombwe T, Young JB. A systematic review of case-mix adjustment models for stroke. Clin Rehabil 2012;26:771–786.
6. Solomon NA, Glick HA, Russo CJ, Lee J, Schulman KA. Patient preferences for stroke outcomes. Stroke 1994;25:1721–1725.
7. Ali M, Bath PMW, Curram J, et al. The virtual international stroke trials archive. Stroke 2007;38:1905–1910.
8. Ali M. The Virtual International Stroke Trials Archive (VISTA): Promulgation of a Clinical Trial Resource [online]. University of Glasgow; 2008. Available at: theses.gla.ac.uk/509/1/2008AliPhD_.pdf. Accessed May 3, 2017.
9. 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.
10. 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.
11. Saposnik G, Fang J, Kapral MK, et al. The iScore predicts effectiveness of thrombolytic therapy for acute ischemic stroke. Stroke 2012;43:1315–1322.
12. Sung SF, Chen YW, Hung LC, Lin HJ. Revised iScore to predict outcomes after acute ischemic stroke. J Stroke Cerebrovasc Dis 2014;23:1634–1639.
13. O'Donnell MJ, Fang J, D'Uva C, et al. The PLAN score: a bedside prediction rule for death and severe disability following acute ischemic stroke. Arch Intern Med 2012;172:1548–1556.
14. 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.
15. Abdul-Rahim AH, Quinn TJ, Alder S, et al. Derivation and validation of a novel prognostic scale (modified-stroke subtype, Oxfordshire Community Stroke Project classification, age, and prestroke modified Rankin) to predict early mortality in acute stroke. Stroke 2016;47:74–79.
16. Kernan WN, Viscoli CM, Brass LM, et al. The Stroke Prognosis Instrument II (SPI-II): a clinical prediction instrument for patients with transient ischemia and nondisabling ischemic stroke. Stroke 2000;31:456–462.
17. Flint AC, Faigeles BS, Cullen SP, et al. THRIVE score predicts ischemic stroke outcomes and thrombolytic hemorrhage risk in VISTA. Stroke 2013;44:3365–3369.
18. Steyerberg E, Moons KGM, van der Windt D, et al. Prognosis research strategy (PROGRESS) series 3: prognostic models. Br Med J 2012;10:e1001381.
19. Quinn TJ, Dawson J, Walters MR, Lees KR. Functional outcome measures in contemporary stroke trials. Int J Stroke 2009;4:200–205.
20. 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.
21. Noel-Storr AH, McCleery JM, Richard E, et al. Reporting standards for studies of diagnostic test accuracy in dementia: the STARDdem Initiative. Neurology 2014;83:364–373.
22. Liao L, Mark DB. Clinical prediction models: are we building better mousetraps? J Am Coll Cardiol 2003;42:851–853.
23. Laupacis A, Sekar NClinical prediction rules. JAMA 1997;277:488–494.