Arch Iran Med. June 2021;24(6):453-460

doi 10.34172/aim.2021.65

http://www.aimjournal.ir

ARCHIVES OF IRANIAN MEDICINE

Open Access

Original Article

A Multicenter Diagnostic Accuracy Study on Prehospital Stroke Screening Scales

Peyman Saberian, MD[1,2]; Hosein Rafiemanesh, PhD[3]; Farhad Heydari, MD[4]; Sahar Mirbaha, MD[5]; Somayeh Karimi, MD[1,6]; Alireza Baratloo, MD[1,6*]

[1]Prehospital and Hospital Emergency Research Center, Tehran University of Medical Sciences, Tehran, Iran
[2]Anesthesiology Department, Imam Khomeini Hospital Complex, Tehran University of Medical Sciences, Tehran, Iran
[3]Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[4]Department of Emergency Medicine, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
[5]Department of Emergency Medicine, Shohadaye Tajrish Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[6]Department of Emergency Medicine, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran

Abstract

Background: Stroke is one of the most common debilitating diseases. Although effective treatment is available, a golden time has been defined in this regard. Therefore, prompt action is needed to identify patients with stroke as soon as possible, even in the pre-hospital stage. In recent years, several clinical scales have been introduced for this purpose. We performed the present study to examine the accuracy of eight clinical scales in terms of stroke diagnosis.

Methods: This multicenter diagnostic accuracy study was conducted in 2019. All patients older than 18 years who were admitted to the emergency department (ED) and underwent brain magnetic resonance imaging (MRI) for a suspected stroke were eligible. All data were gathered through a pre-prepared checklist consisting of three sections, using the clinical records of the patients. The first section of the checklist included basic characteristics and demographic data. The second part included physical examination findings of 19 items related to the 8 scales. The third part was dedicated to the final diagnosis based on the interpretation of brain MRI, which was considered the gold standard for the diagnosis of acute ischemic stroke (AIS) in the current study.

Results: The data from 805 patients suspected of stroke were analyzed. In all, 463 patients (57.5%) were male. The participants' age was 6-95 years with a mean age of 66.9 years (SD = 13.9). Of all the registered patients, 562 (69.8%) had an AIS. The accuracy of screening tests was 63.0% to 84.4%. The sensitivity and specificity were 71.9% to 95.7% and 46.5% to 82.8%, respectively. Among all the screening tests, Los Angeles Pre-Hospital Stroke Screening (LAPSS) had the lowest sensitivity, and Medic Prehospital Assessment for Code Stroke (Med PACS) had the highest sensitivity. In addition, PreHospital Ambulance Stroke Test (PreHAST) had the lowest specificity and LAPSS had the highest specificity.

Conclusion: Based on the findings of the present study, highly sensitive tests that can be used in this regard are Cincinnati Prehospital Stroke Scale (CPSS), Face-Arm-Speech-Time (FAST), and Med PACS, all of which have about 95% sensitivity. On the other hand, none of the studied tools were desirable (specificity above 95%) in any of the examined cut-offs.

Keywords: Decision support techniques, Early diagnosis, Emergency medical services, Stroke

Cite this article as: Saberian P, Rafiemanesh H, Heydari F, Mirbaha S, Karimi S, Baratloo A. A multicenter diagnostic accuracy study on prehospital stroke screening scales. Arch Iran Med. 2021;24(6):453–460. doi: 10.34172/aim.2021.65

Received: July 11, 2020, Accepted: March 3, 2021, ePublished: June 1, 2021

Introduction

Acute ischemic stroke (AIS) is one of the most common and debilitating diseases, ranking third in the most common causes of death following heart disease and cancer. On average, every three seconds, one person dies due to stroke.[1] Brain cells are highly susceptible to ischemia, and if a large vein is blocked by thrombosis, about 1.9 million neurons are lost per minute. Therefore, each hour of delay in stroke treatment results in the loss of numerous brain cells that a human would lose in 3.6 years of his/her normal life.[2] Effective treatments is available for AIS, but a golden time has been defined in this regard. It has been reported that only 1%–8% of stroke patients receive proper treatment and the others face poor outcomes due to delayed referral.[3-6] Therefore, prompt action is vital to identify patients with stroke as soon as possible, even in the pre-hospital stage.[7]

Although public education at the community level plays an undeniable role in promptly calling emergency medical services (EMS) after a stroke, effective interventions in the healthcare system begin from the moment that a patient or his/her companion contacts the EMS. Naturally, paraclinical diagnostic tests before hospitalization are almost impossible, and the diagnosis must be made solely on the basis of clinical presentations. Accurate and timely diagnosis allows the patient to be referred to the right place at the right time.

In recent years, several clinical scales have been

***Corresponding Author:** Alireza Baratloo, MD; Department of Emergency Medicine, Sina Hospital, Hasanabad Square, Tehran, Iran. Tel: +98-912-2884364; Email: arbaratloo@sina.tums.ac.ir

introduced for this purpose, including recognition of stroke in the emergency room (ROSIER), Los Angeles Pre-Hospital Stroke Screening (LAPSS), Face-Arm-Speech-Time (FAST), Cincinnati Pre-hospital Stroke Scale (CPSS), Medic Prehospital Assessment for Code Stroke (Med PACS), Ontario Pre-Hospital Stroke Screening (OPSS), Melbourne Ambulance Stroke Screen (MASS) and PreHAST.[8-15] Choosing a scale depends on both its accuracy and ease of use. Confirming the accuracy of a scale by comparing it to other pre-hospital scales can play an important role in accurate diagnosis of acute stroke and thus, increases the chance of the patient benefiting from successful treatment. Therefore, we performed the present study to examine the accuracy of these criteria in terms of stroke diagnosis in patients admitted to the emergency department (ED) via a multicenter research project. However, in our study, the effectiveness of these scales was not evaluated in the field, and it is recommended for future studies.

Materials and Methods

Study Design

This diagnostic accuracy study was conducted in 2019, using a multi-centric approach, including four major teaching hospitals in Iran (Sina and Shohaday-e-Tajrish hospital in Tehran; Al-Zahra hospital in Isfahan; Golestan General hospital in Ahvaz).

Study Population

All patients who were referred to the ED of the mentioned hospitals, and underwent a brain magnetic resonance imaging (MRI) for a suspicious stroke after the evaluation of an in-charge physician were included in the study. Those with a history of head trauma, previous stroke, known neurological disease or previous neurological surgery, and those who had left the ED against medical advice before undergoing brain MRI were excluded.

The sample size in this study was calculated based on the sample size formula for estimating the difference of the area under the curve (AUC) in the two stroke screening tools. We required 260 positive and negative patients for detection of a 5% difference in the AUC between two stroke screening tools based on the following assumption: 0.5 error type I, 80% test power to detect difference, lower AUC = 0.85 and 0.50 correlation between the two AUCs. Given that we expected 30% of suspected cases to have a stroke, we needed at least 800 suspicious stroke patients. The required sample size for each hospital was determined based on the proportion of patients suspected of stroke admitted in 2017. Then, in each center, all patients meeting the inclusion criteria entered the study from January 2018 until the intended sample size was achieved.

Data Gathering

From almost three years ago, on admission to the ED, an appropriate checklist was included in the patients' hospital files to assess proper neurological examination of patients with any neurological complaints (such as focal neurological deficit, headache, seizure, etc) and all required data for calculating the scales are easily available. Actually, all the findings of neurological examination are routinely recorded when the patient arrives in the ED. The data were gathered through a pre-prepared checklist consisting of three sections, using the patients' clinical records. The first section of the checklist includes basic characteristics and demographic data such as age, gender, past medical history, drug history, and the time of symptom onset. The second part includes physical examination findings of 19 items related to the 8 scales along with other manifestations such as vital signs, blood sugar level, and level of consciousness. The third part is dedicated to the final diagnosis based on the interpretation of brain MRI, which was considered as the gold standard for the diagnosis of AIS in the current study. All data were gathered under the supervision of an emergency medicine resident and three emergency medicine specialists. The required data were collected from the patients' records as well as the MRI images available in the hospital's picture archiving and communication system (PACS). The brain MRI scans were interpreted by both a radiologist and a neurologist.

Statistical Analysis

We described data using frequency and percentage or mean and standard deviation (SD). We used the chi-square test to assess the distribution difference of demographic characteristics and the history of diseases, as well as risk factors between patients with and without a final diagnosis of stroke. Additionally, the independent *t* test was used for assessment of the mean difference in numerical variables such as age, between the two groups of patients. The normality of variables was assessed using the Kolmogorov-Smirnov test and graphical approaches. Also, we checked the homogeneity of variance using Levene's test.

We calculated the sensitivity, specificity, and positive and negative likelihood ratios of all eight screening tests with 95% confidence intervals (CIs) based on their originally defined scoring and cut-off points. The positive and negative predictive values with 95% CI for screening tests were calculated based on the patients' final diagnosis. Also, the prevalence of correct and incorrect diagnoses for each tool is presented [true positive (TP), false positive (FP), true negative (TN), and false negative (FN)]. We used the McNemar's chi-square test to compare the performance of each screening test based on the final diagnosis, and then calculated McNemar's odds ratio (OR) with 95% CI. McNemar's test presented the difference between predicted stroke cases with each screening tool and final diagnosis based on the gold standard. The sensitivities and specificities of screening tests were compared using the McNemar's Chi-square analysis described in previous articles.[16,17] First, the overall test of difference (sensitivity or specificity) between all pairwise comparisons of eight

screening tests were conducted using a 4×4 extension of McNemar's test, and if the difference was significant, then the sensitivity and specificity were compared separately using a 2×2 contingency table of McNemar's test. Finally, we used the Youden's J statistic to compare the performance of the eight screening tests. The receiver operating characteristic (ROC) curve and the AUC with 95% CI of screening tools with a numerical score (ROSIER, LAPSS, FAST, and CPSS) were calculated and their AUCs were compared (as described by DeLong et al).[18] *P* value less than 0.05 was considered statistically significant and all statistical analyses were conducted using Stata version 14 (StataCorp LP, College Station, TX).

Results

Data from 805 patients suspected of stroke, who were transferred to ED by the EMS, were analyzed. In all, 463 patients (57.5%) were male. The participants' age was 6-95 years with a mean age of 66.9 years (SD = 13.9). Of all the registered patients, 562 (69.8%) had an ischemic stroke based on the gold standard.

Table 1 reports the demographic and baseline characteristics of the studied patients. The prevalence of ischemic stroke was higher in males than females (73.9% vs. 64.3%; $P=0.004$). The history of ischemic heart disease was higher in patients with stroke (74.9% vs. 67.1%; $P=0.021$). Also, patients with stroke were older ($P<0.001$).

The prevalence of ischemic stroke based on screening tools ranged between 55.4% and 81.8%, and the lowest and highest prevalence pertained to LAPSS and Med PACS, respectively. The lowest True results (true positive and negative) reported by screening tools were seen in OPSS (74.4%) and the highest were seen in ROSIER (84.4%). The difference between the final diagnosis made in the hospital regarding stroke and the predictions by the screening tests was statistically significant ($P<0.001$), except for OPSS (67.8% vs 69.8% positive; $P=0.531$) and MASS (72.9% vs 69.8% positive; $P=0.424$). Thus, the odds of a positive diagnosis with ROSIER, FAST, CPSS, Med PACS, and PreHAST was about 3.4 to 5.0 times higher than the actual hospital diagnosis (Table 2).

Table 1. Demographic Characteristics and History of Disease and Risk Factors in Patients with and without a Final Diagnosis of Stroke

Variables	Stroke (n = 562)	Non-stroke (n = 243)	*P* value
Age; mean (SD), year	68.1 (13.2)	64.1 (14.9)	<0.001
Gender, n (%)			
Male	342 (73.9)	121 (26.1)	0.004
Female	220 (64.3)	122 (35.7)	
History of hypertension, n (%)			0.420
Positive	348 (68.8)	158 (31.2)	
Negative	213 (71.5)	85 (28.5)	
History of smoking, n (%)			0.956
Smoker	100 (69.9)	43 (30.1)	
Non-smoker	460 (69.7)	200 (30.3)	
History of IHD, n (%)			0.021
Positive	215 (74.9)	72 (25.1)	
Negative	347 (67.1)	170 (32.9)	
History of Diabetic Mellitus, n (%)			0.726
Positive	87 (68.5)	40 (31.5)	
Negative	475 (70.1)	203 (29.9)	

SD, standard deviation; IHD, ischemic heart disease.

The accuracy of screening tests was between 63.0% and 84.4%. Their sensitivity and specificity were between 50.2% to 95.7% and 46.5% to 92.2%, respectively. Also, the positive predictive value and negative predictive value of the screening tests were between 80.1% to 93.7% and 44.4% to 83.9%, respectively (Table 3).

Among all the screening tests, LAPSS had the lowest sensitivity (71.9%) and Med PACS had the highest sensitivity (95.7%). In addition, PreHAST had the lowest specificity (46.5%) and LAPSS had the highest specificity (82.8%) (Figure 1). The test of difference (sensitivity or specificity) showed a statistically significant difference between all pairs of tests in pairwise comparison, except for CPSS and FAST ($P=0.368$); so, the sensitivity and specificity were not significantly different between the two tests. However, the test of difference (sensitivity or specificity) showed a marginally significant difference between ROSIER and FAST ($P=0.060$).

The sensitivity of Med PACS, CPSS, FAST, ROSIER, and Pre-HAST was higher than 90% and significantly higher than the sensitivity of MASS, OPSS, and LAPSS ($P<0.001$). The sensitivity of PreHAST was significantly lower than that of Med PACS (93.2% vs. 95.7%, $P=0.016$), but the sensitivity of PreHAST vs. CPSS ($P=0.076$), PreHAST vs. FAST ($P=0.122$), and also PreHAST vs. ROSIER ($P=0.132$) was not significantly different.

The specificity of ROSIER, MASS, and LAPSS was higher than 60%. The specificity of LAPSS was significantly higher than that of the other tests ($P<0.001$). The specificity of MASS was significantly higher than that of FAST, CPSS, Med PACS, and PreHAST ($P<0.001$), but the specificity of MASS vs. OPSS (61.8% vs. 59.5%, $P=0.311$) and also MASS vs. ROSIER (61.8% vs. 60.1%, $P=0.418$) was not significantly different. Additionally, the specificity of MASS vs. OPSS ($P=0.311$), ROSIER vs. OPSS ($P=0.851$), and FAST vs. OPSS ($P=0.064$) was not significantly different. The specificity of OPSS was higher than that of CPSS (59.5% vs. 54.3%, $P=0.029$) as well as Med PACS and PreHAST ($P<0.001$). The specificity of ROSIER was significantly higher than that of CPSS (60.1% vs. 54.3%, $P=0.016$) as well as Med PACS and PreHAST ($P<0.001$). The specificity of CPSS was significantly higher than that of Med PACS (54.3% vs. 50.6%, $P=0.012$) and PreHAST (54.3% vs. 46.5%, $P=0.002$). The specificities of Med PACS and PreHAST were not significantly different

Table 2. Prevalence of Ischemic Stroke and Frequency of True and False Results by the Studied Screening Tools Compared to the Final Diagnosis in Hospital

Screening Tools	Scoring and Cut-off for Positive Screening	Prevalence (95% CI)	Screening Result				Mcnemar Test Result		
			TP (%)	FP (%)	TN (%)	FN (%)	Diff%[a] (95% CI)	Ratio[b] (95% CI)	OR[c] (95% CI)
ROSIER	Total scores range from -2 to +5 and score >0 is associated with possible stroke	78.4 (75.6, 81.3)	534 (66.3)	97 (12.0)	146 (18.1)	28 (3.5)	8.6 (5.8, 11.3)	1.12 (1.08, 1.2)	3.5 (2.3, 5.5)
LAPSS	If the patient is positive for all six criteria, they are a positive screen for stroke	55.5 (52.1, 59.0)	404 (50.2)	42 (5.2)	201 (25.0)	158 (19.6)	-14.4 (-17.8, -11.0)	0.79 (0.75, 0.84)	0.27 (0.18, 0.38)
FAST	Abnormality demonstrated on one or more of the three items is indicative of suspected stroke	79.8 (77.0, 82.6)	533 (66.2)	109 (14.5)	134 (16.6)	29 (3.6)	9.9 (7.0, 12.8)	1.14 (1.1, 1.2)	3.8 (2.5, 5.9)
CPSS	Abnormality demonstrated on one or more of the three items is indicative of suspected stroke	80.2 (77.4, 82.9)	534 (66.3)	111 (13.8)	132 (16.4)	28 (3.5)	10.3 (7.4, 13.2)	1.15 (1.1, 1.2)	4.0 (2.6, 6.2)
Med PACS	Presence of any physical assessment item (five items) + a "yes" response to at least one eligibility criterion (three items)	81.8 (79.1, 84.4)	538 (66.8)	120 (14.9)	123 (15.3)	24 (3.0)	11.9 (9.0, 14.8)	1.17 (1.1, 1.2)	5.0 (3.2, 8.1)
OPSS	If the patient has at least one of the inclusion criteria and none of the exclusion criteria	68.7 (65.5, 71.9)	454 (56.5)	98 (12.2)	144 (17.9)	108 (13.4)	-1.2 (-4.9, 2.4)	0.98 (0.93, 1.0)	0.91 (0.68, 1.2)
PreHAST	Patients with deficit in any PreHAST item (score 1–19) are defined as Pre-HAST positive	81.3 (78.6, 84.0)	524 (65.1)	130 (16.1)	113 (14.0)	38 (4.7)	11.4 (8.2, 14.6)	1.16 (1.1, 1.2)	3.4 (2.4, 5.0)
MASS	Presence of any of the four physical assessment items + a "yes" response to all four history items	72.9 (69.8, 76.1)	477 (61.9)	84 (10.9)	136 (17.7)	73 (9.5)	1.4 (-1.8, 4.7)	1.02 (0.98, 1.1)	1.1 (0.83, 1.6)

CI, Confidence interval; TP, True positive; FP, False positive; TN, True negative; FN, False negative; ROSIER, Recognition of Stroke in the Emergency Room Scale; LAPSS, Los Angeles Prehospital Stroke Screen; FAST, Face Arm Speech Test; CPSS, Cincinnati Pre-Hospital Stroke Scale; PACS, Prehospital Acute Stroke Severity Scale; Med PACS, Medic Pre-hospital Assessment for Code Stroke; OPSS, Ontario Prehospital Stroke Screen; PreHAST, PreHospital Ambulance Stroke Test; MASS, Melbourne Ambulance Stroke Screen.
[a]Stroke prevalence difference between screening tools and final diagnosis.
[b]Ratio of stroke prevalence in screen tools to final diagnosis.
[c]McNemar odds ratio.

Table 3. Sensitivity, Specificity, Likelihood Ratio, Predictive Value and Other Accuracy Indices of the Eight Studied Screening Tools in Prediction of Acute Ischemic Stroke

Screening Tools	Sensitivity (95% CI)	Specificity (95% CI)	Youden's J	PLR (95% CI)	NLR (95% CI)	PPV (95% CI)	NPV (95% CI)	Accuracy (95% CI)
ROSIER	95.0 (92.9, 96.7)	60.1 (53.6, 66.3)	55.1	2.4 (2.0, 2.8)	0.08 (0.06, 0.1)	84.6 (81.6, 87.4)	83.9 (77.6, 89.0)	84.4% (81.9, 87.0)
LAPSS	71.9 (68.0, 75.6)	82.8 (77.5, 87.3)	54.7	4.2 (3.1, 5.5)	0.34 (0.3, 0.4)	90.6 (87.5, 93.1)	56.1 (50.8, 61.3)	75.2% (72.2, 78.2)
FAST	94.8 (92.7, 96.5)	55.1 (48.7, 61.5)	49.9	2.1 (1.8, 2.4)	0.09 (0.06, 0.1)	83.0 (79.9, 85.8)	82.2 (75.5, 87.7)	82.9% (80.3, 86.5)
CPSS	95.0 (92.9, 96.7)	54.3 (47.8, 60.7)	49.3	2.1 (1.8, 2.4)	0.09 (0.06, 0.1)	82.8 (79.7, 85.6)	82.5 (75.7, 88.0)	82.7% (80.1, 85.4)
Med PACS	95.7 (93.7, 97.2)	50.6 (44.2, 57.1)	46.3	1.9 (1.7, 2.2)	0.08 (0.06, 0.1)	81.8 (78.6, 84.6)	83.7 (76.7, 89.3)	82.1% (79.5, 84.8)
OPSS	80.8 (77.3, 84.0)	59.5 (53.0, 65.7)	40.3	2.0 (1.7, 2.3)	0.32 (0.3, 0.4)	82.2 (78.8, 85.3)	57.1 (50.8, 63.3)	74.4% (71.4, 77.4)
PreHAST	93.2 (90.8, 95.2)	46.5 (40.1, 53.0)	39.7	1.7 (1.5, 2.0)	0.15 (0.1, 0.2)	80.1 (76.9, 83.1)	74.8 (67.1, 81.5)	79.1% (76.3, 81.9)
MASS	86.73 (83.6, 89.5)	61.8 (55.0, 68.3)	48.5	2.3 (1.9, 2.7)	0.21 (0.2, 0.3)	85.0 (81.8, 87.9)	65.1 (58.2, 71.5)	79.6% (76.8, 82.5)

PLR, Positive likelihood ratio; NLR, Negative likelihood ratio; PPV, Positive predictive value; NPV, Negative predictive value; CI, Confidence interval; ROSIER, Recognition of Stroke in the Emergency Room Scale; LAPSS, Los Angeles Prehospital Stroke Screen; FAST, Face Arm Speech Test; CPSS, Cincinnati Pre-Hospital Stroke Scale; PACS, Prehospital Acute Stroke Severity Scale; Med PACS, Medic Pre-hospital Assessment for Code Stroke; OPSS, Ontario Prehospital Stroke Screen; PreHAST, PreHospital Ambulance Stroke Test; MASS, Melbourne Ambulance Stroke Screen.

(50.6% vs. 46.5%, $P=0.132$).

The Youden index for ROSIER and LAPSS was 55.1% and 54.7%, respectively, which was higher than that of the other tests. Therefore, based on this index and assuming that sensitivity and specificity have equal importance, ROSIER and LAPSS had better performance compared to others (Table 3).

Among the screening tools with a numerical score, the AUC of ROSIER and FAST was higher than CPSS and LAPSS. The AUC of both ROSIER and FAST was 0.850, which was significantly higher than the AUC of LAPSS ($P = 0.002$). The pairwise comparison of AUC-ROC was not significantly different for ROSIER vs. CPSS ($P=0.672$), or FAST vs. CPSS ($P=0.245$) (Figure 2).

Discussion

According to the results of the analysis, Med PACS has the highest sensitivity among the 9 assessed tools at cut-off = 1; it also has the highest sensitivity at cut-off = 3. FAST, which is currently used by the Iranian EMS to detect stroke, has a sensitivity of almost 95% at cut-off = 1. Obviously, in pre-hospital settings, the sensitivity of a test is much more important than its specificity, and the tendency to screen correctly and not to miss positive cases is a priority. Therefore, based on the findings of the present study, highly sensitive tests that can be used in this regard are CPSS, FAST, and Med PACS, all of which have about 95% sensitivity. On the other hand, in hospital settings, where diagnoses are expected to be more precise and specialized, examinations should be applied to avoid wasting resources, so tests with higher specificity are required. Unfortunately, none of the studied tools were desirable (specificity above 90%) in any of the examined cut-offs; so, in order to define a criterion for ruling out the diagnosis of stroke in the ED with a clinical rule, it may be necessary to perform more analysis and consider designing a new scoring system for this purpose.

The sensitivity and specificity of these scales have been assessed in various studies, as well as their primary

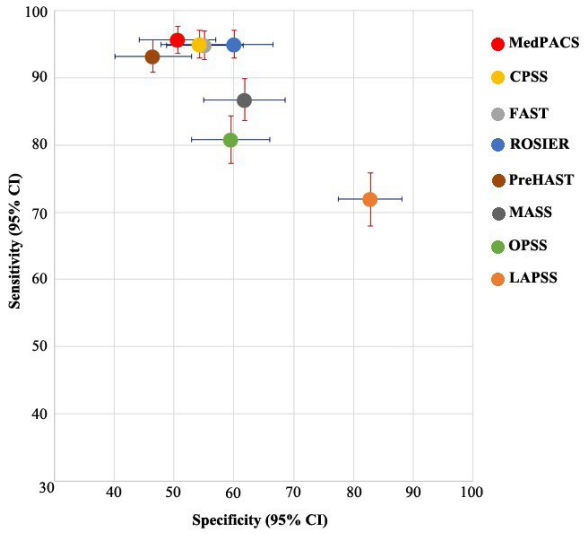

Figure 1. Sensitivity and Specificity of the Eight Studied Screening Tools in Predicting Ischemic Stroke with 95% CI.

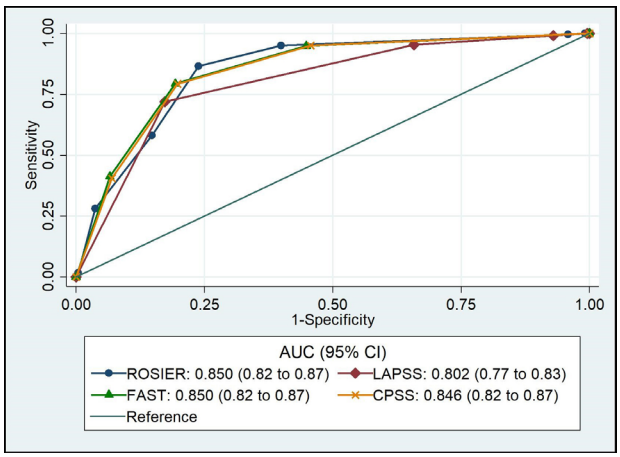

Figure 2. ROC Curve and AUC of the Four Studied Ischemic Stroke Screening Tools Yielding Numerical Scores, with 95% CI. ROC, receiver operating characteristic; AUC, area under the ROC curve.

derivation study. We have tried to summarize the results of some of these papers in Table 4. Based on available reports, the lowest and the highest reported sensitivities pertained to LAPSS (49%, reported by Purrucker et al[19]) and PreHAST (100%, reported by Andsberg et al[11]), respectively. The lowest and highest reported specificity pertained to FAST (13%, reported by Fothergill et al[27]) and LAPSS (99%, reported by Wojner-Alexandrov et al[28]),

Table 4. Characteristics of Prehospital Stroke Diagnostic Scales According to Previous Studies

Author	Tool	Sensitivity	Specificity
Purrucker et al[19]	CPSS	83%	69%
	FAST	85%	68%
	LAPSS	49%	97%
	MASS	63%	94%
	Med PACS	71%	92%
	ROSIER	80%	79%
Bray et al[20]	MASS	90%	74%
	CPSS	95%	56%
	LAPSS	78%	85%
Bray et al[21]	MASS	83%	86%
	CPSS	88%	79%
Bergs et al[22]	FAST	95%	33%
	CPSS	95%	33%
	LAPSS	74%	83%
	MASS	74%	67%
Asimos et al[23]	CPSS	80%	48%
	LAPSS	74%	48%
Chen et al[24]	LAPSS	78%	90%
Andsberg et al[11]	PreHAST	100%	40%
Studenk et al[25]	Med PACS	74%	33%
	CPSS	79%	24%
Kidwell et al[26]	LAPSS	91%	97%
Fothergill et al[27]	FAST	97%	13%
	ROSIER	97%	18%
Wojner-Alexandrov et al[28]	LAPSS	86%	99%
Chenkin et al[29]	OPSS	92%	86%

ROSIER, Recognition of Stroke in the Emergency Room; LAPSS, Los Angeles Pre-Hospital Stroke Screening; FAST, Face-Arm-Speech-Time; CPSS, Cincinnati Pre-hospital Stroke Scale; Med PACS, Medic Pre-hospital Assessment for Code Stroke; OPSS, Ontario Pre-Hospital Stroke Screening; MASS, Melbourne Ambulance Stroke Screen; PreHAST, Pre-Hospital Ambulance Stroke Test.

respectively.[11,19-29] In our study, the lowest and highest sensitivity pertained to LAPSS (71.9%) and Med PACS (95.7%), respectively. In addition, the lowest and highest specificity pertained to PreHAST (46.5%) and LAPSS (82.8%).

Each of these criteria has its strengths and weaknesses. PreHAST, LAPSS, MASS, and OPSS have considered more details, and therefore, completing their checklists is time-consuming and also difficult without specific training.[9-11,15,19] On the other hand, patient assessment with FAST and CPSS is very easy and feasible for almost everyone and does not require any special training. These two tools do not consider lower limbs and eye symptoms. However, it should be mentioned that, given the lack of exclusion criteria, they may declare stroke-mimic cases as false-positive stroke.[11,12]

MASS was indeed designed through integrating LAPSS and CPSS. LAPSS and MASS exclude patients with a history of seizures, those younger than 45 years, bedridden patients and those in a wheelchair. LAPSS has tried to increase specificity and sensitivity by examining blood glucose level and unilateral symptoms. The time of symptom onset has been taken into account by LAPSS but not by MASS. On the other hand, speech difficulty is assessed by MASS but not by LAPSS. In comparison with MASS and LAPSS, Med PACS considers seizure, the onset of symptoms, and blood glucose level, but age has not been taken into account.[14,20,21]

OPSS does not consider the age and eye symptoms but excludes hypoglycemic and terminally ill patients as well as those under palliative care, and those with a transient ischemic attack and Glasgow coma scale < 10.[29]

It is well-known that hypoglycemia is a stroke-mimic diagnosis that could easily be differentiated using a bed side testing of blood glucose, but this is not considered in CPSS, FAST, ROSIER, and PreHAST. It seems that it is an important weak point that leads to an increase in the number of false-positive stroke diagnoses in the pre-hospital setting when these tools are used.[9,25,27]

History of seizure has been considered as a negative point in LAPSS, MASS, Med PACS, OPSS, and ROSIER, but not in FAST, CPSS, and PreHAST. It is known that seizure could occur due to stroke; on the other hand, the post-ictal phase of seizure may mimic stroke. So, it is very challenging to decide to ignore seizure or assign a negative score to it.[9,10,19]

Pre-HAST is a new tool that has been designed based on NIHSS and has tried to cover everything, so completing its checklist is time-consuming and also difficult without training. Age, blood sugar level, history of seizures, and the time of symptom onset are not taken into account. In this scale, all four limbs are examined, so generalized or symmetric weakness can lead to a false positive decision. In general, eliminating those with a history of seizures and those younger than 45 years can cause adverse events, as stroke can also occur in young people, and seizures can be a symptom of a stroke.[11,15]

ROSIER has assigned negative scores to seizure and syncope in order to better differentiate stroke and stroke mimics; also, by adding "new onset of symptoms", it has helped differentiate new stroke cases from old ones.[13,27]

The key point that should be noted regarding the present study is that the instruments were only compared to a gold standard, namely MRI, and their effectiveness in dealing with patients on the scene may differ from the reported findings due to many reasons. For example, the level of knowledge and experience of the emergency medical technicians (EMTs) in this field is very important and a specific scale may not be useful due to difficulty on the scene. Therefore, in future studies, the efficacy of these tools should be examined at the time of dealing with patients in the pre-hospital setting. It will also be understandable if different results are achieved in different communities, as the level of knowledge and educational backgrounds of EMTs obviously vary in different countries.

Overall, the authors of the present article believe that imposing age restrictions might lead to missing young individuals with stroke, who in fact benefit more from treatment compared to old patients. Dealing with seizure is very challenging, as it can either occur due to stroke or stroke-like symptoms may manifest following its occurrence. Monitoring the patients' blood sugar level is definitely important and should be performed as part of routine vital signs examinations so that hypoglycemia cases can be easily eliminated. A previous history of stroke can largely affect the findings of physical examination. It might not lead to false negative results, but it will probably increase the number of false positive cases. Knowledge and skill of EMTs affect the findings of physical examinations; although increasing the number of items that should be considered might increase the accuracy of the screening tool, it might make the evaluation more difficult for EMTs and consequently, hinder the desired outcome. However, the use of calculators and telemedicine can be helpful in this regard.

This study was conducted retrospectively on registered data and was not conducted in the field. On the other hand, the strengths of the work are the number of registered patients as well as a multicenter approach, which adds to its reproducibility.

In conclusion, based on the findings of the present study, highly sensitive tests that can be used in this regard are CPSS, FAST, and Med PACS, all of which have about 95% sensitivity. On the other hand, none of the studied tools were desirable (specificity above 95%) in any of the examined cut-offs.

Authors' Contribution

AB, SK and PS: The conception and design of the work. SK, SM and FH: Data acquisition. HR and AB: Analysis and interpretation of data. AB, HR and SK: Drafting the work. PS, SM and FH: Revising it critically for important intellectual content. All the authors approved the final version to be published and agree to be accountable for all aspects of the work in ensuring

that questions related to the accuracy or integrity of any part of the work.

Conflict of Interest Disclosures

The authors declare that there is no conflict of interest

Ethical Statement

The protocol of the study was approved by the ethics committee of Vice-Chancellor in Research Affairs - Tehran University of Medical Sciences and the code IR.TUMS.VCR.REC.1398.717 has been assigned to it. All the diagnostic and therapeutic processes had been conducted based on in-charge physician's decision; so, no extra charge was imposed on patients or the treatment system. All patients' data were recorded, analyzed and reported anonymously. Consent to participate is not applicable.

Funding

This study was funded with a grant from Tehran EMS Center (98-02-159-43571). The project was commissioned by the organization and the cost was paid to the contractor.

Acknowledgments

This study was conducted with the support of Prehospital and Hospital Emergency Research Center, Tehran University of Medical Sciences, Tehran, Iran. The present article was extracted from the thesis by Dr. Somayeh Karimi to achieve her specialist degree in emergency medicine from Tehran University of Medical Sciences.

References

1. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation. 2019;139(10):e56-e528. doi: 10.1161/cir.0000000000000659.
2. Saver JL. Time is brain--quantified. Stroke. 2006;37(1):263-6. doi: 10.1161/01.STR.0000196957.55928.ab.
3. Fassbender K, Balucani C, Walter S, Levine SR, Haass A, Grotta J. Streamlining of prehospital stroke management: the golden hour. Lancet Neurol. 2013;12(6):585-96. doi: 10.1016/s1474-4422(13)70100-5.
4. Liberman AL, Prabhakaran S. Stroke chameleons and stroke mimics in the emergency department. Curr Neurol Neurosci Rep. 2017;17(2):15. doi: 10.1007/s11910-017-0727-0.
5. Elfil M, Eldokmak M, Baratloo A, Ahmed N, Amin HP, Koo BB. Pathophysiologic mechanisms, neuroimaging and treatment in wake-up stroke. CNS Spectr. 2020;25(4):460-7. doi: 10.1017/s1092852919001354.
6. Baratloo A, Forouzanfar MM, Hashemi B, Safari S, Delavar Kasmaei H, Rouhipour A, et al. Tissue plasminogen activator: a literature review. Arch Neurosci. 2016;3(1):e30452. doi: 10.5812/archneurosci.30452.
7. Baratloo A, Rahimpour L, Abushouk AI, Safari S, Lee CW, Abdalvand A. Effects of telestroke on thrombolysis times and outcomes: a meta-analysis. Prehosp Emerg Care. 2018;22(4):472-84. doi: 10.1080/10903127.2017.1408728.
8. Yew KS, Cheng E. Acute stroke diagnosis. Am Fam Physician. 2009;80(1):33-40.
9. Brandler ES, Sharma M, Sinert RH, Levine SR. Prehospital stroke scales in urban environments: a systematic review. Neurology. 2014;82(24):2241-9. doi: 10.1212/wnl.0000000000000523.
10. Rudd MP, Price CI, Ford GA. Prehospital stroke scales in urban environments: a systematic review. Neurology. 2015;84(9):962. doi: 10.1212/wnl.0000000000001354.
11. Andsberg G, Esbjörnsson M, Olofsson A, Lindgren A, Norrving B, von Euler M. PreHospital Ambulance Stroke Test-pilot study of a novel stroke test. Scand J Trauma Resusc Emerg Med. 2017;25(1):37. doi: 10.1186/s13049-017-0377-x.
12. Saberian P, Tavakoli N, Hasani-Sharamin P, Aghili M, Baratloo A. Accuracy of stroke diagnosis using FAST (Face, Arm, Speech, Time) tool by emergency medical service dispatchers and technicians and its impact on transport time. Arch Neurosci. 2020;7(1):e98691. doi: 10.5812/ans.98691.
13. Zangi M, Karimi S, Mirbaha S, Sotoodehnia M, Rasooli F, Baratloo A. The validity of recognition of stroke in the emergency room (ROSIER) scale in the diagnosis of Iranian patients with acute ischemic stroke in the emergency department. Turk J Emerg Med. 2021;21(1):1-5. doi: 10.4103/2452-2473.301914.
14. Karimi S, Motamed H, Aliniagerdroudbari E, Babaniamansour S, Jami A, Baratloo A. The Prehospital Ambulance Stroke Test vs. the Cincinnati Prehospital Stroke Scale: a diagnostic accuracy study. Australas J Paramedicine. 2020;17:1-7. doi: 10.33151/ajp.17.784.
15. Karimi S, Heydari F, Mirbaha S, Elfil M, Baratloo A. Accuracy of prehospital ambulance stroke test in terms of diagnosis of patients with acute ischemic stroke: a multi-center study. Curr J Neurol. 2020;19(4):196-9. doi: 10.18502/cjn.v19i4.5547.
16. Trajman A, Luiz RR. McNemar chi2 test revisited: comparing sensitivity and specificity of diagnostic examinations. Scand J Clin Lab Invest. 2008;68(1):77-80. doi: 10.1080/00365510701666031.
17. Hawass NE. Comparing the sensitivities and specificities of two diagnostic procedures performed on the same group of patients. Br J Radiol. 1997;70(832):360-6. doi: 10.1259/bjr.70.832.9166071.
18. 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(3):837-45.
19. Purrucker JC, Hametner C, Engelbrecht A, Bruckner T, Popp E, Poli S. Comparison of stroke recognition and stroke severity scores for stroke detection in a single cohort. J Neurol Neurosurg Psychiatry. 2015;86(9):1021-8. doi: 10.1136/jnnp-2014-309260.
20. Bray JE, Martin J, Cooper G, Barger B, Bernard S, Bladin C. Paramedic identification of stroke: community validation of the melbourne ambulance stroke screen. Cerebrovasc Dis. 2005;20(1):28-33. doi: 10.1159/000086201.
21. Bray JE, Coughlan K, Barger B, Bladin C. Paramedic diagnosis of stroke: examining long-term use of the Melbourne Ambulance Stroke Screen (MASS) in the field. Stroke. 2010;41(7):1363-6. doi: 10.1161/strokeaha.109.571836.
22. Bergs J, Sabbe M, Moons P. Prehospital stroke scales in a Belgian prehospital setting: a pilot study. Eur J Emerg Med. 2010;17(1):2-6. doi: 10.1097/MEJ.0b013e32831040ec.
23. Asimos AW, Ward S, Brice JH, Rosamond WD, Goldstein LB, Studnek J. Out-of-hospital stroke screen accuracy in a state with an emergency medical services

protocol for routing patients to acute stroke centers. Ann Emerg Med. 2014;64(5):509-15. doi: 10.1016/j.annemergmed.2014.03.024.
24. Chen S, Sun H, Lei Y, Gao D, Wang Y, Wang Y, et al. Validation of the Los Angeles pre-hospital stroke screen (LAPSS) in a Chinese urban emergency medical service population. PLoS One. 2013;8(8):e70742. doi: 10.1371/journal.pone.0070742.
25. Studnek JR, Asimos A, Dodds J, Swanson D. Assessing the validity of the Cincinnati prehospital stroke scale and the medic prehospital assessment for code stroke in an urban emergency medical services agency. Prehosp Emerg Care. 2013;17(3):348-53. doi: 10.3109/10903127.2013.773113.
26. Kidwell CS, Starkman S, Eckstein M, Weems K, Saver JL. Identifying stroke in the field. Prospective validation of the Los Angeles prehospital stroke screen (LAPSS). Stroke. 2000;31(1):71-6. doi: 10.1161/01.str.31.1.71.
27. Fothergill RT, Williams J, Edwards MJ, Russell IT, Gompertz P. Does use of the recognition of stroke in the emergency room stroke assessment tool enhance stroke recognition by ambulance clinicians? Stroke. 2013;44(11):3007-12. doi: 10.1161/strokeaha.13.000851.
28. Wojner-Alexandrov AW, Alexandrov AV, Rodriguez D, Persse D, Grotta JC. Houston paramedic and emergency stroke treatment and outcomes study (HoPSTO). Stroke. 2005;36(7):1512-8. doi: 10.1161/01.str.0000170700.45340.39.
29. Chenkin J, Gladstone DJ, Verbeek PR, Lindsay P, Fang J, Black SE, et al. Predictive value of the Ontario prehospital stroke screening tool for the identification of patients with acute stroke. Prehosp Emerg Care. 2009;13(2):153-9. doi: 10.1080/10903120802706146.