BMJ Open

Prehospital stroke recognition, severity grading, and prediction of large vessel occlusion using a single clinical scale only

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016893
Article Type:	Research
Date Submitted by the Author:	17-Mar-2017
Complete List of Authors:	Purrucker, Jan; Heidelberg University Hospital, Department of Neurology Härtig, Florian; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Richter, Hardy; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Engelbrecht, Andreas; Heidelberg University Hospital, Department of Neurology Hartmann, Johannes; Eberhard Karls Universitat Tubingen, Department of Medical Informatics Auer, Jonas; Universitat Stuttgart, Department of Computer Science & Software Engineering Hametner, Christian; University of Heidelberg, Neurology Popp, Erik; UniversitatsKlinikum Heidelberg, Department of Anesthesiology Ringleb, Peter; Heidelberg University Hospital, Department of Neurology Nagel, Simon; Heidelberg University Hospital, Department of Neurology Poli, Sven; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Emergency medicine
Keywords:	Stroke < NEUROLOGY, ACCIDENT & EMERGENCY MEDICINE, VASCULAR MEDICINE

Prehospital stroke recognition, severity grading, and prediction of large vessel occlusion using a single clinical scale only

Jan C. Purrucker,[1] Florian Härtig,[2] Hardy Richter,[2] Andreas Engelbrecht,[1] Johannes Hartmann,[3] Jonas Auer,[4] Christian Hametner,[1] Erik Popp,[5] Peter A. Ringleb,[1] Simon Nagel[1] and Sven Poli [2]

Author affiliations:

[1] Department of Neurology, Heidelberg University Hospital, Heidelberg, Germany

[2] Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany

[3] Department of Medical Informatics, University of Tuebingen, Tuebingen, Germany

[4] Department of Computer Science & Software Engineering, University of Stuttgart, Stuttgart, Germany

[5] Department of Anesthesiology, University of Heidelberg, Heidelberg, Germany

Corresponding author:

Sven Poli, MD MSc; Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany; Phone: +49 7071 2968300, Fax: +49 7071 2925047, E-mail: sven.poli@icloud.com

Keywords: Stroke severity grading, triage, large vessel occlusion, emergency medicine

Word-Count: Abstract: 247; Main Manuscript: 2626.

ABSTRACT

Objective: To develop the first fully NIH Stroke Scale (NIHSS)-compatible, all-in-one scale for rapid and comprehensive prehospital stroke assessment including stroke recognition, severity grading and progression monitoring as well as prediction of large vessel occlusion (LVO).

Methods: Emergency medical services (EMS) personnel and stroke physicians (N = 326) rated each item of the NIHSS regarding suitability for prehospital use; best-rated items were included. Stroke recognition was evaluated retrospectively in 689 consecutive patients with acute stroke or stroke mimics, prediction of LVO in 741 consecutive ischemic stroke patients with acute vessel imaging independent of admission-NIHSS score.

Results: Nine of the NIHSS-items were rated as "suitable for prehospital use". After excluding two items in order to increase specificity, the final scale (termed shortened NIHSS for EMS, sNIHSS-EMS), it consists of 'level of consciousness', 'facial palsy', 'motor arm/leg', 'sensory', 'language', and 'dysarthria'. Sensitivity for stroke recognition of the sNIHSS-EMS is 91% (95% confidence interval [CI] 86–94], specificity 52% (95% CI 47–56). Receiver operating curve analysis revealed an optimal cut-off point for LVO prediction of ≥ 6 (sensitivity 70% [95% CI 65–76], specificity 81% [95% CI 76–84], positive predictive value 70 [95% CI 65–75], area under the curve 0.81 [95% CI 0.78–0.84]). Test characteristics were non-inferior to non-comprehensive scales.

Conclusions: The sNIHSS-EMS may overcome the sequential use of multiple emergency stroke scales by permitting parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility allows for evaluation of stroke progression starting at the prehospital phase.

Strengths and limitations of this study

- Prehospital stroke assessment is increasingly gaining relevance in the era of endovascular interventions for large vessel occlusions. Sound triage decisions will have a major impact on patients' outcomes. As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.
- The new clinical scale (sNIHSS-EMS), developed and validated in this study, is the first scale permitting parallel stroke recognition, severity grading, and LVO prediction. Sequential use of multiple emergency stroke scales may thus be avoided.
- A multinational survey among different emergency medical systems and professions was performed to identify items suitable for use in prehospital emergency situations.
- The sNIHSS-EMS shares full compatibility with the in-hospital gold-standard NIHSS, but remains simple and easy to use.
- The scale will be incorporated into a prehospital stroke triage algorithm in a large regional stroke network, but no prospective data are available yet, which is acknowledged as a limitation.

INTRODUCTION

A considerable number of stroke scales for prehospital use have been published over recent years.[1, 2] However, all these scales only focus on single aspects of acute stroke care, i.e. either stroke recognition,[1, 2] severity grading,[3-8] early prediction of outcome,[5] prediction of thrombolysis,[9, 10] or large vessel occlusion (LVO).[11-19] Consequently, to provide a comprehensive prehospital stroke assessment, emergency medical services (EMS) personnel must apply at least two scales. However, this is time and resource consuming. Additionally, communication with receiving hospitals might be complicated by the use of multiple scales. Furthermore, most existing scales lack compatibility with the NIH Stroke Scale (NIHSS), the in-hospital 'gold-standard' for stroke severity grading.[2] This impedes the seamless evaluation of stroke progression from pre- to in-hospital care. In the era of endovascular treatment of LVO, decisions regarding direct emergency referrals to specialized comprehensive stroke centers will have a major impact on patients' outcomes.[20, 21] As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.

We present the development and validation of a novel comprehensive stroke scale, specifically designed for prehospital use with input from EMS. Our aim was to allow for parallel stroke recognition, severity grading and – owing to full NIHSS-compatibility – progression monitoring as well as LVO prediction.

METHODS

International online survey

We invited non-neurologic EMS personnel (paramedics and emergency physicians) and stroke physicians from Austria, Germany, and Switzerland to rate each individual NIHSS item regarding their applicability in a prehospital emergency setting. Invitations were sent out via the German Stroke Society (DSG), the German Society for Neuro-Intensive Care and Emergency Medicine (DGNI), as well as EMS providers.

Participation was voluntary, no financial incentive was offered, and participation was only allowed once. Non-neurologic EMS personnel do not use the NIHSS routinely and did not receive specific NIHSS training before the survey. For each NIHSS item, we created and provided a short video demonstrating in-hospital bedside assessment according to the NIHSS training instructions (a screenshot is shown as Figure 1A in the Appendix). Having watched the video, participants were asked to rate each NIHSS item regarding its suitability for prehospital use on a 6-item scale, ranging from 0 (most suitable) to 5 (most unsuitable). Ratings were automatically entered into a database together with name (optional), profession, professional experience, and place of work. Participation was possible from November 19th 2015, until April 15th 2016, the pre-specified closing date.

Patient cohorts

Test characteristics of the newly designed scale were calculated with regard to performance in stroke recognition and prediction of acute LVO using two distinct clinical cohorts described below.

For stroke recognition, we used a prospectively collected cohort of consecutive patients with acute ischemic or hemorrhagic stroke and stroke mimics, which had already served as a validation cohort in a previous comparison of existing stroke scales.[2] In summary, the database consists of pseudonymized data of consecutive patients (including comatose) with preclinical 'suspected acute CNS disorder' admitted to the Emergency Room of the Department of Neurology, Heidelberg University Hospital, Germany by EMS between November 2007 and August 2010. For all patients, a full-length NIHSS score assessed by certified raters was available at admission. The diagnostic reference standard was the diagnosis at hospital discharge. Cases were dichotomized (by the authors AE and CH) in stroke and non-stroke, i.e. stroke-mimics. AE and CH were

blinded for the admission NIHSS and sNIHSS-EMS scores. Details of the sample size calculation are described in [2].

Test characteristics regarding the prediction of LVO were calculated in a prospectively collected second cohort consisting of consecutive patients with acute ischemic stroke, admitted to the Department of Neurology, Tuebingen University Hospital, Germany between January 2013 and July 2015. In accordance with local standard operating procedures, all received acute vessel imaging on admission independent of stroke severity. Neuroradiological reports and original images were reviewed by the authors HR and SP for presence of acute LVO. HR and SP were blinded to patients' NIHSS scores. Cases were considered as LVO-positive if an acute symptomatic occlusion was present in one of the following arteries: common carotid artery (CCA), internal carotid artery (ICA), carotid T, middle cerebral artery (MCA, including M1/M2 segments), anterior cerebral artery (ACA), basilar artery (BA), or posterior cerebral artery (PCA).

Statistics

To determine suitable items for use in the prehospital phase, we analyzed the online survey response data set; median and interquartile ranges (IQR) were calculated. NIHSS-items receiving median scores of 0 and 1 were – as predefined – regarded eligible for further consideration. Rating differences between the professional groups (i.e. non-neurologic EMS personnel and stroke physicians) were determined using the Mann-Whitney-U test. For the calculation of test performance regarding stroke recognition, the sNIHSS-EMS score was dichotomized as indicative of stroke (score ≥ 1), or not (score = 0). Sensitivity (the proportion of stroke patients who had a positive test, i.e. indicative of stroke) and specificity (the proportion of non-stroke patients who had a negative test), positive predictive value (PPV), and negative predictive value (NPV) were calculated with 95% confidence intervals (CI). To determine the predictive power for LVO detection, we calculated sensitivity, specificity, PPV, and NPV, with

95% CI for each scale score ranging from 0 to 29 for the sNIHSS-EMS, and from 0 to 42 for the original NIHSS. Accuracy is reported additionally. Receiver operating curve (ROC) analysis was performed, area under the curve (AUC) and Youden´s index were calculated. For comparison of the sNIHSS-EMS with existing dedicated LVO prediction scales,[4, 11-13, 15, 16] we calculated the corresponding scores using the NIHSS-equivalents and cut-offs as stated in the original publications. Statistical comparison of AUCs was performed according to DeLong et al.[22] Calculation of the Los Angeles Motor Scale (LAMS) for our LVO cohort was not possible since the item "grip-strength" was not routinely documented. P values were 2-sided with values less than .05 considered statistically significant. SPSS (V23.0.0.2, IBM, New York, USA), MedCalc (V16.8.4, Ostend, Belgium) and GraphPad Prism (V6.0b, San Diego, California, USA) were used for data handling and analysis, and graphic presentation. This study was performed in accordance with the STARD guidelines for studies on diagnostic tests.

RESULTS

Scale development

A total of 326 participants responded to our international online survey (Austria, Germany, and Switzerland), with the majority (57%) representing non-neurologic EMS personnel (33% paramedics and 24% prehospital emergency physicians); 33% stroke physicians, and 10% not specified. Participants reported a high level of professional experience (>10 years, 45%; <5 years, 20%).

Nine of the NIHSS-items received a median score of 0 or 1 (equivalent to most suitable and suitable for prehospital use), whereas the items 'best gaze', 'visual', 'limb ataxia', and 'extinction' were rated as less suitable and thus removed from further analyses (Table 1A in the Appendix). Although rating by stroke physicians was more rigorous, item selection based on median ratings of 0 or 1 was not shifted by the professional vote (Table 1A).

We decided to exclude items 1b (LOC questions) and 1c (LOC commands). Despite being easily assessable and thus rated suitable for prehospital use, these two items are either present in the absence of stroke as frequent features of non-stroke conditions (e.g. dementia, infection or dehydration)[23] or heavily influenced by aphasia[24] and thus redundant for stroke recognition. The new 7-items scale was termed 'shortened NIHSS for emergency medical services' (sNIHSS-EMS; Table 2).

Stroke recognition and severity grading

In our stroke recognition validation cohort of 689 consecutive patients with 'suspected acute CNS disorder', 29% received 'stroke' as discharge diagnosis. Patients with ischemic stroke (n=200) had an admission-NIHSS of 9 (IQR 4–17), patients with hemorrhagic stroke (n=55) of 17 (IQR 5–35). Non-stroke patients (n = 489) had a median admission-NIHSS of 1 (IQR 0–6). The sNIHSS-EMS was found to have 90.5% (95% CI 85.6–94.2) sensitivity and 51.5% (95% CI 47.0–56.1) specificity for stroke

recognition (PPV 43.3% [95% CI 38.5–48.2], NPV 93.0% [95% CI 89.3–95.6]). Cross tabulations are shown in Table 3A in the Appendix. Excluding comatose patients (n = 49), sensitivity was 89.1% (95% CI 83.6–93.3) and specificity 54.2% (95% CI 49.5–58.8).

LVO prediction

In the distinct LVO validation cohort of consecutive 741 ischemic stroke patients with acute vessel imaging independent of their admission-NIHSS score (86.9% CTA; see Table A2 for patient characteristics), a ROC analysis of the sNIHSS-EMS regarding LVO prediction revealed a maximal Youden index at the cut-point of ≥ 6 (sensitivity 70.3% [95% CI 64.7–75.5], specificity 80.7% [95% CI 76.8–84.3]; Figure 1, Table 2). For comparison, in the original NIHSS, the maximal Youden index was calculated for a cut-point of ≥ 9 (Table 2). Combined re-inclusion of the NIHSS items 'visual', 'gaze' and 'extinction' improved test characteristics (AUC 0.826 vs. 0.808, p<0.001). Re-inclusion of singular items did not improve test characteristics.

We validated the sNIHSS-EMS against existing LVO prediction scales through applying them to our cohort and calculation of ROC and Youden indices (Table 3, Figure 1). No statistically significant differences compared to existing scales were found, except for the full-length NIHSS. Notably, due to characteristics of our cohort, external validation based on maximal Youden indices led to cut-points different from those reported in the respective original publications (Table 3).

DISCUSSION

The sNIHSS-EMS is the first comprehensive stroke scale to provide parallel stroke recognition, severity grading, and LVO prediction. Test characteristics with respect to stroke recognition and severity grading as well as identification of patients with large vessel occlusion are non-inferior to other, non-comprehensive, scales. Furthermore,

compatibility with the item assessment in the full-length NIHSS allows for continuous evaluation of the clinical course from pre- to in-hospital care. It may thus represent the ideal stroke scale for routine use in pre-hospital emergency medical care.

As previously shown by our work-group,[2] some of the available stroke severity scales[3, 5] may be used for stroke recognition with similar sensitivity and specificity when compared to scales developed for stroke recognition alone. Existing scales, however, either include items requiring complex assessment (such as extinction[12, 16]) or exclude items highly relevant for evaluation of stroke progression (such as level of consciousness, arm or leg motor function[4, 5, 12]).

Sensitivity of the sNIHSS-EMS regarding stroke recognition (91%) was superior to previously published results for the simpler CPSS (85%) and FAST (87%) evaluated in the same cohort of patients.[2] In contrast, specificity (52%) was lower compared to the CPSS (65%) and FAST (64%).[2] As the overall burden of a missed stroke outweighs the potentially increased workload of emergency departments, higher sensitivity may be considered more relevant. Simpler stroke scales may provide a slightly faster initial assessment, but subsequently require the use of at least one additional scale to determine stroke severity or predict LVO. The use of multiple scales, however, may be error-prone and complicates communication with receiving hospitals.

According to recent European and American recommendations, clinical screening tools may be considered in order to facilitate direct transport of patients with suspected LVO to Comprehensive Stroke Centers (CSC) with endovascular facility.[21, 25] For LVO prediction, our analysis revealed a maximum Youden index for the cut-point of ≥ 6 for the sNIHSS-EMS and, in accordance with previous findings, 9 for the original NIHSS.[26] Importantly, to adjust for hospital capacities and local stroke network requirements, this threshold can be adapted: higher cut-points result in an increased specificity (Table 2) leading to reduced numbers of patients bypassing Acute Stroke

Ready Hospitals (ASRH) or Primary Stroke Centers (PSC) without endovascular facility.

The NIHSS items 'visual', 'gaze' and 'extinction' are part of some dedicated LVO prediction scales,[11, 12, 15, 16] but were not included in the sNIHSS-EMS due to unfavorable ratings regarding prehospital assessability. Re-inclusion of each separate item did not result in the presumed higher predictive value for LVO detection. Only combined re-inclusion of all three rejected items led to marginally enhanced test characteristics, but would result in a significantly increased number of complex-to-assess items and thus an inconvenient scale.

For comparison with existing scales, we externally validated dedicated LVO prediction scales in our cohort by using the cut-points as provided in the original publications and found the sNIHSS-EMS to offer comparable sensitivity and specificity (Table 4). Better test characteristics reported in the original publications for some scales may be due to differences in the definition of LVO (e.g. the 3I-SS focused on carotid T and M1 occlusions only,[11] while the LAMS also included M3/4 occlusions[17]).

LVO prediction by clinical scales has recently been criticized due to the high false-negative rate compared to vessel imaging.[18, 27] The sNIHSS-EMS is not intended to substitute in-hospital acute vessel imaging,[18] and prehospital acute vessel imaging is still an exception.[28] Currently, mainly due to the narrow time window for effective intravenous thrombolysis, patients are transferred to the closest stroke center regardless of LVO suspicion. In the era of interventional thrombectomy however, ASRH or PSC may have to be bypassed in favor of CSC with endovascular facility in sensibly selected cases.

Based on clinical criteria alone, the sNIHSS-EMS identifies the majority of patients with acute LVO, i.e. those patients who might benefit from a direct transfer to CSC with endovascular facility. In addition, the minority of LVO patients not bypassed to

endovascular ready CSC (i.e. total score < 6 despite LVO) are not lost to endovascular therapy since secondary transportation to an endovascular ready CSC is still possible. The sNIHSS-EMS is designed to permit the monitoring of stroke progression from pre- to in-hospital care on the item-level, a feature that has been neglected in other scales. Clinical implications include the earlier recognition of symptom fluctuation with consequences e.g. for blood pressure management or selection of imaging modality. In practice, if a '2' is scored for 'Motor Leg left' on the sNIHSS-EMS, a '4' on the same item during routine NIHSS evaluation in the ER points to early clinical deterioration. Clinical scores using merged items (e.g. 'hemiparesis'[11] or 'language/dysarthria'[9]) or modified item scoring (e.g. motor function scoring from 0 to 2 instead of 0 to 4 [12, 13, 15-17]) impede seamless monitoring of symptom progression.

Despite the positive aspects of the sNIHSS-EMS, some limitations of the present study require further discussion. Test characteristics regarding LVO prediction were calculated in a cohort of patients with confirmed ischemic stroke because determination of the 'true' LVO prediction threshold is only possible in a cohort without stroke mimics or hemorrhagic stroke. However, although this approach is in concordance with methods used in the past in the design of dedicated LVO prediction scales,[13, 15, 17] future prospective validation in the prehospital target population will be necessary to determine prevalence-dependent test characteristics. As patients with stroke mimics (and thus no LVO) exhibit low NIHSS scores, inclusion of these cases into the analyses, would lead to an increased specificity of our cut-points. The sNIHSS-EMS is not able to differentiate between ischemic and hemorrhagic stroke. This might not be a disadvantage as severely affected hemorrhagic stroke patients benefit from direct admission to a CSC with neurological intensive care capacity.[29]

As a strength of this study, LVO was evaluated by CTA or MRA, and not with less accurate duplex sonography as done in previous studies evaluating LVO prediction scales.[12, 17] The sNIHSS-EMS was primarily designed to fulfill requirements for

prehospital use. Although kept simple, additional training on the new scale is recommended. Participation of different EMS systems in design and derivation of the sNIHSS-EMS enhances generalizability to further EMS systems around the world. Moreover, the sNIHSS-EMS may also serve in telemedicine with usually non-neurologic physicians performing the initial patient examination.

CONCLUSION

The sNIHSS-EMS may overcome the need for sequential use of multiple emergency stroke scales by enabling parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility permits evaluation of stroke progression starting from the prehospital phase. Offering comparable test characteristics as dedicated scales, the sNIHSS-EMS may be a promising tool for rapid and comprehensive prehospital stroke assessment and triage.

Acknowledgments

We thank all survey participants for their contribution. Those who gave consent to the publication of their names are listed in the Appendix. We thank Louise Alice Härtig for language revision of the manuscript.

Ethical approval

Ethic approvals were obtained from the ethics committee of the Medical Faculty Heidelberg and the ethics committee at the University Hospital Tuebingen, Germany (protocol-numbers S-109/2013 and 648/2015BO2, respectively). Written informed consent was waived.

Funding

This study was investigator initiated, without funding.

Contributorship

SP and JCP conceived and designed the study. EP provided the EMS data. HR and SP created and validated the LVO prediction database. SP, JCP, AE and CH collected and analyzed the data. JH and JA developed and maintained the online survey. JCP, FH, SP drafted the article. JCP, HR, FH, CH, PAR, SN, SP revised the manuscript.

Guarantor

SP and JCP take the responsibility for the paper as a whole.

Declaration of Conflicting Interests

Personal fees, travel support, speaker honoraria, or research grants were received from Bayer Healthcare (FH, HR, PR, SN, SP), BeneChill (SP), BMS Pfizer (PR, SP), Boehringer-Ingelheim (JCP, PR, SN, SP), Brainomix Ltd. (SN), Covidien (SP), C.R. Bard (SP), EMCOOLS (SP), Helena Laboratories (FH, SP), HVM Medical (SP), Medtronic (SN), Pfizer (JCP, SN), Raumedic (SP), ZOLL (SP), The other authors report no conflicts of interest.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: No additional data are available.

References

1. Brandler ES, Sharma M, Sinert RH, *et al.* Prehospital stroke scales in urban environments: a systematic review. *Neurology* 2014;82:2241-9.

2. Purrucker JC, Hametner C, Engelbrecht A, *et al.* Comparison of stroke recognition and stroke severity scores for stroke detection in a single cohort. *J Neurol Neurosurg Psychiatry* 2015;86:1021-8.

3. Kimura K, Inoue T, Iguchi Y, *et al.* Kurashiki prehospital stroke scale. *Cerebrovasc Dis* 2008;25:189-91.

4. 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.

5. Tirschwell DL, Longstreth WT, Jr., Becker KJ, *et al.* Shortening the NIH Stroke scale for use in the prehospital setting. *Stroke* 2002;33:2801-6.

6. Lyden P, Brott T, Tilley B, *et al.* Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke* 1994;25:2220-6.

7. Meyer BC, Hemmen TM, Jackson CM, *et al.* Modified National Institutes of Health Stroke Scale for use in stroke clinical trials: prospective reliability and validity. *Stroke* 2002;33:1261-6.

8. Whelley-Wilson CM, Newman GC. A stroke scale for emergency triage. *J Stroke Cerebrovasc Dis* 2004;13:247-53.

9. Iguchi Y, Kimura K, Watanabe M, *et al.* Utility of the Kurashiki Prehospital Stroke Scale for hyperacute stroke. *Cerebrovasc Dis* 2011;31:51-6.

10. Hasegawa Y, Sasaki N, Yamada K, *et al.* Prediction of thrombolytic therapy after stroke-bypass transportation: the Maria Prehospital Stroke Scale score. *J Stroke Cerebrovasc Dis* 2013;22:514-9.

11. Singer OC, Dvorak F, du Mesnil de Rochemont R, *et al.* A simple 3-item stroke scale: comparison with the National Institutes of Health Stroke Scale and prediction of middle cerebral artery occlusion. *Stroke* 2005;36:773-6.

12. Perez 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.

13. Katz BS, McMullan JT, Sucharew H, *et al.* Design and validation of a prehospital scale to predict stroke severity: Cincinnati Prehospital Stroke Severity Scale. *Stroke* 2015;46:1508-12.

14. 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;9:122-6.

15. Hastrup S, Damgaard D, Johnsen SP, *et al.* Prehospital Acute Stroke Severity Scale to Predict Large Artery Occlusion: Design and Comparison With Other Scales. *Stroke* 2016;47:1772-6.

16. Lima FO, Silva GS, Furie KL, *et al.* Field Assessment Stroke Triage for Emergency Destination: A Simple and Accurate Prehospital Scale to Detect Large Vessel Occlusion Strokes. *Stroke* 2016;47:1997-2002.

17. Nazliel B, Starkman S, Liebeskind DS, *et al.* A brief prehospital stroke severity scale identifies ischemic stroke patients harboring persisting large arterial occlusions. *Stroke* 2008;39:2264-7.

18. Heldner MR, Hsieh K, Broeg-Morvay A, *et al.* Clinical prediction of large vessel occlusion in anterior circulation stroke: mission impossible? *J Neurol* 2016;263:1633-40.

19. Vanacker P, Heldner MR, Amiguet M, *et al.* Prediction of Large Vessel Occlusions in Acute Stroke: National Institute of Health Stroke Scale Is Hard to Beat. *Crit Care Med* 2016;44:e336-43.

20. Rinaldo L, Brinjikji W, McCutcheon BA, *et al.* Hospital transfer associated with increased mortality after endovascular revascularization for acute ischemic stroke. *J Neurointerv Surg* 2016.

21. 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* 2016.

22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.

23. Nor AM, Davis J, Sen B, *et al.* The Recognition of Stroke in the Emergency Room (ROSIER) scale: development and validation of a stroke recognition instrument. *Lancet Neurol* 2005;4:727-34.

24. Majerus S, Bruno MA, Schnakers C, *et al.* The problem of aphasia in the assessment of consciousness in brain-damaged patients. *Prog Brain Res* 2009;177:49-61.

25. Fiehler J, Cognard C, Gallitelli M, *et al.* European Recommendations on Organisation of Interventional Care in Acute Stroke (EROICAS). *Int J Stroke* 2016;11:701-16.

26. Heldner MR, Zubler C, Mattle HP, *et al.* National Institutes of Health stroke scale score and vessel occlusion in 2152 patients with acute ischemic stroke. *Stroke* 2013;44:1153-7.

27. Turc G, Maier B, Naggara O, *et al.* Clinical Scales Do Not Reliably Identify Acute Ischemic Stroke Patients With Large-Artery Occlusion. *Stroke* 2016;47:1466-72.

28. John S, Stock S, Masaryk T, *et al.* Performance of CT Angiography on a Mobile Stroke Treatment Unit: Implications for Triage. *J Neuroimaging* 2016;26:391-4.

29. Diringer MN, Edwards DF. Admission to a neurologic/neurosurgical intensive care unit is associated with reduced mortality rate after intracerebral hemorrhage. *Crit Care Med* 2001;29:635-40.

Table 1 The shortened NIH Stroke Scale for Emergency Medical Services (sNIHSS-EMS).

No.	sNIHSS-EMS item	Equivalent to the NIHSS-item	Range
1	Level of Consciousness	1a	0 – 3
2	Facial Palsy	4	0 – 3
3	Motor Arm (R+L)	5	0 – 4 / UN
4	Motor Leg (R+L)	6	0 – 4 / UN
5	Sensory	8	0 – 2
6	Best Language	9	0 – 3
7	Dysarthria	10	0 – 2 / UN
	Sum	–	0 – 29

Range indicates possible scores;

Abbreviations: R+L = right and left; UN = untestable (motor items: amputation or joint fusion, dysarthria: intubation or other physical barrier).

Table 2 Cut-off points for prediction of acute large vessel occlusion.

Cut-off Point	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy	*J*
sNIHSS-EMS						
5	74.8 (69.4 – 79.7)	73.4 (69.1 – 77.4)	66.4 (59.0 – 69.5)	81.9 (77.8 – 85.6)	74.0	0.482
6*	70.3 (64.7 – 75.5)	80.7 (76.8 – 84.3)	70.1 (64.5 – 75.3)	80.9 (76.9 – 84.4)	76.7	0.511
7	65.2 (59.4 – 70.6)	85.8 (82.2 – 88.9)	74.7 (68.9 – 79.9)	79.3 (75.4 – 82.8)	77.7	0.510
NIHSS						
8	72.4 (66.9 – 77.5)	80.7 (76.8 – 84.3)	70.7 (65.2 – 75.8)	82.0 (78.1 – 85.4)	77.5	0.531
9*	69.3 (63.7 – 74.6)	85.4 (81.8 – 88.5)	75.3 (69.7 – 80.3)	81.2 (77.4 – 84.6)	79.1	0.547
10	65.9 (60.1 – 71.3)	88.0 (84.7 – 90.9)	78.0 (72.2 – 83.0)	80.0 (76.2 – 83.5)	79.4	0.539

Abbreviations: NIHSS = NIH Stroke Scale; sNIHSS-EMS, shortened NIHSS for emergency medical services.

Data are % (95% CI). *J* indicates Youden's-Index, * indicates the optimal cut-off according to the Youden index.

Table 3 Comparison of clinical scales for prehospital prediction of large vessel occlusions

	3I-SS	**LAMS**	**RACE**	**CPSSS**	**FAST-ED**	**PASS**	**sNIHSS-EMS**
Reference	11	17	12	13	16	15	–
Scale characteristics							
No. of items assessed	3	3	5*	3	5	3	7
Score range	0–6	0–5	0–9	0–4	0–9	0–3	0-29
NIHSS compatible item assessment	–	–	–	–	–	–	●
Stroke Recognition	–	–	–	–	–	–	●
Stroke severity grading	–	(●)	(●)	–	–	–	●
Large vessel occlusion prediction	●	●	●	●	●	●	●
LVO prediction, test characteristics, own cohort (N = 741, 44% LVO)							
Cut-point used †	≥ 4	≥ 4	≥ 5	≥ 2	≥ 4	≥ 2	≥ 6
Sensitivity	40%	–‡	59%	59%	60%	68%	70%

Specificity	95%	–‡	91%	89%	90%	84%	81%
PPV	85%	–‡	81%	77%	80%	74%	70%
NPV	71%	–‡	78%	77%	78%	81%	81%

LVO prediction, test characteristics, original cohorts§

Cohort (N (%LVO))	83 (35%)	119 (62%)	357 (21%)**	303 (73%)	727 (33%)	3127 (35%)††	–
Sensitivity	67%	81%	85%	83%	61%	66%	–
Specificity	92%	89%	68%	40%	89%	83%	–
PPV	74%	nd	42%	nd	72%	68%	–
NPV	89%	nd	94%	nd	82%	81%	–

Abbreviations: nd: no data.

* If right sided hemiparesis, aphasia is assessed, if left sided hemiparesis, agnosia.

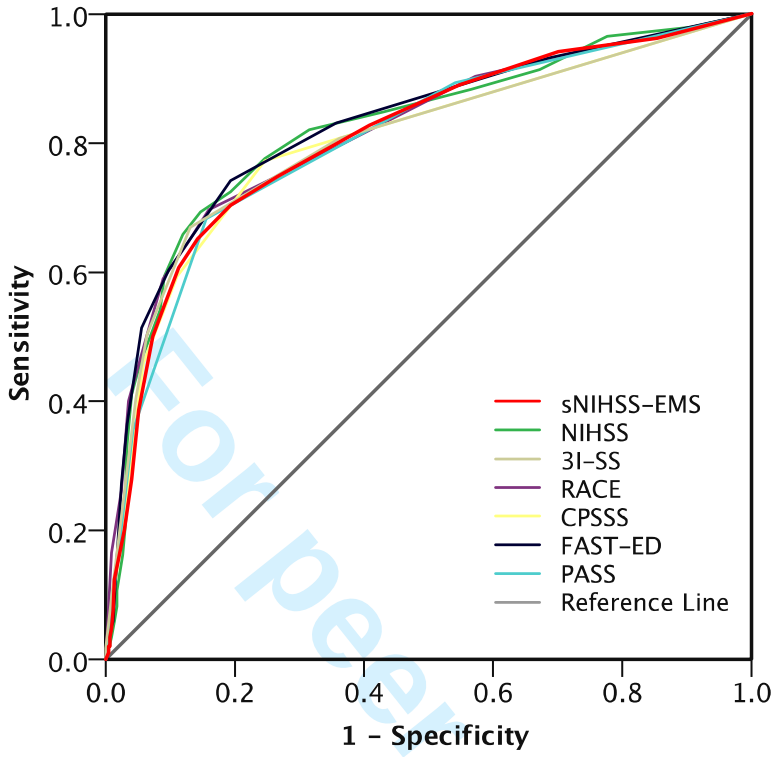
† Cut-points according to original publications. Based on the Youden indices calculated from our data, optimal cut-points are different: 3I-SS ≥ 2, RACE ≥ 3, CPSSS ≥ 1, FAST-ED ≥ 3.

‡ Grip-strength was not routinely documented, therefore external validation of the LAMS was not possible.

§ Definition of large vessel occlusions according to original publications (3I-SS: carotid-T or M1; LAMS: ICA, M1, M2, M3/4, ACA; RACE: terminal ICA, M1, tandem CCA/ICA+M1, BA; CPSSS: ICA, M1, tandem ICA+M2, BA; FAST-ED: ICA, M1, M2, BA; PASS: “visible clot in the anterior or posterior circulation on CTA or MRA“; abbreviations within the main text).

** Including cases assessed by transcranial duplex only (N = 197).

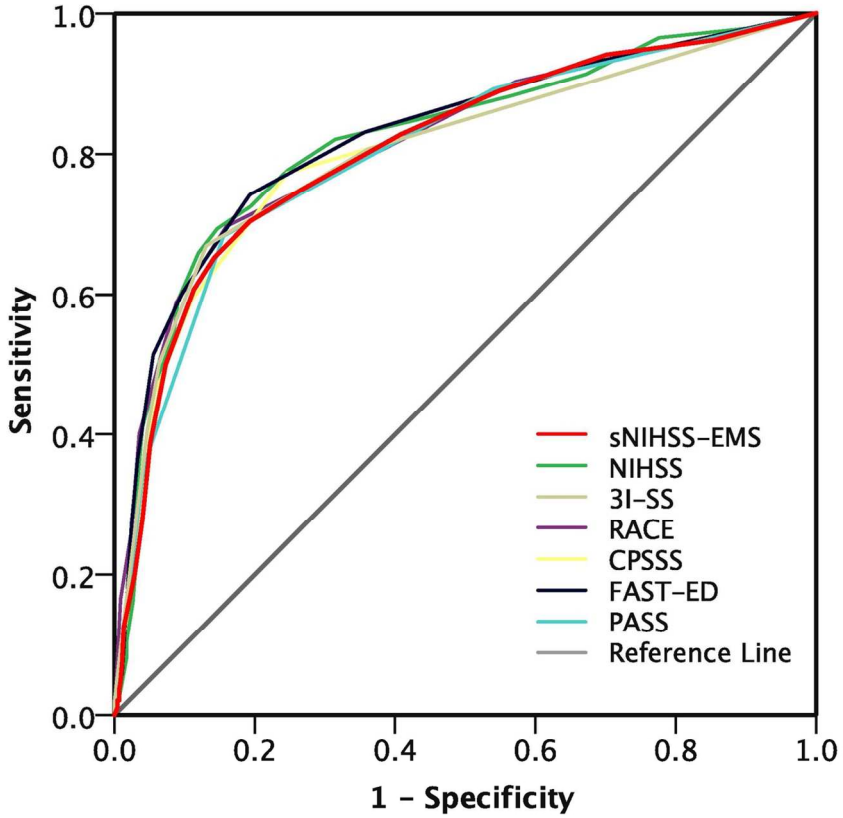
†† Only patients who received intravenous tPA; 2/3 of entire cohort were taken as a random sample for derivation. In the remaining 1/3, sensitivity was 61%, specificity 83%.



Scale	Area Under the Curve (95% CI)	*p*
sNIHSS-EMS	0.808 (0.775 – 0.841)	*ref.*
NIHSS	0.821 (0.788 – 0.854)	0.019
3I-SS	0.804 (0.770 – 0.838)	0.697
RACE	0.816 (0.784 – 0.849)	0.314
CPSSS	0.802 (0.767 – 0.836)	0.605
FAST-ED	0.825 (0.793 – 0.857)	0.078
PASS	0.805 (0.771 – 0.838)	0.766

Figure 1. Receiver operating curves for prediction of acute large vessel occlusion.

Abbreviations: AUC: area under the curve; ref.: reference.



Scale	Area Under the Curve (95% CI)	*p*
sNIHSS-EMS	0.808 (0.775 – 0.841)	*ref.*
NIHSS	0.821 (0.788 – 0.854)	0.019
3I-SS	0.804 (0.770 – 0.838)	0.697
RACE	0.816 (0.784 – 0.849)	0.314
CPSSS	0.802 (0.767 – 0.836)	0.605
FAST-ED	0.825 (0.793 – 0.857)	0.078
PASS	0.805 (0.771 – 0.838)	0.766

116x150mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4/5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5/6
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5/6
	9	Whether participants formed a consecutive, random or convenience series	5/6
Test methods	**10a**	Index test, in sufficient detail to allow replication	4-7
	10b	Reference standard, in sufficient detail to allow replication	5/6
	11	Rationale for choosing the reference standard (if alternatives exist)	5/6, 7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6/7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6/7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6/7
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6/7
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6/7
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6/7
	18	Intended sample size and how it was determined	5/6
RESULTS			
Participants	**19**	Flow of participants, using a diagram	NA
	20	Baseline demographic and clinical characteristics of participants	8, +Appendix.
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	Appendix, previousl. published.
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Appendix
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	8/9
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	3; 11,12,13
OTHER INFORMATION			
	28	Registration number and name of registry	NA (IRB-Prot., 12)
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14

BMJ Open

Design and validation of a clinical scale for prehospital stroke recognition, severity grading, and prediction of large vessel occlusion – the shortened NIH stroke scale for emergency medical services

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016893.R1
Article Type:	Research
Date Submitted by the Author:	29-May-2017
Complete List of Authors:	Purrucker, Jan; Heidelberg University Hospital, Department of Neurology Härtig, Florian; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Richter, Hardy; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Engelbrecht, Andreas; Heidelberg University Hospital, Department of Neurology Hartmann, Johannes; Eberhard Karls Universitat Tubingen, Department of Medical Informatics Auer, Jonas; Universitat Stuttgart, Department of Computer Science & Software Engineering Hametner, Christian; University of Heidelberg, Neurology Popp, Erik; UniversitatsKlinikum Heidelberg, Department of Anesthesiology Ringleb, Peter; Heidelberg University Hospital, Department of Neurology Nagel, Simon; Heidelberg University Hospital, Department of Neurology Poli, Sven; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Emergency medicine
Keywords:	Stroke < NEUROLOGY, ACCIDENT & EMERGENCY MEDICINE, VASCULAR MEDICINE

Design and validation of a clinical scale for prehospital stroke recognition, severity grading, and prediction of large vessel occlusion – the shortened NIH stroke scale for emergency medical services

Jan C. Purrucker,[1] Florian Härtig,[2] Hardy Richter,[2] Andreas Engelbrecht,[1] Johannes Hartmann,[3] Jonas Auer,[4] Christian Hametner,[1] Erik Popp,[5] Peter A. Ringleb,[1] Simon Nagel[1] and Sven Poli [2]

Author affiliations:

[1] Department of Neurology, Heidelberg University Hospital, Heidelberg, Germany

[2] Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany

[3] Department of Medical Informatics, University of Tuebingen, Tuebingen, Germany

[4] Department of Computer Science & Software Engineering, University of Stuttgart, Stuttgart, Germany

[5] Department of Anesthesiology, University of Heidelberg, Heidelberg, Germany

Corresponding author:

Sven Poli, MD MSc; Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany; Phone: +49 7071 2968300, Fax: +49 7071 2925047, E-mail: sven.poli@icloud.com

Keywords: Stroke severity grading, triage, large vessel occlusion, emergency medicine

Word-Count: Abstract: 245; Main Manuscript: 2765.

ABSTRACT

Objective: To develop a NIH Stroke Scale (NIHSS)-compatible, all-in-one scale for rapid and comprehensive prehospital stroke assessment including stroke recognition, severity grading and progression monitoring as well as prediction of large vessel occlusion (LVO).

Methods: Emergency medical services (EMS) personnel and stroke physicians (N = 326) rated each item of the NIHSS regarding suitability for prehospital use; best-rated items were included. Stroke recognition was evaluated retrospectively in 689 consecutive patients with acute stroke or stroke mimics, prediction of LVO in 741 consecutive ischemic stroke patients with acute vessel imaging independent of admission-NIHSS score.

Results: Nine of the NIHSS-items were rated as "suitable for prehospital use". After excluding two items in order to increase specificity, the final scale (termed shortened NIHSS for EMS, sNIHSS-EMS), it consists of 'level of consciousness', 'facial palsy', 'motor arm/leg', 'sensory', 'language', and 'dysarthria'. Sensitivity for stroke recognition of the sNIHSS-EMS is 91% (95% confidence interval [CI] 86–94], specificity 52% (95% CI 47–56). Receiver operating curve analysis revealed an optimal cut-off point for LVO prediction of ≥ 6 (sensitivity 70% [95% CI 65–76], specificity 81% [95% CI 76–84], positive predictive value 70 [95% CI 65–75], area under the curve 0.81 [95% CI 0.78–0.84]). Test characteristics were non-inferior to non-comprehensive scales.

Conclusions: The sNIHSS-EMS may overcome the sequential use of multiple emergency stroke scales by permitting parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility allows for evaluation of stroke progression starting at the prehospital phase.

Strengths and limitations of this study

- Prehospital stroke assessment is increasingly gaining relevance in the era of endovascular interventions for large vessel occlusions. Sound triage decisions will have a major impact on patients' outcomes. As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.
- The new clinical scale (sNIHSS-EMS), developed and validated in this study, is the first scale assessed for parallel stroke recognition, severity grading, and LVO prediction. Sequential use of multiple emergency stroke scales may thus be avoided.
- A multinational survey among different emergency medical systems and professions was performed to identify items suitable for use in prehospital emergency situations.
- The sNIHSS-EMS shares full compatibility with the in-hospital gold-standard NIHSS, but remains simple and easy to use.
- The scale will be incorporated into a prehospital stroke triage algorithm in a large regional stroke network, but no prospective data are available yet, which is acknowledged as a limitation.

INTRODUCTION

A considerable number of stroke scales for prehospital use have been published over recent years.[1, 2] However, most of these scales only focus on single aspects of acute stroke care, i.e. either stroke recognition,[1, 2] early prediction of outcome,[3] prediction of thrombolysis,[4, 5] or severity grading and large vessel occlusion (LVO). [3, 6-18] Consequently, to provide a comprehensive prehospital stroke assessment, emergency medical services (EMS) personnel must apply at least two scales. Furthermore, the majority of existing scales lack compatibility with the NIH Stroke Scale (NIHSS), the in-hospital 'gold-standard' for stroke severity grading.[2] This impedes the seamless evaluation of stroke progression from pre- to in-hospital care. In the era of endovascular treatment of LVO, decisions regarding direct emergency referrals to specialized comprehensive stroke centers will have a major impact on patients' outcomes.[19, 20] As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.

We present the development and validation of a novel comprehensive stroke scale, specifically designed for prehospital use with input from EMS. Our aim was to allow for parallel stroke recognition, severity grading and – owing to full NIHSS-compatibility – progression monitoring as well as LVO prediction.

METHODS

International online survey

We invited non-neurologic EMS personnel (paramedics and emergency physicians) and stroke physicians from Austria, Germany, and Switzerland to rate each individual NIHSS item regarding their applicability in a prehospital emergency setting. Invitations were sent out via the German Stroke Society (DSG), the German Society for Neuro-Intensive Care and Emergency Medicine (DGNI), as well as EMS providers. Participation was voluntary, no financial incentive was offered, and participation was

only allowed once. Non-neurologic EMS personnel do not use the NIHSS routinely and did not receive specific NIHSS training before the survey. For each NIHSS item, we created and provided a short video demonstrating in-hospital bedside assessment according to the NIHSS training instructions (a screenshot is shown as Figure 1A in the Appendix). Having watched the video, participants were asked to rate each NIHSS item regarding its suitability for prehospital use on a 6-item scale, ranging from 0 (most suitable) to 5 (most unsuitable). Ratings were automatically entered into a database together with name (optional), profession, professional experience, and place of work. Participation was possible from November 19th 2015, until April 15th 2016, the pre-specified closing date.

Patient cohorts

Test characteristics of the newly designed scale were calculated with regard to performance in stroke recognition and prediction of acute LVO using two distinct clinical cohorts described below.

For stroke recognition, we used a prospectively collected cohort of consecutive patients with acute ischemic or hemorrhagic stroke and stroke mimics, which had already served as a validation cohort in a previous comparison of existing stroke scales.[2] In summary, the database consists of pseudonymized data of consecutive patients (including comatose) with preclinical 'suspected acute CNS disorder' admitted to the Emergency Room of the Department of Neurology, Heidelberg University Hospital, Germany by EMS between November 2007 and August 2010. For all patients, a full-length NIHSS score assessed by certified raters was available at admission. The diagnostic reference standard was the diagnosis at hospital discharge. Cases were dichotomized (by the authors AE and CH) in stroke and non-stroke, i.e. stroke-mimics. AE and CH were blinded for the admission NIHSS and sNIHSS-EMS scores.

Test characteristics regarding the prediction of LVO were calculated in a prospectively collected second cohort consisting of consecutive patients with acute ischemic stroke, admitted to the Department of Neurology, Tuebingen University Hospital, Germany between January 2013 and July 2015. In accordance with local standard operating procedures, all received acute vessel imaging on admission independent of stroke severity. Neuroradiological reports and original images were reviewed by the authors HR and SP for presence of acute LVO. HR and SP were blinded to patients' NIHSS scores. Cases were considered as LVO-positive if an acute symptomatic occlusion was present in one of the following arteries: common carotid artery (CCA), internal carotid artery (ICA), carotid T, middle cerebral artery (MCA, including M1/M2 segments), anterior cerebral artery (ACA), basilar artery (BA), or posterior cerebral artery (PCA).

Statistics

To determine suitable items for use in the prehospital phase, we analyzed the online survey response data set; median and interquartile ranges (IQR) were calculated. NIHSS-items receiving median scores of 0 and 1 were – as predefined – regarded eligible for further consideration. Rating differences between the professional groups (i.e. non-neurologic EMS personnel and stroke physicians) were determined using the Mann-Whitney-U test. For the calculation of test performance regarding stroke recognition, the sNIHSS-EMS score was dichotomized as indicative of stroke (score ≥ 1), or not (score = 0). Sensitivity (the proportion of stroke patients who had a positive test, i.e. indicative of stroke) and specificity (the proportion of non-stroke patients who had a negative test), positive predictive value (PPV), and negative predictive value (NPV) were calculated with 95% confidence intervals (CI). Details of the sample size calculation are described in the extended methods in the Appendix. To determine the predictive power for LVO detection, we calculated sensitivity, specificity, PPV, and NPV, with 95% CI for each scale score ranging from 0 to 29 for the sNIHSS-EMS, and

from 0 to 42 for the original NIHSS. Accuracy is reported additionally. Receiver operating curve (ROC) analysis was performed, area under the curve (AUC) and Youden´s index were calculated. For comparison of the sNIHSS-EMS with existing dedicated LVO prediction scales,[7, 10-12, 14, 15] we calculated the corresponding scores using the NIHSS-equivalents and cut-offs as stated in the original publications. Statistical comparison of AUCs was performed according to DeLong et al.[21] Calculation of the Los Angeles Motor Scale (LAMS) for our LVO cohort was not possible since the item "grip-strength" was not routinely documented. P values were 2-sided with values less than .05 considered statistically significant. SPSS (V23.0.0.2, IBM, New York, USA), MedCalc (V16.8.4, Ostend, Belgium) and GraphPad Prism (V6.0b, San Diego, California, USA) were used for data handling and analysis, and graphic presentation. This study was performed in accordance with the STARD guidelines for studies on diagnostic tests.

RESULTS

Scale development

326 (13%) of 2562 recipients responded to our international online survey (Austria, Germany, and Switzerland), with the majority (57%) representing non-neurologic EMS personnel (33% paramedics and 24% prehospital emergency physicians); 33% stroke physicians, and 10% not specified. Participants reported a high level of professional experience (>10 years, 45%; <5 years, 20%).

Nine of the NIHSS-items received a median score of 0 or 1 (equivalent to most suitable and suitable for prehospital use), whereas the items 'best gaze', 'visual', 'limb ataxia', and 'extinction' were rated as less suitable and thus removed from further analyses (Table 1A in the Appendix). Although rating by stroke physicians was more rigorous, item selection based on median ratings of 0 or 1 was not shifted by the professional vote (Table 1A).

We decided to exclude items 1b (LOC questions) and 1c (LOC commands). Despite being easily assessable and thus rated suitable for prehospital use, these two items are either present in the absence of stroke as frequent features of non-stroke conditions (e.g. dementia, infection or dehydration)[22] or heavily influenced by aphasia[23] and thus redundant for stroke recognition. The new 7-items scale was termed 'shortened NIHSS for emergency medical services' (sNIHSS-EMS; Table 2).

Stroke recognition and severity grading

In our stroke recognition validation cohort of 689 consecutive patients with 'suspected acute CNS disorder', 29% received 'stroke' as discharge diagnosis. Patients with ischemic stroke (n=200) had an admission-NIHSS of 9 (IQR 4–17), patients with hemorrhagic stroke (n=55) of 17 (IQR 5–35). Non-stroke patients (n = 489) had a median admission-NIHSS of 1 (IQR 0–6). The sNIHSS-EMS was found to have 90.5% (95% CI 85.6–94.2) sensitivity and 51.5% (95% CI 47.0–56.1) specificity for stroke

recognition (PPV 43.3% [95% CI 38.5–48.2], NPV 93.0% [95% CI 89.3–95.6]). Cross tabulations are shown in Table 3A in the Appendix. Excluding comatose patients (n = 49), sensitivity was 89.1% (95% CI 83.6–93.3) and specificity 54.2% (95% CI 49.5–58.8).

LVO prediction

In the distinct LVO validation cohort of consecutive 741 ischemic stroke patients with acute vessel imaging independent of their admission-NIHSS score (86.9% CTA; see Table A2 for patient characteristics), a ROC analysis of the sNIHSS-EMS regarding LVO prediction revealed a maximal Youden index at the cut-point of ≥ 6 (sensitivity 70.3% [95% CI 64.7–75.5], specificity 80.7% [95% CI 76.8–84.3]; Figure 1, Table 2). For comparison, in the original NIHSS, the maximal Youden index was calculated for a cut-point of ≥ 9 (Table 2). Combined re-inclusion of the NIHSS items 'visual', 'gaze' and 'extinction' improved test characteristics (AUC 0.826 vs. 0.808, p<0.001). Re-inclusion of singular items did not improve test characteristics. Exclusion of patients with coma (n=5) did not change the optimal cut-off and test characteristics (sensitivity 70.0% [64.4–75.3], specificity 81.1% [77.1–84.6]).

We validated the sNIHSS-EMS against existing LVO prediction scales through applying them to our cohort and calculation of ROC and Youden indices (Table 3, Figure 1). No statistically significant differences compared to existing scales were found, except for the full-length NIHSS, and the sNIHSS-8. Notably, due to characteristics of our cohort, external validation based on maximal Youden indices led to cut-points different from those reported in the respective original publications (Table 3).

DISCUSSION

The sNIHSS-EMS is the first comprehensive stroke scale assessed for parallel stroke recognition, severity grading, and LVO prediction. Test characteristics regarding identification of patients with large vessel occlusion are non-inferior to existing LVO prediction scales. Furthermore, compatibility with the item assessment in the full-length NIHSS allows for continuous evaluation of the clinical course from pre- to in-hospital care. It may thus represent the ideal stroke scale for routine use in pre-hospital emergency medical care.

As previously shown by our work-group,[2] some of the available stroke severity scales[3, 6] may be used for stroke recognition with similar sensitivity and specificity when compared to scales developed for stroke recognition alone. Existing scales, however, either include items requiring complex assessment (such as extinction[11, 15]) or exclude items highly relevant for evaluation of stroke progression (such as level of consciousness, arm or leg motor function[3, 7, 11]).

Sensitivity of the sNIHSS-EMS regarding stroke recognition (91%) was superior to previously published results for the simpler CPSS (85%) and FAST (87%) evaluated in the same cohort of patients.[2] In contrast, specificity (52%) was lower compared to the CPSS (65%) and FAST (64%).[2] As the overall burden of a missed stroke outweighs the potentially increased workload of emergency departments, higher sensitivity may be considered more relevant. Simpler stroke scales may provide a slightly faster initial assessment, but subsequently require the use of at least one additional scale to determine stroke severity or predict LVO. The use of multiple scales, however, may be error-prone and complicates communication with receiving hospitals.

According to recent European and American recommendations, clinical screening tools may be considered in order to facilitate direct transport of patients with suspected LVO to Comprehensive Stroke Centers (CSC) with endovascular facility.[20, 24] For LVO prediction, our analysis revealed a maximum Youden index for the cut-point of ≥ 6 for the sNIHSS-EMS and, in accordance with previous findings, 9 for the original

NIHSS.[25] Importantly, to adjust for hospital capacities and local stroke network requirements, this threshold can be adapted: higher cut-points result in an increased specificity (Table 2) leading to reduced numbers of patients bypassing Acute Stroke Ready Hospitals (ASRH) or Primary Stroke Centers (PSC) without endovascular facility.

The NIHSS items 'visual', 'gaze' and 'extinction' are part of some dedicated LVO prediction scales,[10, 11, 14, 15] but were not included in the sNIHSS-EMS due to unfavorable ratings regarding prehospital assessability. Re-inclusion of each separate item did not result in the presumed higher predictive value for LVO detection. Only combined re-inclusion of all three rejected items led to marginally enhanced test characteristics, but would result in a significantly increased number of complex-to-assess items and thus an inconvenient scale.

For comparison with existing scales, we externally validated dedicated LVO prediction scales in our cohort by using the cut-points as provided in the original publications and found the sNIHSS-EMS to offer comparable sensitivity and specificity (Table 4). Better test characteristics reported in the original publications for some scales may be due to differences in the definition of LVO (e.g. the 3I-SS focused on carotid T and M1 occlusions only,[10] while the LAMS also included M3/4 occlusions[16]). The sNIHSS-8, which had a higher AUC in the ROC analysis than the sNIHSS-EMS, was not developed for LVO prediction and includes items rejected by EMS personnel in our survey due to the complexity of correct assessment.

LVO prediction by clinical scales has recently been criticized due to the high false-negative rate compared to vessel imaging.[17, 26] The sNIHSS-EMS is not intended to substitute in-hospital acute vessel imaging,[17] and prehospital acute vessel imaging is still an exception.[27] Currently, mainly due to the narrow time window for effective intravenous thrombolysis, patients are transferred to the closest stroke center regardless of LVO suspicion. In the era of interventional thrombectomy however, ASRH or PSC

may have to be bypassed in favor of CSC with endovascular facility in sensibly selected cases.

Based on clinical criteria alone, the sNIHSS-EMS identifies the majority of patients with acute LVO, i.e. those patients who might benefit from a direct transfer to CSC with endovascular facility. In addition, the minority of LVO patients not bypassed to endovascular ready CSC (i.e. total score < 6 despite LVO) are not lost to endovascular therapy since secondary transportation to an endovascular ready CSC is still possible. The sNIHSS-EMS is designed to permit the monitoring of stroke progression from pre- to in-hospital care on the item-level, a feature that has been neglected in other scales. Clinical implications include the earlier recognition of symptom fluctuation with consequences e.g. for blood pressure management or selection of imaging modality. In practice, if a '2' is scored for 'Motor Leg left' on the sNIHSS-EMS, a '4' on the same item during routine NIHSS evaluation in the ER points to early clinical deterioration. Clinical scores using merged items (e.g. 'hemiparesis'[10] or 'language/dysarthria'[4]) or modified item scoring (e.g. motor function scoring from 0 to 2 instead of 0 to 4 [11, 12, 14-16]) impede seamless monitoring of symptom progression.

Despite the positive aspects of the sNIHSS-EMS, some limitations of the present study require further discussion. Test characteristics regarding LVO prediction were calculated in a cohort of patients with confirmed ischemic stroke because determination of the 'true' LVO prediction threshold is only possible in a cohort without stroke mimics or hemorrhagic stroke. However, although this approach is in concordance with methods used in the past in the design of dedicated LVO prediction scales,[12, 14, 16] future prospective validation in the prehospital target population will be necessary to determine prevalence-dependent test characteristics. We were not able to assess LVO prediction of the LAMS because the item 'grip strength' is not part of the NIHSS and thus, was not routinely documented in our cohort. According to a retrospective validation study in anterior circulation stroke, the sensitivity of the LAMS for LVO

prediction was reported as 81% (at a threshold of 4).[16] As patients with stroke mimics (and thus no LVO) exhibit low NIHSS scores, inclusion of these cases into the analyses, would lead to an increased specificity of our cut-points. The sNIHSS-EMS is not able to differentiate between ischemic and hemorrhagic stroke. This might not be a disadvantage as severely affected hemorrhagic stroke patients benefit from direct admission to a CSC with neurological intensive care capacity.[28] Despite involvement of EMS systems from three European countries, generalizability to further EMS systems around the world cannot be concluded. The low response-rate of our online survey makes a non-response bias likely. Due to the participants' high professional experience, one might have expected a shift of the suitability assessment towards more complex items. However, this was not observed. As a strength of this study, LVO was evaluated by CTA or MRA, and not with less accurate duplex sonography as done in previous studies evaluating LVO prediction scales.[11, 16] The sNIHSS-EMS was primarily designed to fulfill requirements for prehospital use. Although kept simple, additional training on the new scale is recommended. Moreover, the sNIHSS-EMS may also serve in telemedicine with usually non-neurologic physicians performing the initial patient examination.

CONCLUSION

The sNIHSS-EMS may overcome the need for sequential use of multiple emergency stroke scales by enabling parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility permits evaluation of stroke progression starting from the prehospital phase. Offering comparable test characteristics as dedicated scales, the sNIHSS-EMS may be a promising tool for rapid and comprehensive prehospital stroke assessment and triage.

Acknowledgments

We thank all survey participants for their contribution. Those who gave consent to the publication of their names are listed in the Appendix. We thank Louise Alice Härtig for language revision of the manuscript.

Ethical approval

Ethic approvals were obtained from the ethics committee of the Medical Faculty Heidelberg and the ethics committee at the University Hospital Tuebingen, Germany (protocol-numbers S-109/2013 and 648/2015BO2, respectively). Written informed consent was waived.

Funding

This study was investigator initiated, without funding.

Contributorship

SP and JCP conceived and designed the study. EP provided the EMS data. HR and SP created and validated the LVO prediction database. SP, JCP, AE and CH collected and analyzed the data. JH and JA developed and maintained the online survey. JCP, FH, SP drafted the article. JCP, HR, FH, CH, PAR, SN, SP revised the manuscript.

Guarantor

SP and JCP take the responsibility for the paper as a whole.

Declaration of Conflicting Interests

Personal fees, travel support, speaker honoraria, or research grants were received from Bayer Healthcare (FH, HR, PR, SN, SP), BeneChill (SP), BMS Pfizer (PR, SP), Boehringer-Ingelheim (JCP, PR, SN, SP), Brainomix Ltd. (SN), Covidien (SP), C.R.

Bard (SP), EMCOOLS (SP), Helena Laboratories (FH, SP), HVM Medical (SP), Medtronic (SN), Pfizer (JCP, SN), Raumedic (SP), ZOLL (SP), The other authors report no conflicts of interest.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: No additional data are available.

References

1. Brandler ES, Sharma M, Sinert RH, *et al.* Prehospital stroke scales in urban environments: a systematic review. *Neurology* 2014;82:2241-9.

2. Purrucker JC, Hametner C, Engelbrecht A, *et al.* Comparison of stroke recognition and stroke severity scores for stroke detection in a single cohort. *J Neurol Neurosurg Psychiatry* 2015;86:1021-8.

3. Tirschwell DL, Longstreth WT, Jr., Becker KJ, *et al.* Shortening the NIH Stroke scale for use in the prehospital setting. *Stroke* 2002;33:2801-6.

4. Iguchi Y, Kimura K, Watanabe M, *et al.* Utility of the Kurashiki Prehospital Stroke Scale for hyperacute stroke. *Cerebrovasc Dis* 2011;31:51-6.

5. Hasegawa Y, Sasaki N, Yamada K, *et al.* Prediction of thrombolytic therapy after stroke-bypass transportation: the Maria Prehospital Stroke Scale score. *J Stroke Cerebrovasc Dis* 2013;22:514-9.

6. Kimura K, Inoue T, Iguchi Y, *et al.* Kurashiki prehospital stroke scale. *Cerebrovasc Dis* 2008;25:189-91.

7. 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.

8. Meyer BC, Hemmen TM, Jackson CM, *et al.* Modified National Institutes of Health Stroke Scale for use in stroke clinical trials: prospective reliability and validity. *Stroke* 2002;33:1261-6.

9. Whelley-Wilson CM, Newman GC. A stroke scale for emergency triage. *J Stroke Cerebrovasc Dis* 2004;13:247-53.

10. Singer OC, Dvorak F, du Mesnil de Rochemont R, *et al.* A simple 3-item stroke scale: comparison with the National Institutes of Health Stroke Scale and prediction of middle cerebral artery occlusion. *Stroke* 2005;36:773-6.

11. Perez 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.

12. Katz BS, McMullan JT, Sucharew H, *et al.* Design and validation of a prehospital scale to predict stroke severity: Cincinnati Prehospital Stroke Severity Scale. *Stroke* 2015;46:1508-12.

13. 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;9:122-6.

14. Hastrup S, Damgaard D, Johnsen SP, *et al.* Prehospital Acute Stroke Severity Scale to Predict Large Artery Occlusion: Design and Comparison With Other Scales. *Stroke* 2016;47:1772-6.

15. Lima FO, Silva GS, Furie KL, *et al.* Field Assessment Stroke Triage for Emergency Destination: A Simple and Accurate Prehospital Scale to Detect Large Vessel Occlusion Strokes. *Stroke* 2016;47:1997-2002.

16. Nazliel B, Starkman S, Liebeskind DS, *et al.* A brief prehospital stroke severity scale identifies ischemic stroke patients harboring persisting large arterial occlusions. *Stroke* 2008;39:2264-7.

17. Heldner MR, Hsieh K, Broeg-Morvay A, *et al.* Clinical prediction of large vessel occlusion in anterior circulation stroke: mission impossible? *J Neurol* 2016;263:1633-40.

18. Vanacker P, Heldner MR, Amiguet M, *et al.* Prediction of Large Vessel Occlusions in Acute Stroke: National Institute of Health Stroke Scale Is Hard to Beat. *Crit Care Med* 2016;44:e336-43.

19. Rinaldo L, Brinjikji W, McCutcheon BA, *et al.* Hospital transfer associated with increased mortality after endovascular revascularization for acute ischemic stroke. *J Neurointerv Surg* 2016.

20. 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* 2016.

21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.

22. Nor AM, Davis J, Sen B, *et al.* The Recognition of Stroke in the Emergency Room (ROSIER) scale: development and validation of a stroke recognition instrument. *Lancet Neurol* 2005;4:727-34.

23. Majerus S, Bruno MA, Schnakers C, *et al.* The problem of aphasia in the assessment of consciousness in brain-damaged patients. *Prog Brain Res* 2009;177:49-61.

24. Fiehler J, Cognard C, Gallitelli M, *et al.* European Recommendations on Organisation of Interventional Care in Acute Stroke (EROICAS). *Int J Stroke* 2016;11:701-16.

25. Heldner MR, Zubler C, Mattle HP, *et al.* National Institutes of Health stroke scale score and vessel occlusion in 2152 patients with acute ischemic stroke. *Stroke* 2013;44:1153-7.

26. Turc G, Maier B, Naggara O, *et al.* Clinical Scales Do Not Reliably Identify Acute Ischemic Stroke Patients With Large-Artery Occlusion. *Stroke* 2016;47:1466-72.

27. John S, Stock S, Masaryk T, *et al.* Performance of CT Angiography on a Mobile Stroke Treatment Unit: Implications for Triage. *J Neuroimaging* 2016;26:391-4.

28. Diringer MN, Edwards DF. Admission to a neurologic/neurosurgical intensive care unit is associated with reduced mortality rate after intracerebral hemorrhage. *Crit Care Med* 2001;29:635-40.

Table 1 The shortened NIH Stroke Scale for Emergency Medical Services (sNIHSS-EMS).

No.	sNIHSS-EMS item	Equivalent to the NIHSS-item	Range
1	Level of Consciousness	1a	0 – 3
2	Facial Palsy	4	0 – 3
3a	Motor Arm (left)	5	0 – 4 / UN
3b	Motor Arm (right)	5	0 – 4 / UN
4a	Motor Leg (left)	6	0 – 4 / UN
4b	Motor Leg (right)	6	0 – 4 / UN
5	Sensory	8	0 – 2
6	Best Language	9	0 – 3
7	Dysarthria	10	0 – 2 / UN
	Sum	–	0 – 29

Range indicates possible scores;

Abbreviations: UN = untestable (motor items: amputation or joint fusion, dysarthria: intubation or other physical barrier).

Table 2 Cut-off points for prediction of acute large vessel occlusion.

Cut-off Point	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy	*J*
sNIHSS-EMS						
5	74.8 (69.4 – 79.7)	73.4 (69.1 – 77.4)	66.4 (59.0 – 69.5)	81.9 (77.8 – 85.6)	74.0	0.482
6*	70.3 (64.7 – 75.5)	80.7 (76.8 – 84.3)	70.1 (64.5 – 75.3)	80.9 (76.9 – 84.4)	76.7	0.511
7	65.2 (59.4 – 70.6)	85.8 (82.2 – 88.9)	74.7 (68.9 – 79.9)	79.3 (75.4 – 82.8)	77.7	0.510
NIHSS						
8	72.4 (66.9 – 77.5)	80.7 (76.8 – 84.3)	70.7 (65.2 – 75.8)	82.0 (78.1 – 85.4)	77.5	0.531
9*	69.3 (63.7 – 74.6)	85.4 (81.8 – 88.5)	75.3 (69.7 – 80.3)	81.2 (77.4 – 84.6)	79.1	0.547
10	65.9 (60.1 – 71.3)	88.0 (84.7 – 90.9)	78.0 (72.2 – 83.0)	80.0 (76.2 – 83.5)	79.4	0.539

Abbreviations: NIHSS = NIH Stroke Scale; sNIHSS-EMS, shortened NIHSS for emergency medical services.

Data are % (95% CI). *J* indicates Youden's-Index, * indicates the optimal cut-off according to the Youden index.

Table 3 Comparison of clinical scales for prehospital prediction of large vessel occlusions

	3I-SS	LAMS	RACE	CPSSS	FAST-ED	PASS	sNIHSS-EMS
Reference	10	16	11	12	15	14	–
Scale characteristics							
No. of items assessed	3	3	5*	3	5	3	7
Score range	0–6	0–5	0–9	0–4	0–9	0–3	0-29
NIHSS compatible item assessment	–	–	–	–	–	–	●
Stroke Recognition	–	–	–	–	–	–	●
Stroke severity grading	–	(●)	(●)	–	–	–	●
Large vessel occlusion prediction	●	●	●	●	●	●	●
LVO prediction, test characteristics, own cohort (N = 741, 44% LVO)							
Cut-point used †	≥ 4	≥ 4	≥ 5	≥ 2	≥ 4	≥ 2	≥ 6
Sensitivity	40%	–‡	59%	59%	60%	68%	70%

Specificity	95%	–‡	91%	89%	90%	84%	81%
PPV	85%	–‡	81%	77%	80%	74%	70%
NPV	71%	–‡	78%	77%	78%	81%	81%

LVO prediction, test characteristics, original cohorts§

Cohort (N (%LVO))	83 (35%)	119 (62%)	357 (21%)**	303 (73%)	727 (33%)	3127 (35%)††	–
Sensitivity	67%	81%	85%	83%	61%	66%	–
Specificity	92%	89%	68%	40%	89%	83%	–
PPV	74%	nd	42%	nd	72%	68%	–
NPV	89%	nd	94%	nd	82%	81%	–

Abbreviations: nd: no data.

* If right sided hemiparesis, aphasia is assessed, if left sided hemiparesis, agnosia.

† Cut-points according to original publications. Based on the Youden indices calculated from our data, optimal cut-points are different: 3I-SS ≥ 2, RACE ≥ 3, CPSSS ≥ 1, FAST-ED ≥ 3.

‡ Grip-strength was not routinely documented, therefore external validation of the LAMS was not possible.

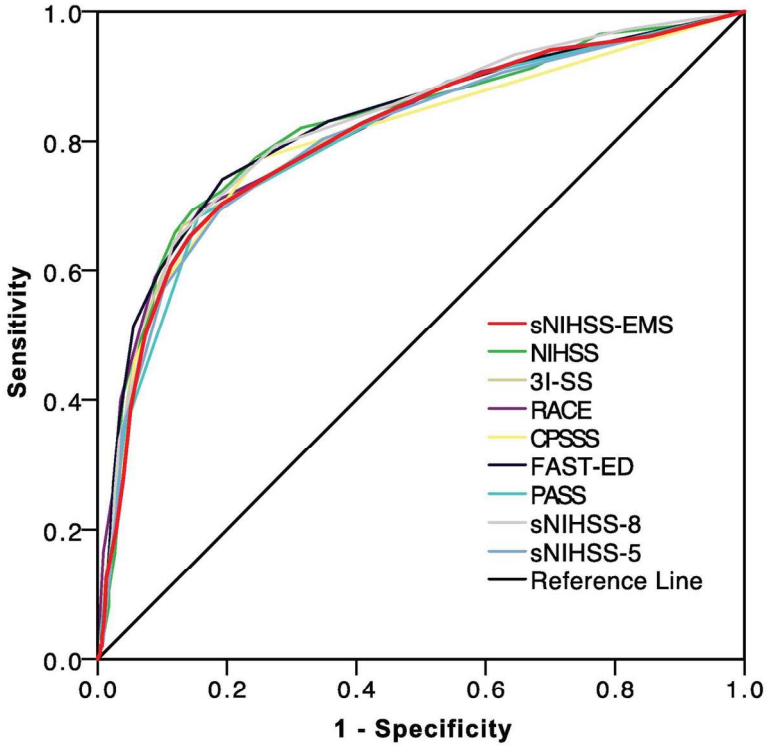
§ Definition of large vessel occlusions according to original publications (3I-SS: carotid-T or M1; LAMS: ICA, M1, M2, M3/4, ACA; RACE: terminal ICA, M1, tandem CCA/ICA+M1, BA; CPSSS: ICA, M1, tandem ICA+M2, BA; FAST-ED: ICA, M1, M2, BA; PASS: “visible clot in the anterior or posterior circulation on CTA or MRA“; abbreviations within the main text).

** Including cases assessed by transcranial duplex only (N = 197).

†† Only patients who received intravenous tPA; 2/3 of entire cohort were taken as a random sample for derivation. In the remaining 1/3, sensitivity was 61%, specificity 83%.

Figure 1. Receiver operating curves for prediction of acute large vessel occlusion.

Abbreviations: AUC: area under the curve; ref.: reference.



Scale	Area Under the Curve (95% CI)	*p*
sNIHSS-EMS	0.808 (0.775 – 0.841)	*ref.*
NIHSS	0.821 (0.788 – 0.854)	0.019
3I-SS	0.804 (0.770 – 0.838)	0.697
RACE	0.816 (0.784 – 0.849)	0.314
CPSSS	0.802 (0.767 – 0.836)	0.605
FAST-ED	0.825 (0.793 – 0.857)	0.078
PASS	0.805 (0.771 – 0.838)	0.766
sNIHSS-8	0.823 (0.791 – 0.855)	0.020
sNIHSS-5	0.803 (0.769 – 0.837)	0.643

Figure 1

124x173mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4/5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5/6
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5/6
	9	Whether participants formed a consecutive, random or convenience series	5/6
Test methods	**10a**	Index test, in sufficient detail to allow replication	4-7
	10b	Reference standard, in sufficient detail to allow replication	5/6
	11	Rationale for choosing the reference standard (if alternatives exist)	5/6, 7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6/7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6/7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6/7
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6/7
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6/7
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6/7
	18	Intended sample size and how it was determined	6, +Appendix
RESULTS			
Participants	**19**	Flow of participants, using a diagram	NA
	20	Baseline demographic and clinical characteristics of participants	8, +Appendix.
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	Appendix, previousl. published.
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Appendix
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	8/9
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	3; 11,12,13
OTHER INFORMATION			
	28	Registration number and name of registry	NA (IRB-Prot., 12)
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14

BMJ Open

Design and validation of a clinical scale for prehospital stroke recognition, severity grading, and prediction of large vessel occlusion – the shortened NIH stroke scale for emergency medical services

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2017-016893.R2
Article Type:	Research
Date Submitted by the Author:	18-Jun-2017
Complete List of Authors:	Purrucker, Jan; Heidelberg University Hospital, Department of Neurology Härtig, Florian; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Richter, Hardy; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research Engelbrecht, Andreas; Heidelberg University Hospital, Department of Neurology Hartmann, Johannes; Eberhard Karls Universitat Tubingen, Department of Medical Informatics Auer, Jonas; Universitat Stuttgart, Department of Computer Science & Software Engineering Hametner, Christian; University of Heidelberg, Neurology Popp, Erik; UniversitatsKlinikum Heidelberg, Department of Anesthesiology Ringleb, Peter; Heidelberg University Hospital, Department of Neurology Nagel, Simon; Heidelberg University Hospital, Department of Neurology Poli, Sven; Eberhard-Karls-Universitat Tubingen Medizinische Fakultat, Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Emergency medicine
Keywords:	Stroke < NEUROLOGY, ACCIDENT & EMERGENCY MEDICINE, VASCULAR MEDICINE

Design and validation of a clinical scale for prehospital stroke recognition, severity grading, and prediction of large vessel occlusion – the shortened NIH stroke scale for emergency medical services

Jan C. Purrucker,[1] Florian Härtig,[2] Hardy Richter,[2] Andreas Engelbrecht,[1] Johannes Hartmann,[3] Jonas Auer,[4] Christian Hametner,[1] Erik Popp,[5] Peter A. Ringleb,[1] Simon Nagel[1] and Sven Poli [2]

Author affiliations:

[1] Department of Neurology, Heidelberg University Hospital, Heidelberg, Germany

[2] Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany

[3] Department of Medical Informatics, University of Tuebingen, Tuebingen, Germany

[4] Department of Computer Science & Software Engineering, University of Stuttgart, Stuttgart, Germany

[5] Department of Anesthesiology, University of Heidelberg, Heidelberg, Germany

Corresponding author:

Sven Poli, MD MSc; Department of Neurology & Stroke, and Hertie Institute for Clinical Brain Research, Tuebingen University Hospital, Tuebingen, Germany; Phone: +49 7071 2968300, Fax: +49 7071 2925047, E-mail: sven.poli@icloud.com

Keywords: Stroke severity grading, triage, large vessel occlusion, emergency medicine

Word-Count: Abstract: 245; Main Manuscript: 2765.

ABSTRACT

Objective: To develop a NIH Stroke Scale (NIHSS)-compatible, all-in-one scale for rapid and comprehensive prehospital stroke assessment including stroke recognition, severity grading and progression monitoring as well as prediction of large vessel occlusion (LVO).

Methods: Emergency medical services (EMS) personnel and stroke physicians (N = 326) rated each item of the NIHSS regarding suitability for prehospital use; best-rated items were included. Stroke recognition was evaluated retrospectively in 689 consecutive patients with acute stroke or stroke mimics, prediction of LVO in 741 consecutive ischemic stroke patients with acute vessel imaging independent of admission-NIHSS score.

Results: Nine of the NIHSS-items were rated as "suitable for prehospital use". After excluding two items in order to increase specificity, the final scale (termed shortened NIHSS for EMS, sNIHSS-EMS), it consists of 'level of consciousness', 'facial palsy', 'motor arm/leg', 'sensory', 'language', and 'dysarthria'. Sensitivity for stroke recognition of the sNIHSS-EMS is 91% (95% confidence interval [CI] 86–94], specificity 52% (95% CI 47–56). Receiver operating curve analysis revealed an optimal cut-off point for LVO prediction of ≥ 6 (sensitivity 70% [95% CI 65–76], specificity 81% [95% CI 76–84], positive predictive value 70 [95% CI 65–75], area under the curve 0.81 [95% CI 0.78–0.84]). Test characteristics were non-inferior to non-comprehensive scales.

Conclusions: The sNIHSS-EMS may overcome the sequential use of multiple emergency stroke scales by permitting parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility allows for evaluation of stroke progression starting at the prehospital phase.

Strengths and limitations of this study

- Prehospital stroke assessment is increasingly gaining relevance in the era of endovascular interventions for large vessel occlusions. Sound triage decisions will have a major impact on patients' outcomes. As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.
- The new clinical scale (sNIHSS-EMS), developed and validated in this study, is the first scale assessed for parallel stroke recognition, severity grading, and LVO prediction. Sequential use of multiple emergency stroke scales may thus be avoided.
- A multinational survey among different emergency medical systems and professions was performed to identify items suitable for use in prehospital emergency situations.
- The sNIHSS-EMS shares full compatibility with the in-hospital gold-standard NIHSS, but remains simple and easy to use.
- The scale will be incorporated into a prehospital stroke triage algorithm in a large regional stroke network, but no prospective data are available yet, which is acknowledged as a limitation.

INTRODUCTION

A considerable number of stroke scales for prehospital use have been published over recent years.[1, 2] However, most of these scales only focus on single aspects of acute stroke care, i.e. either stroke recognition,[1, 2] early prediction of outcome,[3] prediction of thrombolysis,[4, 5] or severity grading and large vessel occlusion (LVO). [3, 6-18] Consequently, to provide a comprehensive prehospital stroke assessment, emergency medical services (EMS) personnel must apply at least two scales. Furthermore, the majority of existing scales lack compatibility with the NIH Stroke Scale (NIHSS), the in-hospital 'gold-standard' for stroke severity grading.[2] This impedes the seamless evaluation of stroke progression from pre- to in-hospital care. In the era of endovascular treatment of LVO, decisions regarding direct emergency referrals to specialized comprehensive stroke centers will have a major impact on patients' outcomes.[19, 20] As those are left entirely to EMS personnel, it is essential to equip them with an effective tool to guide prehospital triage.

We present the development and validation of a novel comprehensive stroke scale, specifically designed for prehospital use with input from EMS. Our aim was to allow for parallel stroke recognition, severity grading and – owing to full NIHSS-compatibility – progression monitoring as well as LVO prediction.

METHODS

International online survey

We invited non-neurologic EMS personnel (paramedics and emergency physicians) and stroke physicians from Austria, Germany, and Switzerland to rate each individual NIHSS item regarding their applicability in a prehospital emergency setting. Invitations were sent out via the German Stroke Society (DSG), the German Society for Neuro-Intensive Care and Emergency Medicine (DGNI), as well as EMS providers. Participation was voluntary, no financial incentive was offered, and participation was

only allowed once. Non-neurologic EMS personnel do not use the NIHSS routinely and did not receive specific NIHSS training before the survey. For each NIHSS item, we created and provided a short video demonstrating in-hospital bedside assessment according to the NIHSS training instructions (a screenshot is shown as Figure 1A in the Appendix). Having watched the video, participants were asked to rate each NIHSS item regarding its suitability for prehospital use on a 6-item scale, ranging from 0 (most suitable) to 5 (most unsuitable). Ratings were automatically entered into a database together with name (optional), profession, professional experience, and place of work. Participation was possible from November 19th 2015, until April 15th 2016, the pre-specified closing date.

Patient cohorts

Test characteristics of the newly designed scale were calculated with regard to performance in stroke recognition and prediction of acute LVO using two distinct clinical cohorts described below.

For stroke recognition, we used a prospectively collected cohort of consecutive patients with acute ischemic or hemorrhagic stroke and stroke mimics, which had already served as a validation cohort in a previous comparison of existing stroke scales.[2] In summary, the database consists of pseudonymized data of consecutive patients (including comatose) with preclinical 'suspected acute CNS disorder' admitted to the Emergency Room of the Department of Neurology, Heidelberg University Hospital, Germany by EMS between November 2007 and August 2010. For all patients, a full-length NIHSS score assessed by certified raters was available at admission. The diagnostic reference standard was the diagnosis at hospital discharge. Cases were dichotomized (by the authors AE and CH) in stroke and non-stroke, i.e. stroke-mimics. AE and CH were blinded for the admission NIHSS and sNIHSS-EMS scores.

Test characteristics regarding the prediction of LVO were calculated in a prospectively collected second cohort consisting of consecutive patients with acute ischemic stroke, admitted to the Department of Neurology, Tuebingen University Hospital, Germany between January 2013 and July 2015. In accordance with local standard operating procedures, all received acute vessel imaging on admission independent of stroke severity. Neuroradiological reports and original images were reviewed by the authors HR and SP for presence of acute LVO. HR and SP were blinded to patients' NIHSS scores. Cases were considered as LVO-positive if an acute symptomatic occlusion was present in one of the following arteries: common carotid artery (CCA), internal carotid artery (ICA), carotid T, middle cerebral artery (MCA, including M1/M2 segments), anterior cerebral artery (ACA), basilar artery (BA), or posterior cerebral artery (PCA).

Statistics

To determine suitable items for use in the prehospital phase, we analyzed the online survey response data set; median and interquartile ranges (IQR) were calculated. NIHSS-items receiving median scores of 0 and 1 were – as predefined – regarded eligible for further consideration. Rating differences between the professional groups (i.e. non-neurologic EMS personnel and stroke physicians) were determined using the Mann-Whitney-U test. For the calculation of test performance regarding stroke recognition, the sNIHSS-EMS score was dichotomized as indicative of stroke (score ≥ 1), or not (score = 0). Sensitivity (the proportion of stroke patients who had a positive test, i.e. indicative of stroke) and specificity (the proportion of non-stroke patients who had a negative test), positive predictive value (PPV), and negative predictive value (NPV) were calculated with 95% confidence intervals (CI). Details of the sample size calculation are described in the extended methods in the Appendix. To determine the predictive power for LVO detection, we calculated sensitivity, specificity, PPV, and NPV, with 95% CI for each scale score ranging from 0 to 29 for the sNIHSS-EMS, and

from 0 to 42 for the original NIHSS. Accuracy is reported additionally. Receiver operating curve (ROC) analysis was performed, area under the curve (AUC) and Youden´s index were calculated. For comparison of the sNIHSS-EMS with existing dedicated LVO prediction scales,[7, 10-12, 14, 15] we calculated the corresponding scores using the NIHSS-equivalents and cut-offs as stated in the original publications. Statistical comparison of AUCs was performed according to DeLong et al.[21] Calculation of the Los Angeles Motor Scale (LAMS) for our LVO cohort was not possible since the item "grip-strength" was not routinely documented. P values were 2-sided with values less than .05 considered statistically significant. SPSS (V23.0.0.2, IBM, New York, USA), MedCalc (V16.8.4, Ostend, Belgium) and GraphPad Prism (V6.0b, San Diego, California, USA) were used for data handling and analysis, and graphic presentation. This study was performed in accordance with the STARD guidelines for studies on diagnostic tests.

RESULTS

Scale development

326 (13%) of 2562 recipients responded to our international online survey (Austria, Germany, and Switzerland), with the majority (57%) representing non-neurologic EMS personnel (33% paramedics and 24% prehospital emergency physicians); 33% stroke physicians, and 10% not specified. Participants reported a high level of professional experience (>10 years, 45%; <5 years, 20%).

Nine of the NIHSS-items received a median score of 0 or 1 (equivalent to most suitable and suitable for prehospital use), whereas the items 'best gaze', 'visual', 'limb ataxia', and 'extinction' were rated as less suitable and thus removed from further analyses (Table 1A in the Appendix). Although rating by stroke physicians was more rigorous, item selection based on median ratings of 0 or 1 was not shifted by the professional vote (Table 1A).

We decided to exclude items 1b (LOC questions) and 1c (LOC commands). Despite being easily assessable and thus rated suitable for prehospital use, these two items are either present in the absence of stroke as frequent features of non-stroke conditions (e.g. dementia, infection or dehydration)[22] or heavily influenced by aphasia[23] and thus redundant for stroke recognition. The new 7-items scale was termed 'shortened NIHSS for emergency medical services' (sNIHSS-EMS; Table 2).

Stroke recognition and severity grading

In our stroke recognition validation cohort of 689 consecutive patients with 'suspected acute CNS disorder', 29% received 'stroke' as discharge diagnosis. Patients with ischemic stroke (n=200) had an admission-NIHSS of 9 (IQR 4–17), patients with hemorrhagic stroke (n=55) of 17 (IQR 5–35). Non-stroke patients (n = 489) had a median admission-NIHSS of 1 (IQR 0–6). The sNIHSS-EMS was found to have 90.5% (95% CI 85.6–94.2) sensitivity and 51.5% (95% CI 47.0–56.1) specificity for stroke

recognition (PPV 43.3% [95% CI 38.5–48.2], NPV 93.0% [95% CI 89.3–95.6]). Cross tabulations are shown in Table 3A in the Appendix. Excluding comatose patients (n = 49), sensitivity was 89.1% (95% CI 83.6–93.3) and specificity 54.2% (95% CI 49.5–58.8).

LVO prediction

In the distinct LVO validation cohort of consecutive 741 ischemic stroke patients with acute vessel imaging independent of their admission-NIHSS score (86.9% CTA; see Table A2 for patient characteristics), a ROC analysis of the sNIHSS-EMS regarding LVO prediction revealed a maximal Youden index at the cut-point of ≥ 6 (sensitivity 70.3% [95% CI 64.7–75.5], specificity 80.7% [95% CI 76.8–84.3]; Figure 1, Table 2). For comparison, in the original NIHSS, the maximal Youden index was calculated for a cut-point of ≥ 9 (Table 2). Combined re-inclusion of the NIHSS items 'visual', 'gaze' and 'extinction' improved test characteristics (AUC 0.826 vs. 0.808, p<0.001). Re-inclusion of singular items did not improve test characteristics. Exclusion of patients with coma (n=5) did not change the optimal cut-off and test characteristics (sensitivity 70.0% [64.4–75.3], specificity 81.1% [77.1–84.6]).

We validated the sNIHSS-EMS against existing LVO prediction scales through applying them to our cohort and calculation of ROC and Youden indices (Table 3, Figure 1). No statistically significant differences compared to existing scales were found, except for the full-length NIHSS, and the sNIHSS-8. Notably, due to characteristics of our cohort, external validation based on maximal Youden indices led to cut-points different from those reported in the respective original publications (Table 3).

DISCUSSION

The sNIHSS-EMS is the first comprehensive stroke scale assessed for parallel stroke recognition, severity grading, and LVO prediction. Test characteristics regarding identification of patients with large vessel occlusion are non-inferior to existing LVO prediction scales. Furthermore, compatibility with the item assessment in the full-length NIHSS allows for continuous evaluation of the clinical course from pre- to in-hospital care. It may thus represent the ideal stroke scale for routine use in pre-hospital emergency medical care.

As previously shown by our work-group,[2] some of the available stroke severity scales[3, 6] may be used for stroke recognition with similar sensitivity and specificity when compared to scales developed for stroke recognition alone. Existing scales, however, either include items requiring complex assessment (such as extinction[11, 15]) or exclude items highly relevant for evaluation of stroke progression (such as level of consciousness, arm or leg motor function[3, 7, 11]).

Sensitivity of the sNIHSS-EMS regarding stroke recognition (91%) was superior to previously published results for the simpler CPSS (85%) and FAST (87%) evaluated in the same cohort of patients.[2] In contrast, specificity (52%) was lower compared to the CPSS (65%) and FAST (64%).[2] As the overall burden of a missed stroke outweighs the potentially increased workload of emergency departments, higher sensitivity may be considered more relevant. Simpler stroke scales may provide a slightly faster initial assessment, but subsequently require the use of at least one additional scale to determine stroke severity or predict LVO. The use of multiple scales, however, may be error-prone and complicates communication with receiving hospitals.

According to recent European and American recommendations, clinical screening tools may be considered in order to facilitate direct transport of patients with suspected LVO to Comprehensive Stroke Centers (CSC) with endovascular facility.[20, 24] For LVO prediction, our analysis revealed a maximum Youden index for the cut-point of ≥ 6 for the sNIHSS-EMS and, in accordance with previous findings, 9 for the original

NIHSS.[25] Importantly, to adjust for hospital capacities and local stroke network requirements, this threshold can be adapted: higher cut-points result in an increased specificity (Table 2) leading to reduced numbers of patients bypassing Acute Stroke Ready Hospitals (ASRH) or Primary Stroke Centers (PSC) without endovascular facility.

The NIHSS items 'visual', 'gaze' and 'extinction' are part of some dedicated LVO prediction scales,[10, 11, 14, 15] but were not included in the sNIHSS-EMS due to unfavorable ratings regarding prehospital assessability. Re-inclusion of each separate item did not result in the presumed higher predictive value for LVO detection. Only combined re-inclusion of all three rejected items led to marginally enhanced test characteristics, but would result in a significantly increased number of complex-to-assess items and thus an inconvenient scale.

For comparison with existing scales, we externally validated dedicated LVO prediction scales in our cohort by using the cut-points as provided in the original publications and found the sNIHSS-EMS to offer comparable sensitivity and specificity (Table 4). Better test characteristics reported in the original publications for some scales may be due to differences in the definition of LVO (e.g. the 3I-SS focused on carotid T and M1 occlusions only,[10] while the LAMS also included M3/4 occlusions[16]). The sNIHSS-8, which had a higher AUC in the ROC analysis than the sNIHSS-EMS, was not developed for LVO prediction and includes items rejected by EMS personnel in our survey due to the complexity of correct assessment.

LVO prediction by clinical scales has recently been criticized due to the high false-negative rate compared to vessel imaging.[17, 26] The sNIHSS-EMS is not intended to substitute in-hospital acute vessel imaging,[17] and prehospital acute vessel imaging is still an exception.[27] Currently, mainly due to the narrow time window for effective intravenous thrombolysis, patients are transferred to the closest stroke center regardless of LVO suspicion. In the era of interventional thrombectomy however, ASRH or PSC

may have to be bypassed in favor of CSC with endovascular facility in sensibly selected cases.

Based on clinical criteria alone, the sNIHSS-EMS identifies the majority of patients with acute LVO, i.e. those patients who might benefit from a direct transfer to CSC with endovascular facility. In addition, the minority of LVO patients not bypassed to endovascular ready CSC (i.e. total score < 6 despite LVO) are not lost to endovascular therapy since secondary transportation to an endovascular ready CSC is still possible. The sNIHSS-EMS is designed to permit the monitoring of stroke progression from pre- to in-hospital care on the item-level, a feature that has been neglected in other scales. Clinical implications include the earlier recognition of symptom fluctuation with consequences e.g. for blood pressure management or selection of imaging modality. In practice, if a '2' is scored for 'Motor Leg left' on the sNIHSS-EMS, a '4' on the same item during routine NIHSS evaluation in the ER points to early clinical deterioration. Clinical scores using merged items (e.g. 'hemiparesis'[10] or 'language/dysarthria'[4]) or modified item scoring (e.g. motor function scoring from 0 to 2 instead of 0 to 4 [11, 12, 14-16]) impede seamless monitoring of symptom progression.

Despite the positive aspects of the sNIHSS-EMS, some limitations of the present study require further discussion. Test characteristics regarding LVO prediction were calculated in a cohort of patients with confirmed ischemic stroke because determination of the 'true' LVO prediction threshold is only possible in a cohort without stroke mimics or hemorrhagic stroke. However, although this approach is in concordance with methods used in the past in the design of dedicated LVO prediction scales,[12, 14, 16] future prospective validation in the prehospital target population will be necessary to determine prevalence-dependent test characteristics. We were not able to assess LVO prediction of the LAMS because the item 'grip strength' is not part of the NIHSS and thus, was not routinely documented in our cohort. According to a retrospective validation study in anterior circulation stroke, the sensitivity of the LAMS for LVO

prediction was reported as 81% (at a threshold of 4).[16] As patients with stroke mimics (and thus no LVO) exhibit low NIHSS scores, inclusion of these cases into the analyses, would lead to an increased specificity of our cut-points. The sNIHSS-EMS is not able to differentiate between ischemic and hemorrhagic stroke. This might not be a disadvantage as severely affected hemorrhagic stroke patients benefit from direct admission to a CSC with neurological intensive care capacity.[28] Despite involvement of EMS systems from three European countries, generalizability to further EMS systems around the world cannot be concluded. The low response-rate of our online survey makes a non-response bias likely. Due to the participants' high professional experience, one might have expected a shift of the suitability assessment towards more complex items. However, this was not observed. As a strength of this study, LVO was evaluated by CTA or MRA, and not with less accurate duplex sonography as done in previous studies evaluating LVO prediction scales.[11, 16] The sNIHSS-EMS was primarily designed to fulfill requirements for prehospital use. Although kept simple, additional training on the new scale is recommended. Moreover, the sNIHSS-EMS may also serve in telemedicine with usually non-neurologic physicians performing the initial patient examination.

CONCLUSION

The sNIHSS-EMS may overcome the need for sequential use of multiple emergency stroke scales by enabling parallel stroke recognition, severity grading, and LVO prediction. Full NIHSS-item-compatibility permits evaluation of stroke progression starting from the prehospital phase. Offering comparable test characteristics as dedicated scales, the sNIHSS-EMS may be a promising tool for rapid and comprehensive prehospital stroke assessment and triage.

Acknowledgments

We thank all survey participants for their contribution. Those who gave consent to the publication of their names are listed in the Appendix. We thank Louise Alice Härtig for language revision of the manuscript. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen.

Ethical approval

Ethic approvals were obtained from the ethics committee of the Medical Faculty Heidelberg and the ethics committee at the University Hospital Tuebingen, Germany (protocol-numbers S-109/2013 and 648/2015BO2, respectively). Written informed consent was waived.

Funding

This study was investigator initiated, without funding.

Contributorship

SP and JCP conceived and designed the study. EP provided the EMS data. HR and SP created and validated the LVO prediction database. SP, JCP, AE and CH collected and analyzed the data. JH and JA developed and maintained the online survey. JCP, FH, SP drafted the article. JCP, HR, FH, CH, PAR, SN, SP revised the manuscript.

Guarantor

SP and JCP take the responsibility for the paper as a whole.

Declaration of Conflicting Interests

Personal fees, travel support, speaker honoraria, or research grants were received from Bayer Healthcare (FH, HR, PR, SN, SP), BeneChill (SP), BMS Pfizer (PR, SP),

Boehringer-Ingelheim (JCP, PR, SN, SP), Brainomix Ltd. (SN), Covidien (SP), C.R. Bard (SP), EMCOOLS (SP), Helena Laboratories (FH, SP), HVM Medical (SP), Medtronic (SN), Pfizer (JCP, SN), Raumedic (SP), ZOLL (SP), The other authors report no conflicts of interest.

Provenance and peer review: Not commissioned; externally peer reviewed.

Data sharing statement: No additional data are available.

References

1. Brandler ES, Sharma M, Sinert RH, *et al.* Prehospital stroke scales in urban environments: a systematic review. *Neurology* 2014;82:2241-9.

2. Purrucker JC, Hametner C, Engelbrecht A, *et al.* Comparison of stroke recognition and stroke severity scores for stroke detection in a single cohort. *J Neurol Neurosurg Psychiatry* 2015;86:1021-8.

3. Tirschwell DL, Longstreth WT, Jr., Becker KJ, *et al.* Shortening the NIH Stroke scale for use in the prehospital setting. *Stroke* 2002;33:2801-6.

4. Iguchi Y, Kimura K, Watanabe M, *et al.* Utility of the Kurashiki Prehospital Stroke Scale for hyperacute stroke. *Cerebrovasc Dis* 2011;31:51-6.

5. Hasegawa Y, Sasaki N, Yamada K, *et al.* Prediction of thrombolytic therapy after stroke-bypass transportation: the Maria Prehospital Stroke Scale score. *J Stroke Cerebrovasc Dis* 2013;22:514-9.

6. Kimura K, Inoue T, Iguchi Y, *et al.* Kurashiki prehospital stroke scale. *Cerebrovasc Dis* 2008;25:189-91.

7. 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.

8. Meyer BC, Hemmen TM, Jackson CM, *et al.* Modified National Institutes of Health Stroke Scale for use in stroke clinical trials: prospective reliability and validity. *Stroke* 2002;33:1261-6.

9. Whelley-Wilson CM, Newman GC. A stroke scale for emergency triage. *J Stroke Cerebrovasc Dis* 2004;13:247-53.

10. Singer OC, Dvorak F, du Mesnil de Rochemont R, *et al.* A simple 3-item stroke scale: comparison with the National Institutes of Health Stroke Scale and prediction of middle cerebral artery occlusion. *Stroke* 2005;36:773-6.

11. Perez 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.

12. Katz BS, McMullan JT, Sucharew H, *et al.* Design and validation of a prehospital scale to predict stroke severity: Cincinnati Prehospital Stroke Severity Scale. *Stroke* 2015;46:1508-12.

13. 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;9:122-6.

14. Hastrup S, Damgaard D, Johnsen SP, *et al.* Prehospital Acute Stroke Severity Scale to Predict Large Artery Occlusion: Design and Comparison With Other Scales. *Stroke* 2016;47:1772-6.

15. Lima FO, Silva GS, Furie KL, *et al.* Field Assessment Stroke Triage for Emergency Destination: A Simple and Accurate Prehospital Scale to Detect Large Vessel Occlusion Strokes. *Stroke* 2016;47:1997-2002.

16. Nazliel B, Starkman S, Liebeskind DS, *et al.* A brief prehospital stroke severity scale identifies ischemic stroke patients harboring persisting large arterial occlusions. *Stroke* 2008;39:2264-7.

17. Heldner MR, Hsieh K, Broeg-Morvay A, *et al.* Clinical prediction of large vessel occlusion in anterior circulation stroke: mission impossible? *J Neurol* 2016;263:1633-40.

18. Vanacker P, Heldner MR, Amiguet M, *et al.* Prediction of Large Vessel Occlusions in Acute Stroke: National Institute of Health Stroke Scale Is Hard to Beat. *Crit Care Med* 2016;44:e336-43.

19. Rinaldo L, Brinjikji W, McCutcheon BA, *et al.* Hospital transfer associated with increased mortality after endovascular revascularization for acute ischemic stroke. *J Neurointerv Surg* 2016.

20. 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* 2016.

21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.

22. Nor AM, Davis J, Sen B, *et al.* The Recognition of Stroke in the Emergency Room (ROSIER) scale: development and validation of a stroke recognition instrument. *Lancet Neurol* 2005;4:727-34.

23. Majerus S, Bruno MA, Schnakers C, *et al.* The problem of aphasia in the assessment of consciousness in brain-damaged patients. *Prog Brain Res* 2009;177:49-61.

24. Fiehler J, Cognard C, Gallitelli M, *et al.* European Recommendations on Organisation of Interventional Care in Acute Stroke (EROICAS). *Int J Stroke* 2016;11:701-16.

25. Heldner MR, Zubler C, Mattle HP, *et al.* National Institutes of Health stroke scale score and vessel occlusion in 2152 patients with acute ischemic stroke. *Stroke* 2013;44:1153-7.

26. Turc G, Maier B, Naggara O, *et al.* Clinical Scales Do Not Reliably Identify Acute Ischemic Stroke Patients With Large-Artery Occlusion. *Stroke* 2016;47:1466-72.

27. John S, Stock S, Masaryk T, *et al.* Performance of CT Angiography on a Mobile Stroke Treatment Unit: Implications for Triage. *J Neuroimaging* 2016;26:391-4.

28. Diringer MN, Edwards DF. Admission to a neurologic/neurosurgical intensive care unit is associated with reduced mortality rate after intracerebral hemorrhage. *Crit Care Med* 2001;29:635-40.

Table 1 The shortened NIH Stroke Scale for Emergency Medical Services (sNIHSS-EMS).

No.	sNIHSS-EMS item	Equivalent to the NIHSS-item	Range
1	Level of Consciousness	1a	0 – 3
2	Facial Palsy	4	0 – 3
3a	Motor Arm (left)	5	0 – 4 / UN
3b	Motor Arm (right)	5	0 – 4 / UN
4a	Motor Leg (left)	6	0 – 4 / UN
4b	Motor Leg (right)	6	0 – 4 / UN
5	Sensory	8	0 – 2
6	Best Language	9	0 – 3
7	Dysarthria	10	0 – 2 / UN
	Sum	–	0 – 29

Range indicates possible scores;

Abbreviations: UN = untestable (motor items: amputation or joint fusion, dysarthria: intubation or other physical barrier).

Table 2 Cut-off points for prediction of acute large vessel occlusion.

Cut-off Point	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	Accuracy	*J*
sNIHSS-EMS						
5	74.8 (69.4 – 79.7)	73.4 (69.1 – 77.4)	66.4 (59.0 – 69.5)	81.9 (77.8 – 85.6)	74.0	0.482
6*	70.3 (64.7 – 75.5)	80.7 (76.8 – 84.3)	70.1 (64.5 – 75.3)	80.9 (76.9 – 84.4)	76.7	0.511
7	65.2 (59.4 – 70.6)	85.8 (82.2 – 88.9)	74.7 (68.9 – 79.9)	79.3 (75.4 – 82.8)	77.7	0.510
NIHSS						
8	72.4 (66.9 – 77.5)	80.7 (76.8 – 84.3)	70.7 (65.2 – 75.8)	82.0 (78.1 – 85.4)	77.5	0.531
9*	69.3 (63.7 – 74.6)	85.4 (81.8 – 88.5)	75.3 (69.7 – 80.3)	81.2 (77.4 – 84.6)	79.1	0.547
10	65.9 (60.1 – 71.3)	88.0 (84.7 – 90.9)	78.0 (72.2 – 83.0)	80.0 (76.2 – 83.5)	79.4	0.539

Abbreviations: NIHSS = NIH Stroke Scale; sNIHSS-EMS, shortened NIHSS for emergency medical services.

Data are % (95% CI). *J* indicates Youden's-Index, * indicates the optimal cut-off according to the Youden index.

Table 3 Comparison of clinical scales for prehospital prediction of large vessel occlusions

	sNIHSS-8	sNIHSS-5	3I-SS	LAMS	RACE	CPSSS	FAST-ED	PASS	sNIHSS-EMS
Reference	3	3	10	16	11	12	15	14	10
Scale characteristics									
No. of items assessed*	7	4	3	3	5†	3	5	3	7
Score range	0–24	0–16	0–6	0–5	0–9	0–4	0–9	0–3	0-29
NIHSS compatible item assessment	●	●	–	–	–	–	–	–	●
Stroke Recognition	●	●	–	–	–	–	–	–	●
Stroke severity grading	●	●	–	(●)	(●)	–	–	–	●
Large vessel occlusion prediction	●	●	●	●	●	●	●	●	●
LVO prediction, test characteristics, own cohort (N = 741, 44% LVO)									
Cut-point used ‡	≥ 6	≥ 3	≥ 4	≥ 4	≥ 5	≥ 2	≥ 4	≥ 2	≥ 6
Sensitivity	64%	69%	40%	–§	59%	59%	60%	68%	70%
Specificity	88%	81%	95%	–§	91%	89%	90%	84%	81%
PPV	78%	70%	85%	–§	81%	77%	80%	74%	70%
NPV	79%	80%	71%	–§	78%	77%	78%	81%	81%
LVO prediction, test characteristics, original cohorts**									
Cohort (N (%LVO))	–	–	83 (35 %)	119 (62%)	357 (21%) ††	303 (73%)	727 (33%)	3127 (35%)‡ ‡	–
Sensitivity	–	–	67%	81%	85%	83%	61%	66%	–

Specificity	–	–	92%	89%	68%	40%	89%	83%	–
PPV	–	–	74%	nd	42%	nd	72%	68%	–
NPV	–	–	89%	nd	94%	nd	82%	81%	–

Abbreviations: nd: no data.

* Motor arm (or leg) scored for each side (left or right) is counted as one item

† If right sided hemiparesis, aphasia is assessed, if left sided hemiparesis, agnosia.

‡ Cut-points according to original publications, with exception of the sNIHSS-8 and -5. Based on the Youden indices calculated from our data, optimal cut-points are different: 3I-SS ≥ 2, RACE ≥ 3, CPSSS ≥ 1, FAST-ED ≥ 3.

§ Grip-strength was not routinely documented, therefore external validation of the LAMS was not possible.

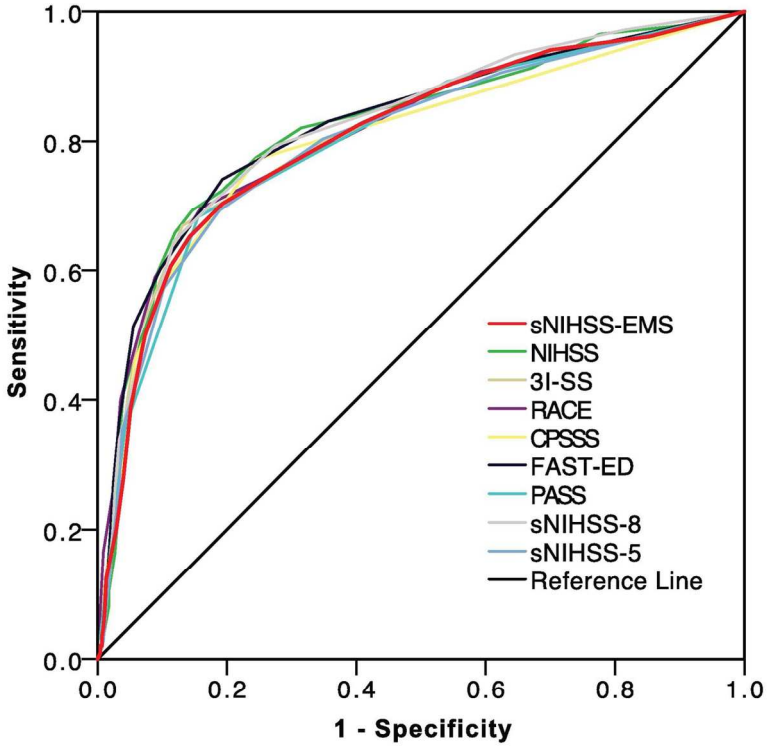
** Definition of large vessel occlusions according to original publications (3I-SS: carotid-T or M1; LAMS: ICA, M1, M2, M3/4, ACA; RACE: terminal ICA, M1, tandem CCA/ICA+M1, BA; CPSSS: ICA, M1, tandem ICA+M2, BA; FAST-ED: ICA, M1, M2, BA; PASS: “visible clot in the anterior or posterior circulation on CTA or MRA“; abbreviations within the main text).

†† Including cases assessed by transcranial duplex only (N = 197).

‡‡ Only patients who received intravenous tPA; 2/3 of entire cohort were taken as a random sample for derivation. In the remaining 1/3, sensitivity was 61%, specificity 83%.

Figure 1. Receiver operating curves for prediction of acute large vessel occlusion.

Abbreviations: AUC: area under the curve; ref.: reference.



Scale	Area Under the Curve (95% CI)	*p*
sNIHSS-EMS	0.808 (0.775 – 0.841)	*ref.*
NIHSS	0.821 (0.788 – 0.854)	0.019
3I-SS	0.804 (0.770 – 0.838)	0.697
RACE	0.816 (0.784 – 0.849)	0.314
CPSSS	0.802 (0.767 – 0.836)	0.605
FAST-ED	0.825 (0.793 – 0.857)	0.078
PASS	0.805 (0.771 – 0.838)	0.766
sNIHSS-8	0.823 (0.791 – 0.855)	0.020
sNIHSS-5	0.803 (0.769 – 0.837)	0.643

Figure 1

124x173mm (300 x 300 DPI)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	2
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4/5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5/6
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5/6
	9	Whether participants formed a consecutive, random or convenience series	5/6
Test methods	**10a**	Index test, in sufficient detail to allow replication	4-7
	10b	Reference standard, in sufficient detail to allow replication	5/6
	11	Rationale for choosing the reference standard (if alternatives exist)	5/6, 7
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6/7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6/7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6/7
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6/7
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6/7
	15	How indeterminate index test or reference standard results were handled	NA
	16	How missing data on the index test and reference standard were handled	NA
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6/7
	18	Intended sample size and how it was determined	6, +Appendix
RESULTS			
Participants	**19**	Flow of participants, using a diagram	NA
	20	Baseline demographic and clinical characteristics of participants	8, +Appendix.
	21a	Distribution of severity of disease in those with the target condition	8
	21b	Distribution of alternative diagnoses in those without the target condition	Appendix, previousl. published.
	22	Time interval and any clinical interventions between index test and reference standard	NA
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	Appendix
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	8/9
	25	Any adverse events from performing the index test or the reference standard	NA
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	12
	27	Implications for practice, including the intended use and clinical role of the index test	3; 11,12,13
OTHER INFORMATION			
	28	Registration number and name of registry	NA (IRB-Prot., 12)
	29	Where the full study protocol can be accessed	NA
	30	Sources of funding and other support; role of funders	14