Received: 18 August 2021 | Accepted: 10 November 2021

DOI: 10.1111/ene.15181

european journal of neurology
the official journal of the european academy of neurology

ORIGINAL ARTICLE

Predicting functional outcome in patients with acute brainstem infarction using deep neuroimaging features

Lingling Ding[1,2,3,4] | Ziyang Liu[5] | Ravikiran Mane[4] | Shuai Wang[4] | Jing Jing[1,2,3] | He Fu[4] | Zhenzhou Wu[4] | Hao Li[1,2] | Yong Jiang[1,2] | Xia Meng[1,2] | Xingquan Zhao[1,2,3] | Tao Liu[4] | Yongjun Wang[1,2,3] | Zixiao Li[1,2,3,6]

[1]Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, Beijing, China

[2]China National Clinical Research Center for Neurological Diseases, Beijing, China

[3]Research Unit of Artificial Intelligence in Cerebrovascular Disease, Chinese Academy of Medical Sciences, Beijing, China

[4]China National Clinical Research Center-Hanalytics Artificial Intelligence Research Centre for Neurological Disorders, Beijing, China

[5]School of Biological Science and Medical Engineering, Beihang University, Beijing, China

[6]Chinese Institute for Brain Research, Beijing, China

Correspondence
Zixiao Li and Yongjun Wang, Department of Neurology, Beijing Tiantan Hospital, Capital Medical University, No. 119 South 4th Ring West Road, Fengtai District, Beijing, 100070, China.
Email: lizixiao2008@hotmail.com and yongjunwang@ncrcnd.org.cn

Funding information
This work was supported by grants from the Beijing Natural Science Foundation (Z200016), Beijing Municipal Committee of Science and Technology (Z201100005620010), Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences (2019-I2M-5-029), Ministry of Science and Technology of the People's Republic of China (National Key R&D Program of China, 2017YFC1310901, 2016YFC0901002, 2017YFC1307905, 2015BAI12B00), Beijing Talents Project (2018000021223ZK03), and National Natural Science Foundation of China (92046016)

Abstract

Background and purpose: Acute brainstem infarctions can lead to serious functional impairments. We aimed to predict functional outcomes in patients with acute brainstem infarction using deep neuroimaging features extracted by convolutional neural networks (CNNs).

Methods: This nationwide multicenter stroke registry study included 1482 patients with acute brainstem infarction. We applied CNNs to automatically extract deep neuroimaging features from diffusion-weighted imaging. Deep learning models based on clinical features, laboratory features, conventional imaging features (infarct volume, number of infarctions), and deep neuroimaging features were trained to predict functional outcomes at 3 months poststroke. Unfavorable outcome was defined as modified Rankin Scale score of 3 or higher at 3 months. The models were evaluated by comparing the area under the receiver operating characteristic curve (AUC).

Results: A model based solely on 14 deep neuroimaging features from CNNs achieved an extremely high AUC of 0.975 (95% confidence interval [CI] = 0.934–0.997) and significantly outperformed the model combining clinical, laboratory, and conventional imaging features (0.772, 95% CI = 0.691–0.847, $p < 0.001$) in prediction of functional outcomes. The deep neuroimaging model also demonstrated significant improvement over traditional prognostic scores. In an interpretability analysis, the deep neuroimaging features displayed a significant correlation with age, National Institutes of Health Stroke Scale score, infarct volume, and inflammation factors.

Conclusions: Deep learning models can successfully extract objective neuroimaging features from the routine radiological data in an automatic manner and aid in predicting the functional outcomes in patients with brainstem infarction at 3 months with very high accuracy.

KEYWORDS
brainstem infarction, CNN, deep learning, neuroimaging, prognosis

INTRODUCTION

Brainstem infarctions constitute 10% of all acute ischemic strokes [1]. Because many vital body functions are controlled by the brainstem, despite their relatively small lesion load, infarcts in this region hold the potential to cause significant neurological deficits and mortality [1–3]. The prediction of long-term functional outcomes in patients with brainstem infarction is important in the management of poststroke care. This can enable clinicians to select appropriate therapeutic strategies, as well as to manage prognostic expectations and support counseling for patients and their families [4].

Previous studies have developed and validated several prognostic scores for predicting outcomes in patients with large vessel occlusion, such as the Acute Stroke Registry and Analysis of Lausanne (ASTRAL) [5,6] and the Houston Intra-arterial Therapy (HIAT) score [7]. However, there is a lack of effective methods for predicting outcomes specifically for patients with brainstem infarctions. Additionally, most published prognostic scores rely only on limited clinical features and laboratory results to predict outcomes of patients with ischemic stroke [4,8]. Recently, a few studies have incorporated more complex topological and morphological features of magnetic resonance (MR) images, which may be more indicative of the underlying pathology of the ischemic stroke [9,10].

Diffusion-weighted imaging (DWI) has high sensitivity and specificity in detecting ischemic lesions, and it can facilitate diagnosis and provide important information associated with prognosis for brainstem infarction [11]. In previous imaging studies, the number and distribution of DWI lesions have been shown to offer vital clues to the underlying stroke mechanisms and prognosis of patients with infratentorial strokes [12,13].

In the era of big data, machine learning algorithms have proven to be useful for medical image processing and analysis. Recent research has shown that machine learning-based prognostic models outperform traditional methods for predicting stroke outcomes [14,15]. In particular, deep learning approaches such as convolutional neural networks (CNNs) have been shown to be useful in the extraction of relevant neuroimaging features from medical imaging data [16]. In this study, we employ deep learning algorithms to predict 3-month functional outcomes in patients with brainstem infarction using clinical and neuroimaging data.

MATERIALS AND METHODS

Study design and participants

This study was conducted based on the Third China National Stroke Registry (CNSR-III) study. CNSR-III is a nationwide, multicenter, prospective registry study containing clinical and neuroimaging data of 15,166 patients with acute ischemic stroke and transient ischemic attacks (TIAs) who presented to hospitals between August 2015 and March 2018 at 201 centers across China. More details concerning the design and data collection protocol of CNSR-III have been published previously [17]. This study was approved by the institutional review board at Beijing Tiantan Hospital, Capital Medical University. Written informed consent was obtained from all participants in CNSR-III.

From the CNSR-III dataset, patients who met the following criteria were included in the first stage of analysis: age ≥ 18 years, acute ischemic stroke within 7 days after presentation, and availability of both DWI (b = 1000 s/mm^2) and apparent diffusion coefficient (ADC) scans. The studies without sufficient quality of DWI or ADC scans (n = 2568), patients with prestroke modified Rankin Scale (mRS) score > 2 (n = 532), those diagnosed with TIAs (n = 801), and those with missing follow-up data (n = 130) were excluded from this analysis. The patients identified with isolated brainstem infarction were included in the analysis (n = 1482). These patients were divided into a training dataset (admitted before August 2017) and a test dataset (admitted after August 2017) based on the date of admission (Supplemental Figure S1). The division of the data based on date of admission was motivated by the TRIPOD guidelines (analysis type 2b) [18].

Clinical and laboratory characteristics

Clinical characteristics, including age, gender, prestroke mRS score, body mass index, and history of prior stroke or TIA, myocardial infarction, hypertension, diabetes mellitus, dyslipidemia, or atrial fibrillation, were obtained through in-person interviews by research coordinators. Stroke severity at admission was assessed using the National Institutes of Health Stroke Scale (NIHSS) score. Various laboratory results including platelet and neutrophil counts were collected. Blood samples were collected in the participating hospitals on the day of enrollment, and transported to and stored at −80°C in the central laboratory. The concentration of interleukin-6 (IL-6) was determined by using an enzyme-linked immunosorbent assay kit (PHS600C), and the concentration of high-sensitivity C-reactive protein (hsCRP) was detected on a Roche Cobas C701 analyzer. All measurements were performed by laboratory staff who were blinded to the outcomes of the patients.

MR imaging

Brain MR images were acquired on either 1.5-T or 3-T scanners in the acute phase of the stroke, and the median time from event onset to MR imaging (MRI) was 2 days (interquartile range [IQR] = 1–4 days). Depending on the clinical requirements, one or more of the following MRI sequences were acquired for each patient: T1-weighted, T2-weighted, fluid-attenuated inversion recovery, DWI, and ADC. All cases were reviewed for the availability of DWI with a b-value of 1000 s/mm^2 and ADC scans.

Outcome measures

The participants' functional outcome at 3 months poststroke, measured using the mRS, was the primary outcome of this study. The follow-up assessment was performed by trained interviewers in the form of a face-to-face interview. The mRS score, ranging from 0 (no symptoms) to 6 (death), was assigned based on a standardized interview protocol. To obtain clinically insightful observations, as done in previous studies [19], the mRS score was dichotomized as indicative of a favorable outcome (mRS scores of 0–2) or an unfavorable outcome (mRS score of 3 or higher).

Deep neuroimaging feature extraction

For the identification of patients with brainstem infarction, an automatic, deep, infarction lesion segmentation model based on a U-Net architecture [20] was developed. To understand the effect of lesion characteristics on the functional outcomes 3 months poststroke, extraction of abstract, compact, and meaningful neuroimaging features was essential. Therefore, to automatically extract relevant neuroimaging features, a deep three-dimensional Residual Neural Network (ResNet) [21] was trained using the DWI scans and predicted an ischemic lesion mask registered to the MNI152 T1 nonlinear asymmetric 1 mm isotropic template. More details are shown in the Supplemental Methods (Figures S1 and S2).

Next, to reduce the dimensionality of deep neuroimaging features, a principal component analysis [22] was performed on the 1536 neuroimaging features extracted from the patients who did not have infarctions in the brainstem region and the first 14 principal components (total explained variance = 80%), and the projected features associated with them were extracted as 14 CNN features (Figure 1a).

In this manner, the trained deep feature extraction models were used to extract the neuroimaging features for patients with brainstem infarction, and these features were projected along the first 14 precomputed principal components to calculate the CNN Features 1–14, which were then used as deep neuroimaging features in the clinical prediction models.

Furthermore, to verify that the feature extraction model focused on the lesioned areas of the brain to extract the deep neuroimaging features, the significant regions in the DWI and ADC scans that the model used as the main focuses for the prediction of functional outcomes were visualized by gradient-weighted class activation mapping (Grad-CAM) technique [23]. Lastly, to understand the relationship between the deep neuroimaging features (CNN Features 1–14) and clinical and laboratory variables, a correlation analysis was conducted.

Functional outcome prediction model for patients with brainstem infarction

Following the extraction of deep neuroimaging features, to predict the 3-month functional outcomes in patients with isolated brainstem infarction, a fully connected neural network model was trained using the training dataset. For inclusion in the prediction model, the baseline clinical and laboratory features were first assessed for missing data, and features with >20% missing values were excluded from the analysis (28.3% missed IL-6, and 29.8% missed hsCRP). For the remaining features, the missing values were replaced with the dataset average value. Next, a univariate analysis assessing between-group

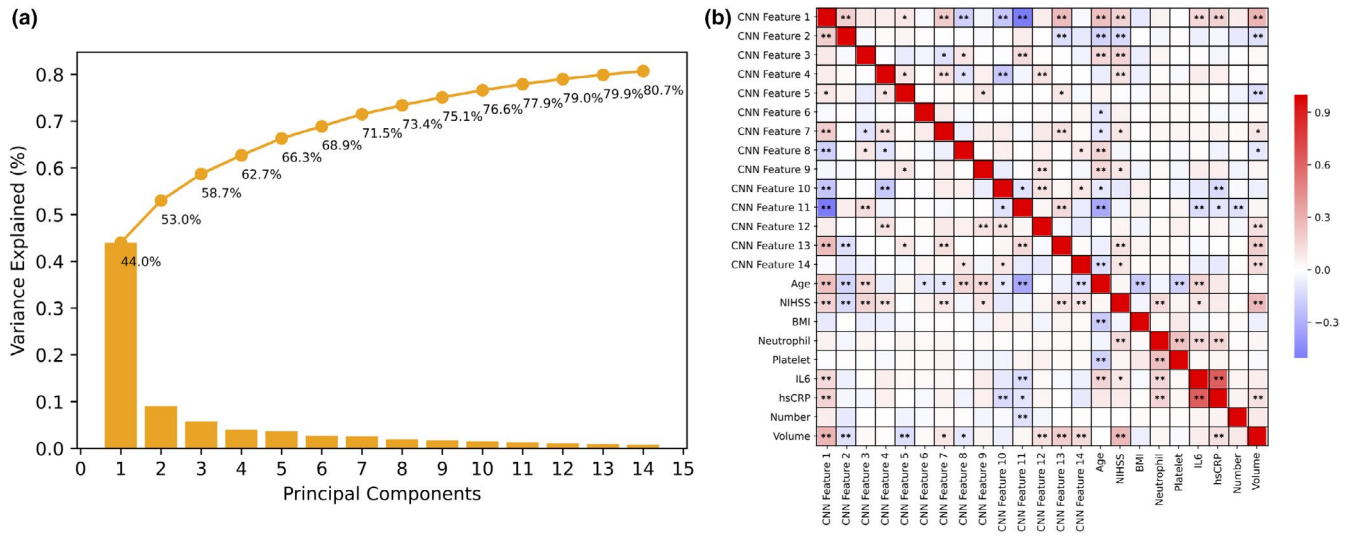

FIGURE 1 The analysis of convolutional neural network (CNN) features. (a) The principal components analyses showed that 14 main components explained 80% of the total variance. (b) The correlations of CNN features with clinical and conventional imaging features. $^{*}p < 0.001$, $^{**}p < 0.0001$. BMI, body mass index; hsCRP, high-sensitivity C-reactive protein; IL6, interleukin-6; NIHSS, National Institutes of Health Stroke Scale

differences was conducted and all variables with p-values < 0.001 were included in the predictive analysis.

To assess the predictive capabilities of various features, five distinct predictive models were trained: Model A, with only baseline clinical features (age, NIHSS); Model B, with clinical and laboratory features (neutrophil count); Model C, with clinical, laboratory, and conventional imaging features (infarct volume, number); Model D, with only the deep neuroimaging features; and Model E, with all features.

Other than the distinct input, all these models shared the same architecture, with five hidden layers (Supplemental Figure S2D), and they were trained identically using focal loss and Adam optimizer. These models accepted a feature vector from a patient as input and provided a probability of the patient of having either favorable or unfavorable functional outcome at 3 months poststroke. The training was performed using the training dataset, it was stopped when the loss on the validation set (20% of the training dataset) did not decrease consecutively for 60 epochs, and the model with the best validation set loss was selected as a final model for testing on the independent test set.

Lastly, to understand the contribution of each of the clinical features in the functional outcome prediction, the trained model was analyzed using the method of Shapley Additive Explanations (SHAP) [24].

Performance evaluation

The performance of the proposed clinical prediction models was assessed using an independent test dataset with the area under the receiver operating characteristic curve (AUC) as a primary performance metric. Furthermore, to assess the effectiveness, the proposed models were also compared against traditional risk scores, namely the HIAT score [7], the stroke prognostication using age and NIHSS (SPAN-100) score [25], the ASTRAL score [5], and the Totaled Health Risks in Vascular Events (THRIVE) score [26]. Lastly, subgroup analyses were performed based on age, NIHSS score, location, Trial of Org 10172 in Acute Stroke Treatment (TOAST) classification, and status of receiving reperfusion therapy.

Statistical analysis

The summary of clinical variables is presented as mean ± SD or median (IQR). Univariate comparisons were done with Wilcoxon rank-sum test for continuous variables and with Fisher exact test for categorical variables. The correlation between the imaging and clinical variables was assessed using Spearman rank correlation. The confidence intervals (CIs) for the AUC were computed using the bootstrapping method with 2000 permutations. The statistical difference between the AUC of various models was assessed using the DeLong test. A two-sided $p < 0.05$ was used as the criterion for statistical significance.

RESULTS

Automatic ischemic lesion segmentation and baseline characteristics

The automatic ischemic lesion segmentation model achieved an overall median Dice score of 0.910 (0.839–0.935) on the segmentation performance evaluation dataset, indicating a good lesion identification performance. Figure 2a,b presents examples of lesion segmentation using this model. Using the trained lesion segmentation model, the ischemic lesion masks for all the patients were predicted. In this manner, in the segmentation test set, the model identified the brainstem infarction patients with an accuracy, sensitivity, and specificity of 97.39%, 78.85%, and 99.99%. The high specificity of the model restricted the inclusion of non-brainstem infarction patients in the predictive analysis.

In the complete CNSR-III cohort, 1482 patients were identified as having an isolated brainstem infarction. These patients were included in the clinical outcome prediction analysis, and among them, 1300 (87.7%) patients had favorable outcomes, whereas 182 (12.3%) had unfavorable outcomes at 3 months. The mean volume of infarctions was 0.6 ± 0.5 ml for the patients with favorable outcomes and 0.9 ± 0.7 ml for the patients with unfavorable outcomes. The baseline demographics and clinical characteristics of these patients are summarized in Table 1. Based on the study design, there were 1169 (78.9%) and 313 (21.1%) isolated brainstem infarction patients in the functional outcome prediction model's training set and independent test set, respectively.

Performance of the prognostic models

In the univariate analysis on the training dataset, age, NIHSS at admission, and neutrophil count were observed to be significantly different between the patient groups ($p < 0.001$) and were included in the functional outcome prediction model. The performance of the proposed outcome model (with and without incorporation of these clinical and laboratory features) and the performance of existing models were evaluated and compared using the independent test dataset (Figure 3a). Model A, which used only the clinical features (age and NIHSS at admission) achieved the test set AUC, sensitivity, and specificity of 0.748 (95% CI, 0.667–0.823), 0.701, and 0.692, respectively. Model B, which added neutrophil count to the clinical features, resulted in an AUC of 0.764 (95% CI = 0.684–0.838). Model C, which added conventional imaging features (infarct volume, number) along with age, NIHSS, and neutrophil count, resulted in an AUC of 0.772 (95% CI = 0.691–0.847). Model D, which only used deep neuroimaging features, achieved the best test AUC of 0.975 (95% CI = 0.934–0.997), and it was significantly better than Models A, B, and C (all $p < 0.05$). Also, the sensitivity and specificity of Model D were 0.985 and 0.872, respectively. Finally, Model E, which included all the features, achieved an AUC of 0.966 (95% CI = 0.930–0.989) in predicting functional outcomes and was not significantly better than

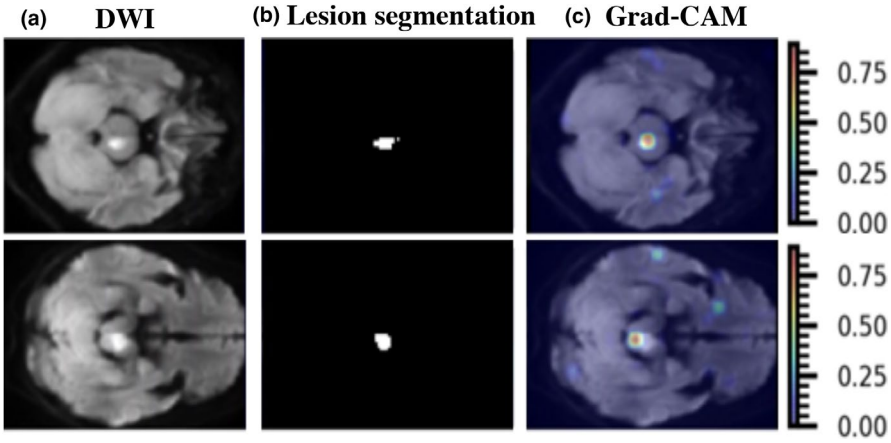

FIGURE 2 Examples of gradient-weighted class activation mapping (Grad-CAM). (a) Acute brainstem ischemic stroke on diffusion-weighted imaging (DWI). (b) Examples of automatic lesion segmentation. (c) The Grad-CAM heatmaps. Regions with high Grad-CAM values were the areas that significantly affect the classification results. The heatmaps were concentrated in the regions of the ischemic core

Model D ($p = 0.577$). Furthermore, Models D and E achieved significantly higher AUCs than all existing prognostic algorithms in this cohort of brainstem infarction patients (AUC [95% CI]; SPAN-100, 0.500 [0.500–0.500]; HIAT, 0.530 [0.460–0.616]; ASTRAL, 0.737 [0.650–0.816]; THRIVE, 0.620 [0.528–0.690]).

Finally, the performance of Model D was subsequently evaluated in selected patient subgroups, including age less than or greater than 65 years, NIHSS less than or greater than 3, location of infarct, whether the patient received any form of reperfusion therapy, and the TOAST class (Table 2). The models showed high prediction accuracy in predicting functional outcomes in each subgroup.

The actual distribution of mRS score at 3 months among patients predicted to have favorable or unfavorable outcome is summarized in Figure 4; 98.8% of patients with mRS 0–2 were predicted to have a favorable outcome, and 75% of patients with mRS 3–6 were predicted to have an unfavorable outcome.

Explanation of deep learning features

A representative Grad-CAM analysis of the deep neuroimaging feature extraction model is presented in Figure 2c. As illustrated, the feature extraction model focuses on the ischemic core lesion area to predict functional outcomes, indicating an extraction of highly relevant neuroimaging features.

The results of the correlation analysis between clinical parameters with the deep neuroimaging features are presented in Figure 1b. Here, CNN Feature 1 showed correlations with age ($r = 0.238$), ischemic lesion volume ($r = 0.287$), NIHSS ($r = 0.167$), IL-6 ($r = 0.148$), and hsCRP ($r = 0.172$). CNN Feature 2 was correlated with age ($r = 0.160$), NIHSS ($r = 0.134$), and infarct lesion volume ($r = 0.112$). CNN Feature 3 was correlated with age ($r = 0.144$) and NIHSS ($r = 0.165$). CNN Feature 11 showed correlations with age ($r = -0.329$) and IL-6 ($r = -0.122$; all $p < 0.0001$).

Lastly, the SHAP analysis of Model E revealed, as shown in Figure 3b, that CNN Feature 3, NIHSS, CNN Feature 2, and CNN Feature 1 were the most important features of the predictive model, followed by number of infarct lesions, other CNN features, age, infarct volume, and neutrophil count.

DISCUSSION

Acute brainstem infarction can result in severe neurological functional deficits, yet outcome prediction is challenging. In this large-scale multicenter cohort of 15,166 patients, we identified 1482 patients with acute brainstem infarction and developed a novel prediction model using clinical, laboratory, and radiological data that are available in routine clinical practice at the time of presentation. Our results showed that deep neuroimaging features automatically learned from a CNN had excellent performance in predicting 3-month functional outcomes (quantified by mRS) in patients with brainstem infarction.

We first developed a predictive model based only on age and NIHSS score, the features that are available immediately at admission (Model A). With the addition of laboratory test results, as well as conventional imaging features extracted on the basis of expert knowledge, the prediction models showed higher prediction accuracy. Most notably, the deep neuroimaging features extracted from DWI were able to achieve an even greater level of accuracy, even without the addition of other clinical or laboratory features. These models, developed based on different types of features, have the potential to play an important role in the clinical workup and care of these patients. The prediction models can provide important prognostic information for clinicians to better treat and manage patients with acute brainstem infarction. A high possibility of unfavorable outcome prompts more rehabilitative evaluation and treatment.

Recently, an increasing number of studies have focused on the predictive capabilities of neuroimaging features [16,27,28]. The traditional prediction models, such as THRIVE and SPAN-100, that do not incorporate many important clinical features, explained only a small proportion of variance in functional outcomes. In contrast, deep learning algorithms have shown excellent performance for solving medical imaging interpretation tasks, including identifying head computed tomography scan abnormalities and classifying brain tumors on histologic images, with accuracy comparable to that of experienced physician adjudicators [29,30]. As a case in point, the currently proposed deep learning model could significantly improve the accuracy of functional outcome prediction after acute brainstem ischemic stroke.

In the present study, most of the ischemic lesions were located in the pons, in agreement with the analysis of Lin et al. [31].

TABLE 1 Baseline characteristics

Characteristic	Train, $n = 1169$			Test, $n = 313$		
	Unfavorable outcome, $n = 143$	Favorable outcome, $n = 1026$	p	Unfavorable outcome, $n = 39$	Favorable outcome, $n = 274$	p
Age, years, median (IQR)	68 (60–75)	62 (55–69)	<0.001	64 (61–73)	62 (55–68)	0.014
Gender, male, *n* (%)	82 (57.3)	683 (66.6)	0.031	20 (51.3)	177 (64.6)	0.114
BMI, mean ± SD	24.4 ± 3.3	25.1 ± 3.1	0.012	24.5 ± 3.7	25 ± 3.1	0.147
Medical history, *n* (%)						
Current smoking	35 (24.5)	313 (30.5)	0.144	7 (17.9)	62 (22.6)	0.68
Regular drinking	11 (7.7)	139 (13.5)	0.061	3 (7.7)	26 (9.5)	1
Prior stroke	39 (27.3)	198 (19.3)	0.034	9 (23.1)	56 (20.4)	0.677
TIA	2 (1.4)	9 (0.9)	0.634	0 (0)	2 (0.7)	1
Coronary artery disease	13 (9.1)	87 (8.5)	0.751	7 (17.9)	21 (7.7)	0.064
Hypertension	105 (73.4)	744 (72.5)	0.92	27 (69.2)	190 (69.3)	1
Diabetes	54 (37.8)	335 (32.7)	0.256	15 (38.5)	87 (31.8)	0.466
Dyslipidemia	17 (11.9)	99 (9.6)	0.374	2 (5.1)	12 (4.4)	0.689
Atrial fibrillation	4 (2.8)	22 (2.1)	0.548	3 (7.7)	8 (2.9)	0.145
Prestroke mRS, median (IQR)	0 (0–0)	0 (0–0)	0.417	0 (0–0)	0 (0–0)	0.129
NIHSS, median (IQR)	7 (4–10)	3 (2–5)	<0.001	6 (5–8)	3 (2–6)	< 0.001
Reperfusion therapy, *n* (%)	6 (4.2)	50 (4.9)	0.837	1 (2.6)	19 (6.9)	0.487
Laboratory measurements, median (IQR)						
Neutrophil, ×10^9/L	4.9 (4–6.3)	4.5 (3.5–5.6)	<0.001	5.2 (3.7–6.8)	4.4 (3.5–5.5)	0.023
Platelet, ×10^9/L	213 (179–255)	212 (179–251)	0.478	181 (160–234)	211 (178–255)	0.019
IL–6, pg/ml	3.5 (2–7)	2.2 (1.4–3.8)	<0.001	2.9 (1.3–4.5)	2.3 (1.4–4.6)	0.275
hsCRP, mg/L	3 (0.8–5.9)	1.2 (0.7–3.2)	<0.001	1.7 (1.1–4.1)	1.3 (0.7–3.5)	0.102
Imaging features						
Number of infarctions, median (IQR)	1 (1–2)	1 (1–1)	< 0.001	1 (1–2)	1 (1–1)	0.077
Infarct volume, mL, mean ± SD	0.9 ± 0.7	0.6 ± 0.5	<0.001	0.8 ± 0.7	0.6 ± 0.5	0.044
Location, *n* (%)			0.065			0.405
Mesencephalon	9 (6.3)	55 (5.4)		2 (5.1)	15 (5.5)	
Pons	118 (82.5)	804 (78.4)		33 (84.6)	225 (82.1)	
Medulla	1 (0.7)	14 (1.4)		0 (0)	4 (1.5)	
Overlapping	15 (10.5)	153 (14.9)		4 (10.3)	30 (10.9)	
TOAST, *n* (%)			0.270			0.163
LAA	29 (20.3)	132 (12.9)		10 (25.6)	39 (14.2)	
CE	4 (2.8)	36 (3.5)		1 (2.6)	8 (2.9)	
SAO	42 (29.4)	457 (44.5)		15 (38.5)	130 (47.4)	
OE	2 (1.4)	18 (1.8)		0 (0)	2 (0.7)	
UE	66 (46.2)	383 (37.3)		13 (33.3)	95 (34.7)	

Abbreviations: BMI, body mass index; CE, cardioembolism; hsCRP, high-sensitivity C-reactive protein; IL-6, interleukin-6; IQR, interquartile range; LAA, large artery atherosclerosis; mRS, modified Rankin Scale; NIHSS, National Institutes of Health Stroke Scale; OE, other determined etiology; SAO, small artery occlusion; TIA, transient ischemic attacks; TOAST, Trial of Org 10172 in Acute Stroke Treatment; UE, undetermined etiology.

According to autoradiographic tracer and lesion studies, the lesion responsible for motor hemiparesis is frequently presumed to be located in the pons, because most ischemic strokes happen following occlusion of the anteromedial pontine artery [32]. Previous imaging studies have also shown that brainstem lesions lead to corticospinal tract dysfunction in patients with motor hemiparesis. The integrity of the corticospinal tract as classified by diffusion tensor tractography (DTT) at the early stage of a pontine infarct has proved to have

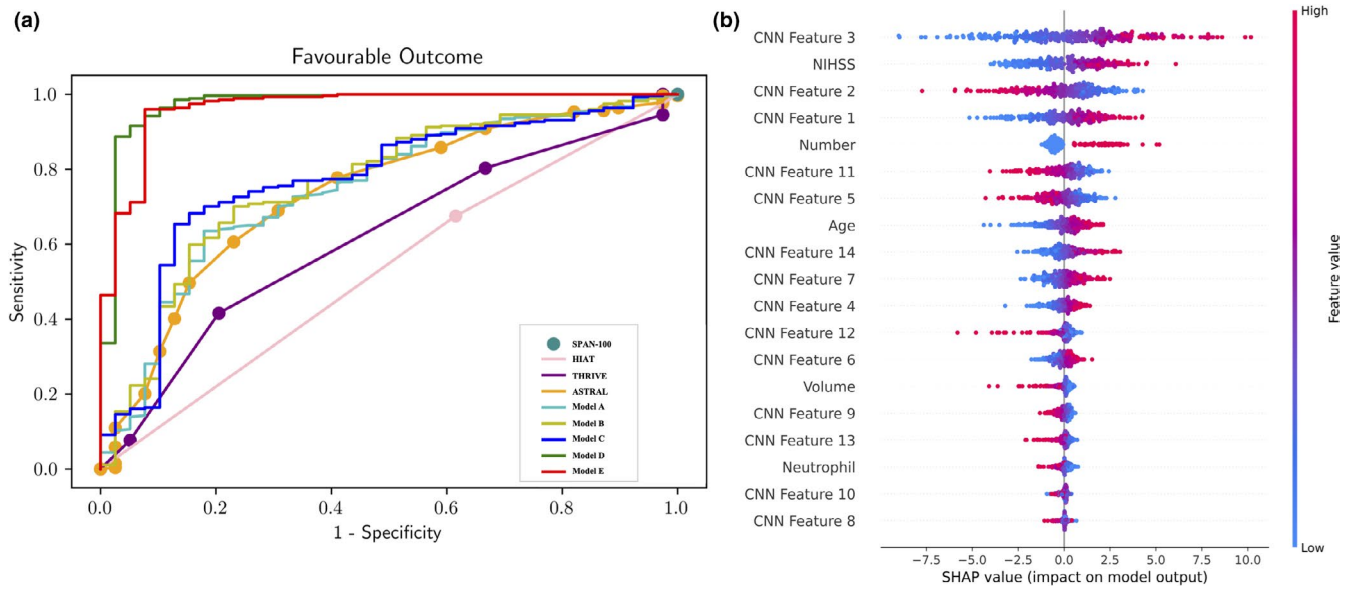

FIGURE 3 (a) The receiver operating characteristic curves of prediction models: Model A, with only baseline clinical features (age, National Institutes of Health Stroke Scale [NIHSS]); Model B, with clinical and laboratory features (neutrophil count); Model C, with clinical, laboratory, and conventional imaging features (infarct volume, number); Model D, with only the deep neuroimaging features; and Model E with all features. (b) Shapley Additive Explanations (SHAP) values-based interpretation of the model, showing importance of the contributing features. The blue and red points in each row represent participants having low to high values for the specific variable. ASTRAL, Acute Stroke Registry and Analysis of Lausanne score; CNN, convolutional neural network; HIAT, Houston Intra-arterial Recanalization Therapy score; SPAN-100, stroke prognostication using age and NIHSS; THRIVE, Totaled Health Risks in Vascular Events score

predictive value for motor function outcome [33]. However, DTT cannot be widely applied in routine clinical practice.

In the current study, we used deep learning-based feature extraction to increase the discriminative power of features for subsequent classification tasks. The data-driven predictive models were superior to the traditional stroke prognostic scores in predicting clinical outcomes. Deep learning algorithms can automatically learn complex data representations and hence avoid the need for prior, potentially arbitrary and biased, definitions by human experts. Deep learning is also robust in protecting against undesired variation, such as interobserver variability, and has the advantages of being applicable to a large variety of clinical conditions and parameters. Dimensionality reduction is frequently used for analyzing and gaining useful insights from high-dimensional data [22,34]. We selected 14 deep learning-based imaging features and subsequently trained classifiers using these features.

Furthermore, we tried to relate and explain deep learning neuroimaging features with respect to traditional quantitative features. Infarct volume, age, and inflammation have typically been identified as risk factors in stroke prognostication. Although not routinely tested, heightened peripheral inflammatory response, including IL-6 and hsCRP concentrations, have been associated with worse functional outcome [5,35,36]. In this study, we found that the CNN features were most strongly associated with infarct volume, age, hsCRP, and IL-6. Thus, CNN features may have the potential to indicate the underlying pathology of the ischemic stroke.

Previous brain MRI studies have shown that brain imaging data can be used to establish the "biological age" of one's brain [37]. During aging, the brain exhibits synapse dysfunction and degeneration, impaired adaptive neuroplasticity and resilience, and aberrant

TABLE 2 AUC for the subgroups

Characteristic	*n* (%)	AUC (95% CI)
Age		
<65 years	185	0.971 (0.900–1.000)
≥65 years	128	0.976 (0.947–0.998)
NIHSS		
≤3	146	0.986 (0.964–1.000)
>3	167	0.955 (0.896–0.991)
Location		
Midbrain	17	1.000 (1.000–1.000)
Pons	258	0.978 (0.932–1.000)
Medulla	4	–
Mix	34	0.938 (0.850–0.984)
Reperfusion therapy		
No	293	0.977 (0.937–0.998)
Yes	20	1.000 (1.000–1.000)
TOAST		
LAA	49	1.000 (1.000–1.000)
CE	9	1.000 (1.000–1.000)
SAO	145	0.984 (0.966–0.998)
Others	2	–
UE	108	0.941 (0.810–1.000)

Abbreviations: AUC, area under the receiver operating characteristic curve; CE, cardioembolism; CI, confidence interval; LAA, large artery atherosclerosis; NIHSS, National Institutes of Health Stroke Scale; SAO, small artery occlusion; TOAST, Trial of Org 10172 in Acute Stroke Treatment; UE, undetermined etiology.

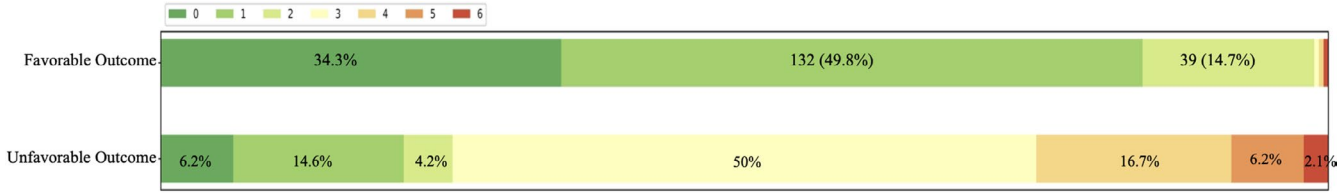

FIGURE 4 The distribution of modified Rankin Scale scores at 3 months for the patients predicted to have favorable or unfavorable outcome

neuronal network activity [38,39]. Brain aging manifests as a decline in the brain's functional network architecture, affecting cognitive performance, executive function, and motor functions. In the aged brain, aberrant activation of microglia likely contributes to the expression of an inducible form of nitric oxide synthase, producing nitric oxide and proinflammatory cytokines, leading to increased CRP levels and IL-6, thus resulting in damage to neurons and functional impairment [40–42]. These alterations render the brain vulnerable to stroke, and neurons suffer compromised ability to withstand and recover from the ischemic stress [43,44]. Interventions that target inflammation and the immune system have been proved to be effective in improving functional outcomes in stroke patients [45].

The current results need to be considered in light of several limitations in this study. First, although the U-Net segmentation model presented high sensitivity and specificity in detecting infarct lesions, it is possible that small infarcts may have been missed. Second, the poor functional outcome can also be a consequence of stroke recurrence. We did not exclude patients who suffered a recurrent stroke during the study, considering that stroke recurrence and functional outcome may share the same risk factors. Removal of these patients might cause selection bias and make the developed prediction model unsuitable for most patients. Third, the prediction model may not fit all cohorts because of differences in MRI scanners and scanning protocols, health care systems, and patient treatments. Accordingly, the model's generalizability should be externally verified by independent institutions. The impact of applying the prediction model as a tool to guide clinical decision-making should be explored in further prospective studies.

In conclusion, the highlight of this work is the prognostic usefulness of high-level neuroimaging features extracted from CNNs in predicting 3-month functional outcomes. Our results suggest that the use of radiological data available in routine clinical practice has the potential to make an accurate prediction of functional outcome in patients with acute brainstem infarction, which may have important clinical value in early evaluation of prognosis and guiding therapeutic decision-making. The presented system can be integrated into clinical workflows and can aid clinicians to effectively manage patients with acute brainstem infarction.

CONFLICT OF INTEREST

None.

AUTHOR CONTRIBUTIONS

Lingling Ding: Conceptualization (equal), investigation (equal), methodology (equal), writing–original draft (lead). **Ziyang Liu:** Investigation (equal), methodology (equal). **Ravikiran Mane:** Methodology (equal), writing–review & editing (equal). **Shuai Wang:** Investigation (equal), methodology (equal). **Jing Jing:** Data curation (equal), project administration (equal), writing–review & editing (equal). **He Fu:** Formal analysis (equal), methodology (equal). **Zhenzhou Wu:** Conceptualization (equal), methodology (equal), writing–review & editing (equal). **Hao Li:** Conceptualization (equal), investigation (equal), methodology (equal), writing–review & editing (equal). **Yong Jiang:** Data curation (equal), formal analysis (equal), project administration (equal), writing–review & editing (equal). **Xia Meng:** Data curation (equal), investigation (equal), project administration (equal), writing–review & editing. **Xingquan Zhao:** Conceptualization (lead), funding acquisition (equal), writing–review & editing (equal). **Tao Liu:** Supervision (equal), writing–review & editing (equal). **Yongjun Wang:** Conceptualization (equal), funding acquisition (equal), supervision (lead), writing–review & editing (equal). **Zixiao Li:** Conceptualization (equal), funding acquisition (equal), supervision (lead), writing–review & editing (equal).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Lingling Ding https://orcid.org/0000-0003-0114-8157
Jing Jing https://orcid.org/0000-0001-9822-5758
Xingquan Zhao https://orcid.org/0000-0001-8345-5147
Yongjun Wang https://orcid.org/0000-0002-9976-2341

REFERENCES

1. de Mendivil AO, Alcalá-Galiano A, Ochoa M, Salvador E, Millán JM. Brainstem stroke: anatomy, clinical and radiological findings. *Semin Ultrasound CT MR*. 2013;34:131-141.
2. Elvsashagen T, Bahrami S, van der Meer D, et al. The genetic architecture of human brainstem structures and their involvement in common brain disorders. *Nat Commun*. 2020;11:4016.
3. Baran G, Gultekin TO, Baran O, et al. Association between etiology and lesion site in ischemic brainstem infarcts: a retrospective observational study. *Neuropsychiatr Dis Treat*. 2018;14:757-766.
4. Jadhav AP, Desai SM, Panczykowski DM, et al. Predicting outcomes after acute reperfusion therapy for basilar artery occlusion. *Eur J Neurol*. 2020;27(11):2176-2184.
5. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the astral score. *Neurology*. 2012;78:1916-1922.
6. Liu G, Ntaios G, Zheng H, et al. External validation of the astral score to predict 3- and 12-month functional outcome in the china national stroke registry. *Stroke*. 2013;44:1443-1445.

7. Hallevi H, Barreto AD, Liebeskind DS, et al. Identifying patients at high risk for poor outcome after intra-arterial therapy for acute ischemic stroke. *Stroke*. 2009;40:1780-1785.
8. Li H, Kang Z, Qiu W, et al. Hemoglobin a1c is independently associated with severity and prognosis of brainstem infarctions. *J Neurol Sci*. 2012;317:87-91.
9. Burger KM, Tuhrim S, Naidich TP. Brainstem vascular stroke anatomy. *Neuroimaging Clin N Am*. 2005;15(2):297-324.
10. Marx JJ, Iannetti GD, Thomke F, et al. Topodiagnostic implications of hemiataxia: an mri-based brainstem mapping analysis. *NeuroImage*. 2008;39:1625-1632.
11. Khaleel NI, Zghair MAG, Hassan QA. Value of combination of standard axial and thin-section coronal diffusion-weighted imaging in diagnosis of acute brainstem infarction. *Open Access Maced J Med Sci*. 2019;7:2287-2291.
12. Payabvash S, Taleb S, Benson JC, McKinney AM. Acute ischemic stroke infarct topology: association with lesion volume and severity of symptoms at admission and discharge. *AJNR Am J Neuroradiol*. 2017;38:58-63.
13. Jang SH. Motor outcome and motor recovery mechanisms in pontine infarct: a review. *NeuroRehabilitation*. 2012;30:147-152.
14. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine learning-based model for prediction of outcomes in acute stroke. *Stroke*. 2019;50:1263-1265.
15. Nishi H, Oishi N, Ishii A, et al. Predicting clinical outcomes of large vessel occlusion before mechanical thrombectomy using machine learning. *Stroke*. 2019;50:2379-2388.
16. Nishi H, Oishi N, Ishii A, et al. Deep learning-derived high-level neuroimaging features predict clinical outcomes for large vessel occlusion. *Stroke*. 2020;51:1484-1492.
17. Wang Y, Jing J, Meng X, et al. The third china national stroke registry (cnsr-iii) for patients with acute ischaemic stroke or transient ischaemic attack: Design, rationale and baseline patient characteristics. *Stroke Vasc Neurol*. 2019;4:158-164.
18. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (tripod): Explanation and elaboration. *Ann Intern Med*. 2015;162:W1-W73.
19. Badhiwala JH, Nassiri F, Alhazzani W, et al. Endovascular thrombectomy for acute ischemic stroke: a meta-analysis. *JAMA*. 2015;314:1832-1843.
20. Ronneberger Olaf FP, Thomas B. U-net: convolutional networks for biomedical image segmentation. *MICCAI*. 2015:234-241.
21. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016:770-778.
22. Tenenbaum JB, de Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science*. 2000;290:2319-2323.
23. Rs R, Cogswell M, Das A, Vedantam R, Parikh D, Grad-cam BD. Visual explanations from deep networks via gradient-based localization. *Int J Comput Vision*. 2020;128(2):336-359.
24. Lundberg S, Lee S-I. A unified approach to interpreting model predictions. arXiv preprint arXiv:1705.07874. 2017.
25. O'Donnell MJ, Fang J, D'Uva C, et al. The plan score: a bedside prediction rule for death and severe disability following acute ischemic stroke. *Arch Intern Med*. 2012;172:1548-1556.
26. Flint AC, Cullen SP, Faigeles BS, Rao VA. Predicting long-term outcome after endovascular stroke treatment: the totaled health risks in vascular events score. *AJNR Am J Neuroradiol*. 2010;31:1192-1196.
27. Appleton JP, Woodhouse LJ, Adami A, et al. Imaging markers of small vessel disease and brain frailty, and outcomes in acute stroke. *Neurology*. 2020;94:e439-e452.
28. Lee WH, Lim MH, Seo HG, Seong MY, Oh BM, Kim S. Development of a novel prognostic model to predict 6-month swallowing recovery after ischemic stroke. *Stroke*. 2020;51:440-448.
29. Chilamkurthy S, Ghosh R, Tanamala S, et al. Deep learning algorithms for detection of critical findings in head ct scans: a retrospective study. *Lancet*. 2018;392:2388-2396.
30. Titano JJ, Badgeley M, Schefflein J, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. *Nat Med*. 2018;24:1337-1341.
31. Lin Y, Zhang L, Bao J, et al. Risk factors and etiological subtype analysis of brainstem infarctions. *J Neurol Sci*. 2014;338:118-121.
32. Marx JJ, Iannetti GD, Thomke F, et al. Somatotopic organization of the corticospinal tract in the human brainstem: a mri-based mapping analysis. *Ann Neurol*. 2005;57:824-831.
33. Jang SH, Bai D, Son SM, et al. Motor outcome prediction using diffusion tensor tractography in pontine infarct. *Ann Neurol*. 2008;64:460-465.
34. Chakraborty R, Yang L, Hauberg S, Vemuri B. Intrinsic grassmann averages for online linear, robust and nonlinear subspace learning. *IEEE Trans Pattern Anal Mach Intell*. 2021;43(11):3904-3917.
35. Whiteley W, Jackson C, Lewis S, et al. Inflammatory markers and poor outcome after stroke: a prospective cohort study and systematic review of interleukin-6. *PLoS Medicine*. 2009;6:e1000145.
36. Zaidi SF, Aghaebrahim A, Urra X, et al. Final infarct volume is a stronger predictor of outcome than recanalization in patients with proximal middle cerebral artery occlusion treated with endovascular therapy. *Stroke*. 2012;43:3238-3244.
37. Jonsson BA, Bjornsdottir G, Thorgeirsson TE, et al. Brain age prediction using deep learning uncovers associated sequence variants. *Nat Commun*. 2019;10:5409.
38. Christman S, Bermudez C, Hao L, et al. Accelerated brain aging predicts impaired cognitive performance and greater disability in geriatric but not midlife adult depression. *Transl Psychiatry*. 2020;10:317.
39. Scheiblich H, Trombly M, Ramirez A, Heneka MT. Neuroimmune connections in aging and neurodegenerative diseases. *Trends Immunol*. 2020;41:300-312.
40. Norden DM, Godbout JP. Review: microglia of the aged brain: primed to be activated and resistant to regulation. *Neuropathol Appl Neurobiol*. 2013;39:19-34.
41. Cribbs DH, Berchtold NC, Perreau V, et al. Extensive innate immune gene activation accompanies brain aging, increasing vulnerability to cognitive decline and neurodegeneration: a microarray study. *J Neuroinflammation*. 2012;9:179.
42. Corlier F, Hafzalla G, Faskowitz J, et al. Systemic inflammation as a predictor of brain aging: contributions of physical activity, metabolic risk, and genetic risk. *NeuroImage*. 2018;172:118-129.
43. Mattson MP, Arumugam TV. Hallmarks of brain aging: adaptive and pathological modification by metabolic states. *Cell Metab*. 2018;27:1176-1199.
44. Prolla TA, Mattson MP. Molecular mechanisms of brain aging and neurodegenerative disorders: lessons from dietary restriction. *Trends Neurosci*. 2001;24:S21-S31.
45. Zhu Z, Fu Y, Tian D, et al. Combination of the immune modulator fingolimod with alteplase in acute ischemic stroke: a pilot trial.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Ding L, Liu Z, Mane R, et al. Predicting functional outcome in patients with acute brainstem infarction using deep neuroimaging features. *Eur J Neurol*. 2021;00:1–9. doi:10.1111/ene.15181