BMJ Neurology Open

Large vessel occlusion prediction scales provide high negative but low positive predictive values in prehospital suspected stroke patients

Kevin J Keenan ![ORCID],[1,2,3] Wade S Smith,[2,3] Sara B Cole,[2] Christine Martin,[2] J Claude Hemphill,[2,3] Debbie Y Madhok[2,3,4]

[1]Department of Neurology, University of California Davis, Sacramento, California, USA
[2]Department of Neurology, Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco, California, USA
[3]Department of Neurology, University of California San Francisco, San Francisco, California, USA
[4]Department of Emergency Medicine, University of California San Francisco, San Francisco, California, USA

Correspondence to
Dr. Kevin J Keenan;
kjkeenan@ucdavis.edu

ABSTRACT

Introduction We studied a registry of Emergency Medical Systems (EMS) identified prehospital suspected stroke patients brought to an academic endovascular capable hospital over 1 year to assess the prevalence of disease and externally validate large vessel occlusion (LVO) stroke prediction scales with a focus on predictive values.

Methods All patients had last known well times within 6 hours and a positive prehospital Cincinnati Prehospital Stroke Scale. LVO prediction scale scores were retrospectively calculated from emergency department arrival National Institutes of Health Stroke Scale scores. Final diagnoses were determined by chart review. Prevalence and diagnostic performance statistics were calculated. We prespecified analyses to identify scale thresholds with positive predictive values (PPVs) ≥80% and negative predictive values (NPVs) ≥95%. A secondary analysis identified thresholds with PPVs ≥50%.

Results Of 220 EMS transported patients, 13.6% had LVO stroke, 15.9% had intracranial haemorrhage, 20.5% had non-LVO stroke and 50% had stroke mimic diagnoses. LVO stroke prevalence was 15.8% among the 184 diagnostic performance study eligible patients. Only Field Assessment Stroke Triage for Emergency Destination (FAST-ED) ≥7 had a PPV ≥80%, but this threshold missed 83% of LVO strokes. FAST-ED ≥6, Prehospital Acute Severity Scale =3 and Rapid Arterial oCclusion Evaluation ≥7 had PPVs ≥50% but sensitivities were <50%. Several standard and lower alternative scale thresholds achieved NPVs ≥95%, but false positives were common.

Conclusions Diagnostic performance tradeoffs of LVO prediction scales limited their ability to achieve high PPVs without missing most LVO strokes. Multiple scales provided high NPV thresholds, but these were associated with many false positives.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The prevalence of large vessel occlusion (LVO) stroke among prehospital suspected stroke patients is not well characterised and affects the positive and negative predictive values (PPVs and NPVs) of the LVO predictions scales used for prehospital triage to endovascular stroke centres.

WHAT THIS STUDY ADDS

⇒ The prevalence of LVO stroke among Emergency Medical Services transported suspected stroke patients was low at 13.6%. Using higher than standard LVO prediction scale thresholds or including intracranial haemorrhages as true positives enabled some scales to reach PPVs ≥50% but false negatives were common. NPVs ≥95% were achieved by several standard and lower than standard score thresholds but false positives were common.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Future prehospital triage studies will need to account for the low prevalence of LVO stroke and currently available LVO predictions scales are more capable of providing high NPVs than PPVs.

treatment delays and worse outcomes associated with initial transportation to a non-EVT centre followed by an interfacility transfer to an EVT centre.[2–5] Most scales assign points based on the absence, presence or gradations of severity of examination findings and provide either a binary positive or negative result or a total score. LVO prediction scales that provide total scores each have a standard scale score threshold, often suggested by the researchers who created the scale, that is typically used to declare it positive or negative. Importantly, alternative lower score thresholds can be used as more sensitive but less specific LVO stroke prediction tests. Conversely, higher score thresholds can serve as less sensitive but more specific tests.

INTRODUCTION

Endovascular therapy (EVT) for large vessel occlusion (LVO) stroke is highly effective and time-dependent.[1] LVO prediction scales use neurological examination findings to identify LVO stroke patients in the prehospital setting and can facilitate their triage directly to EVT centres. This is intended to avoid the EVT

The positive predictive values (PPVs) and negative predictive values (NPVs) of LVO prediction scales are of particular interest to Emergency Medical Service (EMS) providers because they allow for the interpretation of an individual patient's post-test probability of having an LVO stroke. This is in contrast with sensitivity and specificity, which provide evaluations of a test's performance among all of those known to have LVO stroke (sensitivity) or all those known to not have LVO stroke (specificity).[6 7] All four parameters are important, but predictive values are highly dependent on the prevalence of the target disease in the sample of patients tested.[7 8] If the prevalence of LVO stroke in a sample is not similar to the actual prevalence among prehospital suspected stroke patients where the test will be used, the predictive values could be dramatically inaccurate.[7 9]

We set out to retrospectively characterise the prevalence of LVO stroke and other final diagnoses among patients in a prehospital suspected stroke registry. Recent prehospital LVO prediction scale validation studies have focused on standard thresholds and reported PPVs under 40%, suggesting that most patients identified for triage to EVT centres using standard score thresholds will not have LVO stroke.[10 11] We assessed the diagnostic performance of both standard and alternative (higher or lower) LVO prediction scale thresholds, with an emphasis on identifying thresholds that met high PPV or NPV goals, through a retrospective study of emergency department (ED) arrival National Institutes of Health Stroke Scale (NIHSS) subitem scores.

METHODS
Subjects and setting
We studied consecutive suspected stroke patients transported by prehospital EMS as prehospital stroke codes to Zuckerberg San Francisco General Hospital and Trauma Center (ZSFG) between July 2017 and July 2018. ZSFG is an academic county hospital affiliated with the University of California, San Francisco. ZSFG is a Joint Commission certified Primary Stroke Center with full EVT capabilities available 24/7. As a safety net hospital, it provides healthcare for patients regardless of whether or not they have insurance. At the time of this study, there were three EVT capable Primary Stroke Centers (including ZSFG) in San Francisco plus an additional four Primary Stroke Centers without EVT capabilities. However, EMS policies required transportation of all suspected acute stroke patients with positive Cincinnati Prehospital Stroke Scales (CPSS) to the closest stroke centre regardless of EVT capabilities. The San Francisco Fire Department provides EMS transportation for approximately 75 000 patients per year, including 80% of prehospital stroke alert patients. Two private ambulance companies that are also independent of ZSFG contribute as well.

ZSFG patients were retrospectively identified for this study through the prospective Mission Protocol stroke code registry. The Mission Protocol is a quality improvement endeavour at ZSFG to expedite EMS suspected stroke evaluation and treatment.[12] ED nurses at ZSFG would activate the Mission Protocol if the EMS prehospital alert call included suspected stroke, a positive CPSS and a last known well (LKW) time within 6 hours. This prompted a special stroke code activation prior to patient arrival to coordinate efforts between ED providers, neurology and radiology. All adult prehospital stroke code Mission Protocol patients transported by ambulance during the study period were included in the final diagnosis prevalence study. Patients were excluded from the LVO prediction scale diagnostic performance study if the subitems of the arrival NIHSS were not documented.

Patient demographic and clinical characteristics
We ascertained patient demographic and clinical characteristics through chart review. Times between arrival, imaging and treatment were documented using electronic medical record and radiology timestamps. Alteplase administration times were captured from clinical documentation. Whether or not the patient was intubated at the time of initial imaging was determined by review of scout and axial neuroimaging. The sample size was determined by the number of Mission Protocol stroke codes that occurred from July 2017 to July 2018. Study data were collected and managed using REDCap (Research Electronic Data Capture) hosted at the University of California, San Francisco.[13 14]

Index tests
The stroke code neurologist performed a NIHSS examination on ED arrival, typically before imaging, and documented the subitem scores in the medical record. LVO prediction scale scores were retrospectively calculated from the arrival NIHSS subitems using if/then formulas written in StataSE V.16 (Stata Statistical Software: Release 16. StataCorp, College Station, Texas, USA). The LVO prediction scales tested included asymmetric arm weakness (AAW), interval (0–3) scoring of the CPSS, Cincinnati Stroke Triage Assessment Tool (C-STAT), Field Assessment Stroke Triage for Emergency Destination (FAST-ED), Gaze-Face-Arm-Speech-Time (G-FAST), Modified National Institutes of Health Stroke Scale (mNIHSS), NIHSS, Prehospital Acute Severity Scale (PASS), Rapid Arterial oCclusion Evaluation (RACE), Speech Arm Vision Eyes (SAVE) and Vision Aphasia Neglect (VAN).[9 15–23] LVO prediction scales that could not be calculated from the NIHSS were excluded. We prespecified comparisons of standard score thresholds from each scale's publication or the American Heart Association's prediction instrument review, as well as comparisons of alternative (lower and higher) thresholds.[20] The LVO prediction scale scores, except for FAST-ED, were not available to the NIHSS performers as they were retrospectively calculated. FAST-ED was intermittently scored prospectively by ED providers as part of their Mission Protocol assessment and may have been reported to neurologists. We did not
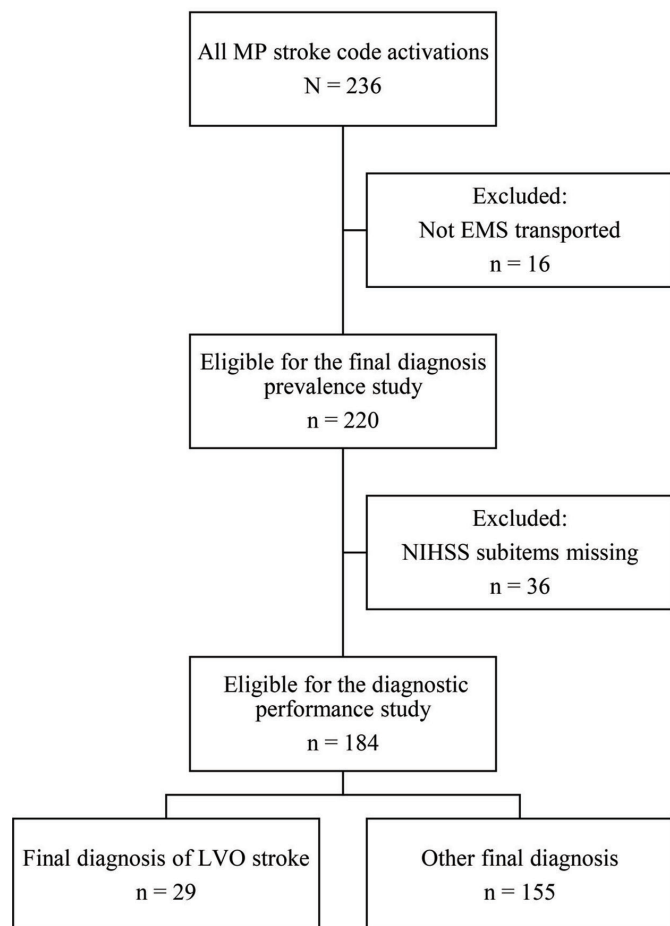
Figure 1 Participant flow diagram for the LVO stroke primary reference standard. EMS, emergency medical services; LVO, large vessel occlusion; MP, Mission Protocol; NIHSS, National Institutes of Health Stroke Scale.

study these prospective FAST-ED assessments as the scores were documented in only a small minority of cases.

Reference standard

Final diagnosis of LVO stroke was the reference standard for the primary diagnostic performance study. This was determined by a retrospective review of all available chart and imaging information by a vascular neurologist (KJK) prior to calculating LVO prediction scale scores. LVO stroke was defined as acute neurological symptoms attributed to an occlusion of the internal carotid artery (ICA), the first segment of the middle cerebral artery (M1), the second segment of the middle cerebral artery (M2) or basilar arteries on acute neurovascular imaging. Involvement of these vessels was prespecified as the definition of LVO stroke due to their candidacy for EVT in clinical practice. Stroke code acute neurovascular imaging included non-contrast CT and CT angiography (CTA), though CTA could be deferred by the clinical stroke code neurologist. Final diagnosis was chosen as the reference standard instead of acute vascular imaging results to avoid the spectrum bias that would accompany excluding patients without vascular imaging, since it is often deferred in patients with intracranial haemorrhage

or low suspicion for LVO stroke. We performed exploratory analyses where the definition of LVO stroke was modified to exclude M2 occlusions as well as separate analyses where intracranial haemorrhage and LVO stroke (with or without M2 occlusions) were combined into a single reference standard.

Analysis

Final diagnosis prevalences were calculated. Diagnostic performance statistics including PPV, NPV, sensitivity, specificity, positive likelihood ratios (PLR) and negative likelihood ratios (NLR) with 95% CIs were calculated for each reference standard of LVO stroke. We specified PPV and NPV goals after the prevalence data were available but prior to the diagnostic performance study. A PPV goal of ≥80% was prespecified to identify LVO scale thresholds that greatly increased the post-test probability of LVO stroke. We also performed a secondary analysis to identify scale thresholds that achieved PPVs ≥50% because the primary analysis PPV goal was only met by one scale threshold. We chose ≥50% to ensure that positive tests would be equally or more likely to be LVO than not. If multiple thresholds in the same scale met these goals, the threshold with the highest NPV was chosen to minimise false negatives.

We also prespecified an assessment for LVO prediction scale thresholds with NPVs ≥95%. This can also be expressed as a post negative test probability of LVO stroke ≤5% and was intended to identify scale thresholds that reduce the risk of missing LVO strokes. Five per cent was thought to be a meaningful reduction of the pre-test probability of LVO stroke of only ~13.6% observed in our prevalence study. If multiple thresholds of the same scale met these criteria, the threshold with the highest simultaneous PPV was selected to minimise false positives. This was repeated for the exploratory analysis that excluded M2 occlusions from the definition of LVO stroke or included intracranial haemorrhage along with LVO stroke in the reference standard.

RESULTS

Figure 1 provides a flow diagram of participants for the prevalence and diagnostic performance studies with LVO stroke as the reference standard. Two hundred and twenty of the 236 patients in the Mission Protocol registry met the prevalence study inclusion criteria during the study time period.

Disease prevalence study

Stroke mimic was the most common final diagnosis category at 50% (table 1). Seizure without status epilepticus was the most common stroke mimic diagnosis but accounted for only 20% of all mimics. Detailed mimic final diagnoses are reported in online supplemental table 1. Non-LVO stroke was the second most common final diagnosis (20.5%), followed by intracranial haemorrhage

Table 1 Final diagnoses and participant characteristics in the prevalence and diagnostic performance studies

Characteristics	Prevalence study (n=220)	Diagnostic performance study subset of patients (n=184)
Final diagnosis		
LVO stroke, n (%)	30 (13.6)	29 (15.8)
M1 occlusions	7	7
Tandem extracranial ICA and M1	4	4
Tandem intracranial ICA and M1	3	3
ICA terminus, M1, and A1	1	1
Extra and intracranial ICA	1	1
Isolated M2 occlusions	10	10
Tandem extracranial ICA and M2	1	1
Tandem extracranial ICA and P2	1	0
Basilar occlusions	2	2
Intracranial haemorrhage, n (%)	35 (15.9)	19 (10.3)
ICH ≥30 mL, n	8	4
Non-LVO ischaemic stroke, n (%)	45 (20.5)	43 (23.4)
Mimics, n (%)	110 (50)	93 (50.5)
Age in years, median (IQR)	70 (61–81)	70 (61–80)
Sex (male), n (%)	118 (54)	94 (51)
Hypertension, n (%)	138 (63)	116 (63)
Hyperlipidaemia, n (%)	74 (37)	63 (34)
Diabetes, n (%)	76 (35)	65 (35)
Atrial fibrillation by history or at presentation, n (%)	33 (15)	28 (15)
Seizure or seizure history, n (%)	25 (11)	19 (10)
Prior ischaemic stroke, n (%)	41 (19)	36 (20)
Prior haemorrhagic stroke, n (%)	12 (5)	10 (5)
Known anticoagulation use, n (%)	17 (8)	13 (7)
Baseline NIHSS, median (IQR)*	8 (4–14)	7 (3–14)
Baseline NIHSS <6, n (%)	91 (46)	75 (41)
Received intravenous alteplase, n (%)	50 (23)	46 (25)
Alteplase given prior to CTA, n (%)	23 (10)	21 (11)
Time between arrival and NCHCT, min, median (IQR)†	11 (9–16)	11 (9–16)
CTA done on arrival, n (%)	166 (75)	145 (79)
Time between arrival and CTA, min, median (IQR)‡	25 (20–32)	25.5 (19–32)
Intubated on arrival imaging, n (%)	10 (5)	4 (2)

Documentation of arrival NIHSS subitem scores was required for inclusion in the diagnostic performance study.

*Total NIHSS scores were known for only 197 of the 220 participants in the prevalence study.
†One patient in the prevalence study did not have an NCHCT, thus time between arrival and NCHCT is based on 219 patients rather than 220.
‡One patient had missing arrival CTA time data, thus time between arrival and CTA is based on 165 rather than 166 patients in the prevalence study and 144 patients rather than 145 in the diagnostic performance study.
A1, first division of the anterior cerebral artery; CTA, CT angiography; ICA, internal carotid artery; ICH, intracerebral haemorrhage; LVO, large vessel occlusion; M1, first division of the middle cerebral artery; M2, second division of the middle cerebral artery; NCHCT, non-contrast head CT; NIHSS, National Institutes of Health Stroke Scale; P2, second division of the posterior cerebral artery.

(15.9%). LVO stroke was the least common final diagnosis at 13.6%.

Of the 30 patients with LVO stroke, there were 7 M1 occlusions, 1 ICA 'T' occlusion, 4 tandem extracranial ICA and M1 occlusions, 3 tandem intracranial ICA/M1 occlusions, 10 M2 occlusions, 1 extracranial ICA/M2 occlusion, 1 tandem extracranial ICA and second segment of the posterior cerebral artery (P2) occlusion, 2 basilar occlusions and 1 extracranial and intracranial ICA occlusion with the ipsilateral middle cerebral artery filled by the posterior communicating artery. The patient with ICA/P2 occlusions was excluded from the diagnostic performance study due to the absence of NIHSS subitem documentation (table 1). Baseline patient demographic, clinical and imaging characteristics for patients in

Table 2 Diagnostic performance statistics for LVO stroke (ICA, M1, M2 or basilar occlusions) including standard and alternative scale thresholds that meet PPV and NPV goals

LVO stroke	PPV % (95% CI)	NPV % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)	PLR (95% CI)	NLR (95% CI)
Standard thresholds						
AAW +	24 (16 to 34)	93 (86 to 98)	79 (60 to 92)	53 (45 to 61)	1.7 (1.3 to 2.2)	0.39 (0.19 to 0.81)
CPSS =3	30 (19 to 43)	91 (85 to 96)	62 (42 to 79)	73 (65 to 80)	2.3 (1.6 to 3.4)	0.52 (0.32 to 0.84)
C-STAT ≥2	32 (20 to 46)	91 (85 to 96)	62 (42 to 79)	76 (68 to 82)	2.5 (1.7 to 3.8)	0.50 (0.31 to 0.81)
FAST-ED ≥4	42 (29 to 56)	95 (90 to 98)	79 (60 to 92)	79 (72 to 85)	3.8 (2.7 to 5.5)	0.26 (0.13 to 0.53)
G-FAST ≥3	30 (20 to 41)	94 (88 to 98)	79 (60 to 92)	65 (56 to 72)	2.2 (1.7 to 3.0)	0.32 (0.16 to 0.66)
mNIHSS ≥6	25 (17 to 36)	94 (87 to 98)	79 (60 to 92)	56 (48 to 64)	1.8 (1.4 to 2.3)	0.37 (0.18 to 0.76)
NIHSS ≥6	23 (15 to 32)	95 (87 to 99)	86 (68 to 96)	46 (38 to 54)	1.6 (1.3 to 2.0)	0.30 (0.12 to 0.76)
NIHSS ≥10	34 (22 to 48)	92 (86 to 96)	66 (46 to 82)	76 (69 to 83)	2.7 (1.9 to 4.0)	0.45 (0.27 to 0.75)
PASS ≥2	25 (16 to 37)	91 (84 to 96)	66 (46 to 82)	64 (56 to 71)	1.8 (1.3 to 2.5)	0.54 (0.32 to 0.90)
RACE ≥5	36 (23 to 52)	91 (85 to 95)	59 (39 to 77)	81 (74 to 87)	3.0 (1.9 to 4.7)	0.51 (0.33 to 0.80)
SAVE ≥2	27 (18 to 37)	96 (89 to 99)	86 (68 to 96)	57 (49 to 65)	2.0 (1.6 to 2.5)	0.24 (0.10 to 0.61)
VAN +	29 (20 to 40)	95 (89 to 98)	83 (64 to 94)	62 (54 to 70)	2.2 (1.7 to 2.8)	0.28 (0.12 to 0.62)
PPV ≥80% with highest simultaneous NPV						
FAST-ED ≥7	83 (36 to 100)	87 (81 to 91)	17 (6 to 36)	99 (97 to 100)	27 (3.2 to 220)	0.83 (0.71 to 0.98)
PPV ≥50% with highest simultaneous NPV						
FAST-ED ≥6	53 (27 to 79)	88 (82 to 92)	28 (13 to 47)	96 (91 to 98)	6.1 (2.4 to 15.5)	0.76 (0.60 to 0.95)
PASS =3	50 (31 to 69)	90 (85 to 95)	48 (29 to 68)	91 (85 to 95)	5.3 (2.9 to 9.9)	0.57 (0.40 to 0.81)
RACE ≥7	55 (32 to 76)	90 (84 to 94)	41 (24 to 61)	94 (89 to 97)	6.4 (3.1 to 13.4)	0.63 (0.46 to 0.85)
NPV ≥95% with highest simultaneous PPV						
CPSS ≥2	23 (15 to 31)	96 (88 to 99)	90 (73 to 98)	43 (35 to 51)	1.6 (1.3 to 1.9)	0.24 (0.08 to 0.72)
FAST-ED ≥4	42 (29 to 56)	95 (90 to 98)	79 (60 to 92)	79 (72 to 85)	3.8 (2.7 to 5.5)	0.26 (0.13 to 0.53)
G-FAST ≥2	22 (15 to 30)	97 (89 to 100)	93 (77 to 99)	38 (30 to 46)	1.5 (1.3 to 1.8)	0.18 (0.05 to 0.70)
mNIHSS ≥2	20 (14 to 27)	100 (90 to 100)	100 (88 to 100)	23 (17 to 31)	1.3 (1.2 to 1.4)	0
NIHSS ≥6	23 (15 to 32)	95 (87 to 99)	86 (68 to 96)	46 (38 to 54)	1.6 (1.3 to 2.0)	0.30 (0.12 to 0.76)
PASS ≥1	19 (13 to 26)	95 (84 to 99)	93 (77 to 99)	25 (19 to 33)	1.2 (1.1 to 1.4)	0.27 (0.07 to 1.1)
RACE ≥3	30 (20 to 41)	95 (89 to 98)	83 (64 to 94)	64 (56 to 71)	2.3 (1.8 to 3.0)	0.27 (0.12 to 0.60)
SAVE ≥2	27 (18 to 37)	96 (89 to 99)	86 (68 to 96)	57 (49 to 65)	2.0 (1.6 to 2.5)	0.24 (0.10 to 0.61)
VAN +	29 (20 to 40)	95 (89 to 98)	83 (64 to 94)	62 (54 to 70)	2.2 (1.7 to 2.8)	0.28 (0.12 to 0.62)

The prevalence of LVO stroke using the ICA, M1, M2 or basilar occlusion definition of LVO stroke was 15.8% in this subset of patients with NIHSS subitems documented.

AAW +, asymmetric arm weakness positive; CPSS, Cincinnati Prehospital Stroke Scale; C-STAT, Cincinnati Stroke Triage Assessment Tool; FAST-ED, Field Assessment Stroke Triage for Emergency Destination; G-FAST, Gaze-Face-Arm-Speech-Time; ICA, internal carotid artery; LVO, large vessel occlusion; M1, first division of the middle cerebral artery; M2, second division of the middle cerebral artery; mNIHSS, modified National Institutes of Health Stroke Scale; NIHSS, National Institutes of Stroke Scale; NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value; RACE, Rapid Arterial oCclusion Evaluation; SAVE, Speech Arm Vision Eyes; VAN +, Vision Aphasia Neglect positive.

both the prevalence study and diagnostic performance study are provided in table 1 and described in additional detail in the online supplemental file.

Diagnostic performance study

The prevalence of LVO stroke was 15.8% in the diagnostic performance study (table 1) because patients without arrival NIHSS subitem documentation were excluded. While this was rare for LVO stroke patients (1/30, 3.3%), it was notably common for intracranial haemorrhage patients (16/35, 45%).

Table 2 provides the diagnostic performance statistics of standard and alternative LVO scale thresholds when predicting the final diagnosis of LVO stroke and alternative (higher or lower) thresholds selected by positive and negative predictive value goals. As expected, PPVs and NPVs were inversely related. Higher sensitivities were also associated with lower specificities, except for FAST-ED ≥4, which was

79% sensitive and specific. CIs were wide, often exceeding a 20% range and sometimes exceeding a 30% range.

None of the standard thresholds met the prespecified PPV ≥80% goal or the secondary ≥50% goal for LVO stroke. FAST-ED ≥7 was the only alternative scale threshold identified by the prespecified analysis searching for PPVs ≥80%; however, this threshold missed 83% of LVO strokes. Alternative scale thresholds of FAST-ED ≥6, PASS =3 and RACE ≥7 met the less stringent goal of PPV ≥50% but were associated with sensitivities less than 50% (28%, 48% and 41%, respectively).

Several standard thresholds (FAST-ED ≥4, NIHSS ≥6, SAVE ≥2 and a positive VAN test) and alternative thresholds (CPSS ≥2, G-FAST ≥2, mNIHSS ≥2 and RACE ≥3) met the prespecified NPV ≥95% goal. Except for FAST-ED ≥4 (PPV 42%), these were all associated with PPVs of 30% or lower (table 2). Scale thresholds that met the high PPV goals had many false negatives. Thresholds that met the high NPV goals had many false positives. The diagnostic statistics for the reference standards of LVO stroke excluding M2s, LVO stroke excluding M2s and intracranial haemorrhage, and LVO stroke and intracranial haemorrhage are provided in online supplemental tables 2–4.

Figure 2 shows the counts of true positives, false negatives, false positives and true negatives for standard LVO prediction scale thresholds with all four studied reference standards. The most restrictive reference standard, LVO stroke without M2s (figure 2, top left), demonstrates that many standard scale thresholds can be sensitive for ICA, M1 and basilar occlusions. Including M2s in the definition of LVO stoke (figure 2, top right), which was our prespecified primary analysis, resulted in comparatively more false negatives and therefore reduced sensitivities and NPVs. However, the addition of M2s increased PPVs and very modestly increased specificities (table 2 and online supplemental table 2). The number of true positives and false negatives increased when intracranial haemorrhage was added to the reference standard (figure 2, bottom left and bottom right), while false positives decreased.

LVO prediction scale thresholds that met our prespecified PPV and NPV goals are shown in figure 3. As the reference standard changed from the most restrictive (LVO stroke excluding M2s) to the most inclusive (LVO stroke including M2s and intracranial haemorrhage), more scales were able to provide thresholds with PPVs ≥50% including several standard thresholds (CPSS =3, C-STAT ≥2, FAST-ED ≥4 and RACE ≥5), though higher alternative thresholds were common. Conversely, the list of scale thresholds that could reach the NPV ≥95% goal was the longest for LVO stroke excluding M2s. Higher alternative thresholds were predominant though a few standard thresholds (CPSS =3, G-FAST ≥3 and VAN +) met this goal. The list of thresholds with NPVs ≥95% was shortest for LVO stroke and intracranial haemorrhage and required very low alternative thresholds that would generate many false positives (figure 3 and online supplemental tables 2–4). Two-by-two tables for each scale threshold and reference standard along with true positive, true false negative, false positive and true negative

counts receiver operating characteristic curves are provided in online supplemental tables 5–8 and online supplemental figures 1–4, respectively. Receiver operating characteristic curves and scale areas under the curve (AUCs) for each reference standard are provided in online supplemental figures 5–8. AUCs were generally similar between scales, increased with the exclusion of M2s and decreased with the addition of intracranial haemorrhages.

DISCUSSION

The American Heart Association/American Stroke Association (AHA/ASA) identified the relative lack of prehospital suspected stroke disease prevalence data as a major research limitation affecting LVO prediction scale assessments.[20] We found that among 220 EMS suspected stroke patients who were prehospital CPSS positive and had LKW times within 6 hours, the most common final diagnosis category was stroke mimic (50%), followed by non-LVO ischaemic stroke (20.5%), intracranial haemorrhage (15.9%) and LVO stroke (13.6%). This distribution is similar to other studies that assessed prehospital suspected stroke patients without excluding significant subgroups. Taken all together, prevalences in the literature have ranged from 25% to 68% for stroke mimics, 24%–38% for non-LVO ischaemic strokes, 4%–16% for intracranial haemorrhages and 4%–15% for LVO strokes.[10 11 24–26] Isolated M2 occlusions accounted for 33% of LVO strokes in our study and over 40% of occlusions in two larger studies with lower overall LVO prevalence than ours.[10 11] Excluding M2 occlusions from our study would reduce the prevalence of LVO stroke to 9.1%. Knowledge of these estimated prevalences of LVO stroke among prehospital suspected stroke patients could help inform future prehospital triage modelling studies and triage test policy decisions when local data are not available.[20 27]

We sought to identify LVO prediction scale thresholds that met high PPV and NPV goals because predictive values are post-test probabilities of LVO stroke when the prevalence of diseases in a study is similar to that of the target population. Predictive values are strongly influenced by disease prevalence but are also affected by variations in test sensitivity and specificity.[6–8] Our prespecified goal of finding thresholds with PPVs ≥80% was too optimistic given the relatively low prevalence of LVO stroke and the diagnostic performance of LVO scales. Only FAST-ED ≥7 met this goal, but its sensitivity was so low (17%) that it missed 83% of LVO strokes. We subsequently lowered our PPV goal to ≥50% in order to find scale thresholds where a positive test result would mean that the patient was equally or more likely to be suffering from LVO stroke than not. PASS =3, RACE ≥7 and FAST-ED ≥6 all met this goal, but none maintained a sensitivity ≥50%. Our prespecified analysis to find scale thresholds with NPVs ≥95% for LVO stroke identified multiple standard and alternative thresholds (table 2).

With high alternative FAST-ED thresholds meeting our PPV goals, although with markedly low sensitivities, and the standard FAST-ED threshold ≥4 standing out as the only threshold to achieve >70% sensitivity and specificity
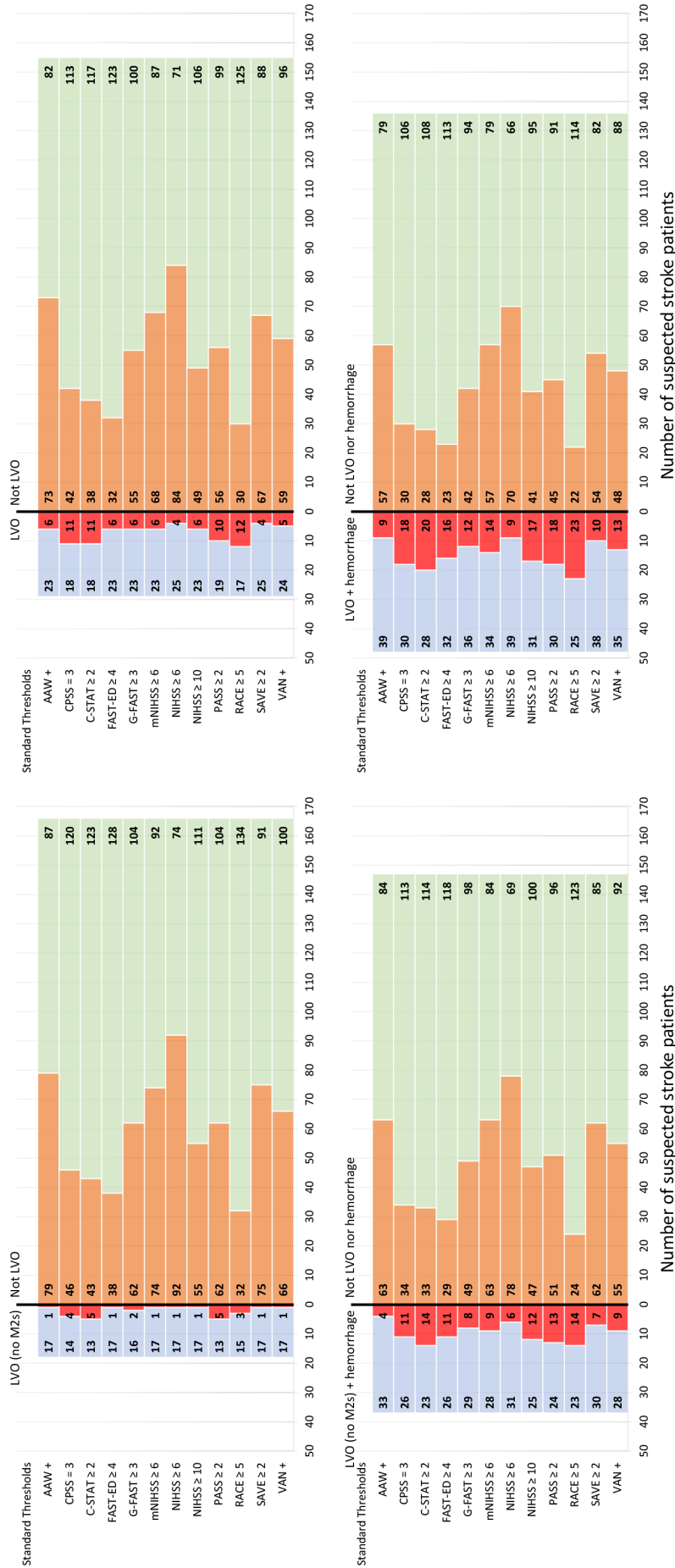
Figure 2 True positive, false negative, false positive and true negative classification counts for standard scale thresholds for the four reference standards. AAW +, asymmetric arm weakness positive; CPSS, Cincinnati Prehospital Stroke Scale; C-STAT, Cincinnati Stroke Triage Assessment Tool; FAST-ED, Field Assessment Stroke Triage for Emergency Destination; G-FAST, Gaze-Face-Arm-Speech-Time; LVO, large vessel occlusion; M2, second division of the middle cerebral artery; mNIHSS, modified National Institutes of Health Stroke Scale; NIHSS, National Institutes of Health Stroke Scale; PASS, Prehospital Acute Severity Scale; RACE, Rapid Arterial oCclusion Evaluation; SAVE, Speech Arm Vision Eyes; VAN +, Vision Aphasia Neglect positive; #, number.
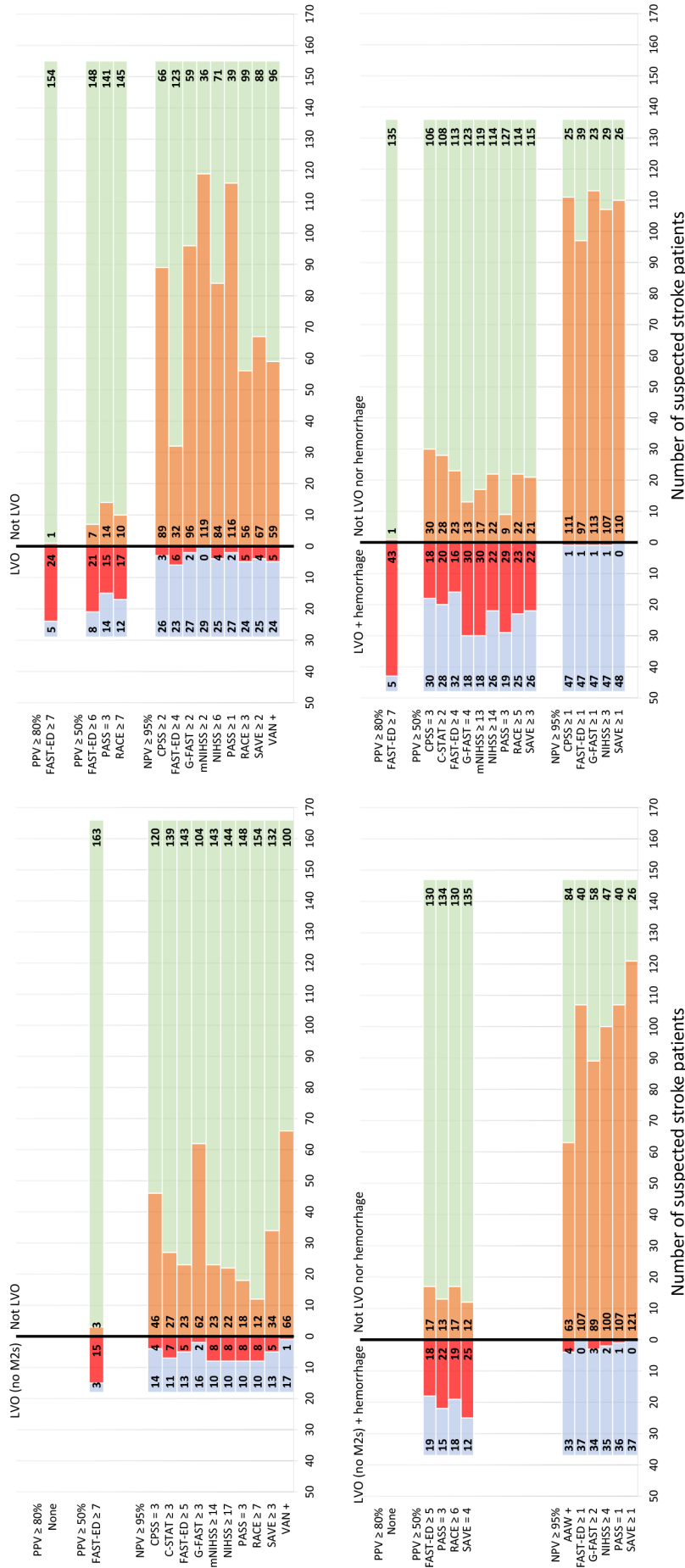
Figure 3 Classification counts for standard or alternative scale thresholds that met positive and negative predictive value goals for the four reference standards. AAW +, asymmetric arm weakness positive; CPSS, Cincinnati Prehospital Stroke Scale; C-STAT, Cincinnati Stroke Triage Assessment Scale; FAST-ED, Field Assessment Stroke Triage for Emergency Destination; G-FAST, Gaze-Face-Arm-Speech–Time; LVO, large vessel occlusion; M2, second division of the middle cerebral artery; mNIHSS, modified National Institutes of Health Stroke Scale; NIHSS, National Institutes of Health Stroke Scale; PASS, Prehospital Acute Severity Scale; RACE, Rapid Arterial oCclusion Evaluation; SAVE, Speech Arm Vision Eyes; VAN +, Vision Aphasia Neglect positive; #, number.

simultaneously, it is tempting to conclude that FAST-ED might outperform other scales. However, these findings should be considered with caution for two reasons. First, Nguyen *et al* studied FAST-ED ≥4 in a larger prospective study in ambulances with EMS providers performing the examination and found that it performed similarly (60% sensitivity and 85% specificity) to C-STAT, PASS, G-FAST and RACE at their standard thresholds.[10] Second, our results could be due to the FAST-ED training given to ED physicians as part of the Mission Protocol quality improvement initiative. The protocol had instructions for ED providers to use FAST-ED as part of their initial assessment and document it. However, this was not routinely done in practice and was very rarely documented. Training was not provided for other scales. The NIHSS subitem scores used in this study to score LVO prediction scales were obtained by neurologists without FAST-ED training; however, their examinations could have been influenced by discussions with ED providers. While this is a limitation of our study, it could also suggest that dedicated training might improve sensitivity and specificity in the ED.

Our results suggest that given the prevalence of LVO stroke among prehospital suspected stroke patients and the sensitivity and specificity tradeoffs of LVO prediction scales, no single scale threshold tested here will likely be able to guide prehospital stroke triage efficiently by itself. Attempts to reach PPVs near 50% by selecting higher than standard thresholds will result in many false-negative missed LVO strokes (figure 2, top right). While several standard thresholds and many lower alternative thresholds can reduce the probability of LVO stroke when tests are negative, this will be accompanied by many false positives. EMS systems that already use the CPSS as a binary initial stroke screen and choose to focus on avoiding missed LVOs could readily adopt CPSS ≥2 as an LVO prediction scale without training personnel to use a new scale. CIs in our study were wide, especially with thresholds that met high PPV goals, but our findings were predominantly in line with larger prospective prehospital validation studies assessing multiple scales simultaneously.[10 11]

Our primary analysis focused on a definition of LVO stroke that included ICA, M1, M2 and basilar occlusions. Our exploratory analysis that excluded M2 occlusions from the definition of LVO stroke (figures 2 and 3, top left) showed that standard thresholds may not miss many ICA, M1 or basilar occlusions (high sensitivity); however, the corresponding proportion of positive tests with these occlusions (PPV) will be low. It is reassuring that excluding M2s increases the sensitivities of some standard thresholds for ICA, M1 or basilar occlusions into the 80%–90% range as more proximal occlusions are more reliably amenable to EVT. ICA and M1 occlusions are also more morbid and benefit from the strongest evidence base for EVT.[28 29] Though more distal MCA occlusions have a higher recanalisation rate with thrombolysis, it is only in the 30% range.[30] Registry study data support the benefit of EVT for M2 occlusions found in clinical practice, however, 2019 AHA/ASA guidelines currently provide only a class IIb recommendation.[29] Patients presenting with non-disabling symptoms or distal M2 occlusions may still represent situations with EVT equipoise. The Australian and New Zealand Clinical Guidelines for Stroke Management suggest EVT may be considered 'based on individual patient and advanced imaging factors'.[31] As noted above, M2s have comprised a significant proportion of LVOs in studies with prehospital suspected stroke patients (33% in our study and over 40% in two others) so the development of prehospital diagnostic tests that are sensitive for them while retaining specificity would be beneficial. Several device-based approaches are under development to address the diagnostic limitations of LVO prediction scales, including by some of the current authors, but no devices have examined M2 occlusions in large numbers nor completed prehospital validation studies.[32–34]

Alternatively, broadening our definition of successful triage to include not only M2 occlusions but also intracranial haemorrhage results in high PPVs but low sensitivities. This may be reassuring to prehospital systems of care decision makers. Many of the 'false positive' non-LVO patients, when only ICA, M1 or basilar occlusions are counted as 'true positives', will be suffering from distal vessel occlusions or intracranial haemorrhage. These patients could also benefit from the frequent colocalisation of vascular and neurosurgical expertise at EVT centres. In health systems that can accommodate the added patient volume, there may be little downside to overtriage, especially where thrombolysis door to needle times are faster at EVT centres than non-EVT centres. Though controversial, map-based modelling studies analysing the USA and Canada suggest that prehospital bypass of non-EVT centres may benefit even thrombolysis eligible suspected LVO patients unless local non-EVT centres can provide door to needle times of 30 min or less.[35]

When examining the exclusion of M2s or addition of intracranial haemorrhages, it is important to note that changing the reference standard does not change patient-level test results and as a result does not affect patient-level triage decisions. For example, in each analysis, the number of patients with positive and negative CPSS =3 test results remains constant at 60 positives and 124 negatives (online supplemental tables 5–8). When the reference standard is changed from LVO stroke alone to LVO stroke and intracranial haemorrhage combined, the same patients still have positive and negative CPSS =3 test results, but the number of true positives, false negatives, false positives and true negatives changes. This in turn leads to different PPV, NPV, sensitivity and specificity results.

There are important limitations to our study. We only included EMS suspected stroke CPSS positive patients with LKW times of 6 hours or less that were brought to ZSFG over the course of 1 year. The prevalence of disease may vary by region, and it is not clear if scale performance would differ in the 6–24-hour stroke time window. At the time of our study, San Francisco was not using a prehospital triage system to divert patients with possible LVO stroke to EVT centres. However, it is possible that the Mission Protocol quality improvement effort at ZSFG, which included EMS outreach and education programmes, could have led to unmeasured

changes in EMS routing patterns for suspected stroke patients that could bias the prevalence estimates described here.

The diagnostic performance portion of our study also excluded patients without NIHSS scores documented on ED arrival. This led to a disproportionate exclusion of patients with intracranial haemorrhage. This should be considered when interpreting our primary LVO stroke analysis and our exploratory analyses that included variations of LVO stroke and intracranial haemorrhage as combined reference standards.

Most importantly, our LVO prediction scales were retrospectively calculated from prospectively recorded NIHSS scores performed by neurologists in the ED rather than prehospital EMS providers prior to transport. While this allowed us to compare many scales at various thresholds, it does not represent the performance of scales as used by EMS providers in the prehospital setting. In addition, neurological examination changes can occur between EMS assessments and ED arrival.[36] Our use of NIHSS scores also precluded us from testing promising scales such as the ambulance clinical triage for acute stroke treatment (ACT-FAST) algorithm due to its stepwise algorithm approach and the Los Angeles Motor Scale (LAMS) due to the absence of handgrip data.[37 38] A recent multisociety consensus statement recommended maximum EMS travel times during LVO stroke triage that are tailored to urban, suburban and rural settings.[39] However, there are no consensus PPV, NPV, sensitivity or specificity goals for prehospital LVO stroke triage. This led us to create PPV and NPV goals that we thought were clinically reasonable, given the prevalence of LVO stroke. We also considered a secondary analysis using an NPV goal of ≥99%. However, only the lowest alternative scale thresholds met this goal, and none did so with a specificity ≥25%.

Our study has several strengths. First, our use of consecutive prehospital stroke alerts over the entire study period allowed us to better estimate the prevalence of disease among suspected stroke patients identified by EMS providers in the prehospital setting. Second, we compared a wide range of LVO prediction scale thresholds rather than limiting our analysis to standard thresholds. In some cases, standard thresholds originally thought to be more sensitive than specific have subsequently been found to be more specific than sensitive.[10 11 22] We propose that selecting thresholds that meet future consensus goals will be more valuable than a focus on standard thresholds. High NPV or sensitivity goals should likely be favoured for LVO stroke given the profound yet time-sensitive benefit of EVT as well as the association of transfers from non-EVT centres to EVT centres with worse outcomes.[1 4 5 7]

Finally, our study did not attempt to derive a new LVO prediction scale and instead focused on the external validation of multiple scales. Although our assessments occurred in an ED, all patients were identified as suspected stroke patients by EMS providers in the prehospital setting.

In conclusion, our data were consistent with others suggesting that the prevalence of LVO stroke among unselected prehospital suspected stroke patients is low. Prevalence is a key factor in the performance of LVO prediction scales. High alternative LVO prediction scale thresholds were required to meet a PPV goal of ≥50%, but these thresholds missed most LVO strokes. Including intracranial haemorrhages as true positives increased the number of scales that could provide PPVs ≥50%. Several standard thresholds and many alternative lower thresholds provided NPVs ≥95%, including CPSS ≥2, though false positives were common. EMS systems already using the CPSS as a binary initial stroke screen could also adopt it as a high NPV LVO triage prediction scale without having to incorporate an additional stroke scale. The limitations of these neurological examination-based tests support the need for further investigation of alternative approaches to prehospital LVO stroke identification, such as Mobile Stroke Units and portable LVO stroke diagnostic devices.[34] In the meantime, implementation of LVO prediction scales still may benefit correctly classified patients more than prehospital systems that do not attempt prehospital LVO stroke triage.

ORCID iD
Kevin J Keenan http://orcid.org/0000-0002-5643-9878

REFERENCES

1 Saver JL, Goyal M, van der Lugt A, *et al*. Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: a meta-analysis. *JAMA* 2016;316:1279–89.
2 Pride GL, Fraser JF, Gupta R, *et al*. Prehospital care delivery and triage of stroke with emergent large vessel occlusion (ELVO): report of the standards and guidelines Committee of the Society of Neurointerventional surgery. *J Neurointerv Surg* 2017;9:802–12.
3 Keenan KJ, Kircher C, McMullan JT. Prehospital prediction of large vessel occlusion in suspected stroke patients. *Curr Atheroscler Rep* 2018;20:x.
4 Prabhakaran S, Ward E, John S, *et al*. Transfer delay is a major factor limiting the use of intra-arterial treatment in acute ischemic stroke. *Stroke* 2011;42:1626–30.
5 Froehler MT, Saver JL, Zaidat OO, *et al*. Interhospital transfer before thrombectomy is associated with delayed treatment and worse outcome in the STRATIS registry (systematic evaluation of patients treated with Neurothrombectomy devices for acute ischemic stroke). *Circulation* 2017;136:2311–21.
6 Trevethan R. Sensitivity, specificity, and predictive values: foundations, Pliabilities, and pitfalls in research and practice. *Front Public Health* 2017;5:307.
7 Akobeng AK. Understanding diagnostic tests 1: sensitivity, specificity and predictive values. *Acta Paediatr* 2007;96:338–41.
8 Montori VM, Wyer P, Newman TB, *et al*. Tips for learners of evidence-based medicine: 5. The effect of spectrum of disease on the performance of diagnostic tests. *CMAJ* 2005;173:385–90.
9 Keenan KJ, Smith WS. The speech arm vision eyes (save) scale predicts large vessel occlusion stroke as well as more complicated scales. *J Neurointerv Surg* 2019;11:659–63.
10 Nguyen TTM, van den Wijngaard IR, Bosch J. Comparison of prehospital scales for predicting large anterior vessel occlusion in the ambulance setting. *JAMA Neurol* 2020.
11 Duvekot MHC, Venema E, Rozeman AD, *et al*. Comparison of eight prehospital stroke scales to detect intracranial large-vessel occlusion in suspected stroke (PRESTO): a prospective observational study. *Lancet Neurol* 2021;20:213–21.
12 Madhok DY, Keenan KJ, Cole SB, *et al*. Prehospital and emergency Department-Focused mission protocol improves thrombolysis metrics for suspected acute stroke patients. *J Stroke Cerebrovasc Dis* 2019;28:104423.
13 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
14 Harris PA, Taylor R, Minor BL, *et al*. The REDCap Consortium: building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
15 Richards CT, Huebinger R, Tataris KL, *et al*. Cincinnati prehospital stroke scale can identify large vessel occlusion stroke. *Prehosp Emerg Care* 2018;22:312–8.
16 McMullan JT, Katz B, Broderick J, *et al*. Prospective prehospital evaluation of the Cincinnati stroke triage assessment tool. *Prehosp Emerg Care* 2017;21:481–8.
17 Lima FO, Silva GS, Furie KL. Field assessment stroke triage for emergency destination: a simple and accurate prehospital scale to detect large vessel occlusion strokes. *Stroke J Cereb Circ* 2016;47:1997–2002.
18 Scheitz JF, Abdul-Rahim AH, MacIsaac RL, *et al*. Clinical selection strategies to identify ischemic stroke patients with large anterior vessel occlusion: results from SITS-ISTR (safe implementation of thrombolysis in stroke international stroke thrombolysis registry). *Stroke* 2017;48:290–7.
19 Mulkerin W, Spokoyny I, Corry M. Standardized prehospital education and modified National Institutes of health stroke scale training for paramedics: a pilot study. *Acad Emerg Med* 2018;25:S259.
20 Smith EE, Kent DM, Bulsara KR, *et al*. Accuracy of prediction instruments for diagnosing large vessel occlusion in individuals with suspected stroke: a systematic review for the 2018 guidelines for the early management of patients with acute ischemic stroke. *Stroke* 2018;49.
21 Hastrup S, Damgaard D, Johnsen SP. Prehospital acute stroke severity scale to predict large artery occlusion: design and comparison with other scales. *Stroke J Cereb Circ* 2016;47:1772–6.
22 Pérez de la Ossa N, Carrera D, Gorchs M, *et al*. Design and validation of a prehospital stroke scale to predict large arterial occlusion: the rapid arterial occlusion evaluation scale. *Stroke* 2014;45:87–91.
23 Teleb MS, Ver Hage A, Carter J, *et al*. Stroke vision, aphasia, neglect (van) assessment-a novel emergent large vessel occlusion screening tool: pilot study and comparison with current clinical severity indices. *J Neurointerv Surg* 2017;9:122-126.
24 Zhao H, Coote S, Pesavento L, *et al*. Large vessel occlusion scales increase delivery to endovascular centers without excessive harm from Misclassifications. *Stroke* 2017;48:568–73.
25 Dozois A, Hampton L, Kingston CW, *et al*. PLUMBER study (prevalence of large vessel occlusion strokes in Mecklenburg County emergency response). *Stroke* 2017;48:3397–9.
26 JL L, McMullan JT, Sucharew H. Potential Impact of C-STAT for Prehospital Stroke Triage up to 24 Hours on a Regional Stroke System. *Prehospital Emerg Care Off J Natl Assoc EMS Physicians Natl Assoc State EMS Dir*. Published online 2019:1–5.
27 Venema E, Burke JF, Roozenbeek B, *et al*. Prehospital triage strategies for the transportation of suspected stroke patients in the United States. *Stroke* 2020;51:3310–9.
28 Smith WS, Lev MH, English JD, *et al*. Significance of large vessel intracranial occlusion causing acute ischemic stroke and TIA. *Stroke* 2009;40:3834–40.
29 Powers WJ, Rabinstein AA, Ackerson T, *et al*. Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: a guideline for healthcare professionals from the American heart Association/American stroke association. *Stroke* 2019;50.
30 Menon BK, Al-Ajlan FS, Najm M, *et al*. Association of clinical, imaging, and thrombus characteristics with recanalization of visible intracranial occlusion in patients with acute ischemic stroke. *JAMA* 2018;320:1017–26.
31 Australian and New Zealand Clinical Guidelines for Stroke Management - Chapter 3 of 8: Acute medical and surgical management. Available: https://app.magicapp.org/#/guideline/QnoKGn/section/jMmGRj [Accessed 20 Apr 2022].
32 Smith WS, Keenan KJ, Lovoi PA. A unique signature of Cardiac-Induced cranial forces during acute large vessel stroke and development of a predictive model. *Neurocrit Care* 2020;33:58–63.
33 Keenan KJ, Lovoi PA, Smith WS. The neurological examination improves cranial Accelerometry large vessel occlusion prediction accuracy. *Neurocrit Care* 2021;35:103–12.
34 Patrick L, Smith W, Keenan KJ. Large vessel occlusion stroke detection in the prehospital environment. *Curr Emerg Hosp Med Rep* 2021;9:64–72.
35 Milne MSW, Holodinsky JK, Hill MD, *et al*. Drip 'n Ship Versus Mothership for Endovascular Treatment: Modeling the Best Transportation Options for Optimal Outcomes. *Stroke* 2017;48:791–4.
36 Shkirkova K, Saver JL, Starkman S, *et al*. Frequency, predictors, and outcomes of prehospital and early Postarrival neurological deterioration in acute stroke. *JAMA Neurol* 2018;75:1364–74.
37 Llanes JN, Kidwell CS, Starkman S, *et al*. The Los Angeles motor scale (LAMS): a new measure to characterize stroke severity in the field. *Prehosp Emerg Care* 2004;8:46–50.
38 Zhao H, Pesavento L, Coote S, *et al*. Ambulance clinical triage for acute stroke treatment: Paramedic triage algorithm for large vessel occlusion. *Stroke* 2018;49:945–51.
39 Jauch EC, Schwamm LH, Panagos PD, *et al*. Recommendations for regional stroke destination plans in rural, suburban, and urban communities from the prehospital stroke system of care consensus conference: a consensus statement from the American Academy of Neurology, American heart Association/American stroke association, American Society of Neuroradiology, National association of EMS physicians, National association of state EMS Officials, society of NeuroInterventional surgery, and society of vascular and Interventional neurology: endorsed by the neurocritical care Society. *Stroke* 2021;52:e133–52.