Prediction of large vessel occlusion for ischaemic stroke by using the machine learning model random forests

Jianan Wang ![ORCID], Jungen Zhang, Xiaoxian Gong, Wenhua Zhang, Ying Zhou, Min Lou

Department of Neurology, Zhejiang University School of Medicine Second Affiliated Hospital Department of Neurology, Hangzhou, Zhejiang, China

Correspondence to
Dr Min Lou; lm99@zju.edu.cn

ABSTRACT

Backgrounds The timely identification of large vessel occlusion (LVO) in the prehospital stage is extremely important given the disease morbidity and narrow time window for intervention. The current evaluation strategies still remain challenging. The goal of this study was to develop a machine learning (ML) model to predict LVO using prehospital accessible data.

Methods Consecutive acute ischaemic stroke patients who underwent CT or MR angiography and received reperfusion therapy within 8 hours from symptom onset in the Computer-based Online Database of Acute Stroke Patients for Stroke Management Quality Evaluation-II dataset from January 2016 to August 2021 were included. We developed eight ML models to integrate National Institutes of Health Stroke Scale (NIHSS) items with demographics, medical history and vascular risk factors to identify LVO and validate its efficiency.

Results Finally, 15 365 patients were included in the training set and 4215 patients were included in the test set. On the test set, random forests (RF), gradient boosting machine and extreme gradient boosting presented area under the curve (AUC) of 0.831 (95% CI 0.819 to 0.843), which were higher than other models, and RF presented the highest specificity (0.827). In addition, the AUC of RF was higher than other scales, and the accuracy of the model was improved by 6.4% compared with NIHSS. We also found the top five items of identifying LVO were total NIHSS score, gaze deviation, level of consciousness (LOC), LOC commands and motor left leg.

Conclusions Our proposed model could be a useful screening tool to predict LVO based on the prehospital accessible medical data.

Trial registration number NCT04487340.

INTRODUCTION

Acute ischaemic stroke (AIS) with large vessel occlusion (LVO) is associated with poor functional outcomes and high mortality.[1 2] Recently, mechanical thrombectomy (MT) for LVO has revolutionised the stroke treatment for most disabling strokes, but this intervention is time critical.[3] Although an estimated 11%–20% of patients with AIS can have LVO, fewer than 2% receive MT in daily practice.[2 4–6] A relevant reason for nontreatment is delayed hospital presentation, often resulting in exclusion from treatment due to the loss of salvageable brain tissue.

Notably, most delays occur in the prehospital phase of acute stroke management.[2] To achieve the goal of rapid diagnosis, transfer and treatment, several scales are currently in use for identifying LVO in the prehospital setting, including the three-item Stroke Scale (3I-SS),[7] gaze–face–arm–speech–time test (G-FAST),[8] Cincinnati Pre-hospital Stroke Severity scale (CPSSS),[9] FAST,[10] Field Assessment Stroke Triage for Emergency Destination scale (FAST-ED),[11] Finnish Prehospital Stroke Scale (FPSS),[12] Los Angeles Motor Scale (LAMS),[13] Pre-hospital Acute Stroke Severity scale (PASS),[14] Rapid Arterial Occlusion Evaluation Scale (RACE),[15] Recognition of Stroke in the Emergency Room (ROSIER),[16] stroke vision, aphasia, neglect assessment (VAN),[17] National Institutes of Health Stroke Scale (NIHSS),[18] abbreviated NIHSS, modified NIHSS (mNIHSS), shortened versions of the NIHSS (sNIHSS), sNIHSS for emergency medical services (EMSs).[19]

However, there are still issues influencing the predictive precision of these scales to recognise LVO. Previous studies showed that the use of published prehospital prediction scales for triage decision-making left 20% of LVO undetected, and no threshold on any scale could detect LVO with both a high sensitivity and a high specificity.[3 20 21] The main reason is that the accuracy of scales for detecting LVO is related to the prevalence of symptoms, which depends on dispatch call identification and local prevalence.[2] Furthermore, the algorithms of scales are all based on the hypothesis of linear correlation between input and output parameters, they ignore the intricate associations among the input parameters, and they are unable to integrate various stroke-related parameters (eg, medical history or family history) in a user-independent fashion.

Recent new and exciting advances in the applications of machine learning (ML) in many healthcare areas have inspired innovations in the development of novel ML-based LVO diagnostic technology. ML, which is non-parametric, does not impose a particular structure on the data, and has been proven to be able to capture

non-linear relationships, may be an appropriate tool for identifying LVO, given that most of the medical data are nonlinear, non-normal, correlation structured and complex in nature. Therefore, we developed a predictive framework for LVO based on ML methods using prehospital accessible data including demographics, NIHSS items, medical history and vascular risk factors and then compared its diagnostic parameters with previously established prehospital prediction scales.

MATERIALS AND METHODS
Study population
This study came from a multicentre prospective registry, Computer-based Online Database of Acute Stroke Patients for Stroke Management Quality Evaluation (CASE-II, NCT04487340). Initiated in 2016, CASE-II was designed to examine the current status of stroke care in China, and the data would be used to help develop strategies to improve stroke care. In-hospital medical documents of consecutive stroke patients were collected through a special electronic data capture system.

We retrospectively reviewed the CASE-II dataset from January 2016 to August 2021, and enrolled consecutive AIS patients who underwent CT angiography (CTA) or time of flight MR angiography (TOF-MRA) and received reperfusion therapy within 8 hours from symptom onset, and had complete information of emergency for analysis. Patients with poor image quality due to motion artefacts were excluded. Then we split the training and test sets by time period. Patients from January 2016 to January 2021 were included in the training set, and patients from January 2021 to August 2021 were included in the test set. Details of patient characteristics and the inclusion and exclusion criteria were given in figure 1. Demographic, clinical, laboratory and imaging data at admission were recorded, including age, gender; prior antiplatelet therapy, prior anticoagulant therapy; risk factors (smoking, hypertension, atrial fibrillation (AF), diabetes mellitus, hyperlipidaemia, hyperhomocysteinaemia, coronary heart disease, congestive heart failure,

history of stroke/transient ischaemic attack (TIA) and family history of cardiovascular disease); blood pressure at admission.

LVO was defined as unilateral occlusion of intracranial internal carotid artery (ICA), or M1/M2 segments of the middle cerebral artery (MCA), or basilar artery (BA) on baseline CTA or TOF-MRA. Each patient's large vessel condition was assessed by two experienced senior neurologists, who were blinded to the study design and independently reviewed, with any disputes settled via reviewed by a third neurologist for consensus decision.

Feature selection
According to published literatures, pathophysiological consideration, and meanwhile based on the availability and convenience of items in the prehospital setting, 15 NIHSS items, age, gender, prior antiplatelet therapy, prior anticoagulant therapy, risk factors (smoking, hypertension, AF, diabetes mellitus, hyperlipidaemia, hyperhomocysteinaemia, coronary heart disease, congestive heart failure, history of stroke/TIA and family history of cardiovascular disease), and blood pressure at admission were selected at first consideration.[22–25] According to the statistical analysis, features with statistical significance, which was defined as $p<0.05$, were then chosen as the model variables. Selection of final parameters was performed using the Scikit-Learn package in Python software.

Data analysis
We chose eight common ML models: random forests (RFs), logistic regression, Extreme Gradient Boosting (XGBoost), K-Nearest Neighbour, Ada Boosting, Gradient Boosting Machine (GBM), LightGBM and artificial neural network (ANN). We developed the models using Scikit-Learn package in Python software.[26–28]

Specifically, model development was based on the dataset of training set with the retained features. Ten-fold cross-validation was performed for the model derivation and internal evaluation through dividing the dataset of training set into ten mutually exclusive parts, nine of which were used as training data for the model derivation
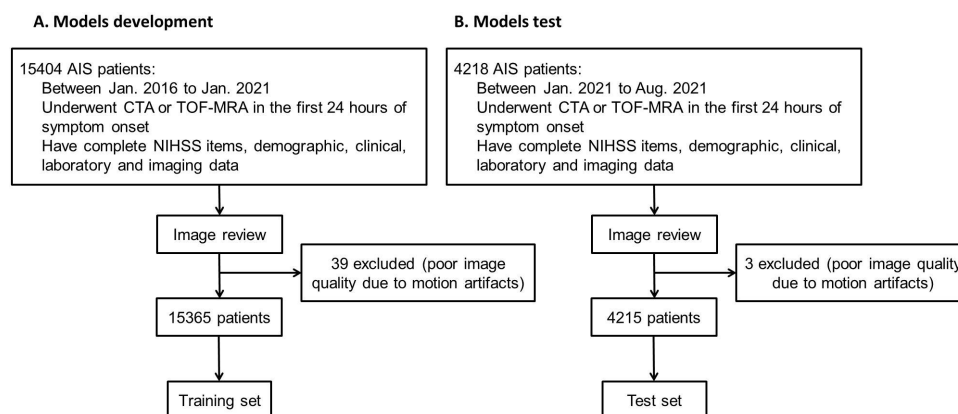
Figure 1 Flow chart of the study population and process. AIS, acute ischaemic stroke; CTA, CT angiography; NIHSS, National Institutes of Health Stroke Scale; TOF-MRA, time of flight MR angiography.

and one for evaluation as inner validation data; this process was repeated ten times to generate 10 different but overlapping training data and 10 unique validation data. In the training step, we optimised model hyperparameters with a grid search algorithm, and we did not choose the hyperparameter value until the model with the highest F1 score. During the searching process, we set the area under the curve (AUC) of receiver operating characteristic (ROC) as the score. Training of the non-linear model, which not only incorporated linear correlation between input and output parameters, but also integrated various stroke-related parameters and included the intricate associations among the input parameters, in the study was based on a dichotomisation of AIS patients into LVO versus non-LVO. By training predictive models, nonlinear combinations of the prehospital accessible data of the AIS patients could be taken into the account to predict the large artery condition. We evaluated each model in the test set and compared the predictive power with the previously established scales.

Code availability
The code used to generate results shown in this study is available from the corresponding author on request.

Statistical analysis
Patients were dichotomised into the LVO and non-LVO group. Clinical characteristics were summarised by computing the median (IQR), and differences between two groups were estimated by the t-test or Mann-Whitney U test if they were continuous variables. Categorical or binary datum was summarised by proportion (n); and differences between two groups were estimated by the Pearson χ^2 test. ROC analysis was used to get the AUC of each prehospital prediction scale. The ROC-derived optimal cut-off was determined at the maximal Youden Index. Finally, we calculated sensitivity, specificity and accuracy for the prediction of LVO. All statistical analysis was performed using SPSS, V.22.0 (IBM). All comparisons were two sided, with statistical significance defined as p<0.05.

RESULTS
Study population
Finally, 15 365 patients were included in the training set from January 2016 to January 2021, and 4215 patients were included in the test set from January 2021 to August 2021. Of the included patients, mean age was 70 (60–79) years, number of female was 7386 (37.7%), median NIHSS on admission was 6 (3–13). Baseline demographics, medical history, NIHSS and risk factors were listed in table 1. Compared with patients in the training set, those in test set were less likely to be female, have AF, have coronary heart disease, have family history of cardiovascular disease, have hyperlipidaemia, have hyperhomocysteinaemia, smoke, and more likely to have LVO and present lower baseline NIHSS score at admission (table 1).

Model performance in the test set
Table 2 shows the AUC, sensitivity, specificity and accuracy of eight models in the testing set. RF, GBM and XGBoost presented higher AUC than other five models, and RF presented higher specificity than GBM and XGBoost. Therefore, we chose the RF model as the final prediction model.

In addition, table 3 shows the comparison of AUC, sensitivity, specificity and accuracy between RF and previously established prehospital prediction scales including 3I-SS, G-FAST, CPSSS, FAST, FAST-ED, FPSS, LAMS, PASS, RACE, ROSIER, VAN, NIHSS, mNIHSS, sNIHSS and sNIHSS-EMS. Both the AUC and accuracy of RF were better than other prehospital prediction scales. Comparing with NIHSS in the study, the accuracy of the model was improved by 6.4%.

Importance of features contributing to identification of LVO
Gini importance is a measurement of the feature importance, which is defined as the total reduction of the criterion brought by that feature. Gini importance of every risk factor was then calculated. As figure 2 shows, items incorporated in NIHSS were identified as significant contributors in the LVO estimation, and the top five items were total NIHSS score (0.2381), gaze deviation (0.1412), level of consciousness (LOC) (0.0759), LOC commands (0.0613) and motor left leg (0.0567), respectively. Interestingly, systolic blood pressure (0.0233), diastolic blood pressure (0.0217) and AF (0.0159) were found to play an important role in determining LVO, suggesting not only symptoms but medical history of patients need to be taken into account together in terms of LVO identification.

DISCUSSION
To our knowledge, this was the first study that used the ML model, RF, to predict LVO based on a large sample of training and test set by far. Our results showed that the diagnostic parameters of RF for the identification of LVO were significant higher than previously established prehospital prediction scales. Furthermore, we also found total NIHSS score, gaze deviation, LOC, LOC commands and motor left leg were important items for identifying LVO.

In 2015, several trials provided compelling evidences that, for stroke due to LVO, MT resulted in significantly better recanalisation and clinical outcomes than did intravenous thrombolysis alone. After that time, LVO management has been revolutionised by the evidences of the MT effectiveness. In addition, guidelines have suggested that patients with potential LVO might benefit from direct transport to a thrombectomycapable centre, regardless of travel times.[29] Unfortunately, however, 20% of LVO remain undetected although several pre-hospital prediction scales have been developed to detect LVO by far.[20] Additionally, scales for detecting LVO are not effective for strokes in the posterior circulation, and patients with right LVO might be misdiagnosed when they present mild to moderate neurological deficits.[8 30] The accuracy of scales for detecting LVO

Table 1 Comparison of clinical characteristics between cohort of the training set and test set

	Cohort of the training set (n=15365)	Cohort of the test set (n=4215)	P value
Female, n (%)	5868 (38.2)	1518 (36.0)	0.010
Age, year, median (IQR)	70 (60–79)	71 (60–79)	0.012
Atrial fibrillation, n (%)	3274 (21.3)	804 (19.1)	0.002
Congestive heart failure, n (%)	338 (2.2)	75 (1.8)	0.104
Coronary heart disease, n (%)	1283 (8.4)	303 (7.2)	0.015
Family history of cardiovascular disease, n (%)	174 (1.1)	25 (0.6)	0.002
Smoking, n (%)	4637 (30.2)	1181 (28.0)	0.007
Hyperlipidaemia, n (%)	890 (5.8)	199 (4.7)	0.008
Diabetes mellitus, n (%)	2497 (16.3)	702 (16.7)	0.526
Hypertension, n (%)	9815 (63.9)	2624 (62.3)	0.053
History of stroke/TIA, n (%)	1968 (12.8)	560 (13.3)	0.422
Hyperhomocysteinaemia, n (%)	1011 (6.6)	237 (5.6)	0.026
Prior anticoagulation therapy, n (%)	338 (2.2)	83 (2.0)	0.400
Prior antiplatelet therapy, n (%)	2191 (14.3)	558 (13.2)	0.094
LVO, n (%)	4417 (28.7)	1279 (30.3)	0.044
Systolic blood pressure, mm Hg, median (IQR)	154 (139–168)	153 (139–167)	0.450
Diastolic blood pressure, mm Hg, median (IQR)	85 (76–94)	84 (76–94)	0.065
NIHSS sum, median (IQR)	6 (3–13)	5 (2–12)	<0.001
NIHSS items			
LOC, median (IQR)	0 (0–0)	0 (0–0)	0.288
LOC questions, median (IQR)	0 (0–1)	0 (0–0)	0.064
LOC commands, median (IQR)	0 (0–0)	0 (0–0)	0.020
Gaze deviation, median (IQR)	0 (0–0)	0 (0–0)	0.001
Visual field test, median (IQR)	0 (0–0)	0 (0–0)	0.341
Facial palsy, median (IQR)	1 (0–1)	1 (0–1)	0.003
Motor left arm, median (IQR)	0 (0–2)	0 (0–2)	<0.001
Motor right arm, median (IQR)	0 (0–2)	0 (0–2)	0.258
Motor left leg, median (IQR)	0 (0–2)	0 (0–2)	<0.001
Motor right leg, median (IQR)	0 (0–2)	0 (0–2)	0.063
Limb ataxia, median (IQR)	0 (0–0)	0 (0–0)	0.002
Sensory, median (IQR)	1 (0–1)	1 (0–1)	0.003
Aphasia, median (IQR)	0 (0–2)	0 (0–2)	0.003
Dysarthria, median (IQR)	1 (0–1)	1 (0–1)	0.014
Extinction and inattention, median (IQR)	0 (0–0)	0 (0–0)	0.678

LOC, level of consciousness; LVO, large vessel occlusion; NIHSS, National Institutes of Health Stroke Scale; TIA, transient ischaemic attack.

is related to the prevalence of symptoms, which depends on dispatch call identification and local prevalence.[2] Furthermore, the algorithms of scales are all based on the hypothesis of linear correlation between input and output parameters, ignoring the intricate associations among the input parameters, so that they are unable to integrate various stroke-related parameters. Meanwhile, most of scales are derived from NIHSS, however, some certain items of NIHSS and specified patterns of combined deficits may carry a high attributable risk of LVO, but not equal to the simple sum of score.

The advantage of the RF algorithm is that it could capture nonlinear relationships, including interactions among the input parameters and outputs until reaching high accuracy. Previously, prehospital triage tools for detection of LVO were as simple as possible, focusing solely on the elements of the neurological examination, in order to be easily memorised by EMSs personnel.[8] However, more and more literatures suggested that not only symptoms but also medical history such as atherosclerotic stenosis[24] and AF[25] were associated with LVO, thus the ignorance of various stroke-related parameters may reduce the accuracy of prediction. Our RF model contained NIHSS items and stroke-related parameters, which might be one reason for its higher predictive performance than previous scales. One might question that the assessments of these input

Table 2	Comparison of eight models to predict LVO in the test set			
	AUC (95% CI)	**SEN**	**SPE**	**Accuracy**
RF	0.831 (0.819 to 0.843)	0.721	0.827	0.772
GBM	0.831 (0.820 to 0.843)	0.721	0.825	0.772
XGBoost	0.831 (0.820 to 0.844)	0.715	0.825	0.770
LightGBM	0.828 (0.816 to 0.840)	0.721	0.826	0.774
Ada boosting	0.828 (0.817 to 0.841)	0.704	0.824	0.765
ANN	0.819 (0.817 to 0.842)	0.740	0.781	0.761
LR	0.790 (0.778 to 0.804)	0.735	0.746	0.740
KNN	0.774 (0.762 to 0.789)	0.685	0.769	0.727

ANN, artificial neural network; AUC, area under the curve; GBM, gradient boosting machine; KNN, K-Nearest Neighbour; LR, logistic regression; LVO, large vessel occlusion stroke; RF, random forests; SEN, sensitivity; SPE, specificity; XGBoost, extreme gradient boosting.

parameters may need more time. But we noticed that most of parameters were included in routine evaluation under emergency conditions, and NIHSS was still the main assessment in clinical practice.[2] Actually, in the mobile internet era when users get the intuitive outputs by easily inputting measured parameters, leaving the complex underlying algorithm to the online calculation tool or local mobile application, the ability to discriminate is more important than the simplicity of usage in LVO prediction tools.

Otherwise, we found gaze deviation, LOC, LOC commands and motor left leg incorporated in NIHSS were the most important items for identifying LVO in the study, especially for gaze deviation, whose importance was much significant than others. This finding was consistent with previous studies. Vidale and Agostoni[31] have reported that the overall accuracy would be improved considering the presence of gaze in the scoring systems, and Scheitz et al[8] have also reported addition of gaze deviation in scale could improve the specificity. In addition, several studies have reported significant correlations between gaze deviation, LOC, LOC questions, LOC commands, facial palsy and LVO.[8 9 14 15 17] We noted that several studies reported neglect had a significant correlation with LVO,[16 17] but our study did not show a significant importance. This may be potentially explained by the low incidence rate of neglect and low consistency of neglect evaluation.[23] Interestingly, we found AF was more important than sensory in terms of determining LVO, which may have some potential

Table 3	Comparison of various published clinical scales with RF model to predict LVO in the test set				
	Cut-off	**AUC (95% CI)**	**SEN**	**SPE**	**Accuracy**
RF	—	0.831 (0.819 to 0.843)	0.721	0.827	0.772
mNIHSS	≥7	0.809 (0.795 to 0.824)	0.760	0.755	0.769
sNIHSS-EMS	≥6	0.809 (0.795 to 0.824)	0.722	0.788	0.764
NIHSS	≥6	0.806 (0.792 to 0.820)	0.727	0.792	0.708
RACE	≥5	0.806 (0.791 to 0.821)	0.712	0.793	0.764
CPSSS	≥2	0.804 (0.789 to 0.819)	0.658	0.826	0.761
FAST-ED	≥4	0.804 (0.790 to 0.819)	0.611	0.850	0.760
s-NIHSS-5	≥4	0.804 (0.790 to 0.819)	0.738	0.763	0.760
3I-SS	≥4	0.798 (0.782 to 0.813)	0.641	0.852	0.759
LAMS	≥4	0.779 (0.764 to 0.795)	0.652	0.808	0.746
PASS	≥2	0.778 (0.763 to 0.794)	0.660	0.823	0.751
s-NIHSS-1	≥2	0.773 (0.757 to 0.789)	0.698	0.761	0.695
FPSS	≥5	0.762 (0.746 to 0.778)	0.781	0.620	0.711
G-FAST	≥3	0.755 (0.740 to 0.771)	0.763	0.653	0.697
VAN	≥2	0.732 (0.715 to 0.748)	0.847	0.534	0.650
ROSIER	≥4	0.730 (0.714 to 0.746)	0.886	0.467	0.694
FAST	≥3	0.706 (0.690 to 0.723)	0.682	0.681	0.685
aNIHSS	≥1	0.689 (0.672 to 0.706)	0.542	0.751	0.416

aNIHSS, abbreviated NIHSS; AUC, area under the curve; CPSSS, Cincinnati Pre-hospital Stroke Severity scale; EMS, emergency medical services; FAST-ED, Field Assessment Stroke Triage for Emergency Destination scale; FPSS, Finnish Prehospital Stroke Scale; G-FAST, gaze–face–arm–speech–time test; 3I-SS, three-item Stroke Scale; LAMS, Los Angeles Motor Scale; LVO, large vessel occlusion stroke; mNIHSS, modified NIHSS; NIHSS, National Institutes of Health Stroke Scale; PASS, Pre-hospital Acute Stroke Severity scale; RACE, Rapid Arterial Occlusion Evaluation Scale; RF, random forests; ROSIER, Recognition of Stroke in the Emergency Room; SEN, sensitivity; sNIHSS, shortened versions of the NIHSS; SPE, specificity; VAN, stroke vision, aphasia, neglect assessment.
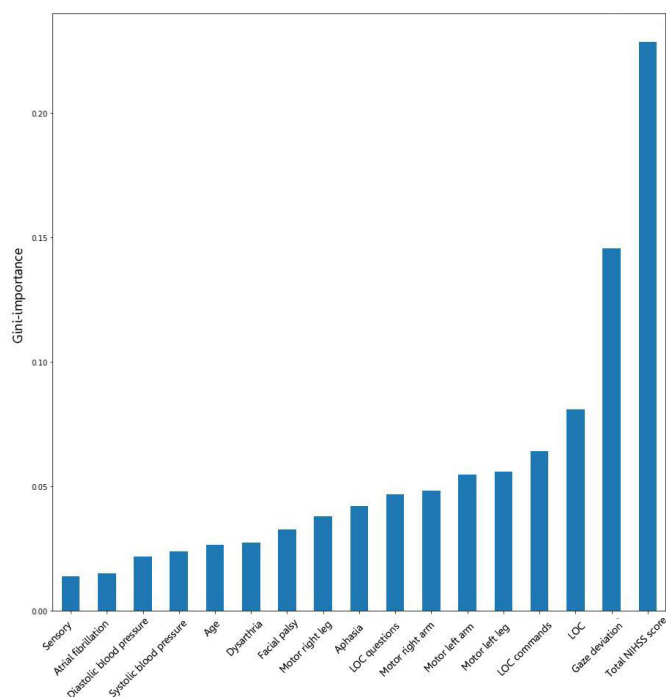
Figure 2 Illustration of features contributing to identification of LVO by Gini importance values. Gini importance is a measurement of the feature importance in the model, the higher the value of Gini importance is, the more important the feature is. LOC, level of consciousness; NIHSS, National Institutes of Health Stroke Scale.

explanations. AF allows blood to stagnate, particularly in the left atrial appendage, and both permanent and paroxysmal AF increase the risk of stroke,[1] and previous literatures have reported LVO was independently associated with AF.[25 32]

Notably, we found the precision of our model was lower than the ANN model reported previously (0.772 vs 0.820).[22] We also performed ANN model in our data to detect LVO, and found its predictive performance inferior to RF (accuracy: 0.761 vs 0.772). There were some potential reasons. First, previous cohort was small (only 600 cases) and used cross-validation to validate its predictive performance in a single cohort, which could overfit dataset and lead to more biases and wrong prediction, while our cohort was very large (19 580 cases) and used an independent cohort to evaluate the model, thus our result was more compelling. Furthermore, the populations between two studies were different. The LVO proportion in previous study was 58.3%, while the proportion in our model was 28.7% in the training set and 30.3% in the test set, which was closer to real prehospital setting (21.3%–35.9%),[8 9 11 15 23] and reflected the practical rather than theoretical predictive performance.

In addition, we noted that the AUC of NIHSS reported in some literatures was higher than that in our paper, and even higher than the AUC of our model.[23 33] This difference may be due to the different characteristics of study populations. First, the study populations in the previous literatures were much smaller than ours (543 and 178 vs 4215) and came from only one centre. However, our study came from

a multicentre prospective registry and had a large amount of subjects, therefore, the generalisability of the conclusion in our study was higher than that reported previously. Furthermore, LVO was defined as occlusion of the ICA and of proximal segments (M1, M2) of the MCA in previous studies. However, the definition of LVO in our study also included the occlusion of BA. Obviously, the prediction performance of NIHSS to detect the occlusion of BA was low, which reduced the AUC of NIHSS to predict LVO in our study. We also noted that the accuracy and specificity of our model was higher than NIHSS, while the sensitivity of ours was slightly lower than NIHSS. Due to a relatively higher specificity, the finding that our model had higher accuracy than NIHSS could be diluted. However, there are some potential reasons to explain the rational of our model. Recently, prehospital suspected LVO screening scales with high sensitivity and low specificity were reported to lead to interfacility transfer-related delays,[34] so-called 'short cuts make long delays', especially for patients with non-LVO. For scales with high sensitivity, too many non-LVO patients would be transferred to comprehensive stroke centres (CSCs), which would lead to CSCs being overburden and delays for thrombolytic therapy, too. Therefore, our model with slightly low sensitivity and high specificity is rational to reduce the non-LVO population mistakenly identified as LVO, and potentially reduce the interfacility transfer-related delays for non-LVO patients. Additionally, from clinical practice, the RFs model, with a relatively high specificity, can not only shunt patients, but also make patients get targeted treatment specifically.

Our study has some limitations. First, all patients were diagnosed with AIS and received reperfusion therapy. Consequently, the sensitivity and specificity of the RFs model for LVO might differ from pre-hospital cohorts with suspected stroke that include stroke mimics and haemorrhagic strokes. Thus, we cannot rule out a selection bias. However, our results still confirmed the feasibility of RFs to identify LVO using prehospital accessible data and gave a future direction for further studies to identify LVO in the prehospital setting. Further studies performed in the preclinical setting are necessary to generalise our results. Second, the retrospective and observational design inherits potential for bias. Our results need to be confirmed prospectively in the prehospital emergency cohorts.

CONCLUSIONS

In conclusion, our study illustrated the ML model, RF, could be a useful screening tool with an excellent accuracy to predict LVO for ischaemic stroke patients based on the prehospital accessible medical data.

Contributors JW, JZ, XG, ML were involved in the design of the study. JW, WZ and YZ did the statistical analysis and wrote the first draft. All authors contributed to the further drafts. ML was responsible for the overall content as the guarantor. JW, JZ and XG contributed equally to this paper. The corresponding author attests that all listed authors meet the authorship criteria. All authors read and approved the final manuscript.

ORCID iD
Jianan Wang http://orcid.org/0000-0003-1566-5518

REFERENCES

1. Campbell BCV, De Silva DA, Macleod MR, *et al*. Ischaemic stroke. *Nat Rev Dis Primers* 2019;5:70.
2. Fassbender K, Walter S, Grunwald IQ, *et al*. Prehospital stroke management in the thrombectomy era. *Lancet Neurol* 2020;19:601–10.
3. Patel P, Yavagal D, Khandelwal P. Hyperacute Management of Ischemic Strokes: JACC Focus Seminar. *J Am Coll Cardiol* 2020;75:1844–56.
4. Rai AT, Seldon AE, Boo S, *et al*. A population-based incidence of acute large vessel occlusions and thrombectomy eligible patients indicates significant potential for growth of endovascular stroke therapy in the USA. *J Neurointerv Surg* 2017;9:722–6.
5. Waqas M, Mokin M, Primiani CT, *et al*. Large vessel occlusion in acute ischemic stroke patients: a Dual-Center estimate based on a broad definition of occlusion site. *J Stroke Cerebrovasc Dis* 2020;29:104504.
6. Aguiar de Sousa D, von Martial R, Abilleira S, *et al*. Access to and delivery of acute ischaemic stroke treatments: a survey of national scientific societies and stroke experts in 44 European countries. *Eur Stroke J* 2019;4:13–28.
7. 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.
8. Scheitz JF, Abdul-Rahim AH, MacIsaac RL, *et al*. Clinical selection strategies to identify ischemic stroke patients with large anterior vessel occlusion: results from SITS-ISTR (safe implementation of thrombolysis in stroke international stroke thrombolysis registry). *Stroke* 2017;48:290–7.
9. 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.
10. Harbison J, Hossain O, Jenkinson D, *et al*. Diagnostic accuracy of stroke referrals from primary care, emergency room physicians, and ambulance staff using the face arm speech test. *Stroke* 2003;34:71–6.
11. 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.
12. Ollikainen JP, Janhunen HV, Tynkkynen JA, *et al*. The Finnish prehospital stroke scale detects thrombectomy and thrombolysis Candidates-A propensity score-matched study. *J Stroke Cerebrovasc Dis* 2018;27:771–7.
13. Noorian AR, Sanossian N, Shkirkova K, *et al*. Los Angeles motor scale to identify large vessel occlusion: prehospital validation and comparison with other screens. *Stroke* 2018;49:565–72.
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. Pérez de la Ossa N, Carrera D, Gorchs M, *et al*. Design and validation of a prehospital stroke scale to predict large arterial occlusion: the rapid arterial occlusion evaluation scale. *Stroke* 2014;45:87–91.
16. Fothergill RT, Williams J, Edwards MJ, *et al*. Does use of the recognition of stroke in the emergency room stroke assessment tool enhance stroke recognition by ambulance clinicians? *Stroke* 2013;44:3007–12.
17. 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.
18. Powers WJ, Rabinstein AA, Ackerson T, *et al*. 2018 guidelines for the early management of patients with acute ischemic stroke: a guideline for healthcare professionals from the American heart Association/American stroke association. *Stroke* 2018;49:e46–110.
19. Purrucker JC, Härtig F, Richter H, *et al*. 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. *BMJ Open* 2017;7:e016893.
20. Turc G, Maïer B, Naggara O, *et al*. Clinical scales do not reliably identify acute ischemic stroke patients with large-artery occlusion. *Stroke* 2016;47:1466–72.
21. Smith EE, Kent DM, Bulsara KR, *et al*. Accuracy of prediction instruments for diagnosing large vessel occlusion in individuals with suspected stroke: a systematic review for the 2018 guidelines for the early management of patients with acute ischemic stroke. *Stroke* 2018;49:e111–22.
22. Chen Z, Zhang R, Xu F, *et al*. Novel prehospital prediction model of large vessel occlusion using artificial neural network. *Front Aging Neurosci* 2018;10:181.
23. Beume L-A, Hieber M, Kaller CP, *et al*. Large vessel occlusion in acute stroke. *Stroke* 2018;49:2323–9.
24. Hendrix P, Sofoluke N, Adams MD, *et al*. Risk factors for acute ischemic stroke caused by anterior large vessel occlusion. *Stroke* 2019;50:1074–80.
25. Inoue M, Noda R, Yamaguchi S, *et al*. Specific factors to predict large-vessel occlusion in acute stroke patients. *J Stroke Cerebrovasc Dis* 2018;27:886–91.
26. Venthur B, Dähne S, Höhne J, *et al*. Wyrm: a brain-computer interface toolbox in python. *Neuroinformatics* 2015;13:471–86.
27. Pedregosa F, Varoquaux G, Gramfort A, *et al*. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011;12:2825–30.
28. Breiman L. Random forests. *Mach Learn* 2001;45:5–32.
29. Pride GL, Fraser JF, Gupta R, *et al*. Prehospital care delivery and triage of stroke with emergent large vessel occlusion (ELVO): report of the standards and guidelines Committee of the Society of Neurointerventional surgery. *J Neurointerv Surg* 2017;9:802–12.
30. Jeffery G. Distribution and trajectory of uncrossed axons in the optic nerves of pigmented and albino rats. *J Comp Neurol* 1989;289:462–6.
31. Vidale S, Agostoni E. Prehospital stroke scales and large vessel occlusion: a systematic review. *Acta Neurol Scand* 2018;138:24–31.
32. Pagola J, Juega J, Francisco-Pascual J, *et al*. Large vessel occlusion is independently associated with atrial fibrillation detection. *Eur J Neurol* 2020;27:1618–24.
33. V Martins-Filho RKdo, Dias FA, Alves FFA, *et al*. Large vessel occlusion score: a screening tool to detect large vessel occlusion in the acute stroke setting. *J Stroke Cerebrovasc Dis* 2019;28:869–75.
34. Schwamm LH, Ntaios G. Short cuts make long delays: getting it right from the start in prehospital stroke triage. *Neurology* 2020;94:341–2.