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

Kevin J Keenan 1,2,3 Wade S Smith,2,3 Sara B Cole,2 Christine Martin,2 J Claude Hemphill,2,3 Debbie Y Madhok2,3,4

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 centres. This is intended to avoid the EVT centre.

CONCLUSIONS
LVO prediction scales used for prehospital triage to endovascular centres.

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ 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.

WHAT THIS STUDY ADDS
⇒ The prevalence of LVO stroke among Emergency Medical Services transported 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 centres.

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.

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 treatment delays and worse outcomes associated with initial transportation to a non-EVT centre followed by an interfacility transfer to an EVT centre.2-3 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.
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). All four parameters are important, but predictive values are highly dependent on the prevalence of the target disease in the sample of patients tested. 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.

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. 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. 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.

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). 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. 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
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
of the 30 patients with LVO stroke, there were 7 M1 occlusions, 4 tandem extracranial ICA and M1 occlusions, 3 tandem intracranial ICA and M1 occlusions, 1 ICA terminus, M1, and A1 occlusion, 1 Extra and intracranial ICA occlusion, 1 isolated M2 occlusions, 1 tandem extracranial ICA and M2 occlusion, 1 tandem extracranial ICA and 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** Final diagnoses and participant characteristics in the prevalence and diagnostic performance studies

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

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, 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
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

**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

<table>
<thead>
<tr>
<th>LVO stroke</th>
<th>PPV % (95% CI)</th>
<th>NPV % (95% CI)</th>
<th>Sensitivity % (95% CI)</th>
<th>Specificity % (95% CI)</th>
<th>PLR (95% CI)</th>
<th>NLR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Standard thresholds</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAW +</td>
<td>24 (16 to 34)</td>
<td>93 (86 to 98)</td>
<td>79 (60 to 92)</td>
<td>53 (45 to 61)</td>
<td>1.7 (1.3 to 2.2)</td>
<td>0.39 (0.19 to 0.81)</td>
</tr>
<tr>
<td>CPSS ≥3</td>
<td>30 (19 to 43)</td>
<td>91 (85 to 96)</td>
<td>62 (42 to 79)</td>
<td>73 (65 to 80)</td>
<td>2.3 (1.6 to 3.4)</td>
<td>0.52 (0.32 to 0.84)</td>
</tr>
<tr>
<td>C-STAT ≥2</td>
<td>32 (20 to 46)</td>
<td>91 (85 to 96)</td>
<td>62 (42 to 79)</td>
<td>76 (68 to 82)</td>
<td>2.5 (1.7 to 3.8)</td>
<td>0.50 (0.31 to 0.81)</td>
</tr>
<tr>
<td>FAST-ED ≥4</td>
<td>42 (29 to 56)</td>
<td>95 (90 to 98)</td>
<td>79 (60 to 92)</td>
<td>79 (72 to 85)</td>
<td>3.8 (2.7 to 5.5)</td>
<td>0.26 (0.13 to 0.53)</td>
</tr>
<tr>
<td>G-FAST ≥3</td>
<td>30 (20 to 41)</td>
<td>94 (88 to 98)</td>
<td>79 (60 to 92)</td>
<td>65 (56 to 72)</td>
<td>2.2 (1.7 to 3.0)</td>
<td>0.32 (0.16 to 0.66)</td>
</tr>
<tr>
<td>mNIHSS ≥6</td>
<td>25 (17 to 36)</td>
<td>94 (87 to 98)</td>
<td>79 (60 to 92)</td>
<td>56 (48 to 64)</td>
<td>1.8 (1.4 to 2.3)</td>
<td>0.37 (0.18 to 0.76)</td>
</tr>
<tr>
<td>NIHSS ≥6</td>
<td>23 (15 to 32)</td>
<td>95 (87 to 99)</td>
<td>86 (68 to 96)</td>
<td>46 (38 to 54)</td>
<td>1.6 (1.3 to 2.0)</td>
<td>0.30 (0.12 to 0.76)</td>
</tr>
<tr>
<td>NIHSS ≥10</td>
<td>34 (22 to 48)</td>
<td>92 (86 to 96)</td>
<td>66 (46 to 82)</td>
<td>76 (69 to 83)</td>
<td>2.7 (1.9 to 4.0)</td>
<td>0.45 (0.27 to 0.75)</td>
</tr>
<tr>
<td>PASS ≥2</td>
<td>25 (16 to 37)</td>
<td>91 (84 to 96)</td>
<td>66 (46 to 82)</td>
<td>64 (56 to 71)</td>
<td>1.8 (1.3 to 2.5)</td>
<td>0.54 (0.32 to 0.90)</td>
</tr>
<tr>
<td>RACE ≥5</td>
<td>36 (23 to 52)</td>
<td>91 (85 to 95)</td>
<td>59 (39 to 77)</td>
<td>81 (74 to 87)</td>
<td>3.0 (1.9 to 4.7)</td>
<td>0.51 (0.33 to 0.80)</td>
</tr>
<tr>
<td>SAVE ≥2</td>
<td>27 (18 to 37)</td>
<td>96 (89 to 99)</td>
<td>86 (68 to 96)</td>
<td>57 (49 to 65)</td>
<td>2.0 (1.6 to 2.5)</td>
<td>0.24 (0.10 to 0.61)</td>
</tr>
<tr>
<td>VAN +</td>
<td>29 (20 to 40)</td>
<td>95 (89 to 98)</td>
<td>83 (64 to 94)</td>
<td>62 (54 to 70)</td>
<td>2.2 (1.7 to 2.8)</td>
<td>0.28 (0.12 to 0.62)</td>
</tr>
</tbody>
</table>

**PPV ≥80% with highest simultaneous NPV**

- FAST-ED ≥7: 83 (36 to 100) Sensitivity 87 (81 to 91) Specificity 17 (6 to 36) PPV 99 (97 to 100) NPV 27 (3.2 to 220) PLR 0.83 (0.71 to 0.98)

**PPV ≥50% with highest simultaneous NPV**

- FAST-ED ≥6: 53 (27 to 79) Sensitivity 88 (82 to 92) Specificity 28 (13 to 47) PPV 96 (91 to 98) NPV 6.1 (2.4 to 15.5) PLR 0.76 (0.60 to 0.95)

**PPV ≥95% with highest simultaneous NPV**

- FAST-ED ≥3: 30 (14 to 27) Sensitivity 100 (90 to 100) Specificity 24 (17 to 31) PPV 100 (88 to 100) NPV 0.63 (0.46 to 0.85)

- FAST-ED ≥1: 22 (15 to 30) Sensitivity 97 (89 to 100) Specificity 30 (23 to 46) PPV 93 (77 to 99) NPV 1.5 (1.3 to 1.8) PLR 0.57 (0.40 to 0.81)

**NPV ≥95% with highest simultaneous PPV**

- PASS ≥1: 19 (13 to 26) Sensitivity 95 (84 to 99) Specificity 25 (19 to 33) PPV 93 (77 to 99) NPV 1.2 (1.1 to 1.4) PLR 0.27 (0.07 to 1.1)

- RACE ≥1: 30 (20 to 41) Sensitivity 95 (89 to 98) Specificity 64 (56 to 71) PPV 83 (64 to 94) NPV 2.3 (1.8 to 3.0) PLR 0.27 (0.12 to 0.60)

- SAVE ≥2: 27 (18 to 37) Sensitivity 96 (89 to 99) Specificity 57 (49 to 65) PPV 86 (68 to 96) NPV 2.0 (1.6 to 2.5) PLR 0.24 (0.10 to 0.61)

- VAN +: 29 (20 to 40) Sensitivity 95 (89 to 98) Specificity 62 (54 to 70) PPV 83 (64 to 94) NPV 2.2 (1.7 to 2.8) PLR 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.
Several standard thresholds (FAST-ED ≥1, 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 stroke (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. 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. 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. 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.

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. 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 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.
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. 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.30 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 neurological 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 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 suspected stroke patients identified by EMS providers in the prehospital setting to our exploratory analyses that included variations of LVO stroke and intracranial haemorrhage as combined reference standards.

Contributors KJK, WSS, JCH and DYM conceived of the study and designed the protocol. KJK, SBC, CM and DYM collected the data. KJK analysed the data and wrote the manuscript, including figures, with input from all coauthors. KJK is the guarantor for the study.

Funding KJK’s research effort was supported by the NIH StrokeNet Fellowship through grants U10NS06482 and U24NS107229, the 2019 Clinician Scientist Development Three-Year Award in Interventional Neurology funded by the Society of Vascular and Interventional Neurology and American Brain Foundation in collaboration with the American Academy of Neurology Institute, and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant number ULTR001860. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Competing interests KJK co-created the SAVE scale and is the principal investigator of a national validation study of a cranial accelerometry device for large vessel occlusion detection funded by the 2019 Clinician Scientist Development Award in Interventional Neurology funded by the Society of Vascular and Interventional Neurology and American Brain Foundation in collaboration with the American Academy of Neurology Institute. WSS co-created the SAVE scale and has ownership interest in MindRhythm, which is developing the above cranial accelerometry device, and consulted for and owns stock in Cerebrotech, Inc.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by University of California, San Francisco. A waiver of consent was approved for this retrospective study of patients identified through a stroke registry.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.
REFERENCES


Open access

BMJ Neurol Open: first published as 10.1136/bmjno-2022-000272 on 11 July 2022. Downloaded from http://neurologyopen.bmj.com/ by guest. Protected by copyright.